Energy 238 (2022) 121981

Contents lists available at ScienceDirect

Energy

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

Multi-step-ahead wind speed forecasting based on a hybrid decomposition method and temporal convolutional networks

Dan Li [a, b], Fuxin Jiang [a, b], Min Chen [c, a, b, *], Tao Qian [c]

[a] Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, China
[b] School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing, 100049, China
[c] Macao Center for Mathematical Sciences, Macau University of Science and Technology, Macau, 999078, China

ARTICLE INFO

Article history:
Received 21 April 2021
Received in revised form
6 August 2021
Accepted 1 September 2021
Available online 7 September 2021

Keywords:
Wind speed forecasting
Ensemble patch transform
Complete ensemble empirical mode decomposition
Temporal convolutional network
Hybrid method

ABSTRACT

Recently, the boom in wind power industry has called for the accurate and stable wind speed forecasting, on which reliable wind power generation systems depend heavily. Due to the intermittency and complexity of wind, an appropriate decomposition is proved as a pivotal part in the precise wind speed prediction. On this account, this paper constructs a hybrid decomposition method coupling the ensemble patch transform (EPT) and the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), where EPT is utilized to extract the trend of wind speed, then CEEMDAN is employed to divide the volatility into several fluctuation components with different frequency characteristics. Subsequently, the proposed decomposition method is combined with temporal convolutional networks (TCN) for the individual prediction of the trend and fluctuation components. Ultimately, the forecasted values for the wind speed prediction are obtained by reconstructing the prediction results of all the components. To evaluate the performance of the proposed EPT-CEEMDAN-TCN model, the historical wind speed data from three wind farms across China are used. The experimental results verify the notable effectiveness and necessity of the proposed EPT-CEEMDAN decomposition. In the meanwhile, the results demonstrate the significant superiority of the proposed EPT-CEEMDAN-TCN model on accuracy and stability.

1. Introduction

With the reduction in global petrochemical resource reserves and the proposal of sustainable development strategy, the development and utilization of renewable energy sources have aroused increasing attention [1]. Due to the economic competitiveness and environmental friendliness of wind energy, the wind power industry has ushered in new opportunities for prosperous development. Following the wind power statistics published by World Wind Energy Association in March 2021 [2], the total capacity of wind farms worldwide has expanded to 744 GW, meeting seven percent of global electricity demand. In spite of the pandemic, the strong growth was realized. In 2020, the global market for annual new turbines has raised by around 50% to 93 GW, where China has created a new world record with 52 GW, contributing 56% of the global market share.

However, the large-scale wind power integration brings great challenges to the regular operation and dispatching of power system. As the wind power depends heavily on the uncontrollable and changeful wind speed, it presents large variation and complex randomness, which cause the fluctuations in current frequency and diminish the reliability of the electrical power system. Specifically, the quantitative relationship between wind speed and the generated wind power is established as Eq. (1) [3]:

$$P_a = \left\{ \frac{\exp\left[-(v_c/c)^k\right] - \exp\left[-(v_r/c)^k\right]}{(v_r/c)^k - (v_c/c)^k} - \exp\left[-\left(v_f/c\right)^k\right] \right\} \times P_r \tag{1}$$

where v_c, v_f, and v_r are the cut-in, cut-off, and nominal wind speed values (m/s), respectively. Additionally, c is the Weibull scale parameter (m/s), while P_a and P_r are the average power output of

* Corresponding author. Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, China.
E-mail address: mchen@amss.ac.cn (M. Chen).

https://doi.org/10.1016/j.energy.2021.121981
0360-5442/

the wind turbine (kW) and rated electrical power of the wind turbine (kW), respectively. Clearly, the accurate and stable wind speed forecasting takes a leading part in wind power generation. Thereby, the accurate and stable wind speed forecasting is vital for improving the wind power utilization level and constructing the smart grid, for it will benefit the optimization of dispatching plan and reduce the system reserve capacity, raising the economic and social profits.

Considering the intermittency and uncertainty of wind speed, it is difficult to conduct an accurate and stable prediction. In that case, numerous forecasting methods have been investigated recently to forecast wind speed series over different time-scales, including physical models, statistical models, machine learning methods, artificial neural networks, etc. Specifically, the complex physical models inevitably rely on the current meteorological and geographical information, such as temperature, pressure, topography structure, obstacles, etc., in which the required numeric weather prediction (NWP) data are hard to acquire in most cases. Besides, the computation is too complicated to forecast wind speed in a short time [4–7]. On the contrary, the statistical models, such as autoregressive (AR) [8], autoregressive moving average (ARMA) [9–11] and generalized autoregressive conditional heteroskedasticity (GARCH) [12], only input the historical wind speed series and run fast, but they often fail to yield good performance in the presence of large uncertainties. Moreover, the machine learning algorithms, such as the support vector regression (SVR) [13–15], have good generalization ability to reach the global solution in a fast manner, while their scalability for the large dataset is limited. To address that, the neural networks are developed owing to their powerful multivariable mapping capability. But the traditional back-propagation neural network easily traps into the local optimal solution and loses sight of the internal influence of time series [16–18]. To solve this problem, the recurrent neural networks (RNN) are introduced that are suitable to take the time series characteristics into account through the self-connection between hidden layers [19]. Considering the gradient vanishment problem of RNNs, a long short-term memory (LSTM) neural network is achieved, mitigating the issue by adding a special unit structure to the basic RNN [20,21]. Marndi et al. [22] compared LSTM with SVM and the extreme learning machine (ELM), verifying the higher forecasting precision of LSTM. Besides, an increasing number of researches have indicated the LSTM can outperform statistical models, machine learning methods and other previous neural networks in various domains, such as wind speed forecasting [23], financial market forecasting [24], COVID-19 forecasting [25]. Whereas the number of model parameters in LSTM is usually quite large, raising the problem of hard-training and the tendency of over-fitting. Consequently, the gated recurrent unit (GRU), as a new type of RNN, is designed to tackle the above issues specifically [26]. Upgraded based on the LSTM, the optimized GRU promotes the forecasting speed effectively without loss of accuracy [27–30]. Peng et al. [31] adopted GRU to predict the multi-step-ahead wind speed values, verifying that GRU can achieve relatively high accuracy, fast forecasting speed, small volatility of errors and good adaptability when it is compared with BPNN, ELM, Simple RNN, LSTM, etc. Yet in 2018, Bai et al. [32] pointed out that the temporal convolutional networks (TCN) with a simple dilated causal convolution can outperform the popular recurrent networks such as LSTMs and GRUs on various tasks and datasets, even possessing longer effective memory. Following that, the temporal convolutional networks (TCN) have exhibited the significant performance benefits in traffic flow forecasting [33], short-term passenger demand prediction [34], and fault diagnosis for power converters [35]. Gan et al. [36] constructed interval prediction for wind speed forecasting based on TCNs, suggesting that TCNs can appreciably enhance the prediction accuracy and reliability as compared with the classic artificial neural networks and the canonical recurrent neural networks. Accordingly, this paper employs TCNs for prediction. Apart from using a single model to predict wind speed series, an increasing number of researches indicate that decomposition-based models can transcend the single models due to the nonstationarity and complicated randomness of wind speed series [37–41]. For instance, the wind speed series is decomposed first, then a single model is used to predict each subsequence respectively, finally the forecasting results are reconstructed to obtain the predicted value of the wind speed. Among the literatures, there exists some prevalent decomposition methods for time series, such as the wavelet transform (WT) [42–44], the variational mode decomposition (VMD) [38,45–47], and the empirical mode decomposition (EMD) [48–50]. Pei et al. [51] combined New Cell Update Long Short-Term Memory network with empirical wavelet transform, which enhanced the prediction accuracy in a shorter training time. Zhang et al. [52] applied VMD to decompose wind speed series into the nonlinear components, linear components and noise, then predicted these components via PCA-RBF model and MCMC-ARMA model respectively, reflecting the characteristics of wind speed series properly and obtaining the veracious prediction.

