Energy Conversion and Management 166 (2018) 120–131

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network

Hui Liu*, Xiwei Mi, Yanfei Li

Institute of Artificial Intelligence & Robotics (IAIR), Key Laboratory of 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 prediction model
Wavelet packet decomposition
Convolutional neural network
Convolutional long short term memory network
Deep learning

ABSTRACT

High precision and reliable wind speed forecasting is important for the management of the wind power. This paper develops a novel wind speed prediction model based on the WPD (*Wavelet Packet Decomposition*), CNN (*Convolutional Neural Network*) and CNNLSTM (*Convolutional Long Short Term Memory Network*). In the proposed WPD-CNNLSTM-CNN model, the WPD is employed to decompose the original wind speed time series into a number of sub-layers; the CNN with 1D convolution operator is used to forecast the obtained high-frequency sub-layers; and the CNNLSTM is adopted to complete the forecasting of the low-frequency sub-layer. To verify and compare the prediction performance of the proposed model, eight models are used. According to the results of four experimental tests, it can be observed that: (1) the proposed model is robust and effective in predicting the 1D wind speed time series, besides, among the involved eight models, the proposed model can perform best in wind speed 1-step to 3-step predictions; (2) when the wind speed experiences sudden change, the proposed model can have better prediction performance than the other involved models.

1. Introduction

Along with the development of the world economy, the energy demand has become one of the most important topics in the world. The past years have witnessed the phenomenon that there are some prevalent problems in the conventional resources, for example, the energy scarcity and environmental pollution. In order to overcome these problems, renewable energy is needed. In recent years, wind energy, as a significant and promising renewable energy, has been developed rapidly [1]. However, the high penetration of the wind power can bring challenges to the stability of the power system operation [2]. The wind power generation is strongly related with the inputting wind speed. Accurate wind speed predictions can benefit the management of the wind power and help reduce the unexpected effects on the power system [3]. If the multi-step high-accuracy wind speed forecasted values can be obtained, it is good for the power dispatching department to control the wind power systems scientifically. So the wind speed forecasting technique is desired in the wind power engineering. Nevertheless, due to the inherently intermittent of the wind, it is difficult to predict the wind speed accurately. Therefore, the wind speed forecasting has attracted much attention of the scholars.

Over the past few years, various wind speed prediction models have been developed. These models can be classified into three types based on the time horizons: the short-term prediction models, medium-term prediction models and long-term prediction models [4]. Generally, in order to establish the wind speed prediction models, various methods can be used. These methods can be classified into two types: the prediction methods and optimization methods. The prediction methods are used as the predictors, and they mainly include the physical methods, conventional machine learning methods, and deep learning methods. The optimization methods are used to improve the prediction performance of the predictors, and they mainly contain the signal processing methods and parameters optimization methods [5].

The physical methods predict the wind speed by using the physical parameters, such as the ambient temperature, atmospheric pressure and terrain conditions [6]. These methods are often suitable for the mid-long term and large-scale areas wind speed forecasting. The NWP (*Numerical Weather Prediction*) and CFD (*Computational Fluid Dynamics*) are the key technologies in these methods [7]. Recently, some valuable wind speed models based on the physical methods have been proposed. Wang et al. [8] presented an extreme wind speed non-stationary model based on the NWP. Galanis et al. [9] presented a hybrid wind speed

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

https://doi.org/10.1016/j.enconman.2018.04.021
Received 25 January 2018; Received in revised form 25 March 2018; Accepted 7 April 2018
Available online 03 May 2018
0196-8904/

prediction model by combining the Bayesian model, nonlinear Kalman filter, and NWP. Allen et al. [10] investigated a BLS (*Boundary Layer Scaling*) method for long-term wind speed forecasting.

The conventional machine learning methods predict the wind speed by using single data or multiple data. These methods are often suitable for the wind speed data which has a clear trend. The technologies of these methods mainly comprise of the persistence methods, time series methods, grey methods, Kalman methods, ANN (*Artificial Neural Networks*) methods and SVM (*Support Vector Machine*) methods. Various effective wind speed prediction models have been presented in the literature based on these technologies. Kiplangat et al. [11] demonstrated a hybrid wind speed prediction model based on the linear time series methods. Maatallah et al. [12] carried out a wind speed forecasting model by using the Hammerstein method and Autoregressive method. Akçay et al. [13] built a Kalman filter based model for wind speed forecasting. Mi et al. [14] predicted the wind speed based on the ARMA (*Auto-Regressive Moving Average*) and ELM (*Extreme Learning Machine*). Zhang et al. [15] investigated the ANN-based models and SVM based models for wind speed forecasting. Sun et al. [16] successfully introduced the RELM (*Regularized Extreme Learning Machine*) into wind speed forecasting. Shrivastava et al. [17] executed a wind speed intervals prediction model based on the SVM. Liu et al. [18] conducted a comparative study on ANN-based models for wind speed forecasting. Feng et al. [19] developed a short-term wind speed prediction model, which exploited the ensemble algorithm for combing the ANN, SVM, GBM (*Gradient Boosting Machine*) and RF (*Random Forest*).

The deep learning methods are the new branch of the machine learning methods. With the booming development of the data science, deep learning methods have been widely used in solving classification and regression problems [20]. According to the literature [21], compared with the shallow methods, the deep learning methods can have better performance in extracting the hidden natural structures and inherent abstract features of the data. Accordingly, the deep learning methods can be the promising methods in wind speed forecasting field. Recently, some deep learning methods, such as the DBN (*Deep Belief Network*) and LSTM (*Long Short Term Memory*), have been applied for establishing the wind speed prediction models. Wang et al. [22] put forward a DBN based wind speed prediction models, and their case studies verified the proposed model was accurate and stable. Hu et al. [23] successfully designed an efficient wind speed model based on the DBN and transfer learning. Liu et al. [24] put up with the LSTM based model for wind speed forecasting, and the proposed model could obtain satisfactory prediction performance.

The signal processing methods can effectively enhance the performance of the wind speed prediction models. Ordinarily, these methods can be used for data denoising, data transformation and data feature extraction. In the last years, these methods have been widely applied in wind speed forecasting. Sun et al. [25] adopted the PSR (*Phase Space Reconstruction*) as the input vectors selection method of the wind speed prediction model. Tascikaraoglu et al. [26] employed the WT (*Wavelet Transform*) for decomposing the raw wind speed data into more stationary components. Wang et al. [27] applied the WPT (*Wavelet Packet Transform*) and PSR for wind speed data feature extraction. Liu et al. [28] proposed a wind speed prediction model based on the WPD and FEEMD (*Fast Ensemble Empirical Mode Decomposition*). Hu et al. [29] used the EWT (*Empirical Wavelet Transform*) for extracting the significant information of the wind speed data. Jiang et al. [30] developed an EEMD based wind speed prediction model. Yu et al. [31] analyzed the performance of three wind speed prediction models, in which the EMD (*Empirical Mode Decomposition*), EEMD (*Ensemble Empirical Mode Decomposition*) and CEEMDAN (*Complete Ensemble Empirical Mode Decomposition with Adaptive Noise*) were exploited as the decomposition methods, while the SSA (*Singular Spectrum Analysis*) was exploited to further handle the highest frequency data. Yu et al. [32] designed a hybrid decomposition method for wind speed forecasting by combining the WT and SSA. Peng et al. [33] exploited the CEEMDAN and VMD (*Variational Mode Decomposition*) for processing the non-linearity of the wind speed data.

The parameters optimization methods consist of many technologies and can be used for optimizing the hyperparameters of the single prediction model and the weight parameters of the multiple prediction models. These methods have attracted extensive attention for they can improve the accuracy and stability of the wind speed prediction models. Xiao et al. [34] established a wind speed prediction model and utilized the improved BA (*Bat Algorithm*) and CG (*Conjugate Gradient*) for optimizing the initial weights of the GRNN (*General Regression Neural Network*). Zheng et al. [35] combined the PSO (*Particle Swarm Optimization*) and GSA (*Gravitational Search Algorithm*) for optimizing the weights and bias of the ORELM (*Outlier Robust Extreme Learning Machine*). Zhang et al. [36] used the CLSFPA (*Flower Pollination Algorithm with Chaotic Local Search*) for optimizing the weight parameters of the combined model. Jiang et al. [37] adopted the CS (*Cuckoo Search*) algorithm for tuning the parameters of the v-SVM. Zhao et al. [38] employed the CS algorithm for optimizing the fuzzy clustering.

In this study, a novel wind speed prediction model is proposed based on the WPD, CNN and CNNLSTM. The WPD has been proved to be one of the most effective data decomposing algorithms in the field of wind speed prediction. Both of the CNN and CNNLSTM are the mainstream neural networks in the field of deep learning. The originality of the study is to investigate the combination performance of the WPD and the CNN and CNNLSTM in the wind speed multi-step prediction. The detailed framework of the study is explained as follows: (a) The WPD is employed to decompose the original wind speed time series into a number of sub-layers. The purpose of executing the WPD decomposition is to decrease the non-stationarity of the original wind speed data and also to provide more sub-layer wind speed data for the deep learning prediction; (b) The CNN is used to complete the forecasting of the high-frequency sub-layers; (c) The CNNLSTM is adopted to complete the forecasting of the low-frequency sub-layer. The reason to use the CNN and CNNLSTM for the forecasting computation is to utilize their nonlinear processing capacity to obtain satisfactory wind speed forecasting results.

