Energy Conversion and Management 148 (2017) 895–904

Contents lists available at ScienceDirect

Energy Conversion and Management

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

An improved Wavelet Transform using Singular Spectrum Analysis for wind speed forecasting based on Elman Neural Network

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 15 May 2017
Accepted 26 May 2017
Available online 28 June 2017

Keywords:
Wavelet transform
Singular Spectrum Analysis
Elman Neural Network
Wind speed prediction

ABSTRACT

To raise the wind speed prediction accuracy, Wavelet Transform (WT) is widely employed to disaggregate an original wind speed series into several sub series before forecasting. However, the highest frequency sub series usually has a great disturbance on the final prediction. In the study, for raising the forecasting accuracy, Singular Spectrum Analysis (SSA) is applied to make further processing on the highest frequency sub series, instead of making no modification on or getting rid of it. So a hybrid decomposition technology called Improved WT (IWT) is proposed. Meanwhile, a new hybrid model IWT-ENN combined with IWT and Elman Neural Network (ENN) is also designed. The procedure of IWT is systematically investigated. Experimental results show that: (1) the performance of the hybrid model IWT-ENN has a great improvement compared to that of others including the persistence method, ENN, Auto-Regressive (AR) model, Back Propagation Neural Network (BPNN) and Empirical Mode decomposition (EMD)-ENN; (2) compared to the two general strategies where the highest frequency sub series is without retreatment or eliminated, the new proposed hybrid model IWT-ENN has the best prediction performance.

1. Introduction

Wind power, as one of the plentiful and renewable energies, has becoming increasingly competitive and universally used [1–4]. Owing to the wind inherent characteristics of intermittence and fluctuation, to protect the safety of the power utilization and integration, it is essential and significant to make high precision forecasting of wind speed [5,6].

In recent years, there have been many researches about hybrid models merged with decomposition technologies, including Empirical Mode decomposition (EMD) [7,8] and Wavelet Transform (WT) etc., to make predictions on wind speed and power. EMD is a data driven method, without any prior knowledge. But WT is still one of the most popular method, whose process has a more clear physical meaning.

Catalão et al. [9,10] proposed two hybrid models based on WT and ANN, including Back Propagation Neural Network (BPNN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) Neural Network optimized by Particle Swarm Optimization (PSO) to make short-term wind power. The time series was fell part into several sub series by WT and then were individually modeled and predicted. Liu et al. [11] applied WT to dissolve the short-term wind power time series into a sequence of sub series, which were respectively forecasted by Support Vector Machines (SVMs) with the wavelet kernel function and radial basis kernel function for two different hybrid models. Liu et al. [12–14] designed some hybrid models based on Wavelet Transform combined with Multilayer Perception (MLP) Neural Network, ANFIS Neural Network and Extreme Learning Machines respectively to make wind speed prediction. Osório et al. [15] proposed a method involved WT and the ANFIS Neural Network, which was optimized by Evolutionary Particle Swarm Optimization (EPSO) to make wind power prediction. Kiplangat et al. [16] came up with a hybrid model with Auto-Regressive (AR) models combined with WT to make wind speed prediction, whose performance was greater than that of AR and Fractionally Autoregressive Integrated Moving Average Model (f-ARIMA) models proved by some experiments. Tascikaraoglu et al. [17] utilized WT for decomposing the wind speed data into more stationary components and then applied spatio-temporal models on each sub series for prediction. The basic idea of these hybrid models mentioned above is that firstly to decompose a time series into a sequence of sub series by WT, then suitable regression models are adopted to make predictions, all of which are added up as the final prediction. More similarity works can be found in

* 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.2017.05.063
0196-8904/

Refs. [18–21]. It can be seen that the performance of these hybrid models combined with WT are more great than those without WT, since the subseries decomposed by WT behave more stationary with less fluctuations.

However, the subseries with the highest frequency band always causes a big disturbance on the final prediction. Resulting from that, it is eliminated to improve prediction accuracy by some researchers. Tascikaraoglu et al. [22] used WT to cut up the original wind speed series into a number of better-behaved subseries which were separately forecasted by BPNN except for the highest frequency band. Liu et al. [23] came up with a model based on WT and SVM. The original wind series was decomposed into an approximation signal and a detail signal by WT. The detail signal was eliminated while the approximation signal was forecasted by SVM, whose parameters were optimized by Genetic Algorithm (GA). Wang et al. [24] proposed a model combined with WT, Seasonal Autoregressive Integrated Moving Average Model (SARIMA) and BPNN. There were two parts decomposed by WT from the short-term load series. Only the approximation signal was forecasted by SARIMA and BPNN respectively at first, which then are combined by weighting to get the final short-term load predictions.

From the literature reviewed, it can be seen that WT can actually assist on improving the hybrid models' performance for wind speed prediction. However, the decomposed sub series with the highest frequency usually bother the final prediction accuracy. In order to enhance the forecasting capacity, it is crucial and required to minimize the impact of the highest frequency sub series. To further raising the forecasting accuracy, different from the two common strategies including making no modification on or getting rid of the highest frequency sub series mentioned above, a new hybrid decomposition technology called Improved WT (IWT) is proposed, where SSA is combined with WT to make further operation on the highest frequency sub series. What's more, a new hybrid model IWT-ENN is came up with to make wind speed prediction, where ENN is extensively adopted and proven to do well in wind speed prediction [25]. In the text that follows, the procedure of IWT is systematically investigated. Meanwhile, the forecasting performance of IWT-ENN is compared to that of others including the persistence method, ENN, AR, BPNN and EMD-ENN.

The rest of paper is organized as follows: Section 2 reviews decomposition technologies and proposes the new hybrid decomposition method; Section 3 introduces the forecasting method; Section 4 brings up the new hybrid model; Section 5 presents some experimental results and makes comparative studies.