Gathering the present knowledge on the widely used wind speed decomposition approaches in literatures [80–82], a concise comparison of these main methods is illustrated in Table 1. Note that each method has its own strengths and limitations. Specifically, the Fourier transform is not suitable for transient and nonstationary signals, like wind speed; wavelet decomposition and wavelet packet decomposition requires presetting the basis function and the order, on which the decomposition results depend primarily; empirical wavelet transform fails to detect the components when the signal embodies two chirps which overlap in both the time and frequency domains [65]; singular spectral analysis may generate several meaningless components and omit the important information, due to the troublesome parameter selection problem; variational mode decomposition takes prior experience or multiple trials to deduce the number of modes for the decomposition. Accordingly, the fully data-driven empirical mode decomposition (EMD) proposed by Huang [69] seems more appropriate for handling the complex and nonstationary time series. Fu et al. [72] constructed a wind speed forecasting framework combining time varying filter-based empirical mode decomposition (TVF-EMD), fuzzy entropy (FE) theory, and singular spectrum analysis (SSA) to decompose the wind speed series adequately which reduced the time consumption as well as promoted forecasting performance. Because the mode mixing problem is encountered frequently in practical applications of EMD, Wu and Huang [73] developed a noise-assisted ensemble empirical mode decomposition (EEMD) method. Despite that the EEMD solves the mode mixing problem, it introduces new modes. Considering that, Torres et al. [77] proposed the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) which adds a specific noise at each stage and calculates an exclusive residue for each mode. In that way, CEEMDAN can separate various modes from the wind speed series more precisely and efficiently [78,79]. Wu et al. [76] coupled CEEMDAN and an optimized extreme learning machine (ELM), which surpasses other traditional models with high accuracy and strong stability. However, all the quasi-EMD methods fail to capture the trend of wind speed series, for every intrinsic mode function (IMF) from any quasi-EMD method has to satisfy the following conditions: in the entire data sequence, the number of extrema and the number of zero crossings in any sampled dataset must either be equal or differ at most by one, while the residual item is extremely smooth. Instead, Kim et al. [83]

Table 1
The brief comparison of the main decomposition methods utilized for wind speed forecasting.

Method	Advantages	Disadvantages	References
Fourier Transform (FT) [53]	It has strict mathematical theory. It is only suitable for stationary signals.	It cannot be applied to transient and nonstationary signals.	[54–56]
Wavelet Decomposition (WD) [57]	It has strict mathematical theory. It is suited for constant frequency and almost periodic signals.	It needs to specify the wavelet basis and parameters beforehand. It may split the modes. It is unsuited for highly nonstationary signals.	[31,42,44,58–60]
Wavelet Packet Decomposition (WPD) [61]	It has strict mathematical theory. It improves forecasting models more than WD.	It needs to specify the wavelet basis and parameters beforehand. It may split the modes. It needs more computational resources.	[62–64]
Empirical Wavelet Transform (EWT) [65]	It has mathematical theory. The wavelets are adapted to the signal. The dilation factors are detected empirically.	It fails to separate the chirps that overlap in both the time and frequency domains.	[39,43,51,66]
Singular Spectral Analysis (SSA) [67]	It has strict mathematical theory. It can decompose a time-series into specific components.	The parameters must be adjusted to extract each component.	[27,41,46]
Variational Mode Decomposition (VMD) [68]	It has rigorous mathematical formulation. It is suited for constant frequency and almost periodic signals.	It needs to deduce the number of modes beforehand. It may split the modes. It is unsuited for highly nonstationary signals.	[21,38,45,47,52]
Empirical Mode Decomposition (EMD) [69]	It is fully data-driven.	It lacks strict mathematical theory. It may encounter mode mixing problem.	[37,49,70–72],
Ensemble Empirical Mode Decomposition (EEMD) [73]	It is fully data-driven. It solves the mode mixing problem.	It lacks strict mathematical theory. The extra noise exists in the reconstructed signal. It needs much computational resources. It is difficult to determine an ensemble mean.	[23,50,74]
Complementary Ensemble Empirical Mode Decomposition (CEEMD) [75]	It is fully data-driven. It solves the mode mixing problem. It eliminates the residual noise. It needs fewer ensemble trials than EEMD.	It lacks strict mathematical theory. It is difficult to determine the ensemble mean.	[18,48,76],
Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) [77]	It is fully data-driven. It solves the mode mixing problem. The reconstruction errors are negligible. It resolves the ensemble mean problem.	It lacks strict mathematical theory. It is implemented in sequence, and cannot be computed in parallel.	[40,78,79],

designed the flexible ensemble patch transformation (EPT) for decomposition and filtering of signals, which upgrades the detection of local patterns embedded in a signal effectively through the particular patches and explicates the temporal variation of the signal based on the adjustable ensemble patches. Consequently, we integrate EPT and CEEMDAN to excavate the potential trend of wind speed series as well as separate the various patterns massed in the volatility of wind speed series.

Furthermore, previous researches always considered the forecasting accuracy simply, overlooking the stability and the adaptability, both of which are worth improving. For one thing, the high forecasting stability indicates that the forecasting results are reliable all the time. For another, a satisfactory forecasting model needs adequate adaptability for the practical application at different sites. Thus, the comprehensive considerations of the accuracy, the stability and the adaptability herein are innovative and noteworthy.