The main contributions of this study are demonstrated as follows: (a) A novel wind speed prediction model is proposed by combing the WPD, CNN and CNNLSTM; (b) The real performance of the CNN algorithm in the WPD based high-frequency wind speed sub-layers has not been studied before; (c) The real performance the CNNLSTM in the WPD based low-frequency wind speed sub-layers has also not been investigated before; and (d) the proposed hybrid model will be compared fully with other eight single or hybrid wind speed forecasting models. The involved comparing models include the ARIMA model, SVM model, WPD-BP model, WPD-GRNN model, WPD-Elman model, WPD-ELM model, WPD-CEEMDAN-RBF model and WPD-CNNLSTM-CNN model. The significant difference of the study to others is to only use the deep learning neural networks in the wavelet decomposing framework to realize the multi-step wind speed forecasting computation. The study will provide a useful reference how to combine the data decomposition algorithm and the deep learning algorithm for the wind speed multi-step prediction.

The contents of this study are arranged as follows. In Section 2, the framework of the proposed model and the principle of the necessary individual methods are presented. In Section 3, the proposed model and contrast models are applied in the experimental data; besides, the comparison and discussion of the prediction results are also provided. In Section 4, the paper is concluded.

2. The WPD-CNNLSTM-CNN model

2.1. The entire process of the proposed model

The framework of the WPD-CNNLSTM-CNN model is depicted in Fig. 1, and the entire process is described in detail as follows:

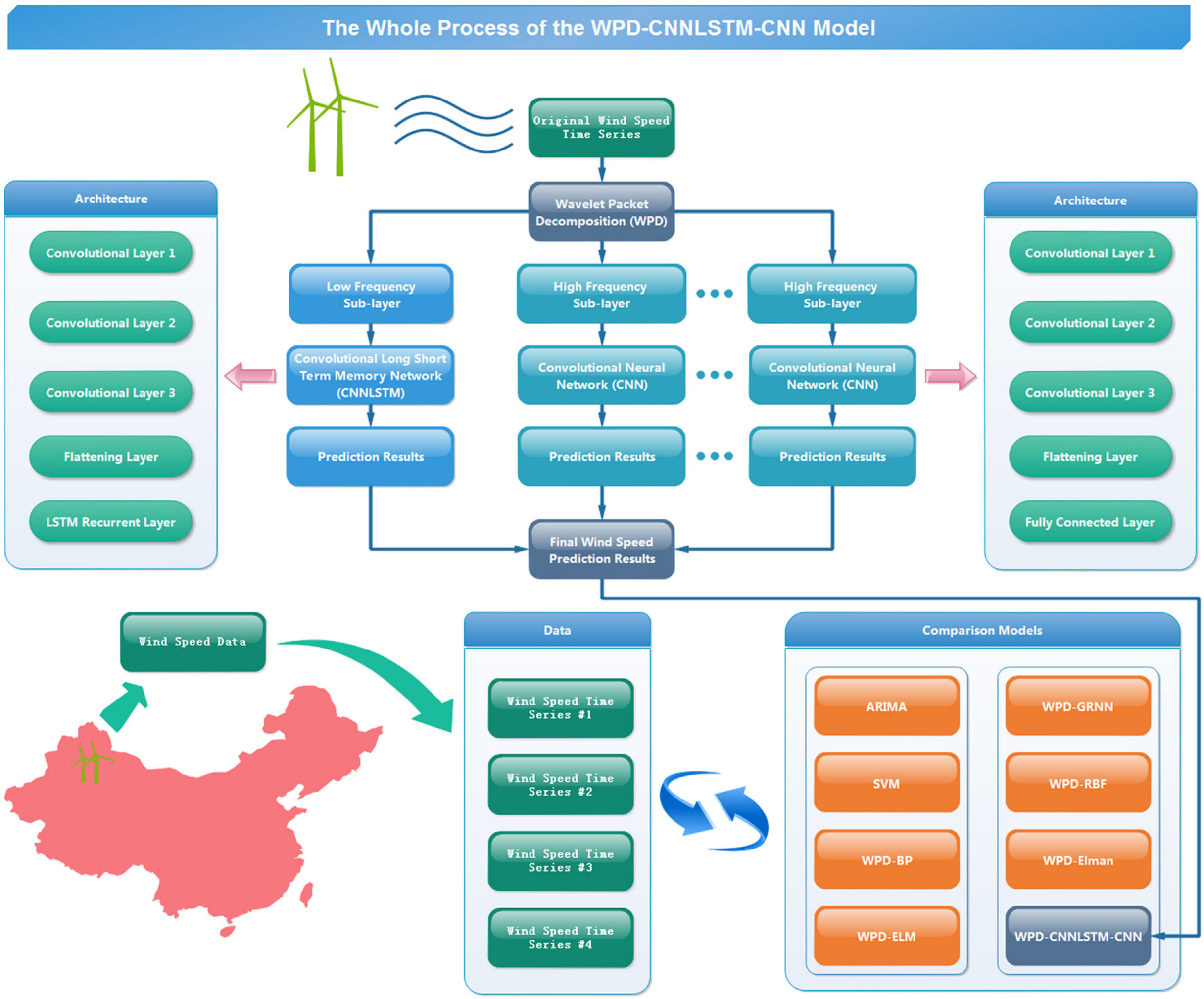

Fig. 1. The entire process of the WPD-CNNLSTM-CNN model.

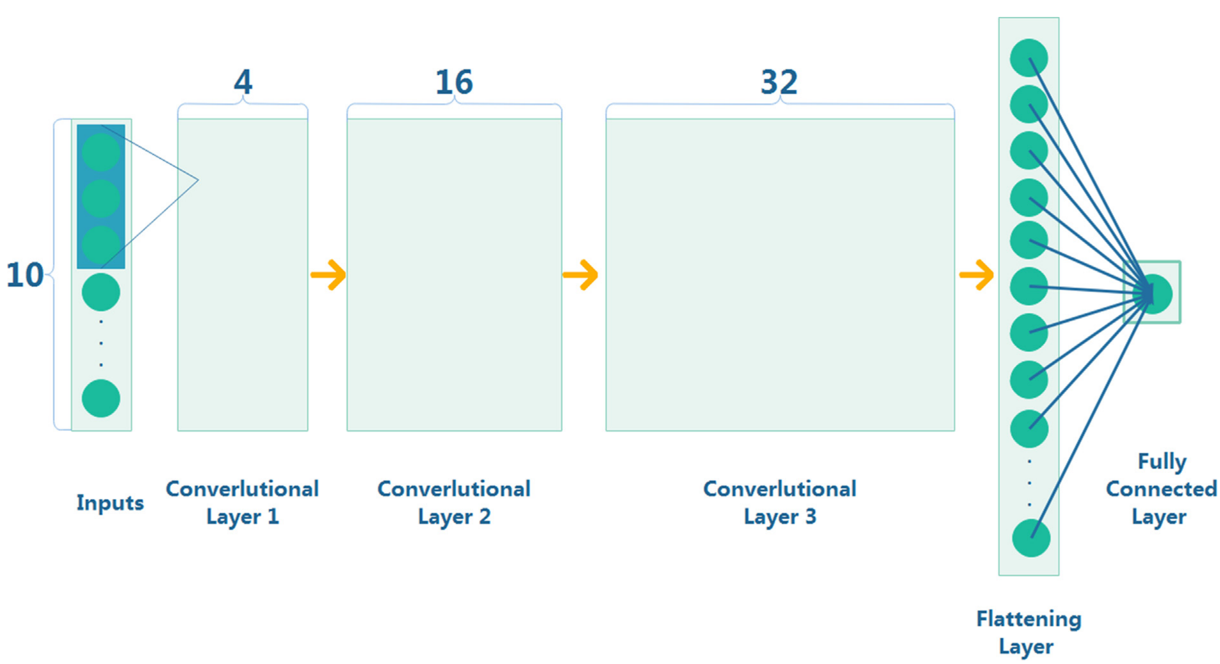

Fig. 2. The structure of the CNN.

