Energy Conversion and Management 156 (2018) 498–514

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network

Hui Liu*, Xi-wei Mi, Yan-fei Li

Institute of Artificial Intelligence and Robotics, Key Laboratory for Traffic Safety on Track of Ministry of Education, School of Traffic and Transportation Engineering, Central South University, Changsha 410075, Hunan, China

ARTICLE INFO

Keywords:
Wind speed forecasting
Deep learning
Empirical wavelet transform
Long short term memory network
Elman neural network

ABSTRACT

The wind speed forecasting plays an important role in the planning, controlling and monitoring of the intelligent wind power systems. Since the wind speed signal is stochastic and intermittent, it is difficult to achieve their satisfactory prediction. In the study, a novel hybrid deep-learning wind speed prediction model, which combines the empirical wavelet transformation and two kinds of recurrent neural network, is proposed. In the proposed new model, the empirical wavelet transformation is adopted to decompose the raw wind speed data into several sub-layers. The long short term memory neural network, a deep learning algorithm based method, is utilized to predict the low-frequency wind speed sub-layers. The Elman neural network, a mainstream recurrent neural network, is built to predict the high-frequency sub-layers. In the executed forecasting experiments, eleven different forecasting models are included to validate the real prediction performance of the proposed model. The experimental results indicate that the proposed model has satisfactory performance in the high-precision wind speed prediction.

1. Introduction

With the development of the world economy, the renewable energy has gradually shown its great importance around the world [1]. As one of the fastest developing renewable energy, wind energy has attracted the extensive attention [2]. However, due to the stochastic and intermittent characteristics of wind speed, the non-stationarity of the wind speed may challenge the security and stability of the power system [3]. To ensure the normal operation of the power networks, the technology of the wind power and wind speed forecasting is expected [4].

Traditionally, the wind power and wind speed forecasting models can be classified into two main types: the physical prediction models and the statistical prediction models [5]. The physical prediction models simulate the wind speed changing by using the physical laws and boundary conditions, which often have good prediction performance in the long-term wind speed forecasting [6]. The CFD (*Computational Fluid Dynamics*) model is one of the main technologies in the physical simulating. Besides, in recent years, some mathematical algorithms are introduced to improve the prediction performance of the physical models. The statistical prediction models forecast the wind power or the wind speed by using the historical samples or spatial-temporal samples, which can produce good estimation results in the short-term wind speed forecasting [7]. The main technologies of the statistical simulating can be classified into three types: the linear methods, the nonlinear methods and the hybrid methods [8]. The linear methods, such as the AR (*Auto Regressive*) model, the ARMA (*Auto Regressive Moving Average*) model and the ARIMA (*Auto Regressive Integrated Moving Average*) model, are reliable and fast [9]. In the past few years, a number of the linear methods have been studied. Lydia et al. [10] investigated the prediction performance of some linear and nonlinear auto-regressive models. Maatallah et al. [11] developed the wind speed forecasting model by combing the Hammerstein algorithms and the auto-regressive approaches. These models have satisfactory performance in the linear prediction performance [12]. The nonlinear methods, such as the BP (*Back Propagation*) neural network, the SVM (*Support Vector Machine*) and the RBF (*Radial Basis Function*) neural network, can always deal with the nonlinear and complex wind power or wind speed series accurately [13]. The benefits of these models have been investigated in some cases. Ren et al. [14] designed the optimal BP neural network for the wind speed forecasting. The proposed model had satisfactory performance. Santamaría-Bonfil et al. [15] built the SVM model. Their results showed that the proposed model was more

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

https://doi.org/10.1016/j.enconman.2017.11.053
Received 17 September 2017; Received in revised form 23 October 2017; Accepted 18 November 2017
Available online 28 November 2017

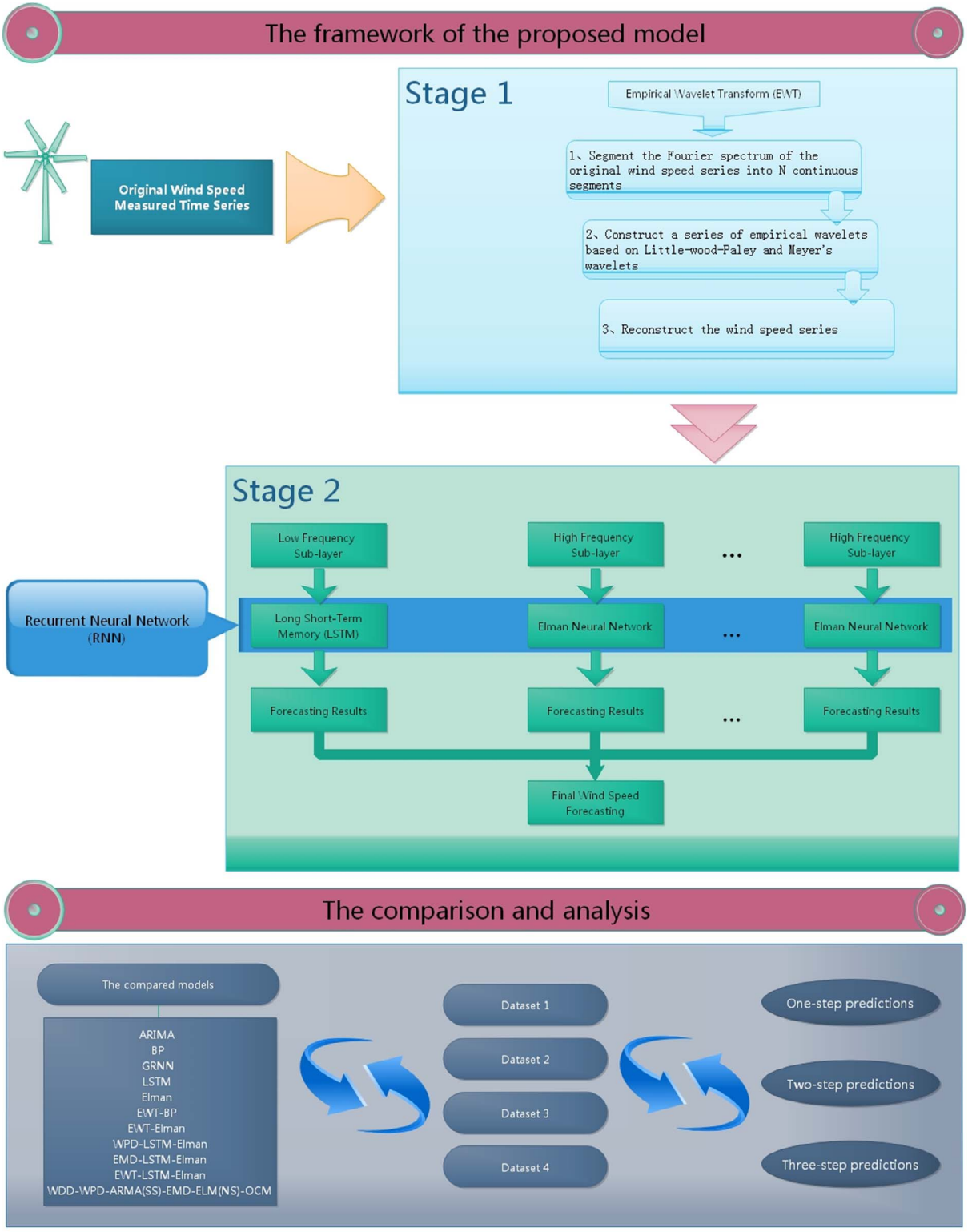

Fig. 1. The whole process of the hybrid EWT-LSTM-Elman model.

accurate than the persistence and auto-regressive models in the medium short term wind speed and wind power forecasting. Zhang et al. [16] adopted the RBF neural network to realize the wind speed interval forecasting, the results indicated that the proposed method had high quality prediction performance. The hybrid methods are the combinations of some mainstream models and algorithms. Due to the good abilities of the hybrid models for the feature-extracting, these models have been widely applied in the wind power and wind speed forecasting [17].

In recent years, many hybrid prediction models in the wind power

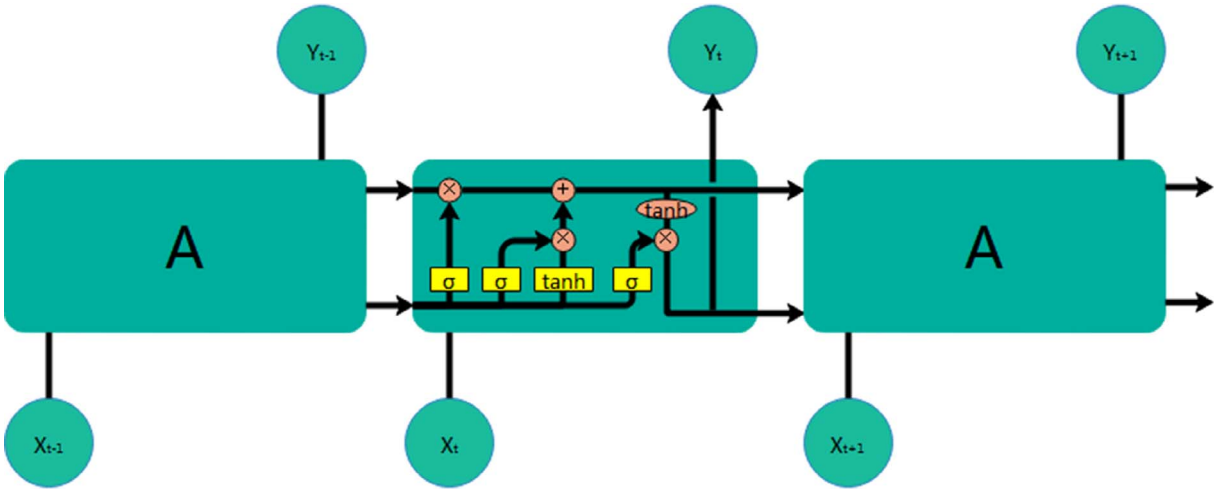

Fig. 2. The architecture of a LSTM model.

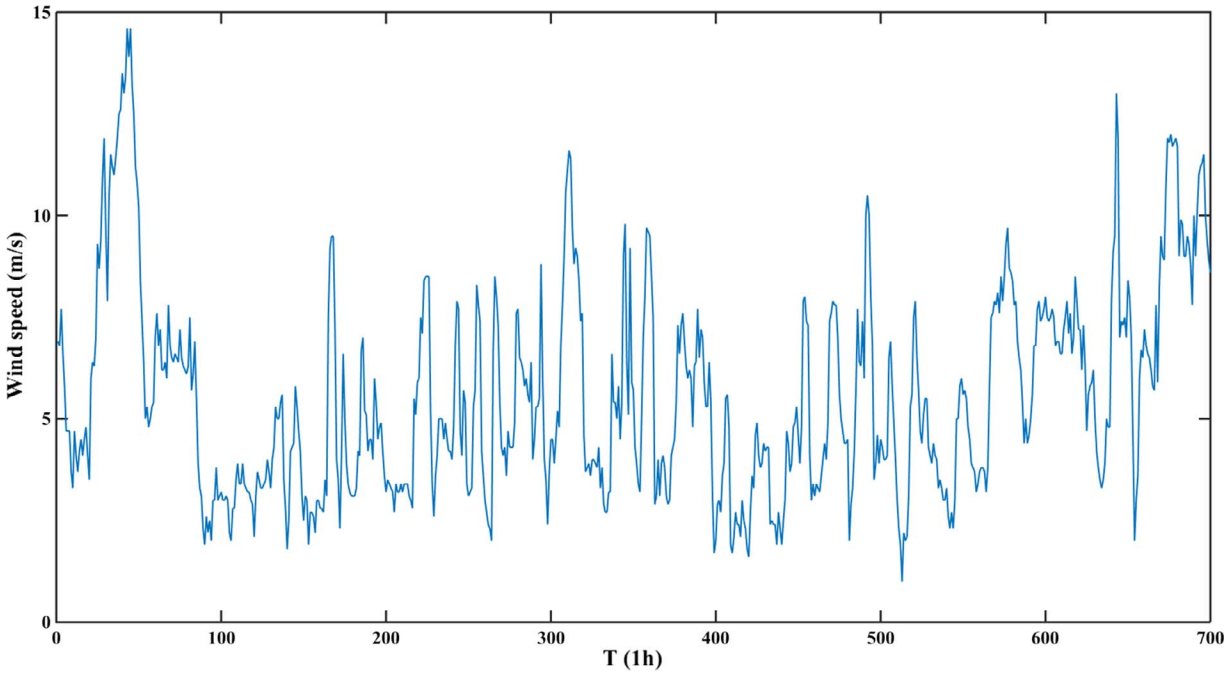

Fig. 3. Original wind speed time series #1.

and wind speed forecasting have been proposed and investigated, which mainly contain the data preprocessing and the forecast modeling [18]. The main technologies for these two modeling types are the signal decomposition algorithms and the prediction algorithms [19].