2. Decomposition technology

2.1. Wavelet transform

Using wavelet transform, the original wind speed series is decomposed into a sequence of sub series with better behavior and more predictable. The Continuous Wavelet Transform (CWT) of a signal $X(t)$ can be illustrated as follows:

$$CWT(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} X(t)\psi^*\left(\frac{t-b}{a}\right)dt \tag{1}$$

where a, b is the scale factor and translation parameter respectively, $\psi(t)$ is the mother wavelet and $*$ denotes the complex conjugate. There is a digitally counterpart of CWT, Discrete Wavelet Transform (DWT), with less computation and approximately accuracy. It is defined as follows:

$$DWT(m,n) = 2^{-(m/2)} \sum_{x=0}^{N-1} X(t)\psi^*\left(\frac{t-n2^m}{2^m}\right) \tag{2}$$

where m and n is the scale factor and translation parameter respectively ($a = 2^m$, $b = n2^m$), and N is the length of the signal $X(t)$.

Based on the Mallat's algorithm [26], the multi-resolution process is achieved where the signal $X(t)$ is disaggregated into some sub series involved with " approximation " and "detail " information respectively. A typical DWT process of the signal $X(t)$ with 3 decomposed level is shown in Fig. 1.

In this study, a wavelet function of type Daubechies of order 6 (abbreviated as Db6) is adopted as the mother wavelet $\psi(t)$ [27,28]. Another is worth mentioning, the characteristic of wavelet decomposition is that the more scales the original signal be decomposed, the better performance the decomposed signals have, but great errors will be generated at the same time [21]. The decomposed level of wavelet transform is 3 in the study by default, which is widely used in many studies [9,10,12].

2.2. Singular Spectrum Analysis

SSA is a popular approach for time series analysis, including trend of quasi-periodic component detection and extraction and denoising [29,30]. The process of SSA is combined with two main stages decomposition and reconstruction. It is detailed described as follows.

Stage I. Decomposition
Step a. Embedding

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

$$Y = \begin{bmatrix} X_1 & \cdots & X_k \\ \vdots & \ddots & \vdots \\ X_L & \cdots & X_N \end{bmatrix} \tag{3}$$

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)

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

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

where $d = \text{rank}(Y)$.

Stage II. Reconstruction
Step c. Grouping

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

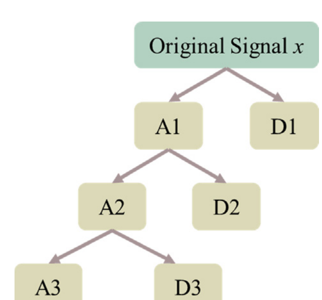


Fig. 1. Wavelet Transform process with 3 decomposed level: A represents the approximate component and D represents the detail component.

Step d. Averaging

It is the Hankelization procedure [30,31] 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:

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

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:

$$X = X_{trend} + X_{noise} \quad (6)$$

2.3. Proposed decomposition method

After the WT decomposition, an original wind speed series is broken down into a sequence of sub series as shown in Fig. 1. The sub series with the highest frequency D1 always has a great disturbance in prediction accuracy. To improve forecasting precision, neither eliminating D1 or make no operation on it, SSA is employed to remove redundant information of D1. So a new decomposition method called IWT is proposed as shown in Fig. 2. There are two main steps:

Step a. utilize WT to decompose an original wind speed series into a sequence of sub series;
Step b. the sub series with the highest frequency is denoised by SSA.

In the SSA process, two vital parameters are what we care about, the window length L and the number r of the selected eigentirples for reconstruction. L is recommended to 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 [32]. About the value of r, in Eq. (3), $\sqrt{\lambda_i}/\sum_{i=1}^{d}\sqrt{\lambda_i}$ reveals the contribution of each eigentirples while the noise components always have very small contribution. Hence, the trend rate (TR) is defined as the total contribution of the r out of d components as follows:

$$TR = \sum_{i=1}^{r}\sqrt{\lambda_i} \Bigg/ \sum_{i=1}^{d}\sqrt{\lambda_i} \quad (7)$$

Through the threshold of TR, the process of SSA is successfully completed with selecting r out of d components.

3. Forecasting method

3.1. Elman Neural Network

Although Multilayer Perceptron neural network (MLP) generally deals with abundant complicated problems, it is limited to only achieve a static mapping of the input space to the output space. ENN, a simple recurrent neural network with dynamic characteristic, was proposed [33]. The structure of ENN is similar to that of a three-layered MLP, except for an addition layer called the context layer involved in the former, as Fig. 3 shows. In ENN, both the input neurons and context neurons activate the hidden neurons, and then the hidden neurons feed forward to activate the output neurons, while they also feed back to activate the context neurons. So the context layer enables ENN adapt to dynamic characteristic.

3.2. Neural Network structure design

The structure of the ENN is critical to the prediction accuracy. Since the overall structure of ENN is defined, only the neuron number in each layer is talked about. Firstly, the number of input neurons is determined by Gradient Boosted Regression Trees (GBRT) [35]. Herein, GBRT is applied to evaluate the importance of input variables measured by one's contribution to the prediction accuracy [36]. For each input variable, the respective importance, denoted as $RI = \{RI_1,\ldots, RI_m\}$ is firstly calculated [37], where m is the whole number of input variables investigated. The biggest RI values of input variables are really different via various time series. So the threshold of cumulative importance (CI) is set to help determine the neuron number. The representation of CI_j is described as follows:

$$CI_j = \sum_{i=1}^{j} RI_j \quad (1 \leqslant j \leqslant m) \quad (8)$$

To avoid "dimensionality curse", the threshold value of CI is set as 90%. Select j out of m input variables where CI_j is just bigger than the threshold. Then, as Ref. [38] indicates, $2n + 1$ hidden neurons are capable of mapping any function for n inputs, which is also adopted in this paper. Finally, there is only one neuron in the output layer.

4. Framework of hybrid model

A new hybrid model IWT-ENN is proposed herein, as Fig. 4 shows. It can been seen that there are about three main steps in the whole forecasting process as the followed.

Step I. An original wind speed series is broken down into four sub series by IWT, including A3, D3, D2 and the trend of D1;
Step II. Each of the sub series will be modeled and forecasted by ENN;
Step III. All the predictions are added up as the final prediction results.