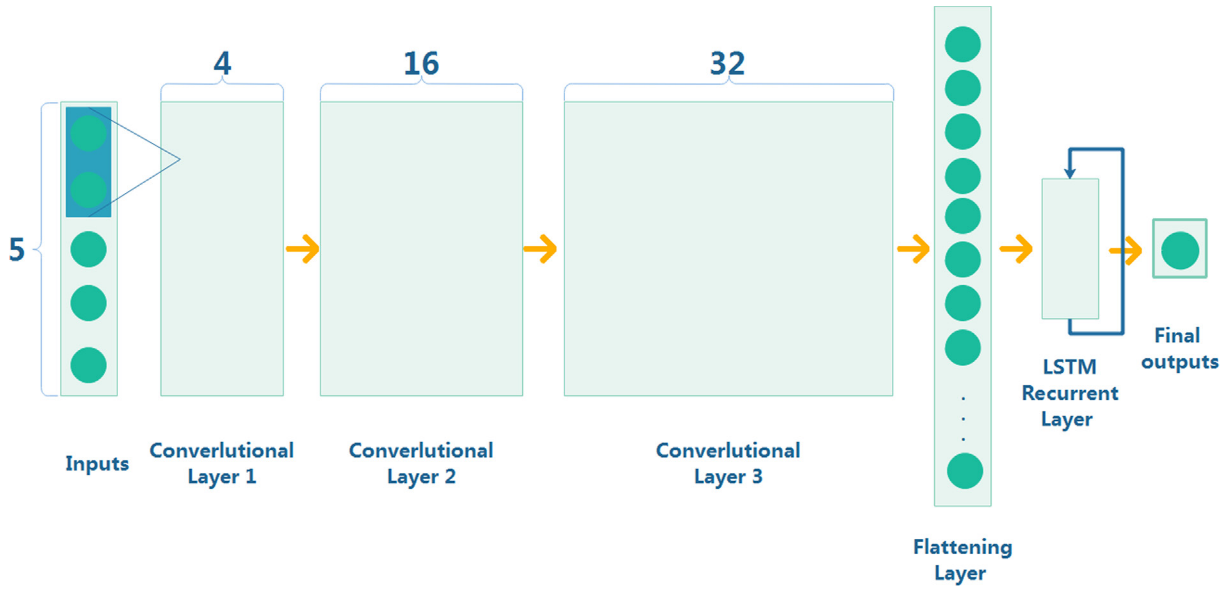

Fig. 3. The structure of the CNNLSTM.

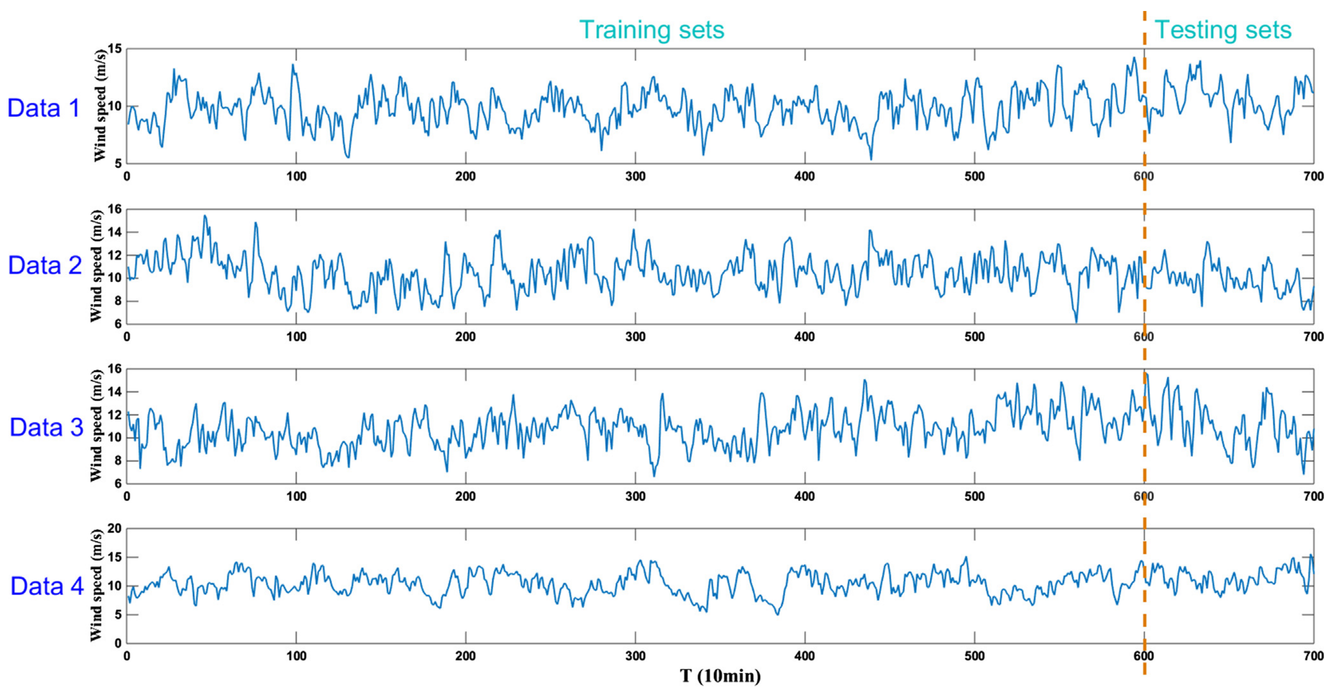

Fig. 4. Four sets of 10-min averaged wind speed time series.

Table 1
Descriptive statistics of the wind speed data.

Data	Min	Max	Mean	Standard Derivation
Data 1	5.3	14.3	9.89	1.55
Data 2	6.1	15.5	10.42	1.55
Data 3	6.6	15.6	10.81	1.62
Data 4	4.9	15.6	10.5	1.94

(1) The WPD is used to decompose the original wind speed time series into several sub-layers. *The details of the WPD method are provided in* Section 2.2.
(2) The CNN is adopted to forecast the high-frequency sub-layers. *The details of the CNN method are provided in* Section 2.3.
(3) The CNNLSTM is employed to predict the low-frequency sub-layer obtained by the WPD. *The details of the CNNLSTM method are provided in* Section 2.4.
(4) To evaluate and compare the prediction performance of the proposed model, some models are used, which consist of the ARIMA model, SVM model, WPD-BP model, WPD-GRNN model, WPD-Elman model, WPD-ELM model, WPD-CEEMDAN-RBF model and WPD-CNNLSTM-CNN model. *The details of the prediction performance criteria are presented in* Section 2.5, *and the experimental results and analysis are presented in* Section 3 and 4.

2.2. Wavelet packet decomposition

The WPD is a classical signal processing method, which can decompose the signal into the appropriate components and detailed components. The decomposition levels and wavelet basis can

Table 2
Analysis of the prediction results for the experimental test #1.

Indexes	1-step	2-step	3-step	1-step	2-step	3-step
	ARIMA			SVM		
MAPE (%)	6.88	9.72	10.47	7.43	10.10	11.64
MAE (m/s)	0.72	1.01	1.10	0.79	1.09	1.28
RMSE (m/s)	0.99	1.29	1.40	1.04	1.43	1.64
	WPD-BP			WPD-GRNN		
MAPE (%)	4.88	7.45	10.72	5.43	8.09	10.16
MAE (m/s)	0.52	0.79	1.16	0.57	0.86	1.09
RMSE (m/s)	0.67	0.98	1.41	0.71	1.08	1.38
	WPD-Elman			WPD-ELM		
MAPE (%)	4.36	7.58	11.17	3.56	6.55	9.28
MAE (m/s)	0.46	0.81	1.20	0.37	0.70	0.99
RMSE (m/s)	0.62	1.02	1.49	0.47	0.86	1.22
	WPD-CEEMDAN-RBF			WPD-CNNLSTM-CNN		
MAPE (%)	3.08	4.50	4.91	1.60	2.89	4.25
MAE (m/s)	0.33	0.48	0.52	0.17	0.31	0.47
RMSE (m/s)	0.42	0.59	0.65	0.22	0.42	0.62

Table 3
Analysis of the prediction results for the experimental test #2.

Indexes	1-step	2-step	3-step	1-step	2-step	3-step
	ARIMA			SVM		
MAPE(%)	6.90	9.50	9.96	6.97	8.52	10.01
MAE(m/s)	0.67	0.91	0.95	0.68	0.92	0.96
RMSE(m/s)	0.88	1.11	1.18	0.90	1.16	1.22
	WPD-BP			WPD-GRNN		
MAPE (%)	3.63	5.66	8.53	4.94	6.97	9.05
MAE (m/s)	0.36	0.55	0.81	0.47	0.66	0.89
RMSE (m/s)	0.51	0.72	1.00	0.61	0.85	1.07
	WPD-Elman			WPD-ELM		
MAPE (%)	5.75	8.46	9.19	3.12	5.50	7.68
MAE (m/s)	0.57	0.83	0.89	0.31	0.53	0.73
RMSE (m/s)	0.73	1.08	1.16	0.41	0.74	0.94
	WPD-CEEMDAN-RBF			WPD-CNNLSTM-CNN		
MAPE (%)	3.35	4.50	5.39	1.52	2.82	4.07
MAE (m/s)	0.32	0.43	0.51	0.15	0.28	0.40
RMSE (m/s)	0.43	0.58	0.71	0.18	0.35	0.51

Table 4
Analysis of the prediction results for the experimental test #3.

Indexes	1-step	2-step	3-step	1-step	2-step	3-step
	ARIMA			SVM		
MAPE (%)	10.07	14.30	15.64	9.82	13.29	14.41
MAE (m/s)	1.07	1.47	1.59	1.07	1.43	1.56
RMSE (m/s)	1.34	1.80	1.92	1.37	1.78	1.92
	WPD-BP			WPD-GRNN		
MAPE (%)	7.70	9.77	13.25	6.69	9.47	12.30
MAE (m/s)	0.85	1.05	1.43	0.72	1.02	1.31
RMSE (m/s)	1.06	1.32	1.78	0.93	1.26	1.59
	WPD-Elman			WPD-ELM		
MAPE (%)	6.55	8.80	11.54	4.78	8.37	10.89
MAE (m/s)	0.70	0.95	1.25	0.53	0.92	1.21
RMSE (m/s)	0.93	1.27	1.68	0.68	1.22	1.70
	WPD-CEEMDAN-RBF			WPD-CNNLSTM-CNN		
MAPE (%)	4.72	6.23	9.13	2.08	3.61	5.27
MAE (m/s)	0.50	0.67	0.98	0.22	0.39	0.58
RMSE (m/s)	0.70	0.84	1.22	0.29	0.51	0.73

profoundly affect the performance of the WPD. The WPD includes two types: the continuous wavelet transform and discrete wavelet transform. The continuous wavelet transform can be described as:

$$\mathrm{CWT}_f(a,b) = \langle f(t),\Psi_{a,b}(t)\rangle = \int_{-\infty}^{+\infty} f(t)\Psi^*((t-b)/a)/\sqrt{a}\,dt \tag{1}$$

Table 5
Analysis of the prediction results for the experimental test #4.