The decomposition algorithms can effectively improve the prediction performance of the built models through decomposing the intermittent raw wind time series into several more stationary sub-layers [20]. Among the decomposition algorithms, the WD (*Wavelet Decomposition*), the WPD (*Wavelet Packet Decomposition*), the EMD (*Empirical Mode Decomposition*), the EEMD (*Ensemble Empirical Mode Decomposition*) and the FEEMD (*Fast Ensemble Empirical Mode Decomposition*) are widely recognized and used in the wind speed prediction. Liu et al. [21] presented a hybrid forecasting model based on the WD, the GA (*Genetic Algorithm*) and the SVM. Liu et al. [22] proposed the hybrid forecasting model using the EEMD and the RARIMA (*Recursive Autoregressive Integrated Moving Average*). Wang et al. [23] provided the short term wind speed forecasting model adopting the EMD, the GA and the BP neural network. Ordinarily, the WD and WPD algorithms have excellent ability of time-frequency analysis, and the EMD, EEMD and FEEMD algorithms have good self-adaptive ability of removing the stochastic volatility. However, the aforementioned decomposition algorithms have some disadvantages: (1) the performance of the WD and WPD algorithms both depends on the wavelet basis and decomposition levels highly; (2) the EMD, EEMD and FEEMD algorithms lack the strict mathematical theory [24]. To overcome these drawbacks, the EWT (*Empirical Wavelet Transform*) has been proposed as shown in Ref. [25]. It can extract a series of modes of a signal by using an appropriate wavelet, so it is very effective in the non-stationary signal processing [26]. In the study, the EWT is adopted to decompose the original wind speed time series.

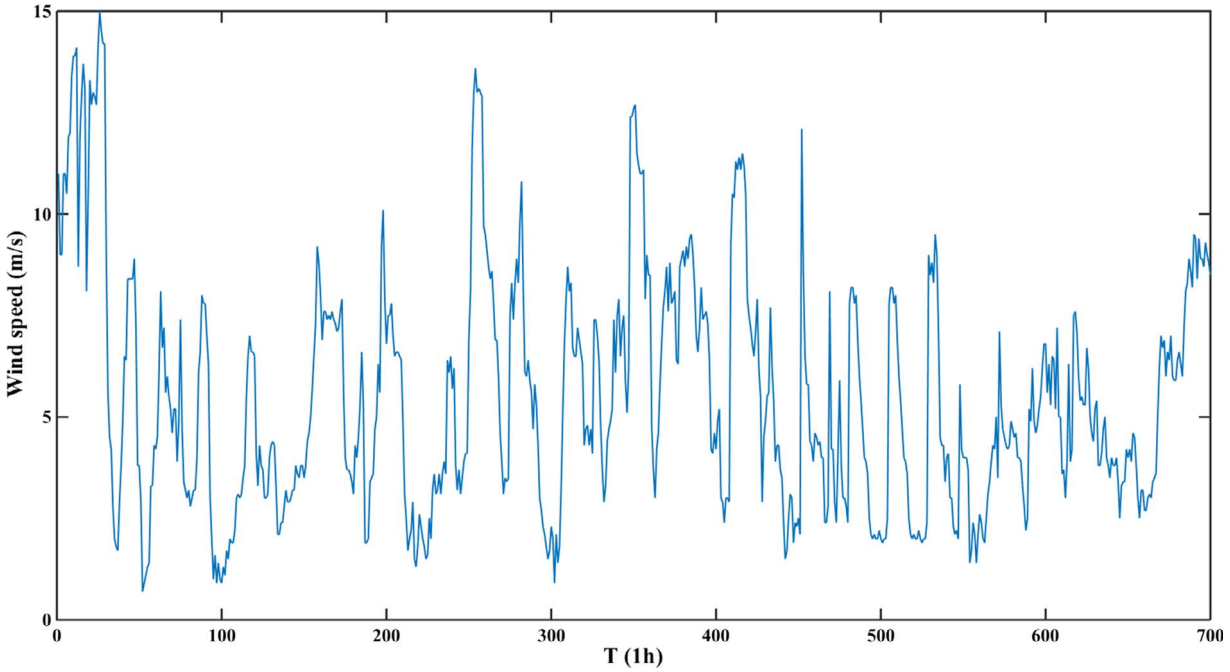

Fig. 4. Original wind speed time series #2.

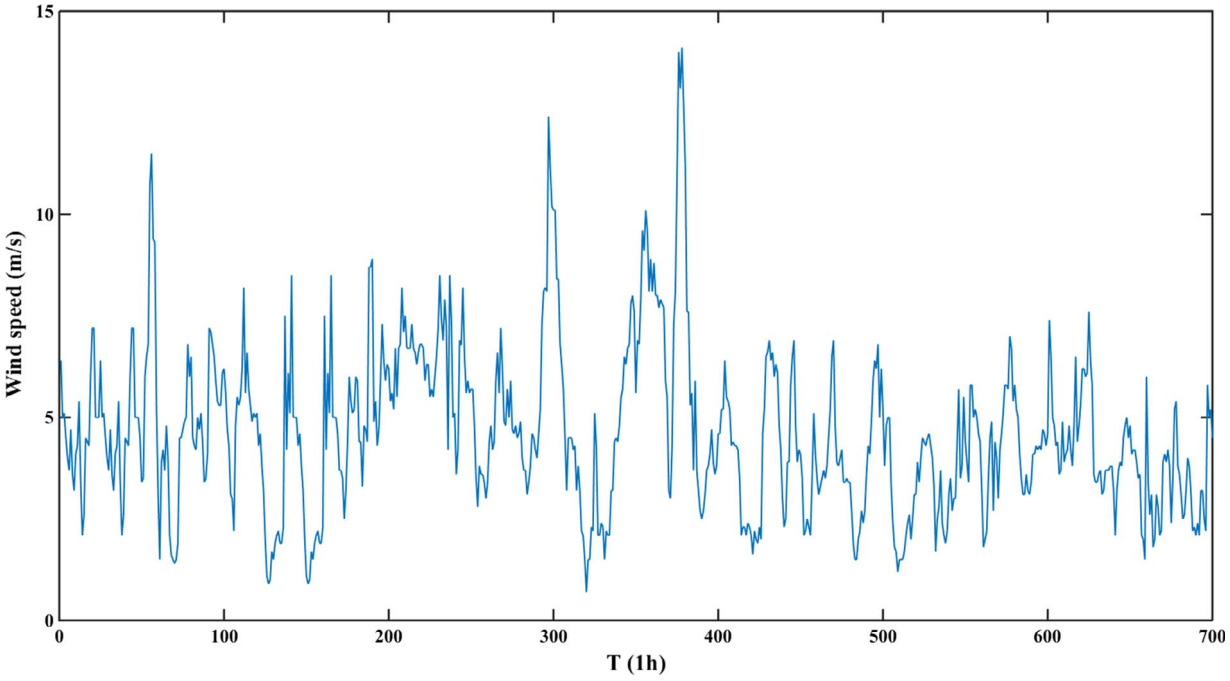

Fig. 5. Original wind speed time series #3.

The prediction algorithms are the core part of the wind speed forecasting. In recent years, some new prediction algorithms have been proposed, among these algorithms, deep learning methods, such as the DBN (*Deep Belief Network*), the CNN (*Convolutional Neural Network*) and the RNN (*Recurrent Neural Network*), have been developed rapidly. Kuremoto et al. [27] built a DBN for the time series forecasting. Kim [28] used the CNN for the sentence level classification tasks. Kim et al. [29] proposed a RNN for the dynamic object recognition. Compared with the shallow models, the deep learning models can extract the deep inherent features in data [30]. Based on the following some forecasting studies, deep learning methods can have good prediction performance in wind forecasting [31]. For example, Wang et al. [32] designed a CNN for the wind power forecasting. Their results demonstrated that the proposed model was effective. Nevertheless, the deep learning methods have not yet been widely used in the wind speed and wind power forecasting. By considering that the wind speed time series often have

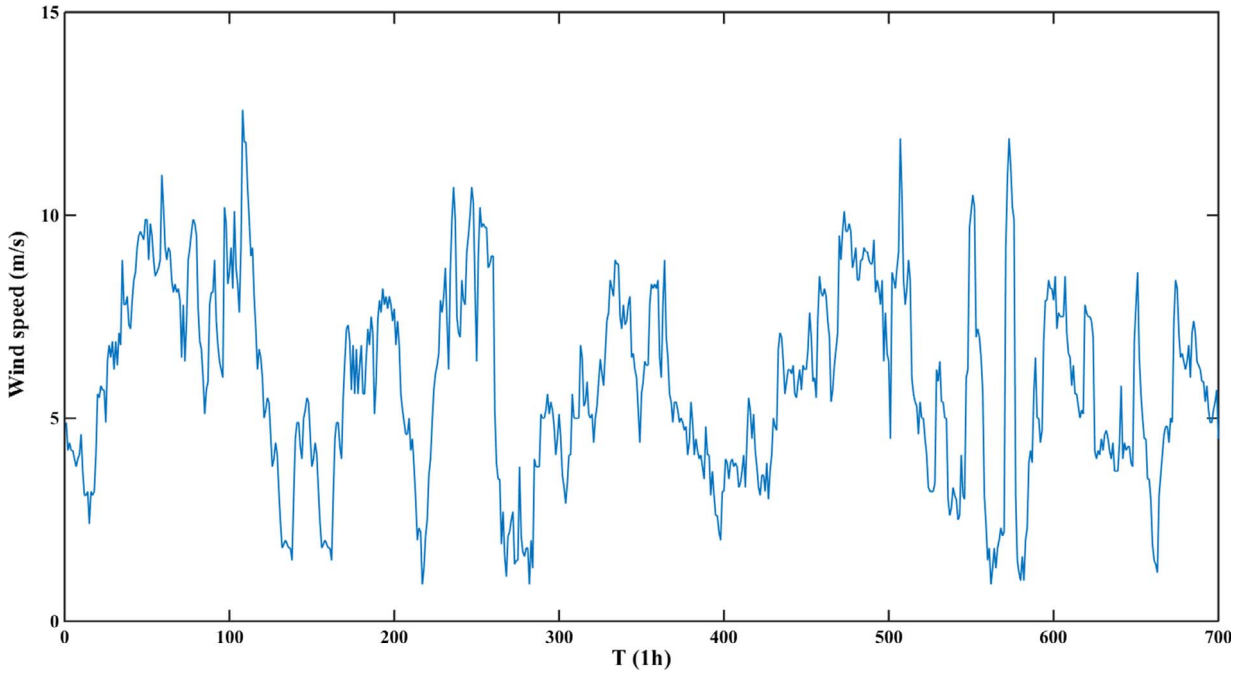

Fig. 6. Original wind speed time series #4.

the long-term and short-term dependency in the low-frequency approximate parts, the LSTM (*Long Short Term Memory*) network [33], which is a special kind of RNN, is employed to predict the decomposed low-frequency sub-layer in this study. Since the high-frequency detailed parts of the wind speed time series have stochastic characteristics and short-term dependency, the prediction performance of the LSTM network does not significantly outperform that of other neural networks. Therefore, to reduce the forecasting computation time, as well as to keep the prediction performance, the Elman neural network, which is a simple kind of RNN, is employed to predict the decomposed high-frequency sub-layers in this study.

Apart from the prediction algorithms, some other technologies are also used in forecast modeling, such as the weighting-based combined algorithms and the parameter optimization algorithms. The weighting-based combined algorithms can improve the prediction performance of the hybrid model by giving an appropriate weight coefficient for each prediction model [34]. The parameter optimization algorithms can improve the performance of the prediction algorithms by searching the optimal parameters [35]. Among the parameter optimization algorithms, the heuristic optimization algorithm is the major part, which mainly includes the PSO (*Particle Swam Optimization*) algorithm, the SA (*Simulated Annealing*) algorithm, the CS (*Cuckoo Search*) algorithm, the CRO (*Coral Reefs Optimization*) algorithm and the bat algorithm [36]. Pandit et al. [37] studied the applications of the PSO based methods for the wind integrated optimal dispatch. Jiang et al. [38] developed a hybrid model for the wind speed forecasting, where the CS algorithm was used to tune the parameters of the prediction model. Salcedo-Sanz et al. [39] designed a hybrid model for the wind speed prediction systems, in which the CRO algorithm was used to select the input parameters of the ELM (*Extreme Learning Machine*). It is recognized that these heuristic optimization algorithms have the ability to search the global optima and obtain the optimal parameters [40]. The weighting-based combined algorithms and the parameter optimization algorithms can both effectively improve the prediction performance. However, on the other hand they may increase the time complexity of the prediction algorithms and undermine the practicability in realistic power systems. To balance the computing performance and the prediction accuracy, only the decomposition algorithms and the prediction algorithms are used in this study.

In this paper, a novel combination wind speed prediction model is proposed based on the EWT, LSTM network and Elman neural network. The model is composed of three steps as: (a) the EWT is adopted to decompose the raw wind speed data into several sub-layers; (b) the LSTM network is employed to predict the low-frequency sub-layer, while the Elman neural network is employed to predict the high-frequency sub-layers; (c) the prediction results of each sub-layer are summarized to obtain the final results for the original wind speed data.