It is noted that the influence of the parameter TR in IWT-ENN will be systematically studied, including three threshold values, 85%, 90% and 95%. What's more, 0% of TR means that D1 is eliminated and 100% of TR stands for that there is no operation on D1, where SSA is not utilized to extract the trend of D1. But for the convenience of description, hybrid models combined with these two strategies after WT are also involved in IWT-ENN.

5. Simulation results and analysis

5.1. Data

There are four actual wind speed series, including S1 and S2 measured in China Sichuan Province, and two more ones S3 and

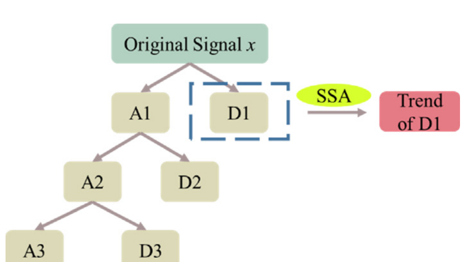


Fig. 2. The proposed decomposition technology (IWT) combined with WT and SSA.

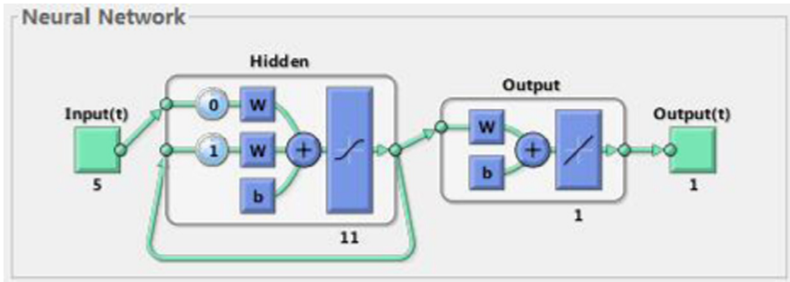


Fig. 3. A typical topology of ENN in Matlab Neural Network Toolbox [34].

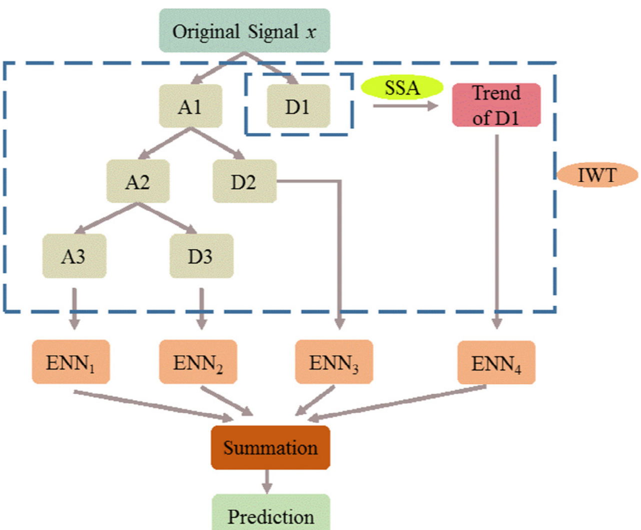


Fig. 4. The proposed prediction method IWT-ENN.

S4 obtained in 43°20′30″N 124°22′30″W as shown in Fig. 5. Their statistical information is listed in Table 1. In each series, the last 168 samples is used to validate the prediction model built through the others.

5.2. Evaluation criteria

Three general error criteria are adopted to evaluate the model performance 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.

5.3. Case study

In this section, the wind speed series S1 will be investigated in details.

5.3.1. Decomposition by IWT

Based on the IWT proposed in the study, S1 is firstly broken down into a sequence of sub series by WT. Two of the sub series A3 and D1 are shown in Fig. 6. Obviously, D1 has the highest frequency, which usually makes a worse impact on the prediction accuracy than other sub series. Since the wind speed almost changes once a day as Fig. 5 shown, so L in IWT is taken as 48, which is enough to map the period of D1. The extracting process of D1 is as Fig. 7 shows with 90% of *TR*.

5.3.2. Prediction model design

As Section 3.2 introduced, GBRT is applied to nail down the number of input neurons in ENN. The *RI* and *CI* of only the first ten input variables evaluated by one's contribution to the

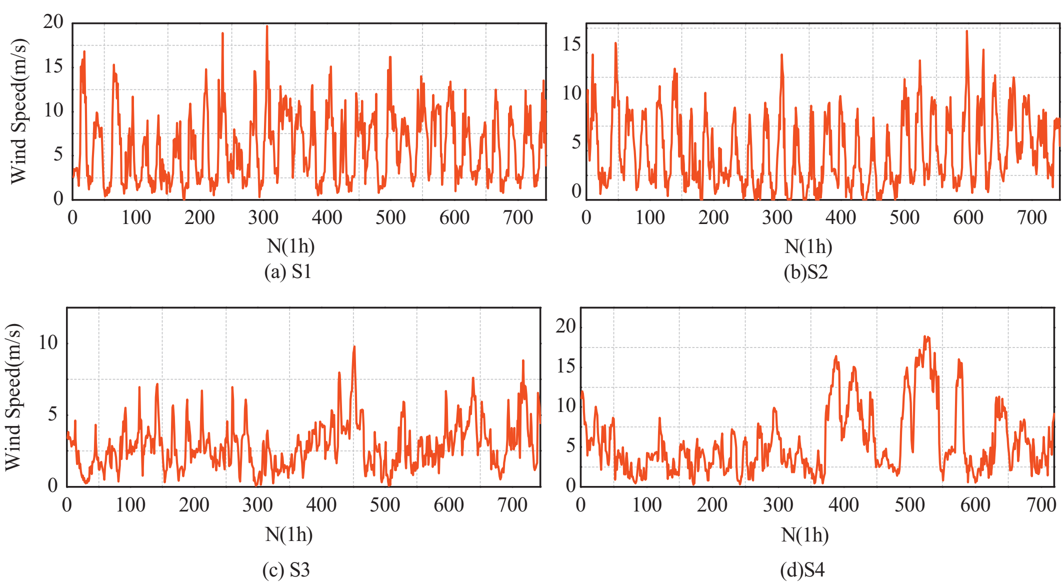


Fig. 5. Four actual wind speed series: S1 ~ S4.