Indexes	1-step	2-step	3-step	1-step	2-step	3-step
	ARIMA			SVM		
MAPE (%)	8.87	12.31	13.17	8.66	12.05	13.43
MAE (m/s)	1.00	1.38	1.47	0.99	1.38	1.56
RMSE (m/s)	1.30	1.75	1.81	1.33	1.82	1.98
	WPD-BP			WPD-GRNN		
MAPE (%)	7.65	9.71	12.74	5.49	8.24	10.92
MAE (m/s)	0.88	1.11	1.42	0.63	0.95	1.26
RMSE(m/s)	1.14	1.36	1.72	0.84	1.25	1.61
	WPD-Elman			WPD-ELM		
MAPE (%)	7.66	9.21	11.72	3.82	7.80	12.18
MAE (m/s)	0.85	1.02	1.30	0.44	0.90	1.40
RMSE (m/s)	1.09	1.29	1.60	0.58	1.18	1.88
	WPD-CEEMDAN-RBF			WPD-CNNLSTM-CNN		
MAPE (%)	3.66	5.19	7.10	1.57	3.15	4.68
MAE (m/s)	0.40	0.57	0.78	0.18	0.36	0.53
RMSE (m/s)	0.57	0.76	0.99	0.22	0.47	0.67

where $f(t)$ denotes the signal, $\Psi(t)$ denotes the mother wavelet function, a denotes the scale coefficient, b denotes the translation coefficient and $*$ denotes the complex conjugate.The a and b in discrete wavelet transform can be described as:

$$\begin{cases} a = 2^j \\ b = k2^j \end{cases} \tag{2}$$

where j and k are the scale coefficient and translation coefficient, respectively.

2.3. Convolutional neural network

Generally, for the wind speed time series, the high-frequency sub-layer has the short-term dependence, while the low-frequency sub-layer has both of the long-term and short-term dependence [39]. For the data with short-term dependence, the fully connected layer can have the similar performance and faster computing speed when compared to the LSTM recurrent layer, thus in the proposed model, the CNN is adopted to predict the high-frequency sub-layers.

Extracting the hidden features of the wind speed time series is a meaningful way to improve the prediction performance. In the past years, the CNN has been proved to be a reliable technology to extract the hidden features, for it can complete the automatic creation of filters [40]. In the same convolutional layer of the CNN framework, there are no connections between the neurons, besides the weights of the filters are shared. Therefore, compared to the MLP (Multilayer Perceptron) with the same layers and neurons, the CNN can be trained more efficiently. Each convolutional layer can be depicted as [41]:

$$h_{ij}^k = f((W^k * x)_{ij} + b_k) \tag{3}$$

where f denotes the activation function, W^k denotes the weights of the kernel connected to kth feature map.

In this paper, the CNN consist of three convolutional layers and a fully connected layer, the channels of the convolutional layers are 4, 16 and 32, respectively. Besides, the *ReLU* and *Adam* [42] are used as the activation function and optimization algorithm, respectively. The *ReLU* can be defined as:

$$f(x) = \max(0,x) \tag{4}$$

In order to simplify the modeling processes and satisfy the real-time implementation of the wind speed forecasting, the 1D convolution operator is used, which can directly predict the 1D wind speed data. The structure of the CNN is illustrated in Fig. 2.

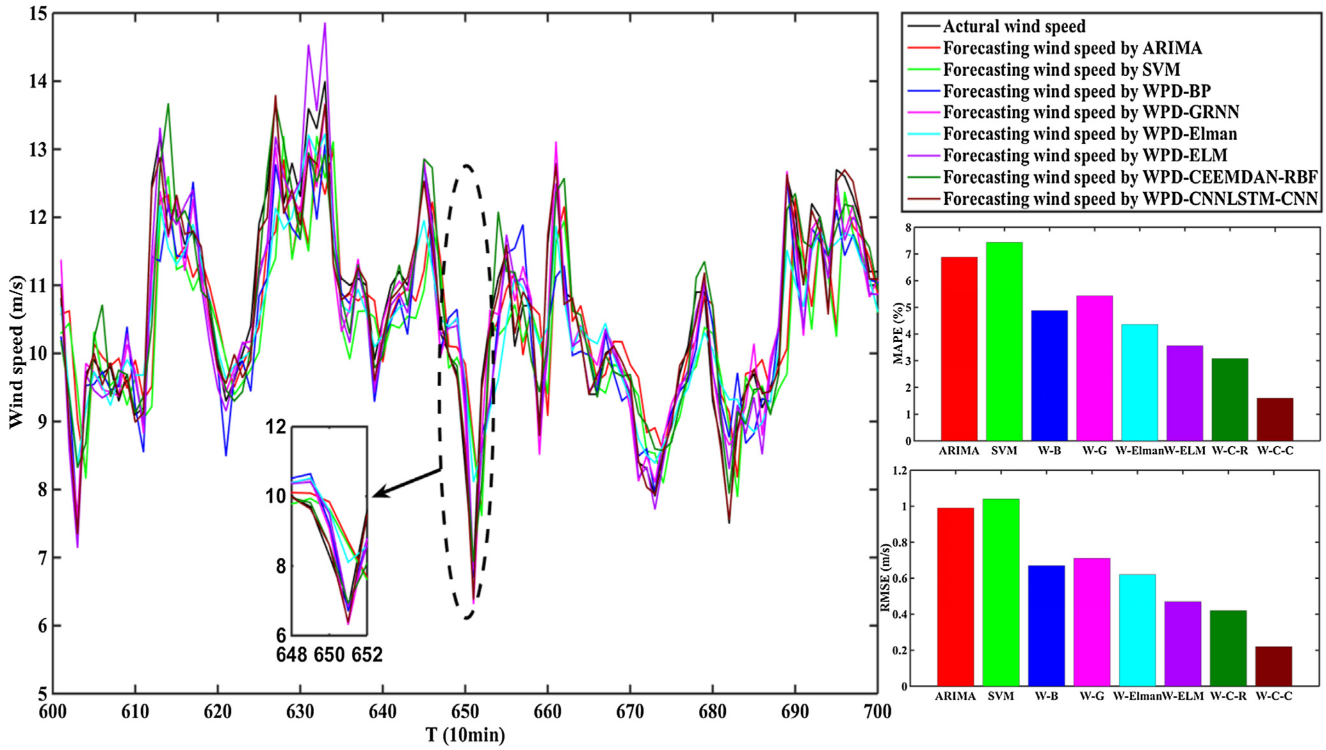

Fig. 5. The 1-step prediction results of the involved models for experimental test #1.

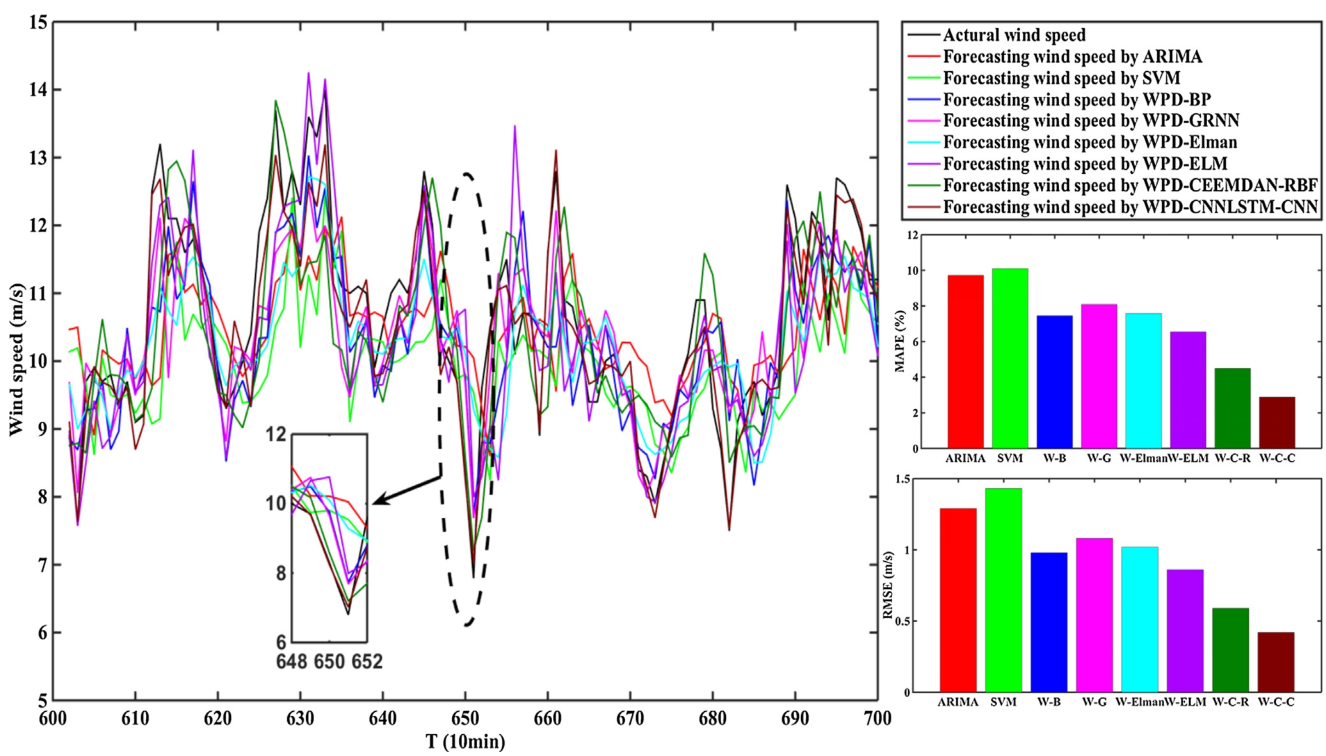

Fig. 6. The 2-step prediction results of the involved models for experimental test #1.