The main contributions of this study are provided as follows: (a) A novel wind speed forecasting framework is proposed by combining with the EWT and two RNNs. The EWT can effectively extract the meaningful components from the wind speed signals, while the two RNNs can explore the hidden high-level nonlinear features of the wind speed signals; (b) Comparing with the traditional wavelet and EMD decomposition, the EWT algorithm has the integrating performance of the wavelet algorithm and the EMD algorithm so that it has good adaption ability in the non-stationary time series decomposition. The potential performance of the EWT algorithm for the wind speed multi-step forecasting has not been studied; (c) The LSTM network, which is a deep learning method, is introduced to predict the low-frequency wind speed sub-layer decomposed by the EWT in the proposed hybrid forecasting model. Since the LSTM network has all of the advantages of the RNN networks and it solves the vanishing gradient disadvantage of the RNN networks, it has the good performance in the nonlinear signal modeling. The combination of the EWT algorithm and the LSTM network for the high-precision wind speed multiple forecasting has also not been investigated; (d) The Elman neural network, which is one type of the classical RNNs, is adopted to predict the high-frequency sub-layers decomposed by the EWT in the proposed hybrid forecasting model. Since the Elman network has good performance in the high-frequency single modeling, it can obtain the short-term trend

information of the jumping wind speed series; and (e) By considering the upper issues, the EWT, the LSTM neural network and the Elman neural network are selected to realize the non-stationary wind speed decomposition, the decomposed wind speed low-frequency forecasting and the decomposed wind speed high-frequency forecasting, respectively. The combination of the EWT, the LSTM network and the Elman neural network for the wind speed multiple forecasting has not been investigated before.

The structure of the study is organized as follows: Section 2 describes the proposed hybrid model and presents the required individual models; Section 3 provides the wind speed forecasting experiments and the prediction results of the proposed model. In the section, the real forecasting results of the comparing models are also analyzed. Section 4 concludes this study.

2. The hybrid EWT-LSTM-Elman model

2.1. The whole process of the proposed model

The whole process of the hybrid EWT-LSTM-Elman model is demonstrated in Fig. 1, which can be further described as follows:

(1) The EWT is adopted to decompose the raw wind speed data into several sub-layers. *The details of the EWT algorithm are presented in* Section 2.2.
(2) The LSTM network is employed to predict the low-frequency sub-layer, while the Elman neural network is employed to predict the high-frequency sub-layers. *The details of the LSTM network and the Elman neural network are presented in* Section 2.3 and Section 2.4, *respectively*.
(3) Several models are used to compare the prediction performance of the proposed model, the compared models include the ARIMA model, the BP model, the GRNN model, the LSTM model, the Elman model, the EWT-BP model, the EWT-Elman model, the WPD-LSTM-Elman model, the EMD-LSTM-Elman model, the EWT -LSTM-Elman model and the WDD-WPD-ARMA(SS)-EMD-ELM(NS)-OCM model. *The details of the precision estimating indexes are provided in* Section 2.5. *The analysis is provided in* Sections 3 and 4.

2.2. Empirical wavelet transform

The EWT can be defined as a set of band pass filters selected according to the spectral characteristics of a signal. In order to determine the frequency ranges of the band pass filters, the Fourier spectrum of the signal is segmented. According to the literature [41], the EWT can effectively identify and extract a finite number of intrinsic modes of a

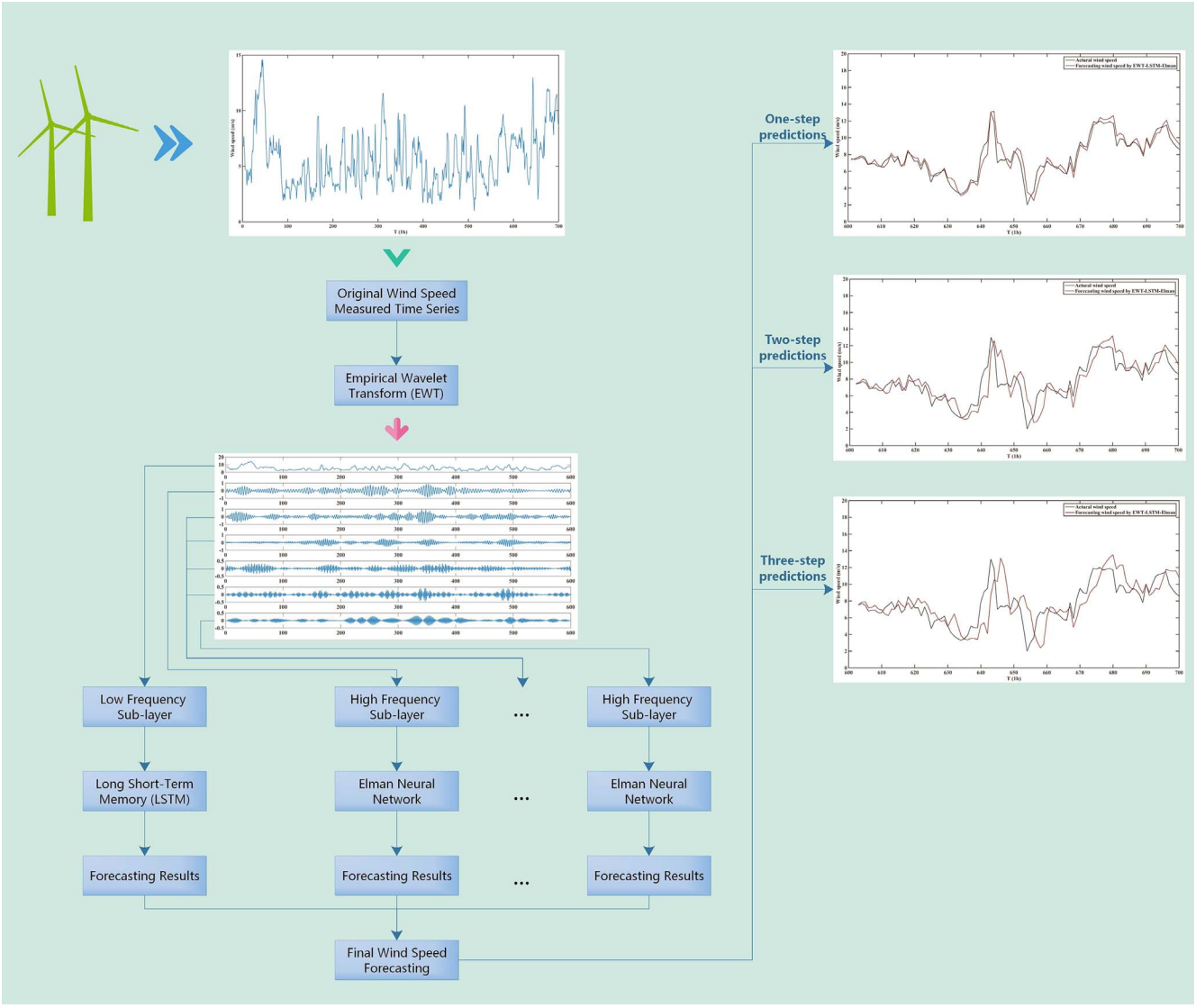

Fig. 7. The framework of the proposed model in case one.

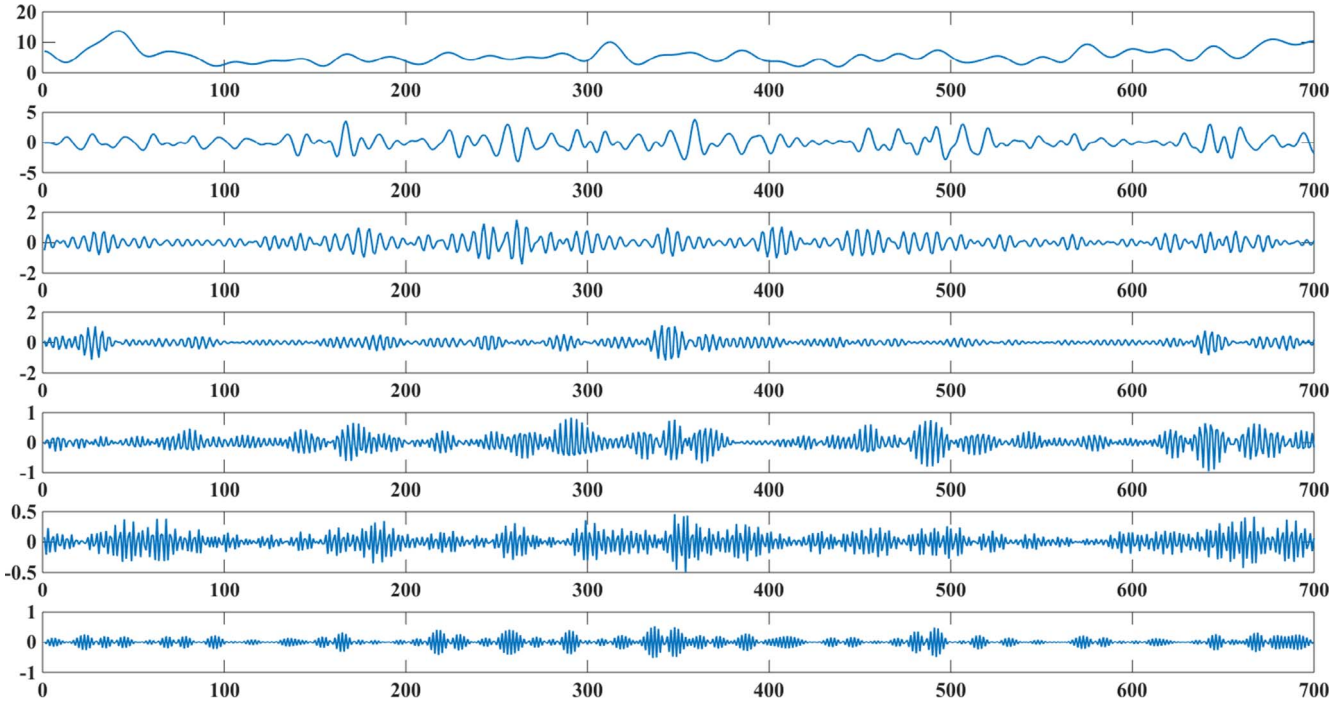

Fig. 8. The EWT decomposed results of the original wind speed time series #1.

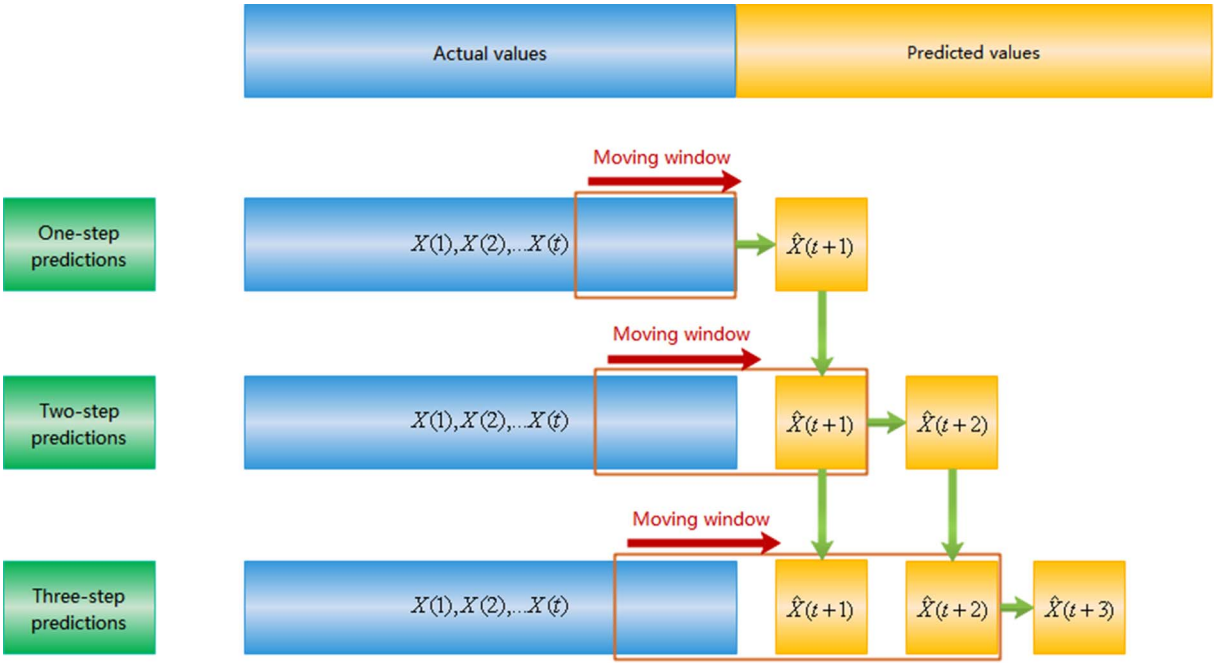

Fig. 9. The structure of the multi-step forecasting strategy.

wind speed time series. The major steps of the EWT algorithm can be described as: (a) extending the signal; (b) executing the Fourier transform; (c) extracting the boundaries; (d) building the filter bank; (e) extracting the sub-bands.