Table 1
Statistical information of S1 ~ S4 (unit: m/s).

Series	Mean	Median	Maximum	Minimum	Standard deviation
S1	5.92	5.5	19.7	0.0	3.81
S2	4.92	3.8	17.2	0.0	3.58
S3	2.86	2.7	9.8	0.1	1.65
S4	5.61	4.3	18.9	0.3	4.19

Table 2
Error criteria adopted in the study.

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

prediction accuracy are calculated. For instance, the input variable importance in A3 and D1 are shown in Fig. 8. As we expected, it can be seen that input variables with smaller lags has greater influence on the prediction. The input neurons number is 5, 2, 7 and 8 for A3, the trend of D1, D2 and D3 respectively, while the trend of D1with different values of *TR* have the same input neurons number after calculation.

5.3.3. Prediction results and analysis

Multistep ahead forecasting results of S1 by IWT-ENN, abbreviated as 1-step, 2-step and 3-step, are shown in Fig. 9. And the error indexes are posted in Table 3. In order to investigate the hybrid models' performances, the persistence method, a popular benchmark model, is also employed to make prediction while a single ENN, AR and BPNN dose. Another hybrid model EMD-ENN combined with EMD and ENN is utilized, too. The error indexes by the persistence method, ENN, AR, BPNN and EMD-ENN are given in Table 4. About the running time of some typical involved forecasting models are shown in Table 5 (the simulation platform is ordinary).

Form Tables 3 and 4, it can be seen that:

(1) the performance of hybrid model IWT-ENN has an enormous boost, even without the help of SSA when *TR* equals 100% or 0%, compared to that of the single models including the persistence method, ENN AR, BPNN and EMD-ENN.

(2) the proposed decomposition method IWT with further processing on D1 by SSA really does help improving the prediction accuracy, while the new hybrid model IWT-ENN has the best prediction performance. About 1-step forecasting, if there is no reprocessing on D1 (*TR* = 100%), the MAE, MAPE and RMSE is 0.4523, 9.84% and 0.6342 respectively. When D1 is eliminated (*TR* = 0%), all the error indexes increase instead. It means that there are some important information involved in D1, which should not be eliminated. However, after the SSA applied on D1, all the error indexes reduce. For example when *TR* equals 90%, the error index is 0.3592, 8.72% and 0.4700 with respect of the MAE, MAPE and RMSE respectively. About multistep forecasting, the performance of the hybrid models combined with IWT also have a substantial increase.

5.3.4. Influence of TR

To investigate how the SSA helps the hybrid model IWT-ENN, the performance between two forecasting models is compared by percentage error criterions, which adopted as follows:

$$P_E = \left|\frac{E_{\mathrm{ben}} - E_{\mathrm{pro}}}{E_{\mathrm{ben}}}\right| \tag{9}$$

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 model and proposed model. IWT-ENN with 100% *TR* is taken as the benchmark model. The P_{MAE}, P_{MAPE} and P_{RMSE} of S1 are got as Fig. 10 shows.

The value of *TR* makes a difference on the prediction performance of the hybrid model IWT-ENN.

For different forecasting steps, the P_{MAE}, P_{MAPE} and P_{RMSE} are highest in 2-step forecasting. But in general, there are some real performance gains for each forecasting step. For different error indexes, most of lines firstly ascend then descend in Fig. 10. It means that the performance of the hybrid model IWT-ENN with 90% of *TR* is the best. Hence, *TR* is recommended as 90%.

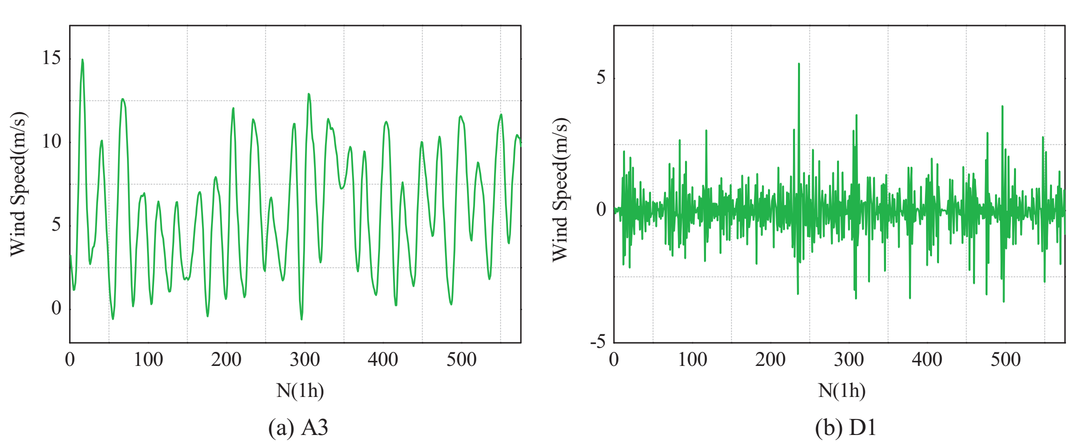
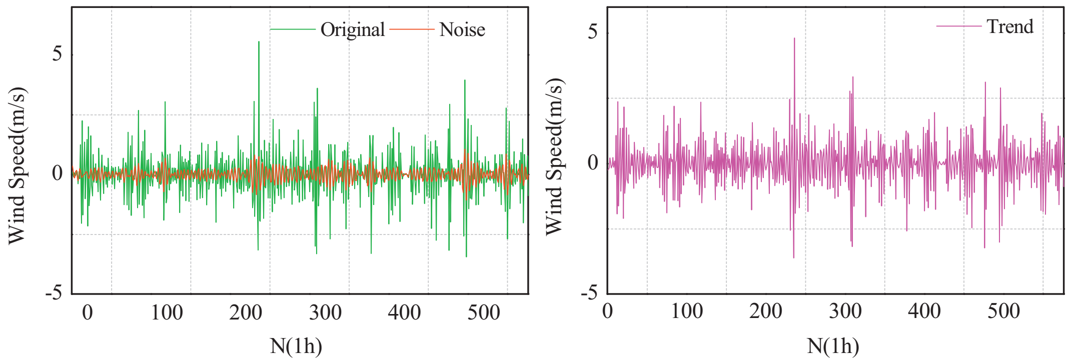
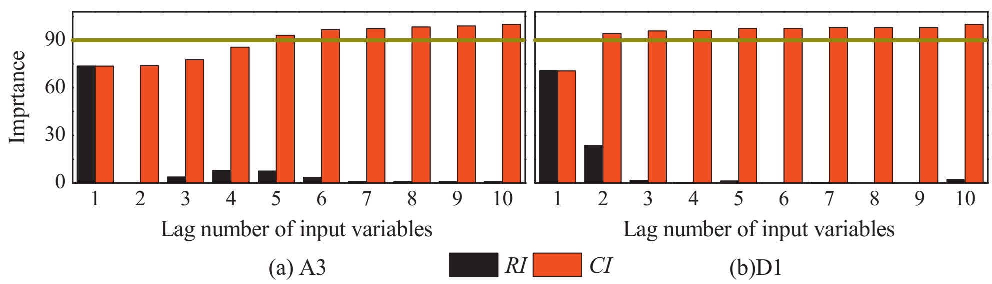