Additionally, as the decomposition process should be executed step by step with the new arrival of data in practical [80], the decomposition-based prediction models will confront several challenges, including: (i) the subseries are constantly changing with the new data; (ii) several illusive components may emerge, decreasing the decomposition validness; (iii) the end effect gets worse, augmenting the subseries volatility [84]. To address these challenges, we develop the real-time EPT-CEEMDAN-TCN model as well, which holds the effectiveness in the practical application of the short-term wind speed forecasting.

Synthetically speaking, a novel hybrid approach coupling the EPT-CEEMDAN decomposition and the TCNs prediction is proposed in this paper to extract the trend of wind speed series exactly, further decompose the volatility of wind speed series completely as well as improve forecasting performance. Correspondingly, the main contributions of this work are stated as follows:

(1) The established ensemble patch transformation (EPT) extracts the essential trend component of wind speed series exactly. As the trend is verified to take the dominant part in the temporal variation of wind speed series, the precise trend extraction by EPT contributes largely to the wind speed forecasting results.
(2) The proposed hybrid decomposition method integrating EPT with CEEMDAN enhances the forecasting performance effectively, where CEEMDAN decomposes the complex volatility of wind speed series into several uncorrelated fluctuation components completely and efficiently.
(3) By virtue of a simple and clear convolutional architecture, the adopted temporal convolutional networks (TCN) forecast the trend and fluctuation components of wind speed series more accurately and stably, transcending the statistical models, the traditional machine learning algorithm, the canonical back propagation neural networks and the prevalent recurrent neural networks.
(4) The proposed prediction approach combining the EPT-CEEMDAN decomposition and TCNs is applied to the multi-step-ahead forecasting of three real wind speed datasets from diverse areas of China. The contrastive experiments verify the accuracy and stability of the proposed method from comprehensive perspectives.

(5) Following an approximated forecasting strategy, the real-time EPT-CEEMDAN-TCN forecasting model is constructed for the practical wind speed forecasting with satisfactory performance in both accuracy and stability, mitigating the existing challenges in decomposition-based models.

This paper is organized as follows. Section 2 introduces the framework of the proposed approach, where the EPT-CEEMDAN decomposition method and the temporal convolutional network are described. Section 3 explicates the performance of the proposed model in comparison experiments in light of multiple evaluation criteria. Section 4 discuss the computational efficiency and complexity of the proposed model. Finally, Section 5 draws the major conclusions.

2. The proposed approach

The proposed method based on EPT-CEEMDAN decomposition method and temporal convolutional networks is termed EPT-CEEMDAN-TCN. It consists of three major processes: (1) decomposing the wind speed via EPT-CEEMDAN decomposition and obtain the uncorrelated components, including the daily trend, fluctuation components and residue; (2) forecasting all the sub-series by TCNs individually; (3) reconstructing the wind speed predicted values by integrating the forecasting results. The structure of EPT-CEEMDAN-TCN model is presented in Fig. 1.

2.1. EPT-CEEMDAN decomposition

The wind speed is influenced by multiple factors such as temperature, humidity and topography, giving rise to its complex nonstationary and nonlinear characteristics. It is difficult for a single model to predict wind speed directly and precisely. Consequently, a suitable decomposition of original series plays a vital role in prediction. On this account, the hybrid EPT-CEEMDAN decomposition is proposed in this paper. Firstly, we utilize ensemble patch transform (EPT) to extract the trend from wind speed series with the volatility separated out in the meantime. Subsequently, we apply complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) to the volatility and procure its fluctuation components with different frequency characteristics, and then analyze them individually.

The ensemble patch transformation (EPT) [83] is a flexible framework for decomposition and filtering of signal. It comprises two primary processes. The first is “patch process”, which is a data-dependent patch of data at a given time point t designed for specifying the local structures with the elastic sizes of patches. The second is termed “ensemble process”, where the time point t of patch is shifted to produce the ensemble patch, which can illustrate the temporal variation of data effectively via adjustable temporal resolution. Due to the adaptable parameters, EPT is appropriate for exploring particular patterns of signals, such as trends, seasonalities and abrupt changes.

The complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) [77] is an improved algorithm based on the fully data-driven ensemble empirical mode decomposition (EEMD) [73]. Comparing with the original empirical mode decomposition (EMD) [69] and the following EEMD [73], CEEMDAN usually achieves a more precise and computationally efficient decomposition by adding a specific noise at each stage and calculating a unique residue for each mode.

Specifically, the proposed EPT-CEEMDAN decomposition for wind speed signal can be done via the following steps:

Stage 1: EPT for wind speed.

(1) **Patch Process.** The rectangle patch is adopted in this paper. Given the period τ, the patch $P_t^\tau(X_t)$ for the wind speed X_t at time t is centered at the point (t, X_t) is a closed rectangle

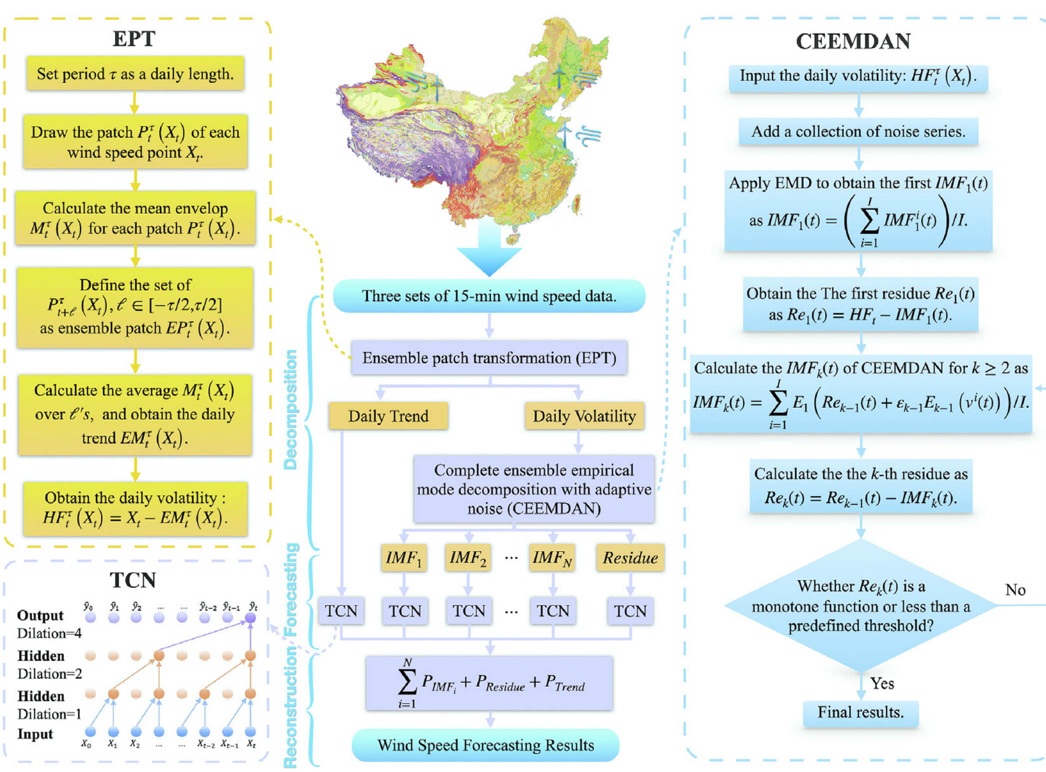


Fig. 1. The framework of EPT-CEEMDAN-TCN model.