The computation of the EWT can be summarized as [42]:

(1) Segment the Fourier spectrum of the original wind speed series into N continuous segments. The limits between each segments can be defined as ω_n, where $\omega_0 = 0$ and $\omega_N = 0$, respectively. Each segment can be defined as $\Lambda_n = [\omega_{n-1},\omega_n]$. $\cup_{n=1}^{N}\Lambda_n = [0,\pi]$. For each ω_n, a transition phase T_n with width $2\tau_n$ is used. $\tau_n = \gamma\omega_n$. The range of γ is defined as:

$$\gamma < \min_n\left(\frac{\omega_{n+1}-\omega_n}{\omega_{n+1}+\omega_n}\right) \quad (1)$$

(2) Construct a series of empirical wavelets based on the Little-wood-

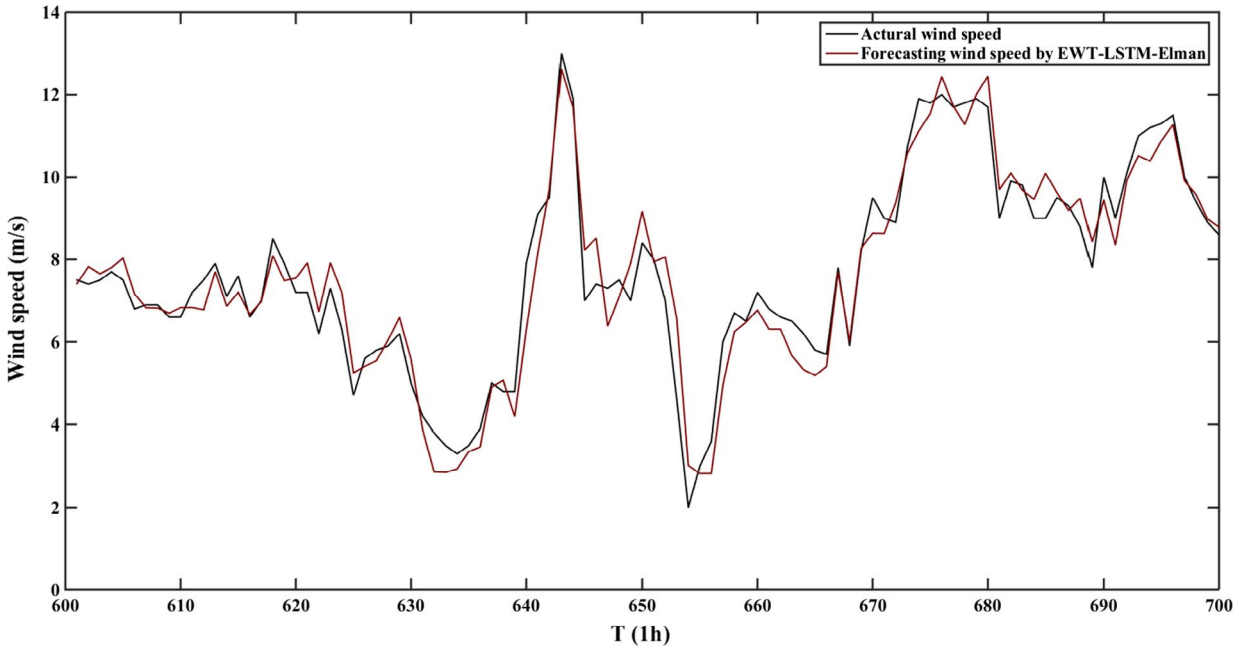

Fig. 10. The results of one-step predictions of the wind speed series #1.

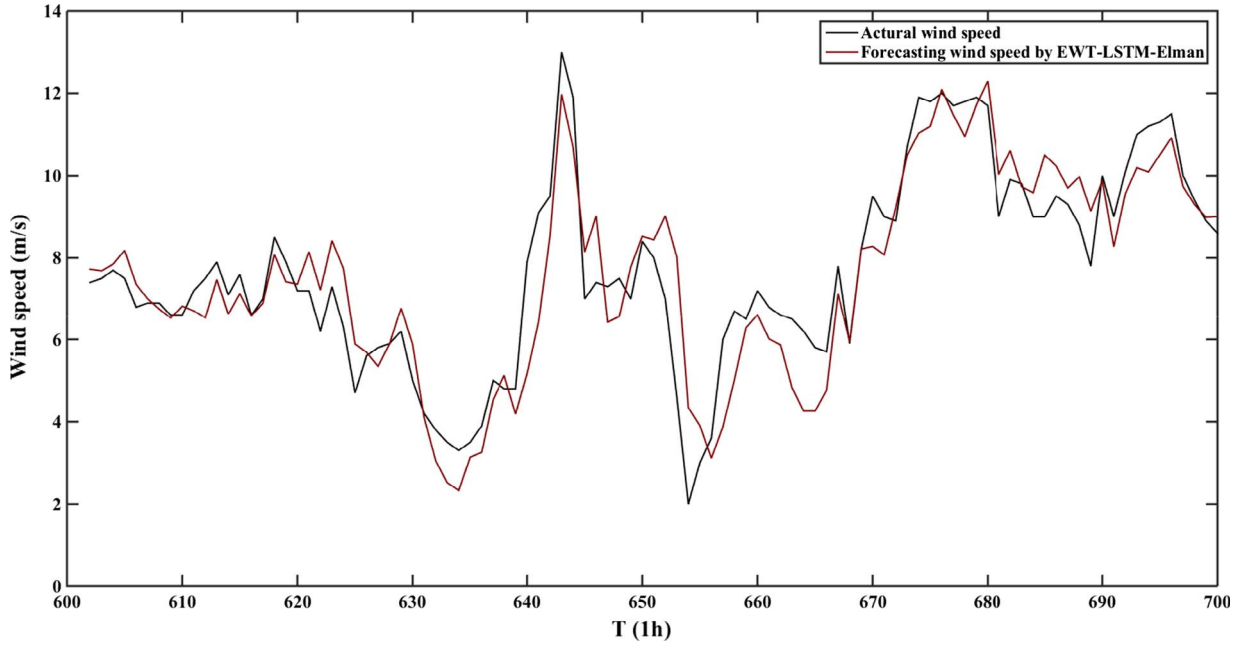

Fig. 11. The results of two-step predictions of the wind speed series #1.

Paley and Meyer's wavelets. For $\forall\, n > 0$, the empirical scaling function and empirical wavelets are defined by Eqs. (2) and (3), respectively:

$$\hat{\phi}_n(\omega) = \begin{cases} 1 & |\omega| \leqslant (1-\gamma)\omega_n \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}|\omega|-\omega_n+\tau_n\right)\right] & (1-\gamma)\omega_n \leqslant |\omega| \leqslant (1+\gamma)\omega_n \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$\hat{\varphi}_n(\omega) = \begin{cases} 1 & \omega_n(1+\gamma) \leqslant |\omega| \leqslant (1-\gamma)\omega_{n+1} \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_{n+1}}|\omega|-\omega_{n+1}+\tau_{n+1}\right)\right] & (1-\gamma)\omega_{n+1} \leqslant |\omega| \leqslant (1+\gamma)\omega_{n+1} \\ \sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau}|\omega|-\omega_n+\tau_n\right)\right] & (1-\gamma)\omega_n \leqslant |\omega| \leqslant (1+\gamma)\omega_n \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

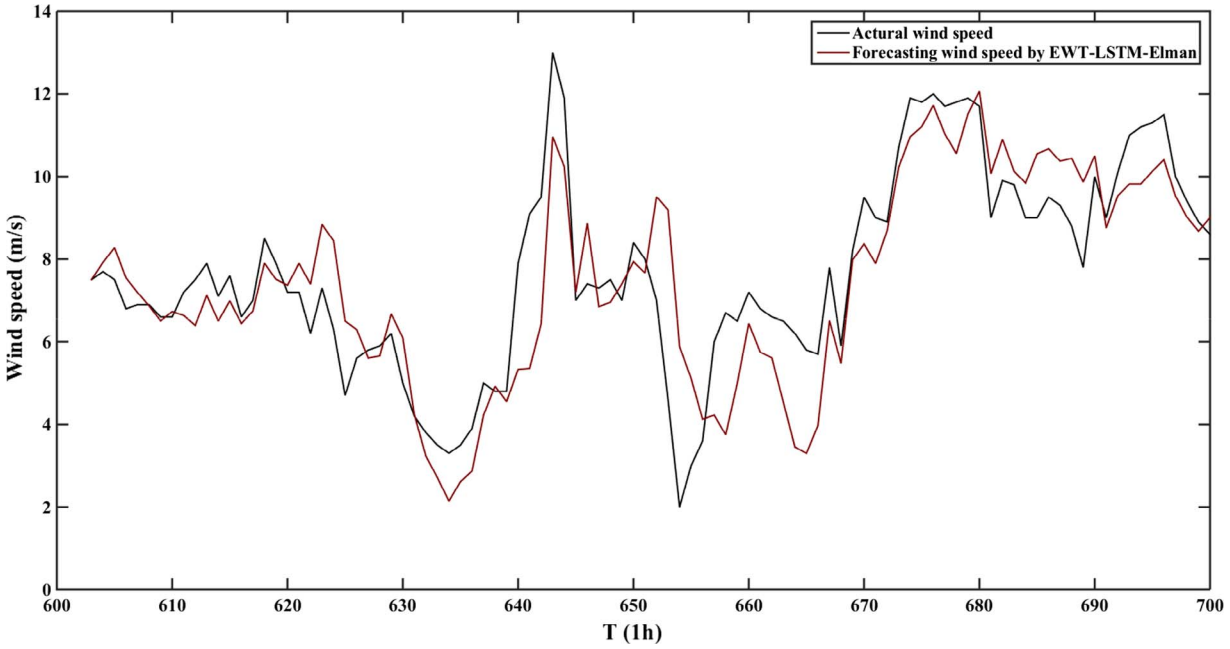

Fig. 12. The results of three-step predictions of the wind speed series #1.

Table 1
Error estimated results of the predictions of the wind speed series#1.

Indexes	1-step	2-step	3-step	1-step	2-step	3-step
	ARIMA			BP		
MAPE (%)	7.17	8.85	15.73	13.22	13.51	13.58
MAE (m/s)	0.93	1.15	2.05	1.72	1.76	1.77
RMSE (m/s)	1.21	1.58	2.46	2.35	2.40	2.41
	GRNN			LSTM		
MAPE (%)	6.88	10.25	13.91	6.70	10.08	13.33
MAE (m/s)	0.89	1.33	1.81	0.87	1.31	1.73
RMSE (m/s)	1.27	1.78	2.34	1.18	1.86	2.47
	Elman			EWT-BP		
MAPE (%)	7.31	11.34	13.58	5.10	10.03	15.43
MAE (m/s)	0.95	1.47	1.77	0.66	1.30	2.01
RMSE (m/s)	1.29	1.95	2.28	0.85	1.65	2.51
	EWT-Elman			WPD-LSTM-Elman		
MAPE (%)	5.08	9.00	13.80	3.89	6.94	9.92
MAE (m/s)	0.66	1.17	1.79	0.51	0.90	1.29
RMSE (m/s)	0.83	1.45	2.20	0.63	1.14	1.65
	EMD-LSTM-Elman			EWT-LSTM-Elman		
MAPE (%)	5.06	6.37	8.23	3.55	5.84	7.80
MAE (m/s)	0.66	0.83	1.07	0.46	0.76	1.01
RMSE (m/s)	0.85	1.03	1.39	0.59	1.00	1.36

The function $\beta(x)$ is defined as:

$$\beta(x) = x^4(35-84x + 70x^2-20x^3) \tag{4}$$

The approximation coefficients $W_f^{\varepsilon}(0,t)$ are obtained by the inner products with the empirical scaling function as follows:

$$W_f^{\varepsilon}(0,t) = \langle f,\phi_1\rangle = \int f(\tau)\overline{\phi_1(\tau-t)}d\tau \tag{5}$$

The detailed coefficients are obtained by the inner products with the empirical wavelets as follows:

$$W_f^{\varepsilon}(n,t) = \langle f,\psi_n\rangle = \int f(\tau)\overline{\psi_n(\tau-t)}d\tau \tag{6}$$

(3) Reconstruct the wind speed series. The reconstruction series and empirical modes are defined as:

$$f(t) = W_f^{\varepsilon}(0,t)^*\phi_1(t) + \sum_{n=1}^{N} W_f^{\varepsilon}(n,t)^*\psi_n(t) \tag{7}$$

$$x_0(t) = W_f^{\varepsilon}(0,t)^*\phi_1(t) \tag{8}$$

$$x_k(t) = W_f^{\varepsilon}(k,t)^*\psi_k(t) \tag{9}$$

2.3. Long short term memory network

The LSTM network is a special kind of the RNN. It is stable and powerful for solving the long-term and short-term dependency problems. The key parameter of the LSTM network is the memory cell, which can memorize the temporal state. The LSTM network can add or remove information to the cell state through three controlling gates as: input gate, forget gate and output gate. The calculation of the LSTM network is described as follows [33]: (1) When a new input comes, if the input gate is activated, the input information can be accumulated to the cell. (2) If the forget gate is activated, the past cell status can be forgotten in the process. (3) The output gate can control whether the latest cell output can be propagated to the ultimate state. The LSTM architecture is illustrated in Fig. 2.