Fig. 6. The WT process operated on S1.



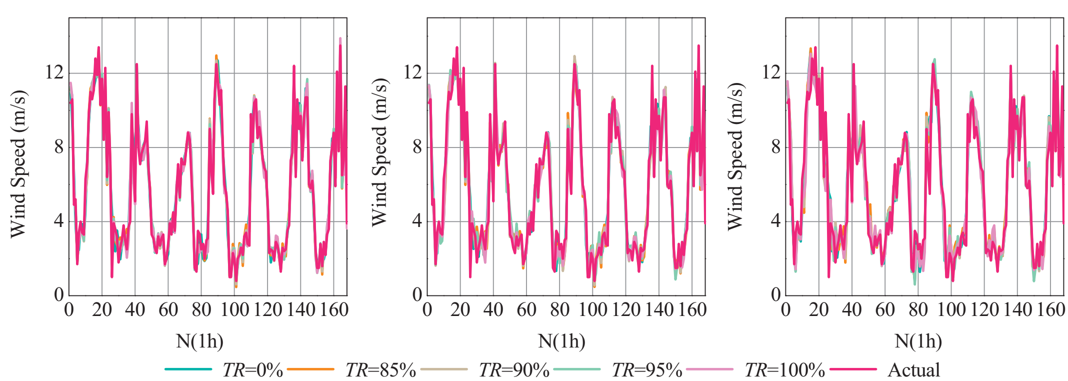
(a) D1 and its noise component (b)The trend component of D1

Fig. 7. The noise and trend component of D1 calculated by SSA (*TR* = 90%).



(a) A3 (b)D1

Fig. 8. Importance of input variables in A3 and D1.



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

Fig. 9. Prediction results of S1 through IWT and ENN.

Table 3
Error indexes of forecasting results of S1 through IWT and ENN.

	Error criteria	*TR* = 0%	*TR* = 85%	*TR* = 90%	*TR* = 95%	*TR* = 100%
1-step	MAE (m/s)	0.8316	0.3586	0.3593	0.4021	0.4523
	MAPE (%)	18.61%	8.92%	8.72%	9.29%	9.84%
	RMSE (m/s)	1.1630	0.4618	0.4700	0.5372	0.6342
2-step	MAE (m/s)	0.8347	0.5515	0.5166	0.6388	0.6942
	MAPE (%)	19.99%	14.53%	12.82%	15.70%	15.81%
	RMSE (m/s)	1.1848	0.7834	0.7260	0.9094	0.9981
3-step	MAE (m/s)	1.0211	0.7932	0.7728	0.8346	0.8982
	MAPE (%)	23.51%	18.67%	17.56%	19.19%	20.61%
	RMSE (m/s)	1.4132	1.1256	1.1047	1.1904	1.2536

Table 4
Error indexes of multistep forecasting results of S1 by the other methods.

	Error criteria	Persistence method	ENN	AR	BPNN	EMD-ENN
1-step	MAE (m/s)	1.6607	1.5664	1.6253	2.0096	0.9233
	MAPE (%)	37.34%	37.44%	41.53%	47.46%	19.67%
	RMSE (m/s)	2.3951	2.1465	2.2289	2.7627	1.2493
2-step	MAE (m/s)	1.7750	1.6483	1.9816	2.2765	1.0954
	MAPE (%)	44.95%	45.07%	57.86%	61.75%	24.48%
	RMSE (m/s)	2.5170	2.1581	2.4728	2.8585	1.4925
3-step	MAE (m/s)	1.9881	1.8269	2.3813	2.5783	1.1546
	MAPE (%)	48.90%	49.23%	71.05%	67.57%	28.06%
	RMSE (m/s)	2.6968	2.3134	2.8606	3.1941	1.5148

Table 5
Running time of the involved forecasting models.

Method	IWT-ENN (*TR* = 90%)	ENN	AR	BPNN	EMD-ENN
Time (s)	11.9	3.8	0.2	4.9	23.7

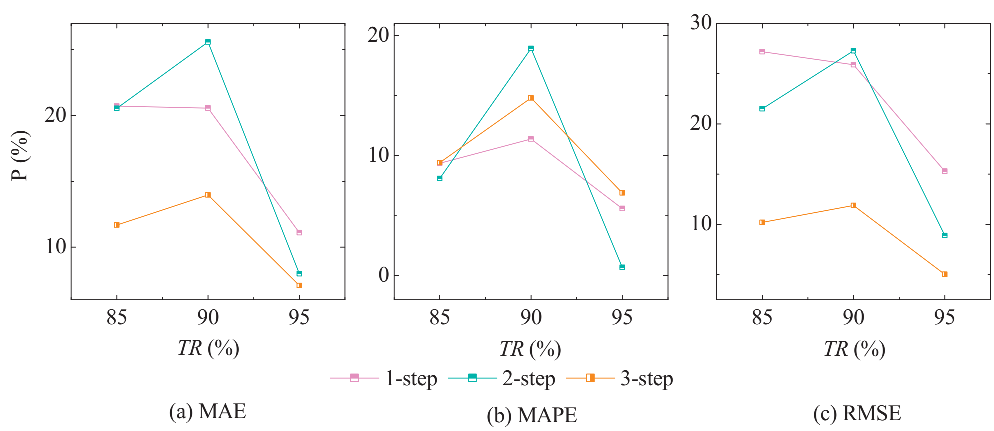


Fig. 10. The influence of *TR* in the proposed methods on P_{MAE} of the $\{X_{1t}\}$.