2.4. Convolutional long short term memory network

The CNN can effectively extract the deep features of the time series. Also, compared to the fully connected layer, the LSTM recurrent layer can better sequentially process the temporal data with long-term and short-term dependence [43]. Therefore, in the proposed model, the CNNLSTM is adopted to predict the low-frequency sub-layer obtained by the WPD.

In the CNNLSTM, the CNN components are employed to extract the high dimensional features of the wind speed low-frequency sub-layer, while the LSTM components use these features for time series forecasting. The CNN components here are similar to the above mentioned CNN in Section 2.3. In the LSTM, the memory cell is the core for processing the long-term dependence, and the cell state can be controlled

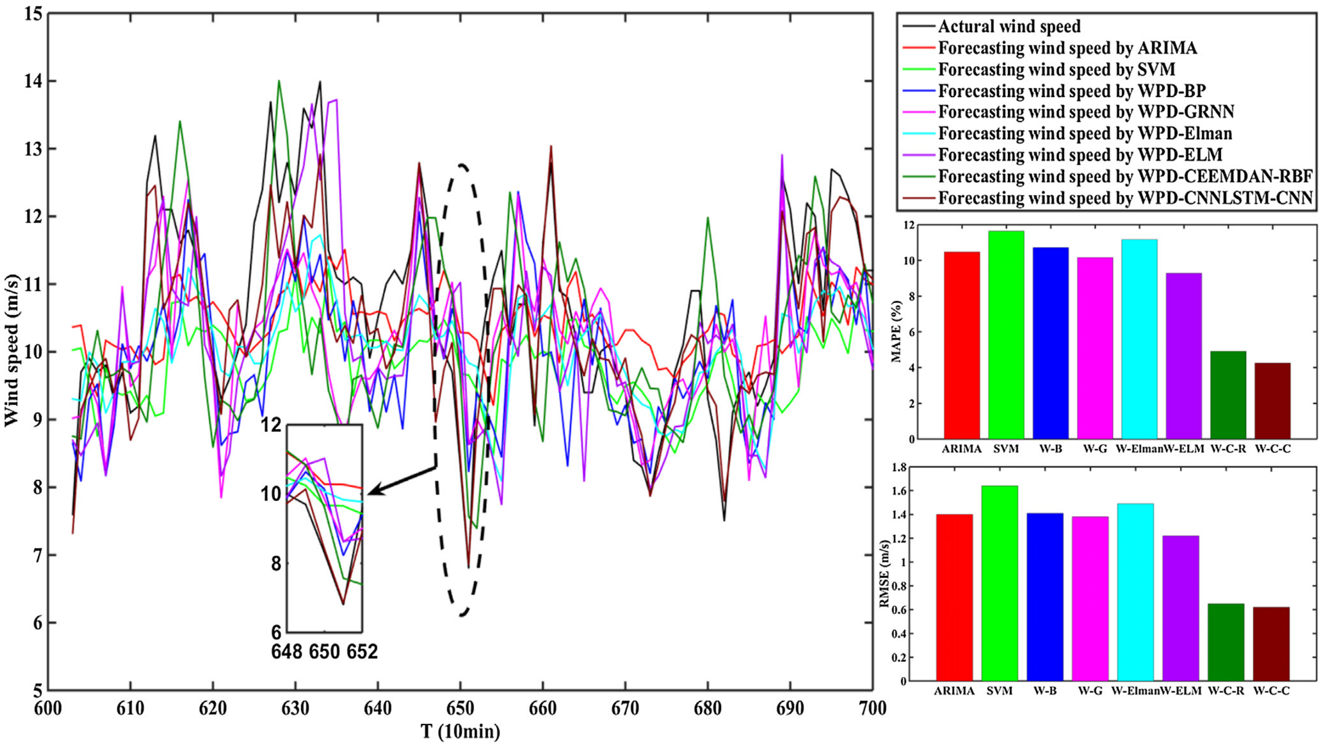

Fig. 7. The 3-step prediction results of the involved models for experimental test #1.

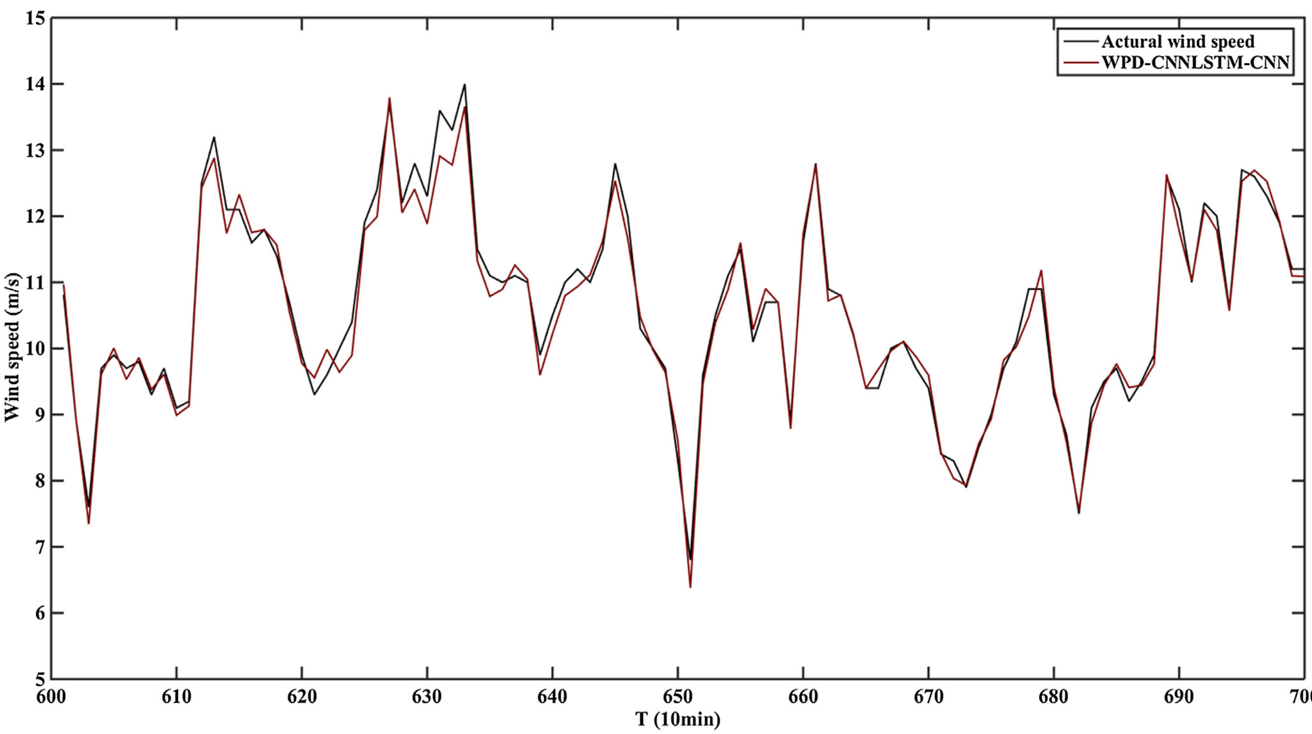

Fig. 8. The 1-step predictions of the proposed model for experimental test #1.

and updated by the input gate, forget gate and output gate. The computation of the LSTM components can be defined as follows [44]:

$$i_t = \sigma(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_i) \tag{5}$$

$$f_t = \sigma(W_{fx}x_t + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f) \tag{6}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tanh(W_{cx}x_t + W_{cm}m_{t-1} + b_c) \tag{7}$$

$$o_t = \sigma(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_t + b_o) \tag{8}$$

$$m_t = o_t \circ \tanh(c_t) \tag{9}$$

$$y_t = W_{ym}m_t + b_y \tag{10}$$

where $x = (x_1, x_2, \ldots, x_T)$ implicates the flattening layer data; i_t, f_t, c_t, o_t implicate the input gate, forget gate, memory cell vectors and output gate, respectively; m_t, W, b, '$\circ$' implicate the activation vectors for each