In the context of wind speed forecasting, $x = (x_1,x_2,...,x_T)$ is the historical input data and $y = (y_1,y_2,...,y_T)$ is the predicted data. The predicted wind speed series can be computed as [43]:

$$i_t = \sigma(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_i) \tag{10}$$

$$f_t = \sigma(W_{fx}x_t + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f) \tag{11}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ g(W_{cx}x_t + W_{cm}m_{t-1} + b_c) \tag{12}$$

$$o_t = \sigma(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_t + b_o) \tag{13}$$

$$m_t = o_t \circ h(c_t) \tag{14}$$

$$y_t = W_{ym} m_t + b_y \tag{15}$$

where i_t denotes the input gate, f_t denotes the forget gate, c_t denotes the activation vectors for each cell, o_t denotes the output gate, m_t denotes the activation vectors for each memory block, W denotes the weigh matrices, b denotes the bias vectors and '$\circ$' denotes the scalar product.

$\sigma(\cdot)$ is the standard logistic function as:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{16}$$

$g(\cdot)$ is the centered logistic function as:

$$g(x) = \frac{4}{1 + e^{-x}} - 2 \quad x \in [-2,2] \tag{17}$$

$h(\cdot)$ is the centered logistic function as:

$$h(x) = \frac{2}{1 + e^{-x}} - 1 \quad x \in [-1,1] \tag{18}$$

2.4. Elman neural network

The Elman neural network, which has the structure of feed-forward connections, is a simple kind of a RNN model. It can store the output information from the hidden layer and relay the previous state in the next iteration, so it has a short-term memory [44]. Since the wind speed time series often have short-term dependency, the Elman neural network is a good choice for the wind speed forecasting.

The computation of the Elman neural network can be shown as:

$$h_t = \sigma_h(W_h x_t + U_h h_{t-1} + b_h) \tag{19}$$

$$y_t = \sigma_y(W_y h_t + b_y) \tag{20}$$

where x_t represents the historical input data, h_t represents the hidden layer vector, y_t represents the output data, W and U represent the weigh matrices, b represents the bias vectors, and σ_h and σ_y represent the activation functions.

2.5. Precision estimating indexes

To assess the prediction performance of the involved models, three error measures, which include the MAPE (*Mean Absolute Percentage Error*), the MAE (*Mean Absolute Error*) and the RMSE (*Root Mean Square Error*), are all utilized in the forecasting experiments.

These indexes can be defined as:

$$MAE = \left(\sum_{t=1}^{N} |X(t) - \hat{X}(t)| \right) / N \tag{21}$$

$$MAPE = \left(\sum_{t=1}^{N} |(X(t) - \hat{X}(t))/\hat{X}(t)| \right) / N \tag{22}$$

$$RMSE = \sqrt{\left(\sum_{t=1}^{N} [X(t) - \hat{X}(t)]^2 \right) / (N-1)} \tag{23}$$

where $X(t)$ is the wind speed series, $\hat{X}(t)$ is the forecast values, and N is the number of the $X(t)$. To compare the prediction performance of the involved models, three percentage error measures are used. They can be described as follows:

$$P_{MAE} = |(MAE_1 - MAE_2)/MAE_1| \tag{24}$$

$$P_{MAPE} = |(MAPE_1 - MAPE_2)/MAPE_1| \tag{25}$$

Table 2
Error estimated results of the predictions of the wind speed series #2.

Indexes	1-step	2-step	3-step	1-step	2-step	3-step
	ARIMA			BP		
MAPE (%)	10.74	11.54	13.51	9.28	10.43	10.5
MAE (m/s)	1.02	1.10	1.28	0.88	0.99	1.00
RMSE (m/s)	1.20	1.33	1.60	1.14	1.26	1.27
	GRNN			LSTM		
MAPE (%)	7.76	10.45	12.69	6.96	9.51	11.52
MAE (m/s)	0.74	0.99	1.21	0.66	0.90	1.09
RMSE (m/s)	0.97	1.29	1.51	0.87	1.20	1.50
	Elman			EWT-BP		
MAPE (%)	6.74	8.73	10.23	3.81	6.48	13.62
MAE (m/s)	0.64	0.83	0.97	0.36	0.62	1.29
RMSE (m/s)	0.88	1.12	1.24	0.47	0.80	1.60
	EWT-Elman			WPD-LSTM-Elman		
MAPE (%)	5.20	7.28	11.26	4.61	6.62	8.51
MAE (m/s)	0.49	0.69	1.07	0.43	0.63	0.81
RMSE (m/s)	0.65	0.88	1.31	0.60	0.87	1.06
	EMD-LSTM-Elman			EWT-LSTM-Elman		
MAPE (%)	4.85	6.48	7.92	3.46	5.47	7.58
MAE (m/s)	0.46	0.62	0.75	0.33	0.52	0.72
RMSE (m/s)	0.63	0.79	0.95	0.44	0.67	0.91

Table 3
Error estimated results of the predictions of the wind speed series #3.

Indexes	1-step	2-step	3-step	1-step	2-step	3-step
	ARIMA			BP		
MAPE (%)	8.94	11.71	12.48	11.88	12.63	13.04
MAE (m/s)	0.68	0.89	0.95	0.90	0.96	0.99
RMSE (m/s)	0.98	1.19	1.24	1.16	1.19	1.25
	GRNN			LSTM		
MAPE (%)	9.31	12.05	12.76	9.39	14.59	17.40
MAE (m/s)	0.71	0.92	0.97	0.71	1.11	1.32
RMSE (m/s)	0.99	1.19	1.22	1.02	1.44	1.75
	Elman			EWT-BP		
MAPE (%)	9.15	11.78	12.35	4.39	9.70	13.96
MAE (m/s)	0.70	0.90	0.94	0.33	0.74	1.06
RMSE (m/s)	0.99	1.19	1.21	0.43	0.96	1.37
	EWT-Elman			WPD-LSTM-Elman		
MAPE (%)	6.86	10.48	13.13	5.52	8.9	11.51
MAE (m/s)	0.52	0.80	1.00	0.42	0.68	0.87
RMSE (m/s)	0.65	0.98	1.24	0.58	0.88	1.11
	EMD-LSTM-Elman			EWT-LSTM-Elman		
MAPE (%)	6.83	8.15	9.02	3.06	4.56	6.75
MAE (m/s)	0.52	0.62	0.69	0.23	0.35	0.51
RMSE (m/s)	0.71	0.85	0.94	0.29	0.44	0.65

$$P_{RMSE} = |(RMSE_1 - RMSE_2)/RMSE_1| \tag{26}$$

3. Case study

In this section, to compare the prediction performance of the hybrid EWT -LSTM-Elman model with other comparison models, the original wind speed data collected from a wind farm in China are adopted. The comparison models are listed as follows: the ARIMA model, the BP model, the GRNN model, the LSTM model, the Elman model, the EWT-BP model, the EWT-Elman model, the WPD-LSTM-Elman model, the EMD-LSTM-Elman model and the WDD-WPD-ARMA(SS)-EMD -ELM (NS)-OCM model [45].

Table 4
Error estimated results of the predictions of the wind speed series #4.

Indexes	1-step	2-step	3-step	1-step	2-step	3-step
	ARIMA			BP		
MAPE (%)	6.16	8.85	11.64	10.47	13.85	18.53
MAE (m/s)	0.53	0.76	1.00	0.90	1.19	1.59
RMSE (m/s)	0.81	1.14	1.41	1.19	1.51	2.00
	GRNN			LSTM		
MAPE (%)	7.02	9.99	12.56	7.10	11.4	12.56
MAE (m/s)	0.60	0.86	1.08	0.61	0.98	1.08
RMSE (m/s)	0.88	1.22	1.50	0.86	1.36	1.50
	Elman			EWT-BP		
MAPE (%)	7.17	10.26	12.83	4.80	6.83	10.15
MAE (m/s)	0.62	0.88	1.10	0.41	0.59	0.87
RMSE (m/s)	0.87	1.21	1.46	0.57	0.77	1.15
	EWT-Elman			WPD-LSTM-Elman		
MAPE (%)	7.53	7.84	10.58	5.46	9.93	13.30
MAE (m/s)	0.65	0.67	0.91	0.47	0.85	1.14
RMSE (m/s)	0.85	0.86	1.28	0.64	1.13	1.50
	EMD-LSTM-Elman			EWT-LSTM-Elman		
MAPE (%)	6.62	10.04	13.32	3.24	5.86	8.55
MAE (m/s)	0.57	0.86	1.15	0.28	0.50	0.74
RMSE (m/s)	0.74	1.12	1.44	0.37	0.67	0.98

3.1. Wind speed time series

Four sets of original wind speed series including 700 samples are shown in Figs. 3–6. The 1st-600th samples of each wind speed series will be used as the training dataset, while the leaving 601st-700th samples of each wind speed series will be used as the testing dataset.

3.2. Simulation

This section gives the modeling procedure of the proposed model based on the original wind speed time series #1. Fig. 7 depicts the framework of the proposed model in case one. Fig. 8 illustrates the EWT results of the original wind speed time series #1. Fig. 9 displays the structure of multi-step forecasting strategy. Figs. 10–12 show the prediction results of the original wind speed time series #1 by adopting the EWT-LSTM-Elman model. Tables 1–4 give the error estimated results of the predictions for the wind speed series #1, #2, #3 and #4.

3.3. The comparisons and analysis

Figs. 13–21 demonstrate the comparing results between the EWT -LSTM-Elman model and the other nine models.

From Tables 1–4 and Figs. 13–21, it can be found that:

(a) The prediction accuracy of the EWT-LSTM-Elman model is higher than that of the ARIMA model significantly. For example, in case two, the promoting MAPE percentages of the ARIMA model in the multi-step results by the EWT-LSTM-Elman model are 67.78%, 52.60% and 43.89%, respectively; the promoting MAE percentages of the ARIMA model in the multi-step results by the EWT-LSTM-Elman model are 67.65%, 52.73% and 43.75%, respectively; and the promoting RMSE percentages of the ARIMA model in the multi-step results by the EWT-LSTM-Elman model are 63.33%, 49.62% and 43.13%, respectively.

(b) The prediction accuracy of the EWT-LSTM-Elman model is higher than that of the BP model significantly. For example, in case two, in case two, the promoting MAPE percentages of the BP model in the multi-step results by the EWT-LSTM-Elman model are 62.72%, 47.56% and 27.81%, respectively; the promoting MAE percentages of the BP model in the multi-step results by the EWT-LSTM-Elman model are 62.50%, 47.47% and 28.00%, respectively; and the promoting RMSE percentages of the BP model in the multi-step results by the EWT-LSTM-Elman model are 61.40%, 46.83% and 28.35%, respectively.

(c) The prediction accuracy of the EWT-LSTM-Elman model is higher than that of the GRNN model significantly. For example, in case two, the promoting MAPE percentages of the GRNN model in the multi-step results by the EWT-LSTM-Elman model are 55.41%, 47.66% and 40.27%, respectively; the promoting MAE percentages of the GRNN model in the multi-step results by the EWT-LSTM-Elman model are 55.41%, 47.47% and 40.50%, respectively; and

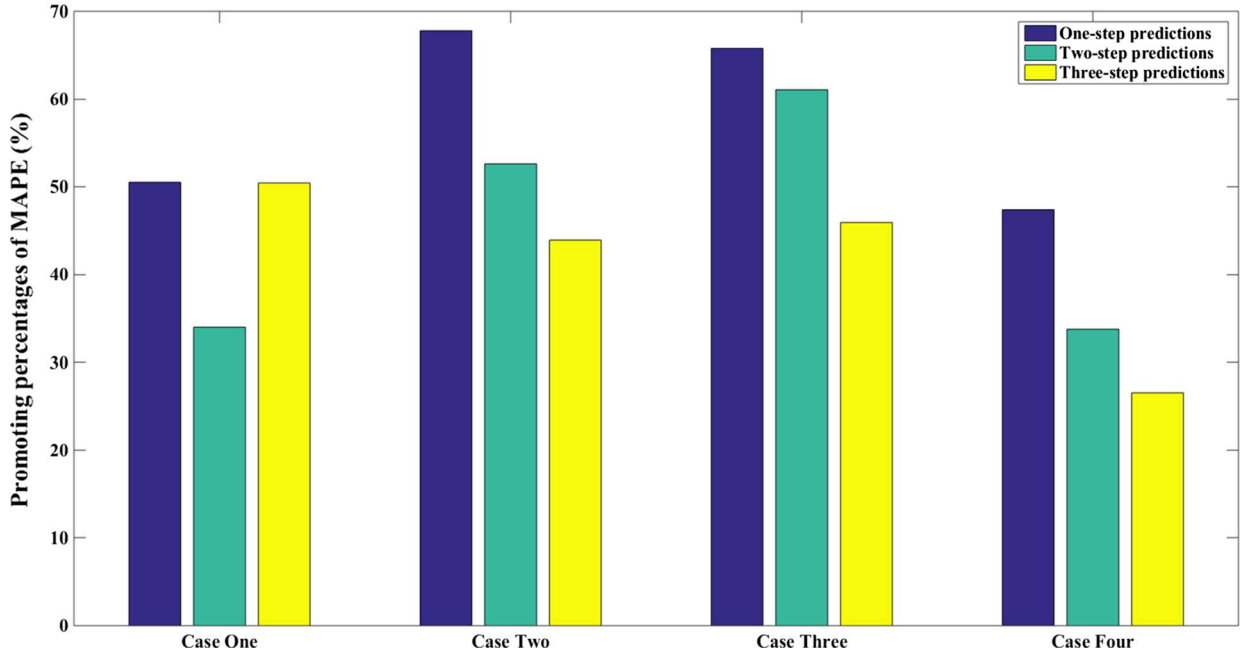

Fig. 13. The promoting percentages of the ARIMA model by the EWT-LSTM-Elman model.