5.4. Further validation

In order to further investigate the performance of the hybrid model IWT- ENN, other actual wind speed series are under study. The forecasting results of S2 ~ S4 by IWT- ENN, including 1-step, 2-step and 3-step are shown from Figs. 11–13 and the error indexes of them are given in Tables 6–11.

From Tables 6–11, some same conclusions can be drawn:

(1) the performance of the hybrid model IWT-ENN has a great improvement compared to that of others including the persistence method, ENN, AR, BPNN and EMD-ENN.
(2) compared to the two general strategies where the highest frequency sub series is without retreatment or eliminated, the new proposed hybrid model IWT-ENN with the further processing on D1 by SSA has the best prediction performance.

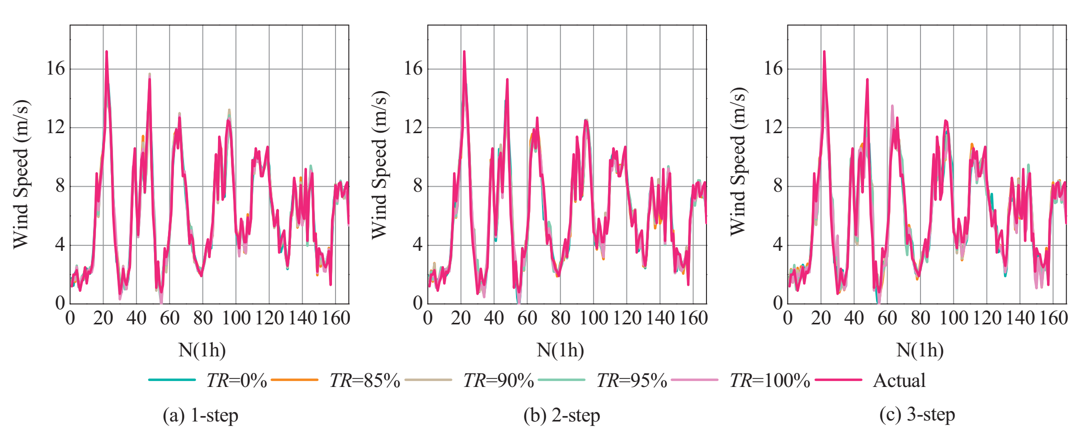


Fig. 11. Prediction results of S2 through IWT-ENN.

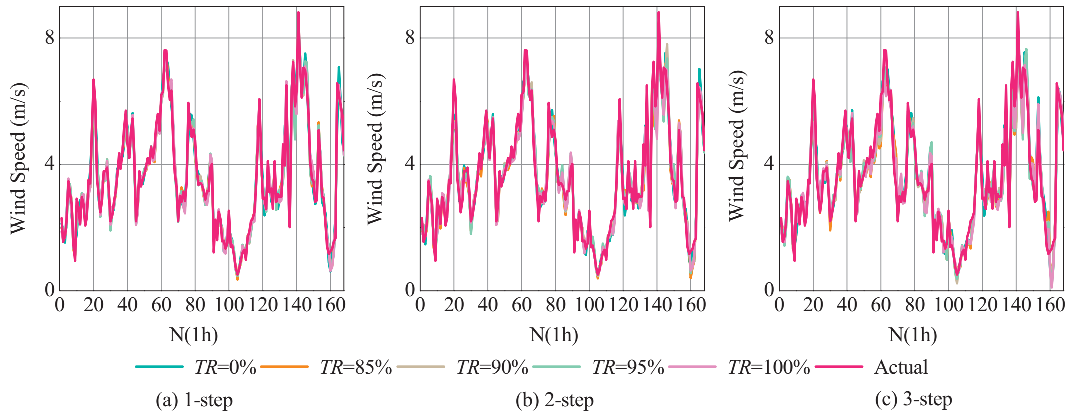


Fig. 12. Prediction results of S3 through IWT-ENN.

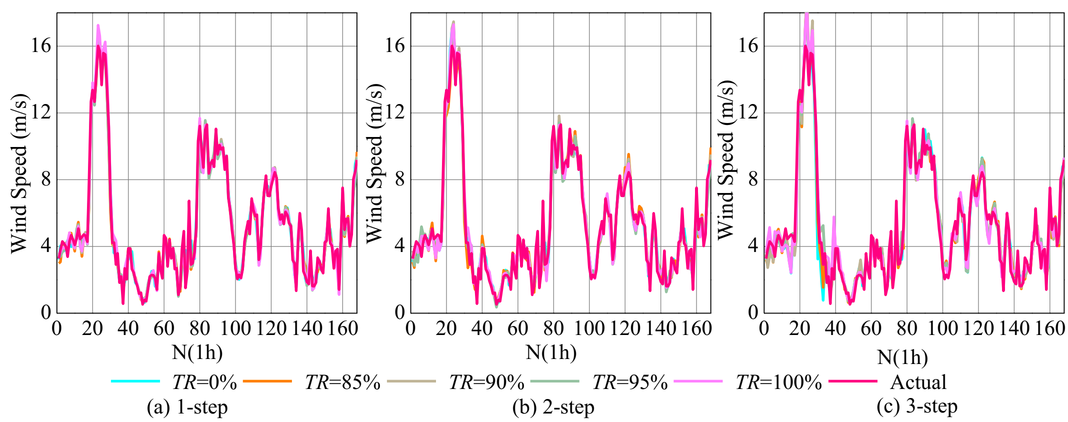


Fig. 13. Prediction results of S4 through IWT-ENN.