formed by the points $\left(t+k, \min_{k\in[-\tau/2,\tau/2]}\ \{X_{t+k}\} - 0.5\gamma\tau\right)$ and $\left(t+k, \max_{k\in[-\tau/2,\tau/2]}\{X_{t+k}\} + 0.5\gamma\tau\right)$ for $k \in [-\tau/2, \tau/2]$. For the rectangle patch, the width is τ and height h_t^τ is

$$h_t^\tau = \max_{k\in[-\tau/2,\tau/2]}\{X_{t+k}\} - \min_{k\in[-\tau/2,\tau/2]}\{X_{t+k}\} + \gamma\tau \tag{2}$$

where γ is a scale factor.

The lower envelope $L_t^\tau(X_t)$ and upper envelope $U_t^\tau(X_t)$ of the rectangle patch $P_t^\tau(X_t)$ are

$$L_t^\tau(X_t) = \min_{k\in[-\tau/2,\tau/2]}\{X_{t+k}\} - 0.5\gamma\tau \tag{3}$$

$$U_t^\tau(X_t) = \max_{k\in[-\tau/2,\tau/2]}\{X_{t+k}\} + 0.5\gamma\tau. \tag{4}$$

Then, we can obtain the mean envelop $M_t^\tau(X_t)$ for each patch $P_t^\tau(X_t)$:

$$M_t^\tau(X_t) = \frac{1}{2}\left(L_t^\tau(X_t) + U_t^\tau(X_t)\right) \tag{5}$$

(2) **Ensemble Process.** For any fixed period τ, the l-th shifted patch at time point t is defined as $P_{t+\ell}^\tau(X_t), \ell\in[-\tau/2, \tau/2]$. Accordingly, a collection of all possible shifted patches at time point t is defined as an ensemble patch:

$$EP_t^\tau(X_t) := \left\{P_{t+\ell}^\tau(X_t) : \ell\in[-\tau/2, \tau/2]\right\} \tag{6}$$

Then, we get the low-frequency mode:

$$EM_t^\tau(X_t) = average\left(M_{t+\ell}^\tau(X_t)\right) over\ \ell's \tag{7}$$

and the high-frequency mode:

$$HF_t^\tau(X_t) = X_t - EM_t^\tau(X_t) \tag{8}$$

Given that period τ is so flexible that can be selected from a diverse range, it is applicable to capture the temporal characters of signal effectively. In this paper, the period τ is fixed as a daily length, thereby $EM_t^\tau(X_t)$ can be regarded as the daily trend of wind speed, while $HF_t^\tau(X_t)$ as the daily volatility.

Stage 2: CEEMDAN for daily volatility.

(1) A collection of Gaussian white noise series is added to daily volatility as:

$$S^i(t) = HF_t + \epsilon_i v^i(t) \tag{9}$$

where $S^i(t)$ is the time series with the additional noise in the i-th trial ($i = 1, 2, \ldots,$ I), HF_t is the daily volatility of wind speed, ϵ_i is the ratio of the additional noise to the signal, and $v^i(t)$ is the Gaussian white noise series.

(2) EMD is used to obtain the first intrinsic mode function $IMF_1(t)$ as:

$$IMF_1(t) = \left(\sum_{i=1}^{I} IMF_1^i(t)\right) \Big/ I \tag{10}$$

where $IMF_1^i(t)$ is the first intrinsic mode function obtained in the i-th trial.

(3) The first residue $Re_1(t)$ is:

$$Re_1(t) = HF_t - IMF_1(t) \tag{11}$$

(4) For $k = 2, \ldots, N$, the k-th residue is:

$$Re_k(t) = Re_{k-1}(t) - IMF_k(t) \tag{12}$$

where $IMF_k(t)$ is the k-th intrinsic mode function of CEEMDAN.

(5) The $IMF_{k+1}(t)$ of CEEMDAN is:

$$IMF_{k+1}(t) = \sum_{i=1}^{I} E_1\left(Re_k(t) + \epsilon_k E_k\left(v^i(t)\right)\right) \Big/ I \tag{13}$$

where $E_k(\bullet)$ is the k-th mode of EMD, ϵ_k is the ratio of the additional noise to the signal in the k-th stage of CEEMDAN.

(6) Repeat steps 4) and 5) until all the intrinsic mode functions are found.

Finally, we decompose nonstationary and nonlinear wind speed into a finite number of components as Eq. (14):

$$X(t) = \sum_{i=1}^{N} IMF_i(t) + \mathrm{Re}(t) + Trend(t) \tag{14}$$

where $Trend(t) := EM_t^\tau(X_t)$.

Considering that the decomposition methods may cause the edge effect in practice, we mirror the periods (in EPT stage) and extrema (in CEEMDAN stage) close to the edges before the corresponding decompositions, so as to minimize error propagations over the finite observation lengths [83,85].

2.2. Temporal convolutional network

Bai et al. constructed temporal convolutional networks (TCN) for sequence modeling tasks under causal constraint [32]. In a sequence modeling task, the model output is a prediction sequence $\{\hat{y}_0, \hat{y}_1, \ldots, \hat{y}_T\}$ with the corresponding input sequence $\{x_0, x_1, \ldots, x_T\}$ given. The typical causal constraint is that only those previous observations $x_0, x_1, \ldots, x_t$ can be taken as inputs for predicting the output y_t at the time t. Accordingly, the temporal convolutional network is designed as a nonlinear function $f : \mathcal{X}^{T+1} \rightarrow \mathcal{Y}^{T+1}$ that yields the mapping:

$$\hat{y}_0, \ldots, \hat{y}_T = f(x_0, \ldots, x_T) \tag{15}$$

where $\hat{y}_t$ depends only on known $x_0, \ldots, x_t$ instead of any unknown $x_{t+1}, \ldots, x_T$ at the time t. The network is trained by supervised learning to find the function f that minimizes the loss function $L(y_0, \ldots, y_T, f(x_0, \ldots, x_T))$ between the actual outputs and the predictions. In the case of wind speed forecasting, the input sequence is the historical wind speed observations over the past few moments, while the actual output is the current wind speed observation.