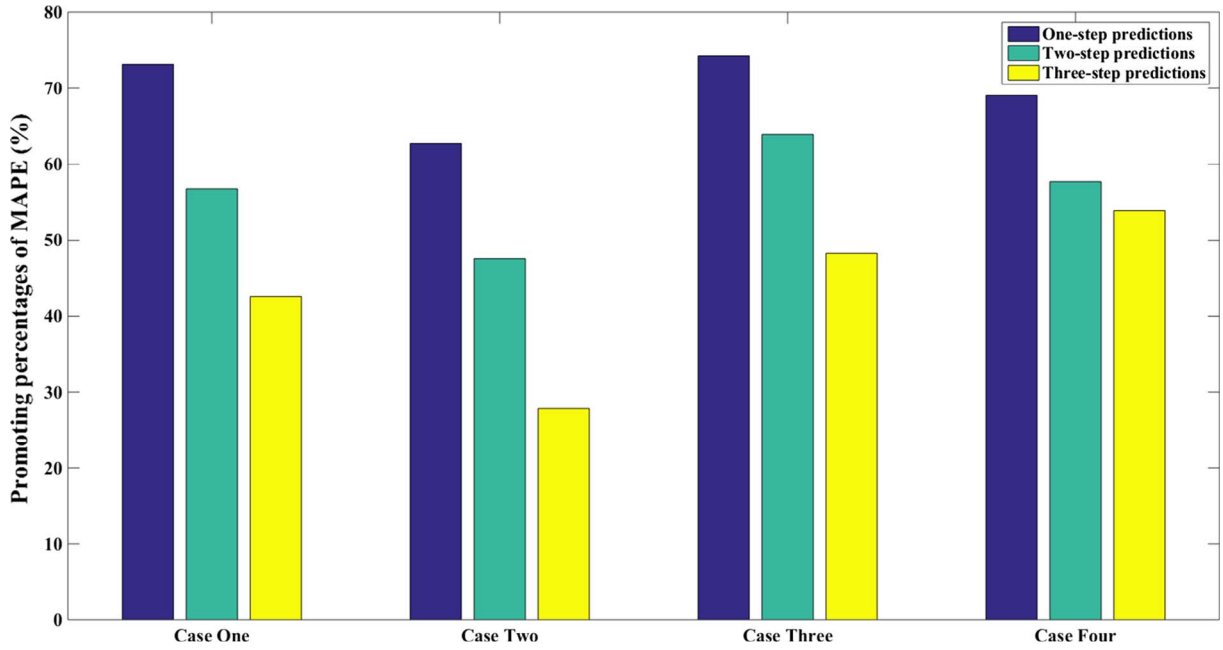

Fig. 14. The promoting percentages of the BP model by the EWT-LSTM-Elman model.

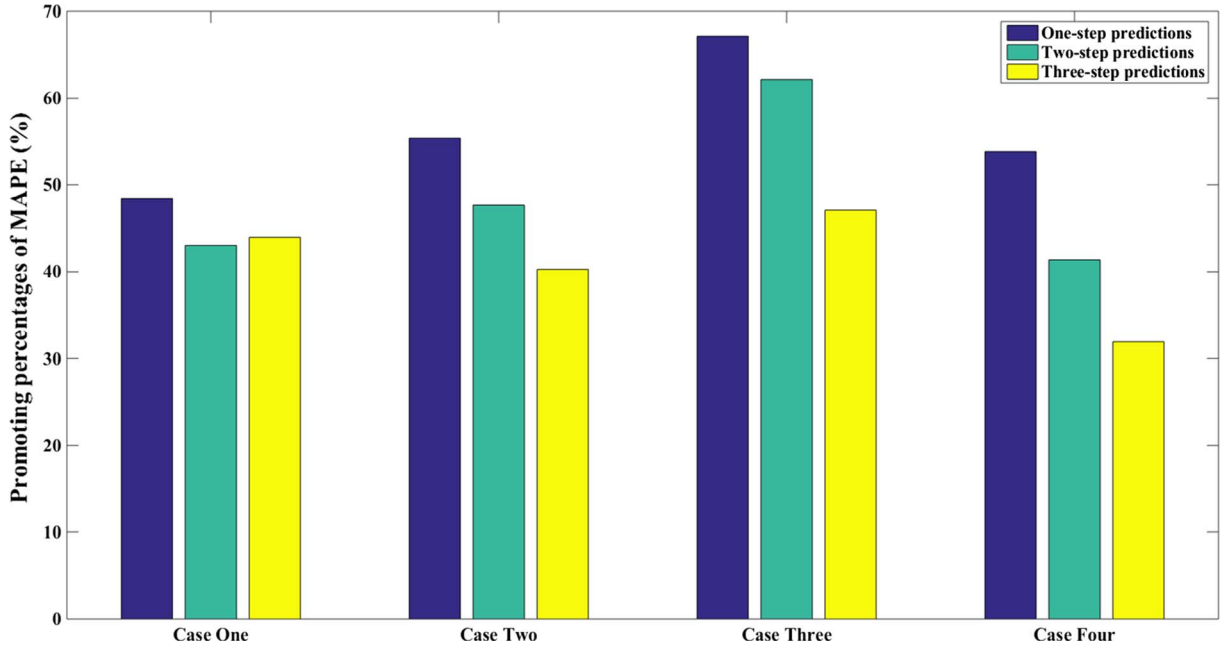

Fig. 15. The promoting percentages of the GRNN model by the EWT-LSTM-Elman model.

the promoting RMSE percentages of the GRNN model in the multi-step results by the EWT-LSTM-Elman model are 54.64%, 48.06% and 39.74%, respectively.

(d) The prediction accuracy of the EWT-LSTM-Elman model is higher than that of the LSTM model significantly. For example, in case two, the promoting MAPE percentages of the LSTM model in the multi-step results by the EWT-LSTM-Elman model are 50.29%, 42.48% and 34.20%, respectively; the promoting MAE percentages of the LSTM model in the multi-step results by the EWT-LSTM-Elman model are 50.00%, 42.22% and 33.94%, respectively; and the promoting RMSE percentages of the LSTM model in the multi-step results by the EWT-LSTM-Elman model are 49.43%, 44.17% and 39.33%, respectively.

(e) The prediction accuracy of the EWT-LSTM-Elman model is higher than that of the Elman model significantly. For example, in case two, the promoting MAPE percentages of the Elman model in the multi-step results by the EWT-LSTM-Elman model are 48.66%, 37.34% and 25.90%, respectively; the promoting MAE percentages of the Elman model in the multi-step results by the EWT-LSTM-Elman model are 48.44%, 37.35% and 25.77%, respectively; and

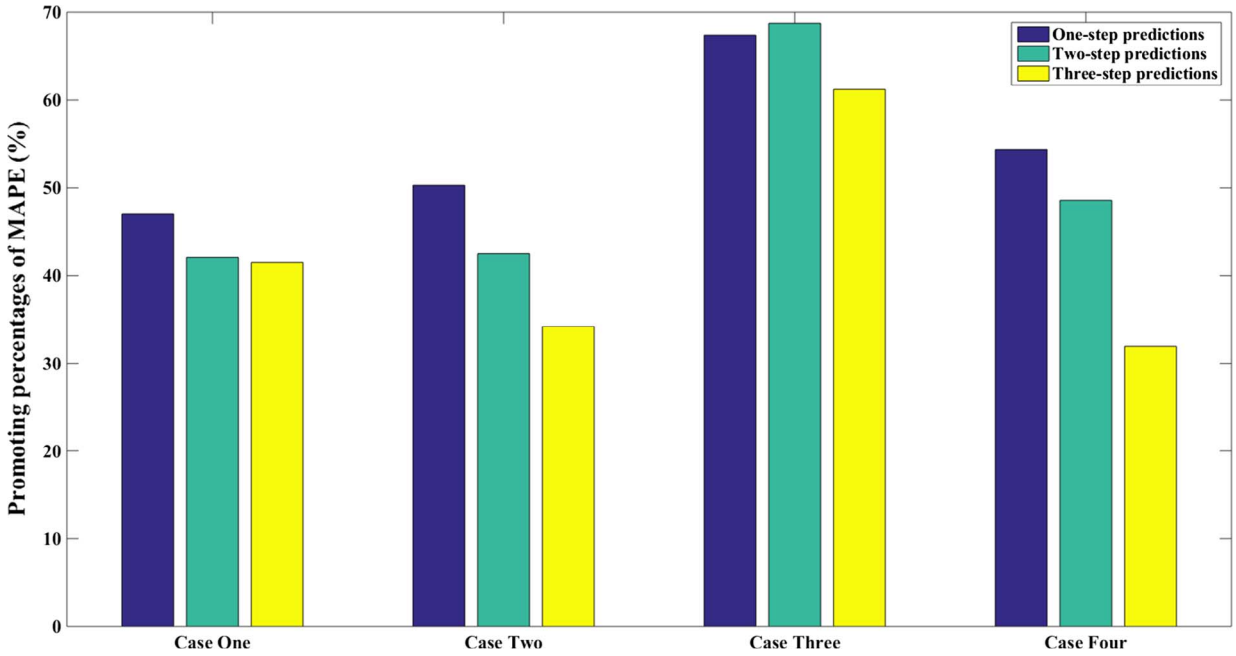

Fig. 16. The promoting percentages of the LSTM model by the EWT-LSTM-Elman model.

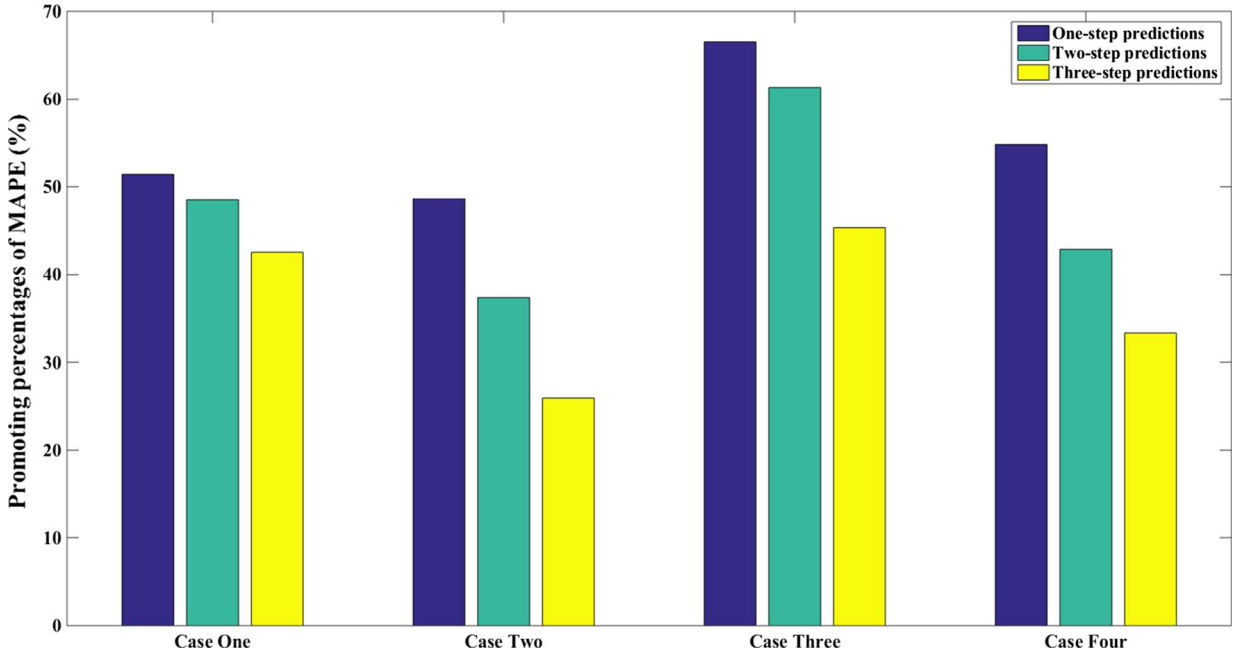

Fig. 17. The promoting percentages of the Elman model by the EWT-LSTM-Elman model.

the promoting RMSE percentages of the Elman model in the multi-step results by the EWT-LSTM-Elman model are 50.00%, 40.18% and 26.61%, respectively.

(f) The prediction accuracy of the EWT-LSTM-Elman model is higher than that of the EWT-BP model significantly. For example, in case two, the promoting MAPE percentages of the EWT-BP model in the multi-step results by the EWT-LSTM-Elman model are 9.19%, 15.59% and 44.35%, respectively; the promoting MAE percentages of the EWT-BP model in the multi-step results by the EWT-LSTM-Elman model are 8.33%, 16.13% and 44.19%, respectively; and the promoting RMSE percentages of the EWT-BP model in the multi-step results by the EWT-LSTM-Elman model are 6.39%, 16.25% and 43.13%, respectively.