Table 6
Error indexes of forecasting results of S2 through IWT-ENN.

	Error criteria	*TR* = 0%	*TR* = 85%	*TR* = 90%	*TR* = 95%	*TR* = 100%
1-step	MAE (m/s)	0.5827	0.3803	0.3532	0.3874	0.3905
	MAPE (%)	13.35%	8.78%	7.81%	9.09%	9.16%
	RMSE (m/s)	0.7778	0.5094	0.4598	0.5160	0.5001
2-step	MAE (m/s)	0.6846	0.5984	0.5421	0.6050	0.6320
	MAPE (%)	16.48%	13.92%	12.92%	15.05%	15.35%
	RMSE (m/s)	0.9089	0.8087	0.7328	0.8168	0.8462
3-step	MAE (m/s)	0.8550	0.7790	0.7640	0.8064	0.7858
	MAPE (%)	19.25%	18.80%	18.06%	18.26%	19.15%
	RMSE (m/s)	1.1387	1.0589	1.0580	1.1057	1.0705

Table 7
Error indexes of multistep forecasting results of S2 by the other methods.

	Error criteria	Persistence method	ENN	AR	BPNN	EMD-ENN
1-step	MAE (m/s)	1.4619	1.3632	1.3858	1.7722	0.7463
	MAPE (%)	31.28%	27.46%	30.05%	34.66%	15.44%
	RMSE (m/s)	1.8773	1.7897	1.8186	2.3461	1.0512
2-step	MAE (m/s)	1.7435	1.5702	1.9334	1.8295	0.9320
	MAPE (%)	38.04%	32.81%	41.70%	47.41%	20.41%
	RMSE (m/s)	2.3204	2.1253	2.4936	2.3183	1.3532
3-step	MAE (m/s)	2.1256	1.7867	2.2743	2.3926	1.1316
	MAPE (%)	44.03%	33.86%	51.80%	48.60%	23.73%
	RMSE (m/s)	2.8823	2.4469	2.9749	3.1489	1.5691

Table 8
Error indexes of forecasting results of S3 through IWT-ENN.

	Error criteria	*TR* = 0%	*TR* = 85%	*TR* = 90%	*TR* = 95%	*TR* = 100%
1-step	MAE (m/s)	0.3435	0.2199	0.2181	0.2270	0.2202
	MAPE (%)	11.16%	7.50%	7.45%	7.63%	7.60%
	RMSE (m/s)	0.4661	0.2836	0.2825	0.2938	0.2942
2-step	MAE (m/s)	0.3661	0.3359	0.3225	0.3329	0.3645
	MAPE (%)	11.87%	10.68%	10.51%	10.71%	11.65%
	RMSE (m/s)	0.5046	0.4776	0.4397	0.4677	0.5187
3-step	MAE (m/s)	0.4353	0.3879	0.3833	0.3736	0.3918
	MAPE (%)	13.95%	12.76%	12.82%	12.43%	12.79%
	RMSE (m/s)	0.6080	0.5428	0.5297	0.5231	0.5529

Table 9
Error indexes of multistep forecasting results of S3 by the other methods.

	Error criteria	Persistence method	ENN	AR	BPNN	EMD-ENN
1-step	MAE (m/s)	0.7332	0.7197	0.6991	1.0320	0.4588
	MAPE (%)	22.79%	22.78%	22.81%	30.69%	16.36%
	RMSE (m/s)	1.0441	1.0378	1.0014	1.3885	0.6163
2-step	MAE (m/s)	0.8996	0.8641	0.9756	1.1561	0.5538
	MAPE (%)	28.65%	28.11%	33.17%	36.67%	18.08%
	RMSE (m/s)	1.2986	1.2476	1.3183	1.6108	0.7712
3-step	MAE (m/s)	1.0526	0.9628	1.1121	1.2393	0.6206
	MAPE (%)	33.95%	31.57%	38.65%	39.37%	20.55%
	RMSE (m/s)	1.3980	1.3006	1.4279	1.6707	0.8996

Table 10
Error indexes of forecasting results of S4 through IWT-ENN.

	Error criteria	*TR* = 0%	*TR* = 85%	*TR* = 90%	*TR* = 95%	*TR* = 100%
1-step	MAE (m/s)	0.4803	0.3462	0.3178	0.3246	0.3407
	MAPE (%)	14.34%	9.70%	9.09%	8.89%	9.59%
	RMSE (m/s)	0.6419	0.4513	0.4119	0.4261	0.4580
2-step	MAE (m/s)	0.5208	0.4891	0.4711	0.4933	0.4987
	MAPE (%)	14.69%	13.74%	13.53%	14.01%	14.68%
	RMSE (m/s)	0.6961	0.6309	0.5998	0.6313	0.6404
3-step	MAE (m/s)	0.6768	0.6004	0.5914	0.5744	0.5966
	MAPE (%)	18.72%	16.89%	16.05%	16.55%	17.80%
	RMSE (m/s)	0.9228	0.8237	0.8187	0.7947	0.8251

Table 11
Error indexes of multistep forecasting results of S4 by the other methods.

	Error criteria	Persistence method	ENN	AR	BPNN	EMD-ENN
1-step	MAE (m/s)	1.1031	1.1002	1.1267	1.4355	0.6269
	MAPE (%)	29.57%	31.49%	34.05%	38.68%	18.65%
	RMSE (m/s)	1.5114	1.4657	1.4894	1.9669	0.8278
2-step	MAE (m/s)	1.3677	1.3069	1.6259	1.8848	0.7584
	MAPE (%)	34.90%	35.86%	48.45%	48.84%	21.93%
	RMSE (m/s)	1.9019	1.8221	2.1135	2.6614	0.9902
3-step	MAE (m/s)	1.4411	1.3596	1.8545	2.0320	0.9723
	MAPE (%)	34.38%	36.02%	59.99%	53.92%	30.17%
	RMSE (m/s)	1.9943	1.8654	2.4810	2.7621	1.2597

(3) when *TR* equals 90%, the hybrid model IWT-ENN has the best prediction performance.