Under the causal constrain, the TCN develops dilated causal convolutions to expand the receptive field exponentially, taking more historical information into consideration. The dilated convolution operation F on element s of the 1-D sequence $x\in\mathbb{R}^n$ for a filter $f : \{0, 1, \ldots, k-1\} \rightarrow \mathbb{R}$ is formulated as Eq. (16).

$$F(s) = (\mathbf{x}_d^* f)(s) = \sum_{i=0}^{k-1} f(i) \cdot \mathbf{x}_{s-d \cdot i}, \tag{16}$$

where k is the filter size, d is the dilation factor, and $*$ is the convolution operator. The dilated causal convolution used in this paper is illustrated in Fig. 2 (a) with dilation factors $d = 1, 2, 4, 8$ and filter size $k = 2$. Definitely, the larger filter sizes k or dilation factor d is set, the broader receptive field of the network is.

Furthermore, a generic residual module is employed for feature extraction at each layer in the TCN. As is depicted in Fig. 2 (b), the residual block consists of two layers of dilated causal convolution and nonlinearity, where the rectified linear unit (ReLU) is taken as the activation function. In each layer, batch normalization [86] is adopted to the convolutional filters and a spatial dropout [87] is added subsequent to each dilated convolution for regularization. On account of the incompatible input-output widths, an additional 1×1 convolution is utilized to ensure that elementwise addition $\oplus$ takes in tensors of the consistent shape.

2.3. Reconstruction

According to Eq. (14), the prediction results of each component are aggregated to obtain the final wind speed predicted results, as shown in the Eq. (17).

$$P_{\text{Wind}} = \sum_{i=1}^{N} P_{IMF_i} + P_{Residue} + P_{Trend} \tag{17}$$

where P_{Wind} is the final wind speed predicted value, and P_{IMF_i}, $P_{Residue}$, and P_{Trend} are the predicted values for $IMF_i(t)$, $\text{Re}(t)$, and $Trend(t)$ in Eq. (14) respectively, with $i = 1, 2, \ldots, N$.

3. Experimental analysis

In this section, the wind speed series collected from three sites are used to evaluate the performance of the proposed EPT-CEEMDAN-TCN model comparing with several benchmark models. Section 3.1 describes the experimental design. Section 3.2 demonstrates the experimental decomposition results. Section 3.3 analyses the forecasting results from multiple perspectives. Section 3.4 compares the prediction performance with different decomposition methods. Section 3.5 constructs the real-time forecasting model.

3.1. Experimental design

3.1.1. Data description

Up till now, China has constructed wind capacity of 289 GW, 39% of global capacity, ranking the first in the world [2]. Thereupon, the models are performed on the historical short-term wind speed data collected from three wind farms across China. They are located in Gansu Province, Liaoning Province and Jiangsu Province respectively, varying considerably in longitudes and latitudes. As is illustrated in Fig. 3, Gansu lies in northwest inland area, Liaoning lies in northeastern coastal area, and Jiangsu lies in southeastern coastal area. Gansu and Liaoning are located in the “Three-North” (Northwest China, North China, and Northeast China) areas in which wind power resources are concentrated, while Jiangsu faces the Yellow Sea to the east with installed capacity of offshore wind power ranking first in China for consecutive years.

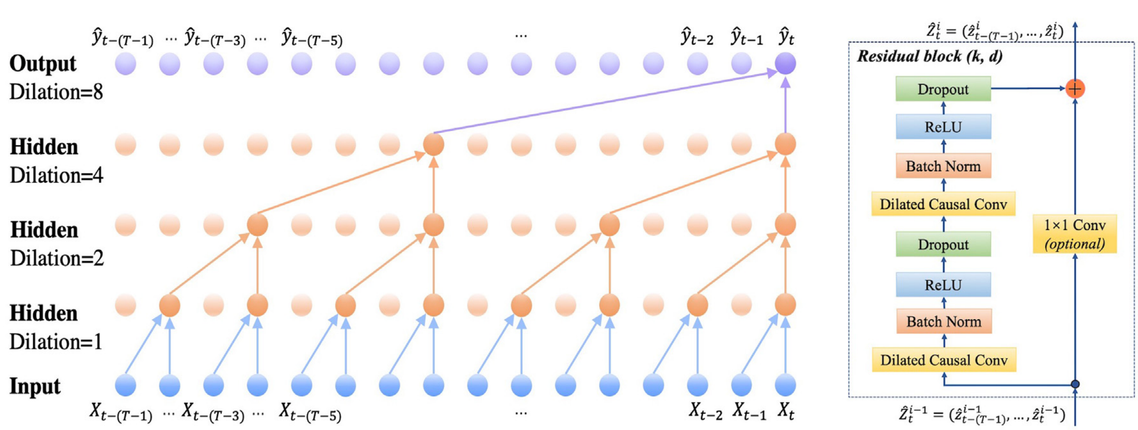
All the wind speeds are measured every 15 min at 10 m at the level of the ground, with 96 times scanning frequency per day. As is listed in Table 2, the detailed time periods of the three datasets are August 1st-31st in 2017 (Gansu), December 1st-31st in 2017 (Liaoning), and April 1st-30th in 2017 (Jiangsu). Each the wind speed dataset is divided into three parts: training sets, validation sets and testing sets, which are utilized for model training, hyper parameters selection and model verification correspondingly.

The original wind speed series are shown in Fig. 4. Their statistical information is shown in Table 3. It is noticeable that each wind speed series contains a significant trend and complicated high-frequency components, where the trends, the fluctuations and the distributions of the wind speed series vary from site to site.

3.1.2. Evaluation criteria

We adopt mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE) and normalized mean absolute percentage error (NMAPE) to evaluate the forecasting accuracy. The following are their formulas.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |\hat{y}_i - y_i| \tag{18}$$



(a) A dilated causal convolution (b) TCN residual block.

Fig. 2. Temporal convolutional network. (a) A dilated causal convolution with dilation factors $d = 1, 2, 4, 8$ and filter size $k = 2$. (b) A 1×1 convolution is added when residual input and output have different dimensions.

Fig. 3. The location of the three wind farms.

Table 2
The descriptions of the three wind speed datasets.

Dataset	Location	Samples	Time interval	# of samples
1	Gansu	All	August 1st - August 31st	2976
		Training	August 1st - August 24th	2304
		Validation	August 25th - August 26th	192
		Testing	August 27th - August 31st	480
2	Liaoning	All	December 1st - December 31st	2976
		Training	December 1st - December 24th	2304
		Validation	December 25th - December 26th	192
		Testing	December 27th - December 31st	480
3	Jiangsu	All	April 1st - April 30th	2880
		Training	April 1st - April 23rd	2208
		Validation	April 24th - April 25th	192
		Testing	April 26th - April 30th	480

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{19}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{20}$$

$$NRMSE = 100\%\frac{RMSE}{y_{\max} - y_{\min}} \tag{21}$$

where n is the number of testing samples, $\hat{y}_i$ is the predicted value of the actual value y_i. A more precise prediction is achieved when MAE, MSE, MAPE and NMAPE are smaller.