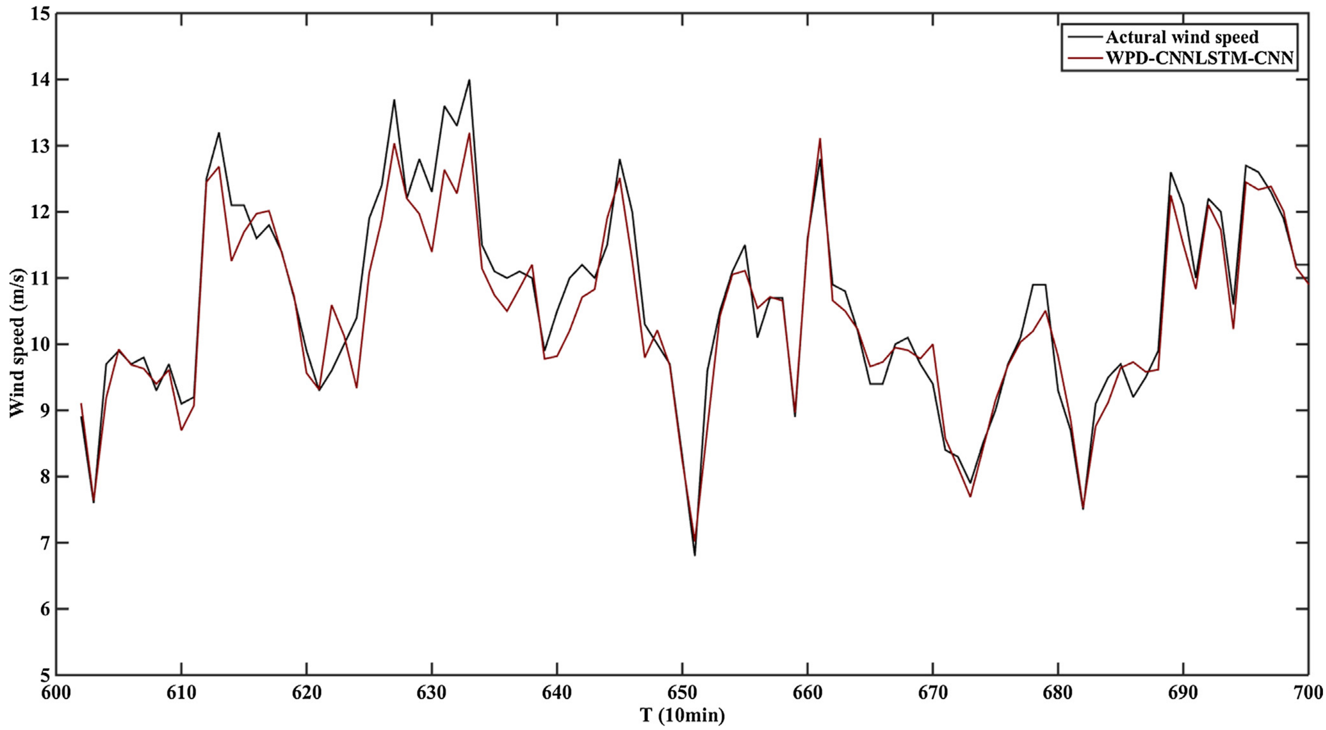

Fig. 9. The 2-step predictions of the proposed model for experimental test #1.

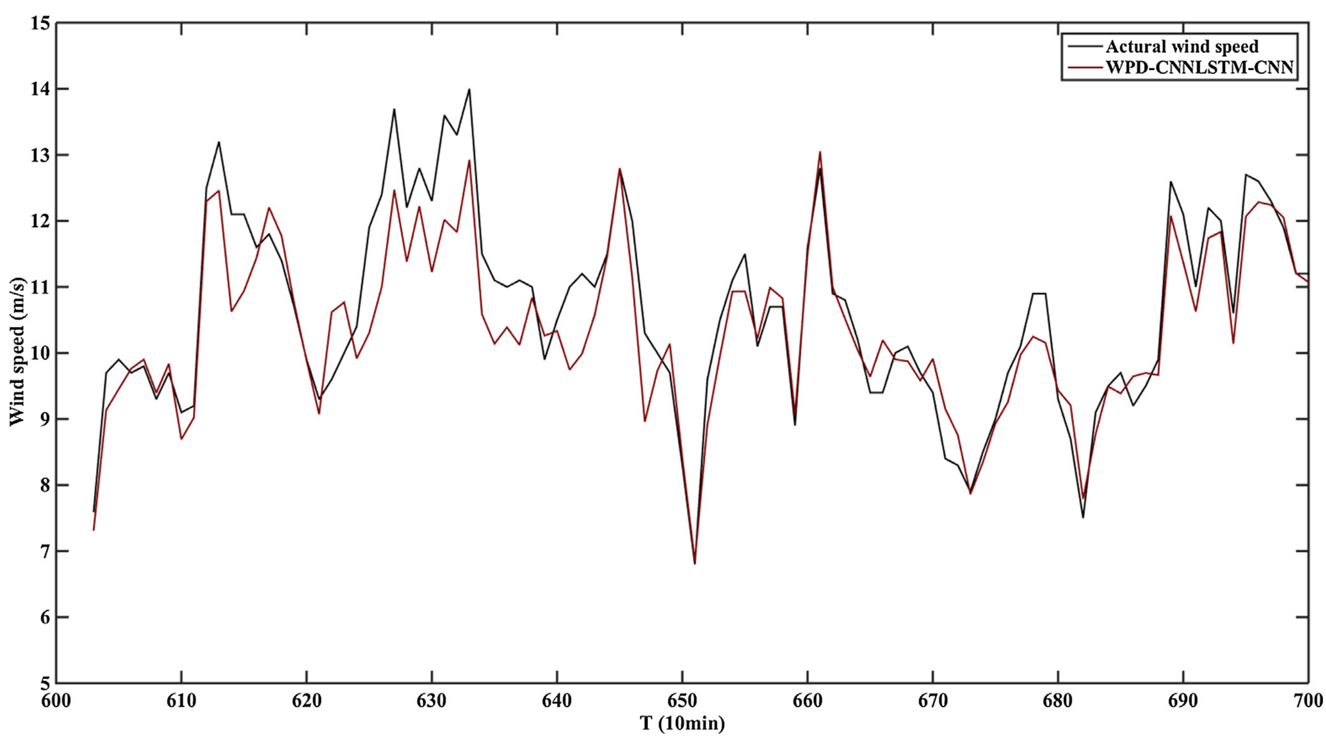

Fig. 10. The 3-step predictions of the proposed model for experimental test #1.

memory block, weigh matrices, bias vectors and scalar product, respectively; and $y = (y_1, y_2, ..., y_T)$ implicates the final outputs.$\sigma(\cdot)$ and $\tanh(\cdot)$ are the two activation functions, which can be defined as:

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{11}$$

$$\tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{12}$$

In this paper, the CNNLSTM comprises of three convolutional layers with no pooling operations, and the channels of the convolutional layers are 4, 16 and 32, respectively. The outputs of the flattening layer are given to the inputs of the LSTM recurrent layer, and the LSTM recurrent layer is connected to the final outputs. Further details of the CNNLSTM are described as: the convolutional layers apply *ReLU* activation function, the LSTM recurrent layer drops the outputs with the probability of p = 0.7, and the *Adam* is selected as the optimization

Table 6
The order of the comparison models based on the prediction accuracy.

Comparison models	Test #1	Test #2	Test #3	Test #4
ARIMA	7	7	8	7
SVM	8	8	7	7
WPD-BP	4	4	6	3
WPD-GRNN	4	5	5	3
WPD-Elman	4	6	4	3
WPD-ELM	3	3	3	3
WPD-CEEMDAN-RBF	2	2	2	2
WPD-CNNLSTM-CNN	1	1	1	1

algorithm. The structure of the CNNLSTM is illustrated in Fig. 3.

2.5. Prediction performance criteria

To evaluate the prediction results, six prediction performance criteria are exploited in this study. Among the six prediction performance criteria, MAPE is the mean absolute percentage error, MAE is the mean absolute error, RMSE is the root mean square error, P_{MAPE} is the promoting percentages of mean absolute percentage error, P_{MAE} is the promoting percentages of mean absolute error, and P_{RMSE} is the promoting percentages of root mean square error. They can be computed as:

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

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

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

$$P_{MAPE} = |(MAPE_1-MAPE_2)/MAPE_1| \tag{24}$$

$$P_{MAE} = |(MAE_1-MAE_2)/MAE_1| \tag{25}$$

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

where $X(t)$, $\hat{X}(t)$, N represent the actual values, prediction values, and the number of the actual values, respectively.

3. Case study

3.1. Wind speed data description

Xinjiang Uygur Autonomous Region, which has abundant wind resources, is in the northwest of China. In this paper, the wind speed data gathered from one wind farm in Xinjiang region in 2016. In the experiments, four sets of 10-min averaged wind speed time series are used for forecasting, each of which includes 700 samples, the 1st–600th samples are targeted for training, and the 601st–700th samples are targeted for testing. These wind speed time series are shown in Fig. 4. The descriptive statistics of the wind speed data are shown in Table 1.