6. Conclusion

A new hybrid decomposition technology called IWT is proposed, where an original wind speed series is broken down into a sequence of sub series by WT and the one with the highest frequency is denoised by SSA. Then based on IWT and ENN, a new hybrid model IWT-ENN is also proposed to make wind speed prediction. Through the above prediction experiments and comparative analysis, some conclusions can be drawn as follows:

(1) the performance of the hybrid model IWT-ENN has a great improvement compared to that of others including the persistence method, ENN, AR, BPNN and EMD-ENN.
(2) compared to the two general strategies where the highest frequency sub series is without retreatment or eliminated, the new proposed hybrid model IWT-ENN with the further processing on D1 by SSA has the best prediction performance.
(3) the *TR* threshold in IWT-ENN is recommended as 90%.

In this study, the performance of ENN in the proposed model has not been optimized. The way to enhance the ENN model to improve overall performance of IWT-ENN will be also searched in the future work. In the while, the applicability of the proposed decomposition technology IWT with other forecasting methods will be also studied.

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] Aghajani A, Kazemzadeh R, Ebrahimi A. A novel hybrid approach for predicting wind farm power production based on wavelet transform, hybrid neural networks and imperialist competitive algorithm. Energy Convers Manag 2016;121:232–40.
[2] Azimi R, Ghofrani M, Ghayekhloo M. A hybrid wind power forecasting model based on data mining and wavelets analysis. Energy Convers Manag 2016;127:208–25.
[3] Zameer A, Arshad J, Khan A, Raja MAZ. Intelligent and robust prediction of short term wind power using genetic programming based ensemble of neural networks. Energy Convers Manag 2017;134:361–72.
[4] Chitsaz H, Amjady N, Zareipour H. Wind power forecast using wavelet neural network trained by improved Clonal selection algorithm. Energy Convers Manag 2015;89:588–98.
[5] Noorollahi Y, Jokar MA, Kalhor A. Using artificial neural networks for temporal and spatial wind speed forecasting in Iran. Energy Convers Manag 2016;115:17–25.
[6] Okumus I, Dinler A. Current status of wind energy forecasting and a hybrid method for hourly predictions. Energy Convers Manag 2016;123:362–71.
[7] Hong YY, Yu TH, Liu CY. Hour-ahead wind speed and power forecasting using empirical mode decomposition. Energies 2013;6:6137–52.
[8] 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.
[9] 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.
[10] 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.
[11] Liu Y, Shi J, Yang Y, Lee WJ. Short-term wind-power prediction based on wavelet transform-support vector machine and statistic-characteristics analysis. IEEE Trans Ind Appl 2012;48(4):1136–41.
[12] Liu H, Tian HQ, Pan DF, Li Y-F. Forecasting models for wind speed using wavelet, wavelet packet, time series and Artificial Neural Networks. Appl Energy 2013;107:191–208.
[13] Liu H, Tian HQ, Li YF. Comparison of new hybrid FEEMD-MLP, FEEMD-ANFIS, Wavelet Packet-MLP and Wavelet Packet-ANFIS for wind speed predictions. Energy Convers Manage 2015;89:1–11.
[14] Liu H, Tian HQ, Li YF. Four wind speed multi-step forecasting models using extreme learning machines and signal decomposing algorithms. Energy Convers Manag 2015;100:16–22.
[15] 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.
[16] Kiplangat DC, Asokan K, Kumar KS. Improved week-ahead predictions of wind speed using simple linear models with wavelet decomposition. Renew Energy 2016;93:38–44.
[17] 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.
[18] Amjady N, Keynia F. Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm. Energy 2009;34(1):46–57.
[19] 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.
[20] Khan AA, Shahidehpour M. One day ahead wind speed forecasting using wavelets. In: Power systems conference and exposition, 2009. PSCE'09. IEEE/PES; 2009. p. 1–5.
[21] Lei C, Ran L, "Short-term wind speed forecasting model for wind farm based on wavelet decomposition". In: Electric utility deregulation and restructuring and power technologies, 2008. DRPT 2008. Third International Conference on; 2008. p. 2525–2529.
[22] 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.
[23] Liu D, Niu D, Wang H, Fan L. Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. Renew Energy 2014;62:592–7.
[24] Wang J, Wang J, Li Y, Zhu S, Zhao J. Techniques of applying wavelet de-noising into a combined model for short-term load forecasting. Int J Electr Power Energy Syst 2014;62:816–24.
[25] 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.
[26] Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 1989;11(7):674–93.
[27] Siddiqi AH, Khan S, Rehman S. Wind speed simulation using wavelets. Am J Appl Sci 2005;2(2):557–64.
[28] Saleh SA, Ahshan R, Moloney CR. Wavelet-based signal processing method for detecting ice accretion on wind turbines. IEEE Trans Sustain Energy 2012;3(3).
[29] 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.
[30] Briceño H, Rocco CM, Zio E. Singular spectrum analysis for forecasting of electric load demand. In: Prognostics and system health management conference-PHM-2013. 2013: 33: 919–24.
[31] Hassani H. Singular spectrum analysis: methodology and comparison singular spectrum analysis: methodology and comparison. J Data Sci 2007;5:239–57.
[32] 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.
[33] Elman JL. Finding structure in time. Cogn Sci 1990;14(2):179–211.
[34] Hudson M, Martin B, Hagan T, Demuth HB. R2016b, "Neural Network ToolboxTM Reference".
[35] Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat 2001;29(5):1189–232.
[36] Natekin A, Knoll A. Gradient boosting machines, a tutorial. Front Neurorobot 2013.
[37] Kusiak A, Zheng H, Song Z. Short-term prediction of wind farm power: a data mining approach. IEEE Trans Energy Convers 2009;24(1):125–36.
[38] 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.