Besides, the improvement percentage is introduced for quantificational comparison between the proposed model and the benchmark models [29]. The improvement percentage of RMSE (P_{RMSE}), the improvement percentage of MAE (P_{MAE}), the improvement percentage of MAPE (P_{MAPE}) and the improvement percentage of NRMSE (P_{NRMSE}) are the descent rates of the proposed model compared with the benchmark model in terms of RMSE, MAE, MAPE and NRMSE respectively. P_{RMSE}, P_{MAE}, P_{MAPE} and P_{NRMSE} are calculated as follows:

$$P_{RMSE} = \frac{RMSE_1 - RMSE_2}{RMSE_1} \times 100\% \tag{22}$$

$$P_{MAE} = \frac{MAE_1 - MAE_2}{MAE_1} \times 100\% \tag{23}$$

$$P_{MAPE} = \frac{MAPE_1 - MAPE_2}{MAPE_1} \times 100\% \tag{24}$$

$$P_{NRMSE} = \frac{NRMSE_1 - NRMSE_2}{NRMSE_1} \times 100\% \tag{25}$$

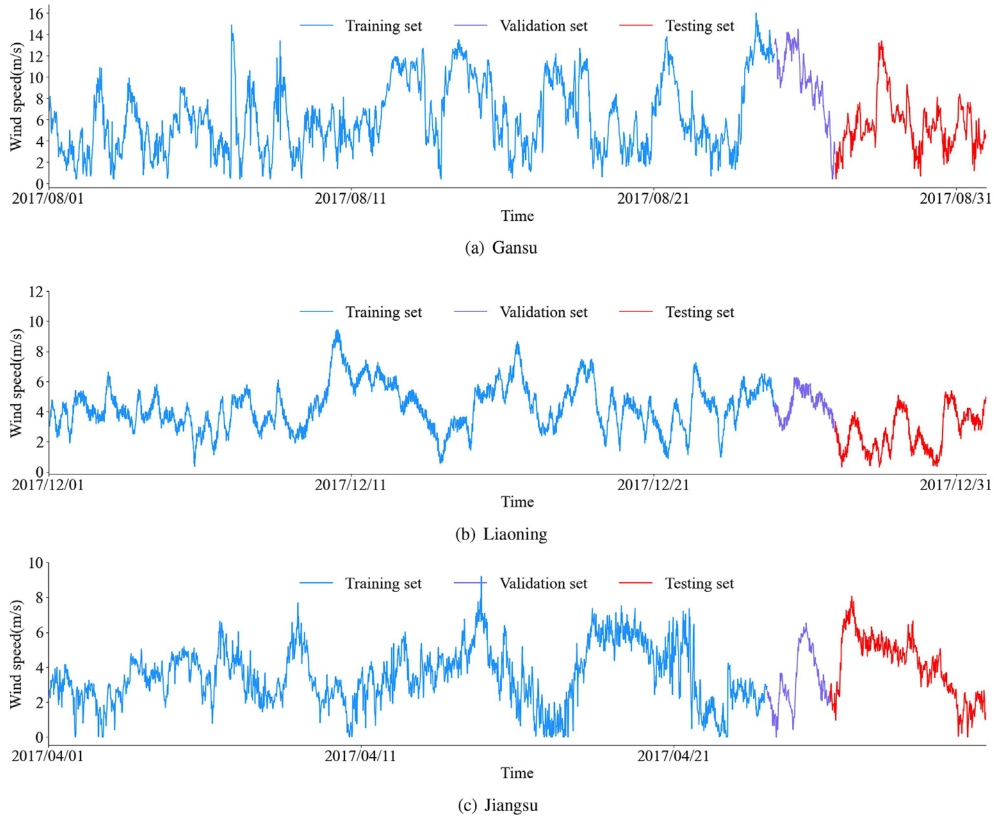
where $RMSE_1$, MAE_1, $MAPE_1$ and $NRMSE_1$ are the errors of the benchmark model, and $RMSE_2$, MAE_2, $MAPE_2$ and $NRMSE_2$ are the errors of the proposed model. A large positive value of P_{RMSE}, P_{MAE}, P_{MAPE} and P_{NRMSE} indicates that the proposed model performs much more accurately than the benchmark model.

Furthermore, the variance of absolute error (VAE) is employed to evaluate the forecasting stability. The VAE between the predicted value $\hat{y}_i$ and the actual value of y_i is as follow:

$$VAE = Var(|y_t - \hat{y}_t|) \tag{26}$$

The improvement percentage of VAE (P_{VAE}) of the proposed model compared with the benchmark model is defined in the same way as follows.

$$P_{VAE} = \frac{VAE_1 - VAE_2}{VAE_1} \times 100\% \tag{27}$$



(a) Gansu

(b) Liaoning

(c) Jiangsu

Fig. 4. The wind speed time series.

Table 3
The statistical information of the three wind speed datasets.

Site	Location	Minimum	Mean	Maximum	Standard Deviation	Skewness	Kurtosis
1	Gansu	0.400	6.302	16.000	3.303	0.487	2.339
2	Liaoning	0.32	4.070	9.420	1.585	0.217	3.048
3	Jiangsu	0.000	3.557	9.200	1.668	0.097	2.430

where VAE_1 is the VAE of the benchmark model, and VAE_2 is the VAE of the proposed model. Note that a large positive value of P_{VAE} means that the proposed model performs much more stably than the benchmark model.

3.1.3. Model development

To evaluate the forecasting performance of the proposed EPT-CEEMDAN-TCN model, we introduce several benchmark models for comparison:

(1) the traditional statistical model, i.e. autoregressive integrated moving average model (ARIMA);
(2) a modified statistical model which can outperform the recently developed neural networks in wind speed forecasting, i.e. seasonal autoregression integrated moving average (SARIMA) [11];
(3) the traditional machine learning algorithm, i.e. support vector regression (SVR);
(4) the benchmark neural networks, including the canonical back propagation neural network (BPNN) [18], the popular long short-term memory network (LSTM) [88], the newly emerging gated recurrent unit network (GRU) [46] and the adopted temporal convolutional networks (TCN);
(5) the other three benchmark neural networks with EPT-CEEMDAN decomposition, i.e. EPT-CEEMDAN-BPNN, EPT-CEEMDAN-LSTM, EPT-CEEMDAN-GRU, for investigating the effectiveness of EPT-CEEMDAN decomposition.