3.2. Experiments

In this section, to compare and investigate the prediction performance of the proposed model, four experimental tests are conducted for 1-step, 2-step, and 3-step predictions. And the actual wind speed series #1, #2, #3 and #4 are used in these experimental tests, respectively. Each experiment consists of eight wind speed prediction models, including the ARIMA model, SVM model, WPD-BP model, WPD-GRNN model, WPD-Elman model, WPD-ELM model, WPD-CEEMDAN-RBF model and WPD-CNNLSTM-CNN model. The WPD-ELM model is the best model in the literature [45], and the WPD-CEEMDAN-RBF model is the best model in the literature [18]. Among these models, the ARIMA model, SVM model, WPD-BP model, WPD-GRNN model, WPD-Elman model, WPD-ELM model, WPD-CEEMDAN-RBF model are implemented in Matlab R2014b, while the WPD-CNNLSTM-CNN model is implemented in TensorFlow 1.3.0.

To better compare the prediction performance of the proposed model, all the involved models have the similar parameters. In Tables 2–5, the MAPE, MAE, and RMSE of all the eight involved models for the four experimental tests are reported.

3.3. The comparisons and analysis

From Tables 2–5, it can be found that the prediction results of the four experimental tests have the similar law. Figs. 5–7 show the 1-step, 2-step and 3-step prediction results of the involved models for the experimental test #1. Figs. 8–10 show the 1-step, 2-step and 3-step prediction results of the proposed model for experimental test #1.

Based on Figs. 5–10, it can be observed that:

(1) The WPD-CNNLSTM-CNN model can have satisfactory prediction performance in wind speed 1-step to 3-step predictions. Besides, among the involved eight models for experimental test #1, the WPD-CNNLSTM-CNN model can perform best in wind speed 1-step to 3-step predictions.
(2) When the wind speed experiences sudden change, the WPD-CNNLSTM-CNN model can have better prediction performance than the other involved models.

Table 6 gives the order of the comparison models based on the prediction accuracy. In Table 6, some different models are given the same order, for the prediction accuracy of which is similar. From Table 6, it can be found that:

(1) Among all the involved models, the WPD-CNNLSTM-CNN model has the highest prediction precision, and the WPD-CEEMDAN-RBF model has the second.
(2) The ARIMA model and SVM model have the similar prediction precision. The WPD-BP model, WPD-GRNN model, and WPD-Elman

Table 7
Promoting percentages of the comparison models by the WPD-CNNLSTM-CNN model for the experimental test #1.

Comparison models	P_{MAPE} (%)			P_{MAE} (%)			P_{RMSE} (%)		
	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
ARIMA	76.74	70.27	59.41	76.39	69.31	57.27	77.78	67.44	55.71
SVM	78.47	71.39	63.49	78.48	71.56	63.28	78.85	70.63	62.20
WPD-BP	67.21	61.21	60.35	67.31	60.76	59.48	67.16	57.14	56.03
WPD-GRNN	70.53	64.28	58.17	70.18	63.95	56.88	69.01	61.11	55.07
WPD-Elman	63.30	61.87	61.95	63.04	61.73	60.83	64.52	58.82	58.39
WPD-ELM	55.06	55.88	54.20	54.05	55.71	52.53	53.19	51.16	49.18
WPD-CEEMDAN-RBF	48.05	35.78	14.44	48.48	35.42	9.62	47.62	28.81	4.62

Table 8
Promoting percentages of the comparison models by the WPD-CNNLSTM-CNN model for the experimental test #2.

Comparison models	P_{MAPE} (%)			P_{MAE} (%)			P_{RMSE} (%)		
	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
ARIMA	77.97	70.32	59.14	77.61	69.23	57.89	79.55	68.47	56.78
SVM	78.19	66.90	59.34	77.94	69.57	58.33	80.00	69.83	58.20
WPD-BP	58.13	50.18	52.29	58.33	49.09	50.62	64.71	51.39	49.00
WPD-GRNN	69.23	59.54	55.03	68.09	57.58	55.06	70.49	58.82	52.34
WPD-Elman	73.57	66.67	55.71	73.68	66.27	55.06	75.34	67.59	56.03
WPD-ELM	51.28	48.73	47.01	51.61	47.17	45.21	56.10	52.70	45.74
WPD-CEEMDAN-RBF	54.63	37.33	24.49	53.13	34.88	21.57	58.14	39.66	28.17

Table 9
Promoting percentages of the comparison models by the WPD-CNNLSTM-CNN model for the experimental test #3.

Comparison models	P_{MAPE} (%)			P_{MAE} (%)			P_{RMSE} (%)		
	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
ARIMA	79.34	74.76	66.30	79.44	73.47	63.52	78.36	71.67	61.98
SVM	78.82	72.84	63.43	79.44	72.73	62.82	78.83	71.35	61.98
WPD-BP	72.99	63.05	60.23	74.12	62.86	59.44	72.64	61.36	58.99
WPD-GRNN	68.91	61.88	57.15	69.44	61.76	55.73	68.82	59.52	54.09
WPD-Elman	68.24	58.98	54.33	68.57	58.95	53.60	68.82	59.84	56.55
WPD-ELM	56.49	56.87	51.61	58.49	57.61	52.07	57.35	58.20	57.06
WPD-CEEMDAN-RBF	55.93	42.05	42.28	56.00	41.29	40.82	58.57	39.29	40.16

Table 10
Promoting percentages of the comparison models by the WPD-CNNLSTM-CNN model for the experimental test #4.

Comparison models	P_{MAPE} (%)			P_{MAE} (%)			P_{RMSE} (%)		
	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
ARIMA	82.30	74.41	64.46	82.00	73.91	63.95	83.08	73.14	62.98
SVM	81.87	73.86	65.15	81.82	73.91	66.03	83.46	74.18	66.16
WPD-BP	79.48	67.56	63.27	79.55	67.57	62.68	80.70	65.44	61.05
WPD-GRNN	71.40	61.77	57.14	71.43	62.11	57.94	73.81	62.40	58.39
WPD-Elman	79.50	65.80	60.07	78.82	64.71	59.23	79.82	63.57	58.13
WPD-ELM	58.90	59.62	61.58	59.09	60.00	62.14	62.07	60.17	64.36
WPD-CEEMDAN-RBF	57.10	39.31	34.08	55.00	36.84	32.05	61.40	38.16	32.32

model have the similar prediction precision.

To further investigate the prediction performance of the proposed model, the P_{MAPE}, P_{MAE}, and P_{RMSE} of the experimental tests are employed to make the comparisons and analysis. Tables 7–10 give the comparative analysis between the WPD-CNNLSTM-CNN model and other involved models for the four experimental tests, respectively.

Based on Tables 7–10, it can be known that:

(1) The WPD-CNNLSTM-CNN model is robust and effective in predicting the 1D wind speed time series.
(2) The WPD-CNNLSTM-CNN model can significantly outperform the ARIMA model. For example, in experimental test #4, compared to the ARIMA model, the MAPE of the WPD-CNNLSTM-CNN model are reduced by 82.30%, 74.41% and 64.46%, respectively; the MAE of the WPD-CNNLSTM-CNN model are reduced by 82.00%, 73.91% and 63.95%, respectively; the RMSE of the WPD-CNNLSTM-CNN model are reduced by 83.08%, 73.14% and 62.98%, respectively.
(3) The WPD-CNNLSTM-CNN model can significantly outperform the SVM model. For example, in experimental test #4, compared to the SVM model, the MAPE of the WPD-CNNLSTM-CNN model are reduced by 81.87%, 73.86% and 65.15%, respectively; the MAE of the WPD-CNNLSTM-CNN model are reduced by 81.82%, 73.91% and 66.03%, respectively; the RMSE of the WPD-CNNLSTM-CNN model are reduced by 83.46%, 74.18% and 66.16%, respectively.
(4) The WPD-CNNLSTM-CNN model can significantly outperform the WPD-BP model. For example, in experimental test #4, compared to the WPD-BP model, the MAPE of the WPD-CNNLSTM-CNN model are reduced by 79.48%, 67.56% and 63.27%, respectively; the MAE of the WPD-CNNLSTM-CNN model are reduced by 79.55%, 67.57% and 62.68%, respectively; the RMSE of the WPD-CNNLSTM-CNN model are reduced by 80.70%, 65.44% and 61.05%, respectively.
(5) The WPD-CNNLSTM-CNN model can significantly outperform the WPD-GRNN model. For example, in experimental test #4, compared to the WPD-GRNN model, the MAPE of the WPD-CNNLSTM-CNN model are reduced by 71.40%, 61.77% and 57.14%, respectively; the MAE of the WPD-CNNLSTM-CNN model are reduced by 71.43%, 62.11% and 57.94%, respectively; the RMSE of the WPD-CNNLSTM-CNN model are reduced by 73.81%, 62.40% and 58.39%, respectively.
(6) The WPD-CNNLSTM-CNN model can significantly outperform the WPD-Elman model. For example, in experimental test #4, compared to the WPD-Elman model, the MAPE of the WPD-CNNLSTM-CNN model are reduced by 79.50%, 65.80% and 60.07%, respectively; the MAE of the WPD-CNNLSTM-CNN model are reduced by 78.82%, 64.71% and 59.23%, respectively; the RMSE of the WPD-CNNLSTM-CNN model are reduced by 79.82%, 63.57% and 58.13%, respectively.
(7) The WPD-CNNLSTM-CNN model can significantly outperform the WPD-ELM model. For example, in experimental test #4, compared to the WPD-ELM model, the MAPE of the WPD-CNNLSTM-CNN model are reduced by 58.90%, 59.62% and 61.58%, respectively;

the MAE of the WPD-CNNLSTM-CNN model are reduced by 64.00%, 62.11% and 64.90%, respectively; the RMSE of the WPD-CNNLSTM-CNN model are reduced by 65.08%, 60.83% and 66.16%, respectively.