(g) The prediction accuracy of the EWT-LSTM-Elman model is higher than that of the EWT-Elman model significantly. For example, in case two, the promoting MAPE percentages of the EWT-Elman model in the multi-step results by the EWT-LSTM-Elman model are 33.46%, 24.86% and 32.68%, respectively; the promoting MAE

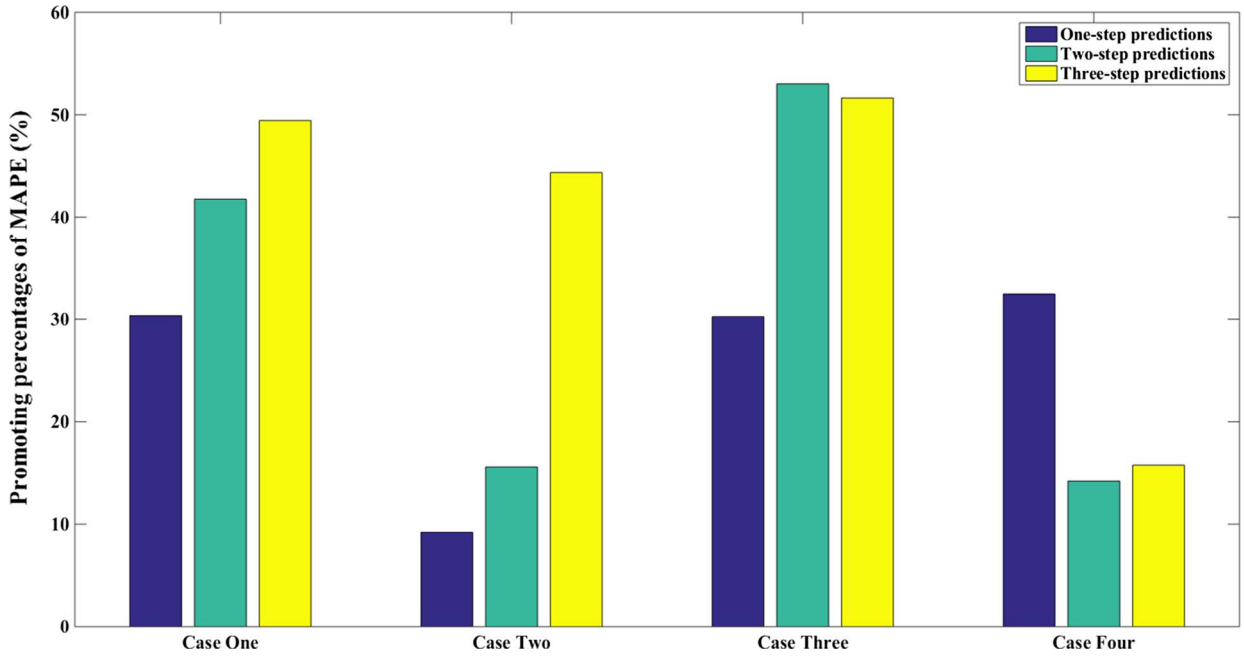

Fig. 18. The promoting percentages of the EWT-BP model by the EWT-LSTM-Elman model.

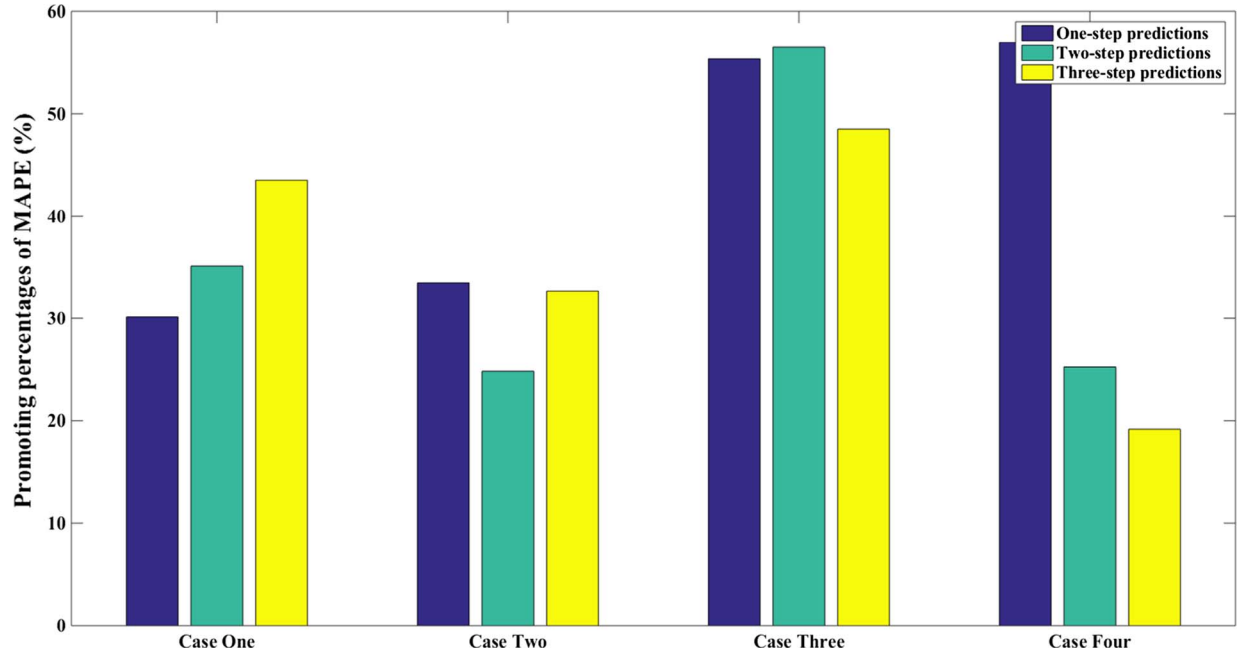

Fig. 19. The promoting percentages of the EWT-Elman model by the EWT-LSTM-Elman model.

percentages of the EWT-Elman model in the multi-step results by the EWT-LSTM-Elman model are 32.65%, 24.64% and 32.71%, respectively; and the promoting RMSE percentages of the EWT-Elman model in the multi-step results by the EWT-LSTM-Elman model are 32.31%, 23.86% and 30.53%, respectively.

(h) The prediction accuracy of the EWT-LSTM-Elman model is higher than that of the WPD-LSTM-Elman model significantly. For example, in case two, the promoting MAPE percentages of the WPD-LSTM-Elman model in the multi-step results by the EWT-LSTM-Elman model are 24.95%, 17.37% and 10.93%, respectively; the promoting MAE percentages of the WPD-LSTM-Elman model in the multi-step results by the EWT-LSTM-Elman model are 23.26%, 17.46% and 11.11%, respectively; and the promoting RMSE percentages of the WPD-LSTM-Elman model in the multi-step results by the EWT-LSTM-Elman model are 26.67%, 22.99% and 14.15%, respectively.

(i) The prediction accuracy of the EWT-LSTM-Elman model is higher than that of the EMD-LSTM-Elman model significantly. For example, in case two, the promoting MAPE percentages of the EMD-LSTM-Elman model in the multi-step results by the EWT-LSTM-

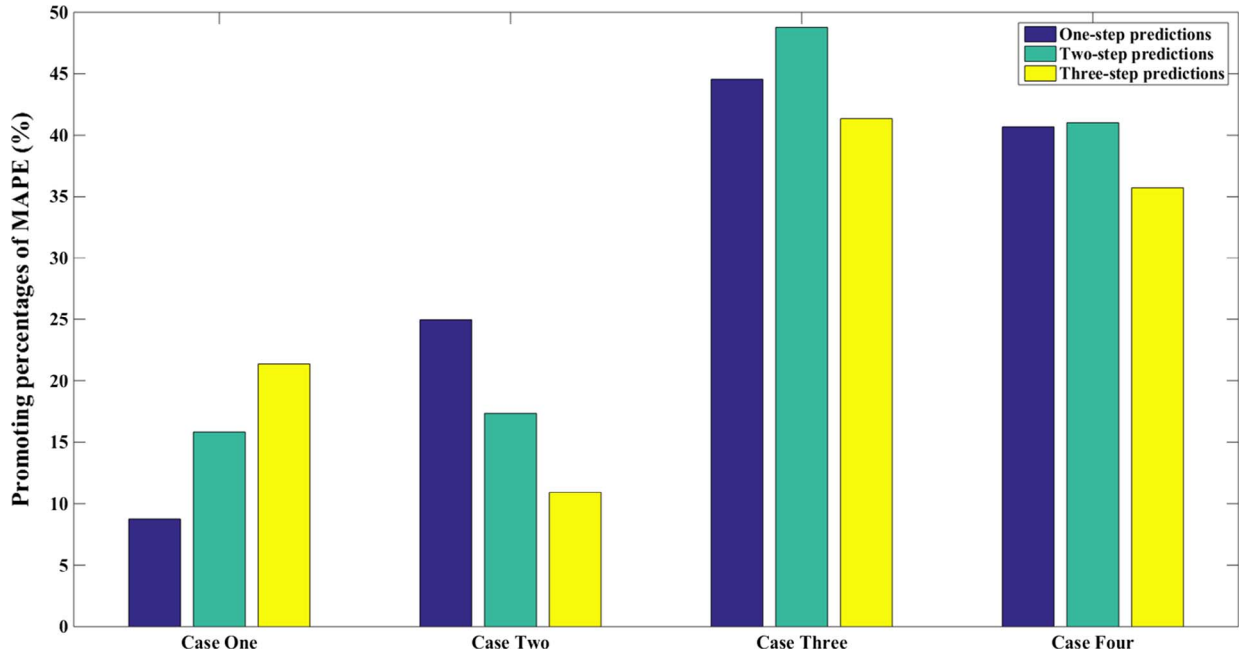

Fig. 20. The promoting percentages of the WPD-LSTM-Elman model by the EWT-LSTM-Elman model.

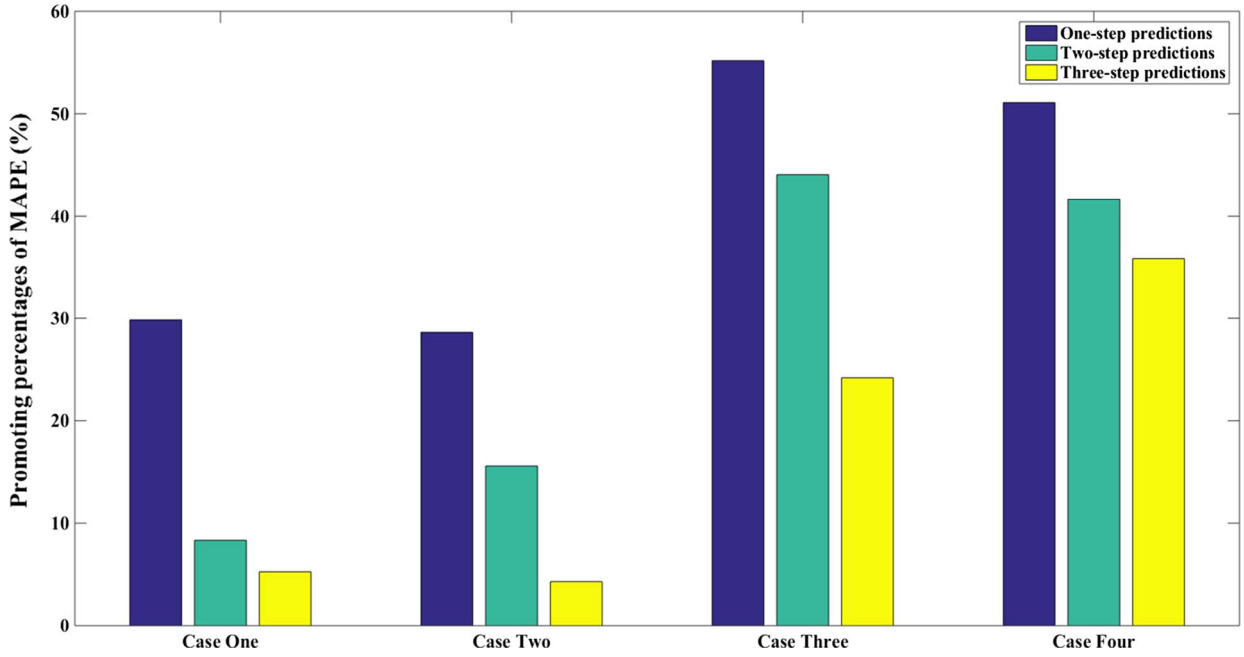

Fig. 21. The promoting percentages of the EMD-LSTM-Elman model by the EWT-LSTM-Elman model.

Elman model are 28.66%, 15.59% and 4.29%, respectively; the promoting MAE percentages of the EMD-LSTM-Elman model in the multi-step results by the EWT-LSTM-Elman model are 28.26%, 16.13% and 4.00%, respectively; and the promoting RMSE percentages of the EMD-LSTM-Elman model in the multi-step results by the EWT-LSTM-Elman model are 30.16%, 15.19% and 4.21%, respectively.

(j) Among all the aforementioned wind speed prediction models, the EWT-LSTM-Elman model has the best estimated performance. The reasons of such a significant forecasting promotion can be explained as: Firstly, comparing to the WPD/EMD decomposing algorithms, the EWT has better decomposing adaption due to its combination of the empirical ability and the wavelet computation. So the EWT can provide the deeper decomposing wind speed sub-layers for the latter forecasting computation; secondly, in the proposed hybrid forecasting model, the LSTM neural network is adopted to forecast the low-frequency wind speed sub-layers. The LSTM neural network is an improved model of the RNN network, which has all of the advantages of the deep learning RNN network but at the same time overcome the vanishing gradient problem of

Table 5
The prediction results of the WDD-WPD-ARMA(SS)-EMD-ELM(NS)-OCM model for the wind speed series #1, series #2, series #3 and series #4.