For each model, the input is the previous sixteen observations (4h ahead), while the output is the forecasting values. Table 1 in Appendix presents the specific parameter settings of these models, including the determination approaches. In the light of the forecasting accuracy on the divided validation sets, hyper parameters in forecasting models are generally predetermined by grid search and trial and error approach [89], while the hyper parameters in decomposition techniques are mainly preset [72]. In addition, we employ the mean square error as the loss function, and use the adaptive momentum estimation method (Adam) [90] to optimize the weights. All the models are implemented on the Keras platform using Tensorflow as backend in Python 3.7.4 with CPU 2.3 GHz and GPU NVIDIA TITAN RTX 24G RAM.

3.2. Decomposition results

Based on complex nonstationary characteristics of wind speed series, a proper decomposition of the original series plays a critical role in the enhancement of the forecasting performance. First, we set the period $\tau = 96$ in the ensemble patch transform (EPT) stage to extract the daily wind speed trend, for the data measured every 15 min with 96 times scanning frequency per day. Subtracting the daily trend from the original wind speed, we get the daily volatility. After that we use complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) to divide the daily volatility into several fluctuation components IMF_i with different frequency patterns. Then we analyze and forecast these components individually. To evaluate the performance of our proposed EPT-CEEMDAN, we employ the classic CEEMDAN on the wind speed series for comparison.

Fig. 5 demonstrates the decomposition results of CEEMDAN and EPT-CEEMDAN. In Fig. 5 (a), (c) and (e), it presents IMFs from high-frequency to low-frequency, residue and the original wind speed signal in order from top to bottom. In Fig. 5 (b), (d) and (f), it presents IMFs, residue, trend and the wind speed signal sequentially. Intuitively, it is clear that the traditional CEEMDAN decomposes the wind speed into several *IMFs* and a residual item, leaving out the fundamental trend item, while the EPT-CEEMDAN captures the trend of wind speed exactly.

Subsequently, the correlations for the components sets obtained by CEEMDAN and EPT-CEEMDAN are exhibited in Fig. 6. Note that the trends extracted by EPT always achieve the highest positive correlation with the original wind speed series, that is 0.75 in Gansu, 0.79 in Liaoning, and 0.77 in Jiangsu. That is to say, the trend contributes largely to wind speed variation, thereby it is necessary to extract it out precisely before prediction. On the contrary, the correlations between the components obtained by CEEMDAN and original wind speed series lack apparent rule. In Fig. 6 (a)–(c), the most correlated components with original wind speed series from CEEMDAN is IMF_7 in Gansu, IMF_6 in Liaoning and IMF_8 in Jiangsu, as the corresponding correlations are 0.60, 0.51, and 0.70. They are all smaller than the correlations between trends and wind speed series. In other words, the components from CEEMDAN take a more trivial part in the reconstruction of wind speed forecasting values according to Eq. (17).

Furthermore, there is less correlation between components obtained by EPT-CEEMDAN. It implies that EPT-CEEMDAN exerts the advantage of CEEMDAN to decompose the nonstationary volatility completely, for it performs on the nonlinear volatility instead of the complex wind speed series.

3.3. Forecasting results

In this section, each component obtained by EPT-CEEMDAN is forecasted by the temporal convolutional network respectively. Summing up the prediction results of all the components, we procure the final wind speed predicted values, according to Eq. (17). Note that the testing data should be 'unknown' in the training phase [71], thereupon the decomposition is applied only on the training sets firstly, then prediction models are trained on the subseries obtained from the training sets, avoiding the leakage of testing data during the training period. Subsequently, the validation sets are decomposed alongside training sets for hyper parameters selection. Finally, the trained prediction models are performed on the subseries from the whole datasets for model testing.

To adequately evaluate the performance of the proposed model, we also employ the other ten benchmark models aforementioned in Section 3.1.3 for comparison, including ARIMA, SARIMA, SVR, BPNN, LSTM, GRU, TCN, EPT-CEEMDAN-BPNN, EPT-CEEMDAN-LSTM, and EPT-CEEMDAN-GRU.

In practical applications, it is also worthwhile to forecast wind speeds 1 h ahead. Thus, the multi-step-ahead forecasting is implemented. Given the time series $\{y_1, y_2, \ldots, y_T\}$, we calculate the k-step-ahead forecasting value $\hat{y}_{t+k}$ directly as follows:

$$\hat{y}_{t+k} = f\left(y_t, y_{t-1}, \ldots, y_{t-(p-1)}\right), t = 1, 2, \ldots, T \tag{28}$$

where $\hat{y}_{t+k}$ is the forecasted value at time $t+k$, y_t is the actual value at time t, and p is the lag order of the input features. The lag order p determines the input features of the prediction model, which is important for the model performance. In this paper, the lag order p is set as 16, taking historical wind speed series over the past 4 h as input. The prediction horizon k is set as 1, 2, 3, 4 sequentially.

3.3.1. The accuracy of forecasting

The 1-step-ahead (15 min ahead) to 4-step-ahead (1 h ahead) forecasting accuracy in the three provinces are shown in Tables 4–6. In all the experiments, the proposed EPT-CEEMDAN-TCN method reaches all the best values of all the accuracy evaluation metrics including the least MAE (m/s), RMSE (m/s), MAPE and NRMSE values. It implies that EPT-CEEMDAN-TCN outperforms the other ten models on accuracy.

Take the case of Gansu as an example. The MAE, RMSE, MAPE and NRMSE obtained by EPT-CEEMDAN-TCN in one-step-ahead forecasting are 0.28890 m/s, 0.40157 m/s, 0.07595, and 3.08901, all of which are quite smaller than the corresponding errors of the other ten contrast models. Specifically, the RMSE values obtained by ARIMA, SARIMA, SVR, BPNN, LSTM, GRU, TCN, EPT-CEEMDAN-BPNN. EPT-CEEMDAN-LSTM, EPT-CEEMDAN-GRU, and EPT-CEEMDAN-TCN are 0.99677 m/s, 0.77507 m/s, 0.86547 m/s, 0.81122 m/s, 0.79083 m/s, 0.77943 m/s, 0.77494 m/s, 0.50294 m/s, 0.45949 m/s, 0.47798 m/s and 0.40157 m/s respectively, where EPT-CEEMDAN-TCN reaches the minimum 0.40157 m/s. Besides that, the RMSE values obtained by EPT-CEEMDAN-TCN in one-step-ahead, two-step-ahead, three-step-ahead, and four-step-ahead forecasting are 0.40157 m/s, 0.55225 m/s, 0.56956 m/s, and 0.63539 m/s in ascending order, and all of them reach the corresponding minimums. Likewise, the metrics MAE, MAPE and NRMSE follow the similar pattern.