(8) The WPD-CNNLSTM-CNN model can significantly outperform the WPD-CEEMDAN-RBF model. For example, in experimental test #4, compared to the WPD-CEEMDAN-RBF model, the MAPE of the WPD-CNNLSTM-CNN model are reduced by 57.10%, 39.31% and 34.08%, respectively; the MAE of the WPD-CNNLSTM-CNN model are reduced by 55.00%, 36.84% and 32.05%, respectively; the RMSE of the WPD-CNNLSTM-CNN model are reduced by 61.40%, 38.16% and 32.32%, respectively.

4. Conclusions

In this paper, a novel wind speed prediction model is proposed based on the WPD, CNN and CNNLSTM. In the WPD-CNNLSTM-CNN model, the WPD is employed to decompose the original wind speed time series into a number of sub-layers; the CNN with 1D convolution operator is used to forecast the obtained high-frequency sub-layers; and the CNNLSTM is adopted to complete the forecasting of the low-frequency sub-layer. In order to verify and compare the prediction performance of the proposed model, some models are used, which consist of the ARIMA model, SVM model, WPD-BP model, WPD-GRNN model, WPD-Elman model, WPD-ELM model, WPD-CEEMDAN-RBF model and WPD-CNNLSTM-CNN model. According to the results of four experimental tests, it can be observed that: (1) in the proposed model, the WPD can effectively extract the features of the signal, the CNNGRU and CNN can have good prediction precision in the main trend component and detail components forecasting; (2) the proposed model is robust and effective in predicting the 1D wind speed time series, besides, among the involved eight models, the proposed model can perform best in wind speed 1-step to 3-step predictions; (3) when the wind speed experiences sudden change, the proposed model can have better prediction performance than the other involved models.

Acknowledgements

This study is fully supported by the National Natural Science Foundation of China (Grant No. 51308553), the Changsha Science & Technology Project (Grant No. KQ1707017), 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 Fundamental Research Funds for the Central Universities of Central South University. This study is partly supported by the National Natural Science Foundation of China (Grant No. U1534210, U1334205).

References

[1] Liu H, Wu H, Li Y. Smart wind speed forecasting using EWT decomposition, GWO evolutionary optimization, RELM learning and IEWT reconstruction. Energy Convers Manage 2018;161:266–83.
[2] Liu H, Mi X, Li Y. An experimental investigation of three new hybrid wind speed forecasting models using multi-decomposing strategy and ELM algorithm. Renew Energy 2018;123:694–705.
[3] Dong Q, Sun Y, Li P. A novel forecasting model based on a hybrid processing strategy and an optimized local linear fuzzy neural network to make wind power forecasting: a case study of wind farms in China. Renew Energy 2017;102:241–57.
[4] Liu H, Duan Z, Han F-Z, Li Y-F. Big multi-step wind speed forecasting model based on secondary decomposition, ensemble method and error correction algorithm. Energy Convers Manage 2018;156:525–41.
[5] Zhao Y, Ye L, Li Z, Song X, Lang Y, Su J. A novel bidirectional mechanism based on time series model for wind power forecasting. Appl Energy 2016;177:793–803.
[6] González-Aparicio I, et al. Simulating European wind power generation applying statistical downscaling to reanalysis data. Appl Energy 2017;199:155–68.
[7] Wang J, Hu J, Ma K, Zhang Y. A self-adaptive hybrid approach for wind speed forecasting. Renew Energy 2015;78:374–85.
[8] Wang L, Li J. Estimation of extreme wind speed in SCS and NWP by a non-stationary model. Theor Appl Mech Lett 2016;6(3):131–8.
[9] Galanis G, Papageorgiou E, Liakatas A. A hybrid Bayesian Kalman filter and applications to numerical wind speed modeling. J Wind Eng Ind Aerodyn 2017;167(April):1–22.
[10] Allen DJ, Tomlin AS, Bale CSE, Skea A, Vosper S, Gallani ML. A boundary layer scaling technique for estimating near-surface wind energy using numerical weather prediction and wind map data. Appl Energy 2017;208(May):1246–57.
[11] 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.
[12] Ait O, Achuthan A, Janoyan K, Marzocca P. Recursive wind speed forecasting based on Hammerstein Auto-Regressive model. Appl Energy 2015;145:191–7.
[13] 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.
[14] 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.
[15] Zhang C, Wei H, Zhao J, Liu T, Zhu T, Zhang K. Short-term wind speed forecasting using empirical mode decomposition and feature selection. Renew Energy 2016;96:727–37.
[16] Sun W, Liu M. Wind speed forecasting using FEEMD echo state networks with RELM in Hebei, China. Energy Convers Manage 2016;114:197–208.
[17] Shrivastava NA, Lohia K, Panigrahi BK. A multiobjective framework for wind speed prediction interval forecasts. Renew Energy 2016;87:903–10.
[18] Liu H, Mi X, Li Y. Comparison of two new intelligent wind speed forecasting approaches based on wavelet packet decomposition, complete ensemble empirical mode decomposition with adaptive noise and artificial neural networks. Energy Convers Manage 2018;155(October 2017):188–200.
[19] Feng C, Cui M, Hodge BM, Zhang J. A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. Appl Energy 2017;190:1245–57.
[20] Coelho IM, Coelho VN, Luz EJ da S, Ochi LS, Guimarães FG, Rios E. A GPU deep learning metaheuristic based model for time series forecasting. Appl Energy 2017;201:1–7.
[21] 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.
[22] 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.
[23] Hu Q, Zhang R, Zhou Y. Transfer learning for short-term wind speed prediction with deep neural networks. Renew Energy 2016;85:83–95.
[24] H. Liu, X. wei Mi, Y. fei Li. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. Energy Convers Manage 2018;156(November 2017):498–514.
[25] Sun S, Qiao H, Wei Y, Wang S. A new dynamic integrated approach for wind speed forecasting. Appl Energy 2017;197(June 2016):151–162.
[26] 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.
[27] 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.
[28] Liu H, Tian HQ, Liang XF, Li YF. Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks. Appl Energy 2015;157:183–94.
[29] Hu J, Wang J, Ma K. A hybrid technique for short-term wind speed prediction. Energy 2015;81:563–74.
[30] Jiang Y, Huang G. Short-term wind speed prediction: hybrid of ensemble empirical mode decomposition, feature selection and error correction. Energy Convers Manage 2017;144:340–50.
[31] Yu C, Li Y, Zhang M. 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. Energy Convers Manage 2017;147:75–85.
[32] Yu C, Li Y, Zhang M. An improved wavelet transform using singular spectrum analysis for wind speed forecasting based on elman neural network. Energy Convers Manage 2017;148:895–904.
[33] Peng T, Zhou J, Zhang C, Zheng Y. Multi-step ahead wind speed forecasting using a hybrid model based on two-stage decomposition technique and AdaBoost-extreme learning machine. Energy Convers Manage 2017;153(August):589–602.
[34] 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.
[35] Zheng W, et al. Composite quantile regression extreme learning machine with feature selection for short-term wind speed forecasting: a new approach. Energy Convers Manage 2017;151(2):737–52.
[36] 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.
[37] Jiang P, Wang Y, Wang J. Short-term wind speed forecasting using a hybrid model. Energy 2017;119:561–77.
[38] Zhao J, Guo Y, Xiao X, Wang J, Chi D, Guo Z. Multi-step wind speed and power forecasts based on a WRF simulation and an optimized association method. Appl Energy 2017;197:183–202.

[39] Liu H, Mi X, Li Y. Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM. Energy Convers Manage 2018;159(January):54–64.
[40] Oehmcke S, Zielinski O, Kramer O. Input quality aware convolutional LSTM networks for virtual marine sensors. Neurocomputing 2018;275:2603–15.
[41] Ding L, Fang W, Luo H, Love PED, Zhong B, Ouyang X. A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory. Autom Construct 2018;86(March 2017):118–124.
[42] Kingma DP, Ba J. Adam: a method for stochastic optimization. CoRR; 2014. abs/1412.6980.
[43] Tsironi E, Barros P, Weber C, Wermter S. An analysis of convolutional long short-term memory recurrent neural networks for gesture recognition. Neurocomputing 2017;268:76–86.
[44] Ke J, Zheng H, Yang H, (Michael) Chen X. Short-term forecasting of passenger demand under on-demand ride services: a spatio-temporal deep learning approach. Transport Res Part C: Emerg Technol 2017;85(October):591–608.
[45] Liu H, Tian HQ, Li YF. Four wind speed multi-step forecasting models using extreme learning machines and signal decomposing algorithms. Energy Convers Manage 2015;100:16–22.