Indexes	WDD-WPD-ARMA(SS)-EMD-ELM(NS)-OCM					
	Case one (Series #1)			Case two (Series #2)		
	1-step	2-step	3-step	1-step	2-step	3-step
MAPE (%)	4.43	4.74	5.19	4.67	4.99	5.75
MAE (%)	0.58	0.62	0.68	0.44	0.47	0.55
RMSE (%)	0.79	0.83	0.90	0.57	0.62	0.72
	Case Three (Series #3)			Case Four (Series #4)		
	1-step	2-step	3-step	1-step	2-step	3-step
MAPE (%)	6.68	6.89	7.97	7.04	8.29	10.49
MAE (%)	0.51	0.52	0.61	0.61	0.71	0.90
RMSE (%)	0.68	0.73	0.81	0.78	0.92	1.11

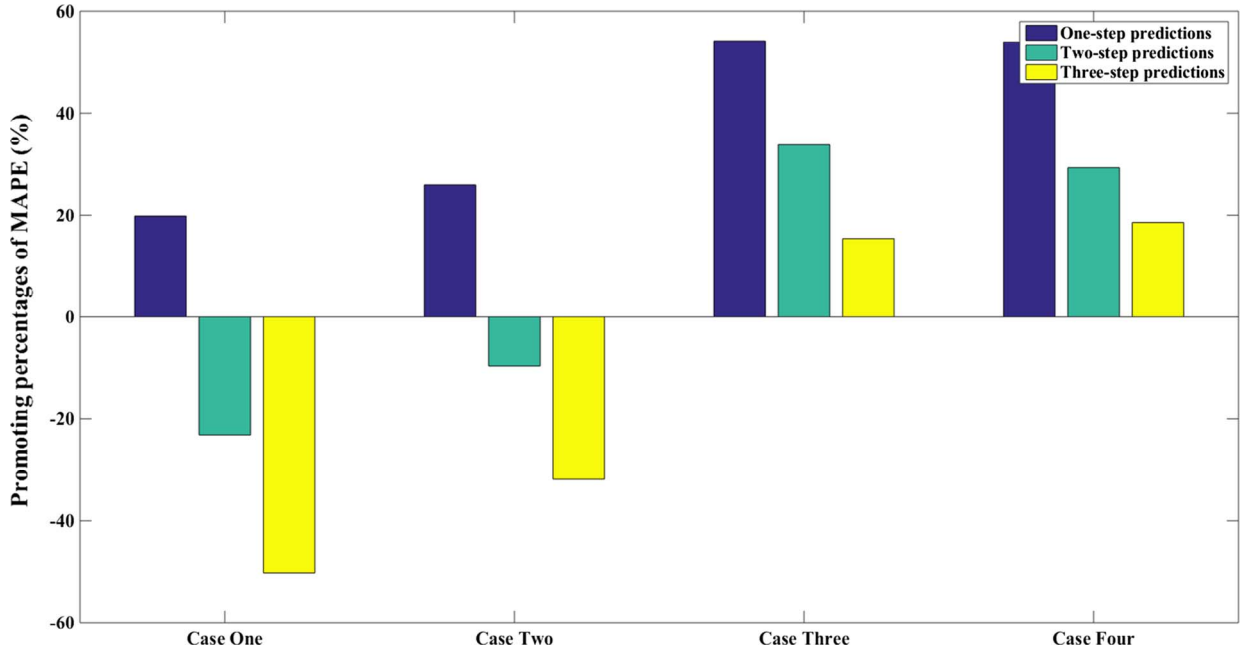

Fig. 22. The promoting percentages of the WDD-WPD-ARMA(SS)-EMD -ELM(NS)-OCM model by the EWT-LSTM-Elman model.

the RNN network. So the LSTM network has the strong nonlinear processing ability and it is suitable for the non-stationary wind speed forecasting computation; thirdly, in the proposed hybrid forecasting model, the Elman neural network is built to forecast the high-frequency wind speed sub-layers. The Elman network is also one kind of the RNN networks so that it has the good performance in the memory and the nonlinear data processing. Based on the successful combination of these upper algorithms, the proposed EWT-LSTM-Elman model has the satisfactory multi-step forecasting results.

3.4. Additional comparisons

To further investigate the performance of the proposed model, a comparison of the WDD-WPD-ARMA(SS)-EMD-ELM(NS)-OCM model in Ref. [45] and the proposed model is provided. Table 5 gives the prediction results of the WDD-WPD -ARMA(SS)-EMD-ELM(NS)-OCM model for the wind speed series #1, series #2, series #3 and series #4, respectively. Fig. 22 demonstrates the comparing results of the EWT-LSTM-Elman model and the WDD-WPD-ARMA(SS)-EMD-ELM(NS)-OCM model.

From Tables 1–5 and Fig. 22, it can be observed that: both of the two models have good prediction performance.

4. Conclusions

In this paper, a novel hybrid wind speed prediction model based on the empirical wavelet transform, the long short memory network and the Elman neural network is proposed. In the proposed hybrid EWT-LSTM-Elman model, the EWT is adopted to decompose the raw wind speed data into several sub-layers, and the LSTM network and the Elman neural network are employed to predict the low-frequency sub-layer and high-frequency sub-layers, respectively. To validate the prediction performance of the proposed model, eleven mainstream forecasting models are included as: the ARIMA model, the BP model, the GRNN model, the LSTM model, the Elman model, the EWT-BP model, the EWT-Elman model, the WPD-LSTM-Elman model, the EMD-LSTM-Elman model, the EWT-LSTM-Elman model and the WDD-WPD -ARMA(SS)-EMD-ELM(NS)-OCM model. Based on the real forecasting cases, it can be concluded that the proposed model has satisfactory performance in the wind speed multiple forecasting.

Acknowledgements

This study is fully supported by the National Natural Science Foundation of China (Grant No. 51308553), the Shenghua Yu-ying Talents Program of the Central South University (*Principle Investigator: Dr. Hui Liu*), the innovation driven project of the Central South University (*Project No. 502501002, Principle Investigator: Dr. Hui Liu*), the engineering college 'double first-rate' supporting project of the Central South University and the Changsha Outstanding Innovative Youth Project (*Principle Investigator: Dr. Hui Liu*). This study is partly supported by the National Natural Science Foundation of China (Grant No. U1534210, U1334205).

References

[1] Chang T, Liu F, Ko H, Huang M. Oscillation characteristic study of wind speed, global solar radiation and air temperature using wavelet analysis. Appl Energy 2017;190:650–7.
[2] Jiang Y, Song Z, Kusiak A. Very short-term wind speed forecasting with Bayesian structural break model. Renew Energy 2013;50:637–47.
[3] Kou P, Liang D, Gao F, Gao L. Probabilistic wind power forecasting with online model selection and warped gaussian process. Energy Convers Manage 2014;84:649–63.
[4] Zuluaga CD, Álvarez MA, Giraldo E. Short-term wind speed prediction based on robust Kalman filtering: an experimental comparison. Appl Energy 2015;156:321–30.
[5] Wang J, Xiong S. A hybrid forecasting model based on outlier detection and fuzzy time series – a case study on Hainan wind farm of China. Energy 2014;76:526–41.
[6] Su Z, Wang J, Lu H, Zhao G. A new hybrid model optimized by an intelligent optimization algorithm for wind speed forecasting. Energy Convers Manage 2014;85:443–52.
[7] Wang JZ, Wang Y, Jiang P. The study and application of a novel hybrid forecasting model – a case study of wind speed forecasting in China. Appl Energy 2015;143:472–88.
[8] Akçay H, Filik T. Short-term wind speed forecasting by spectral analysis from long-term observations with missing values. Appl Energy 2017;191:653–62.
[9] 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.
[10] Lydia M, Suresh Kumar S, Immanuel Selvakumar A, Edwin Prem Kumar G. Linear and non-linear autoregressive models for short-term wind speed forecasting. Energy Convers Manage 2016;112:115–24.
[11] Ait Maatallah O, Achuthan A, Janoyan K, Marzocca P. Recursive wind speed forecasting based on Hammerstein Auto-Regressive model. Appl Energy 2015;145:191–7.
[12] Zhou Y, Huang M. Lithiumion batteries remaining useful life prediction based on a mixture of empirical mode decomposition and ARIMA model. Microelectron Reliab 2016;65:265–73.
[13] Wu X, et al. A study of single multiplicative neuron model with nonlinear filters for hourly wind speed prediction. Energy 2015;88:194–201.
[14] Ren C, An N, Wang J, Li L, Hu B, Shang D. Optimal parameters selection for BP neural network based on particle swarm optimization: a case study of wind speed forecasting. Knowl-Based Syst 2014;56:226–39.
[15] Santamaría-Bonfil G, Reyes-Ballesteros A, Gershenson C. Wind speed forecasting for wind farms: a method based on support vector regression. Renew Energy 2016;85:790–809.
[16] Zhang C, Wei H, Xie L, Shen Y, Zhang K. Direct interval forecasting of wind speed using radial basis function neural networks in a multi-objective optimization framework. Neurocomputing 2016;205:53–63.
[17] Ata R. Artificial neural networks applications in wind energy systems: a review. Renew Sustain Energy Rev 2015;49:534–62.
[18] Sun S, Qiao H, Wei Y, Wang S. A new dynamic integrated approach for wind speed forecasting. Appl Energy 2017;197(June):151–62. June 2016.
[19] 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.
[20] Wang Y, Wu L. On practical challenges of decomposition-based hybrid forecasting algorithms for wind speed and solar irradiation. Energy 2016;112:208–20.
[21] 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.
[22] Liu H, qi Tian H, fei Li Y. An EMD-recursive ARIMA method to predict wind speed for railway strong wind warning system. J Wind Eng Ind Aerodyn 2015;141:27–38.
[23] 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;94:629–36.
[24] Liu H, Tian H qi, Li Y fei. 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.
[25] Gilles J. Empirical wavelet transform. IEEE Trans Signal Process 2013;61(16):3999–4010.
[26] Hu J, Wang J. Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression. Energy 2015;93:1456–66.
[27] Kuremoto T, Kimura S, Kobayashi K, Obayashi M. Time series forecasting using a deep belief network with restricted Boltzmann machines. Neurocomputing 2014;137:47–56.
[28] Kim Y. Convolutional neural networks for sentence classification. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP 2014); 2014. p. 1746–1751.
[29] Visin F, Kastner K, Cho K, Matteucci M, Courville A, Bengio Y. ReNet: a recurrent neural network based alternative to convolutional networks; 2015. p. 1–9.
[30] Chen Y, Lin Z, Zhao X, Wang G, Gu Y. Deep learning-based classification of hyperspectral data. IEEE J Sel Top Appl Earth Obs Remote Sens 2014;7(6):2094–107.
[31] Wang HZ, Wang GB, Li GQ, Peng JC, Liu YT. Deep belief network based deterministic and probabilistic wind speed forecasting approach. Appl Energy 2016;182:80–93.
[32] Wang H, Li G, Wang G, Peng J, Jiang H, Liu Y. Deep learning based ensemble approach for probabilistic wind power forecasting. Appl Energy 2017;188:56–70.
[33] Shi X, Chen Z, Wang H, Yeung D-Y, Wong W, Woo W. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. Adv Neural Inform Process Syst 2015;28:802–10.
[34] 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.
[35] Zhang W, Qu Z, Zhang K, Mao W, Ma Y, Fan X. A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. Energy Convers Manage 2017;136:439–51.
[36] Xiao L, Qian F, Shao W. Multi-step wind speed forecasting based on a hybrid forecasting architecture and an improved bat algorithm. Energy Convers Manage 2017;143:410–30.
[37] Pandit M, Chaudhary V, Dubey HM, Panigrahi BK. Multi-period wind integrated optimal dispatch using series PSO-DE with time-varying Gaussian membership function based fuzzy selection. Int J Electr Power Energy Syst 2015;73:259–72.
[38] Jiang P, Wang Y, Wang J. Short-term wind speed forecasting using a hybrid model. Energy 2017;119:561–77.
[39] 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.
[40] Zahraee SM, Khalaji Assadi M, Saidur R. Application of artificial intelligence methods for hybrid energy system optimization. Renew Sustain Energy Rev 2016;66:617–30.
[41] Hu J, Wang J, Xiao L. A hybrid approach based on the Gaussian process with t -observation model for short-term wind speed forecasts. Renew Energy 2017;114:670–85.
[42] Cao H, Fan F, Zhou K, He Z. Wheel-bearing fault diagnosis of trains using empirical wavelet transform. Measurement 2016;82:439–49.
[43] Liu H, Tian H qi, Liang X feng, Li Y fei. Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks. Appl Energy 2015;157:183–94.
[44] 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.
[45] Mi X, Liu H, Li Y. Wind speed forecasting method using wavelet, extreme learning machine and outlier correction algorithm. Energy Convers Manage 2017;151(September):709–22.