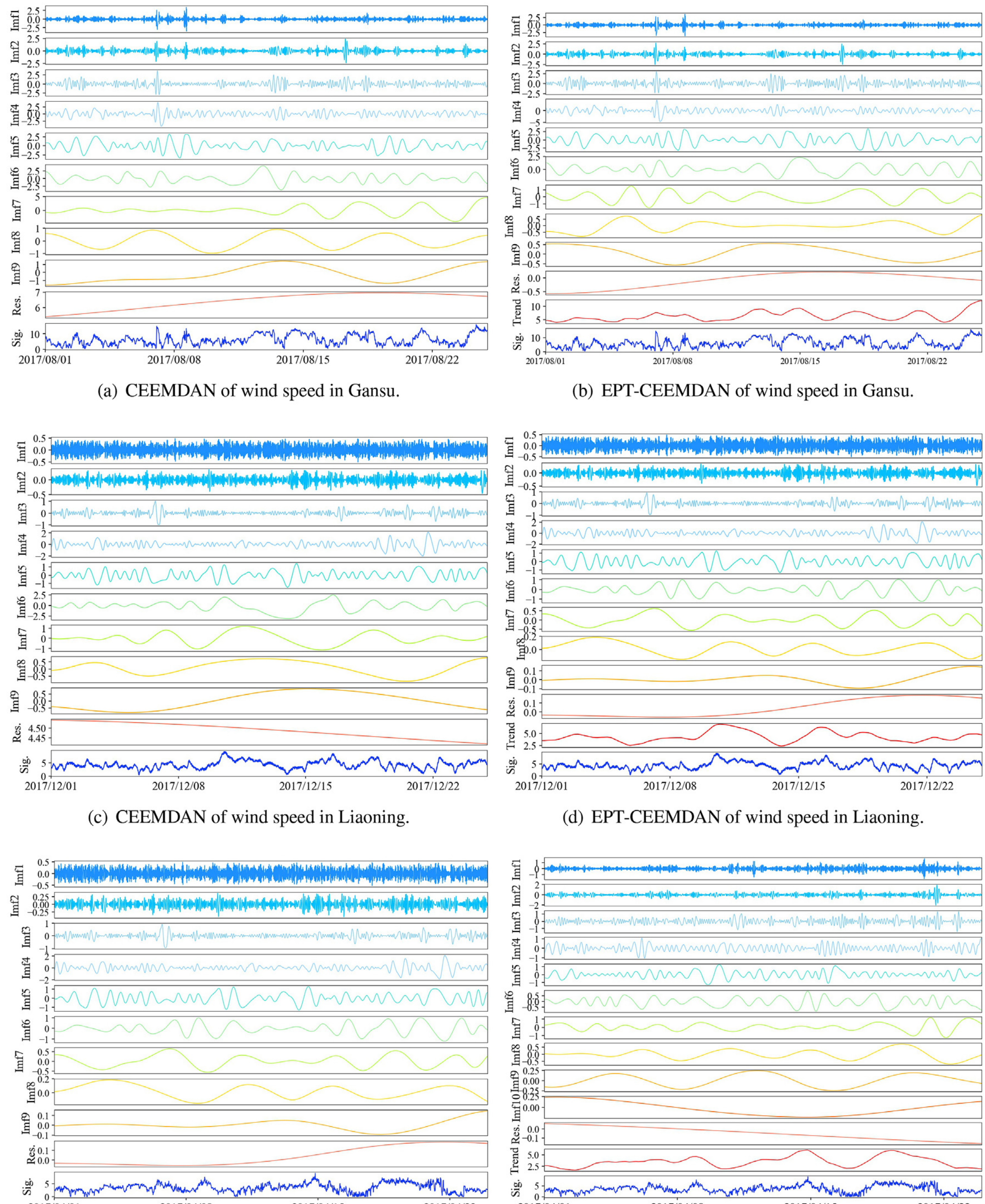
In addition, the hybrid models based on EPT-CEEMDAN decomposition perform appreciably better than these single models with the much lower errors. Take the one-step-ahead forecasting in Gansu as an instance. The MAPE of TCN is 0.15420 while the MAPE of EPT-CEEMDAN-TCN is 0.07595, which is less than a half of 0.15420. Thereby, it is obvious that the EPT-CEEMDAN decomposition possesses the essential capability of assisting the prediction model to grasp the valuable patterns within original series, improving the forecasting accuracy considerably.

The one-step-ahead prediction curves and the area graph of actual wind speed series in three datasets are shown in Fig. 7. Note that the fitting curve obtained by EPT-CEEMDAN-TCN (red lines) approximates to the edge of pink region representing actual values most closely in each dataset. It is convincingly verified the superiority of the high-precision prediction owned by EPT-CEEMDAN-TCN.

Moreover, in each dataset, the fitting curves obtained by hybrid models based on EPT-CEEMDAN decomposition, i.e. EPT-CEEMDAN-BPNN, EPT-CEEMDAN-LSTM, EPT-CEEMDAN-GRU and EPT-CEEMDAN-TCN, approximate to the edge of pink region more closely than those from single models, such as ARIMA, SARIMA, SVR, BPNN, LSTM, GRU and TCN. It reveals the virtue and necessity of the proposed EPT-CEEMDAN decomposition as well.

3.3.2. The improvement percentages on accuracy

The improvement percentages on accuracy of the proposed EPT-CEEMDAN-TCN compared with the benchmark models are shown in Tables 7–9. As the proposed EPT-CEEMDAN-TCN cuts down the



(a) CEEMDAN of wind speed in Gansu.

(b) EPT-CEEMDAN of wind speed in Gansu.

(c) CEEMDAN of wind speed in Liaoning.

(d) EPT-CEEMDAN of wind speed in Liaoning.

(e) CEEMDAN of wind speed in Jiangsu.

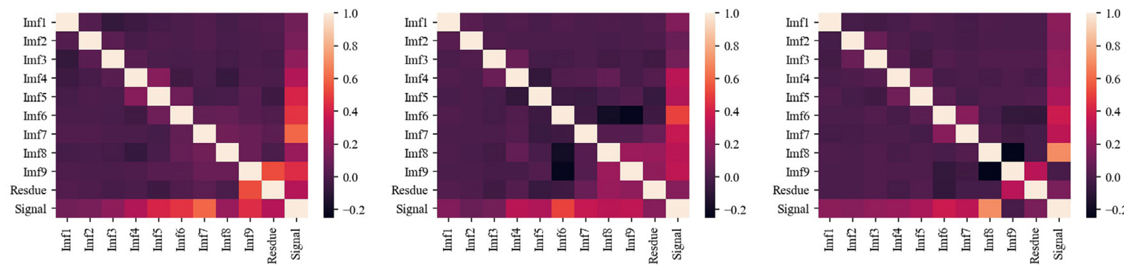
(f) EPT-CEEMDAN of wind speed in Jiangsu.

Fig. 5. The components sets from CEEMDAN and EPT-CEEMDAN at three sites.

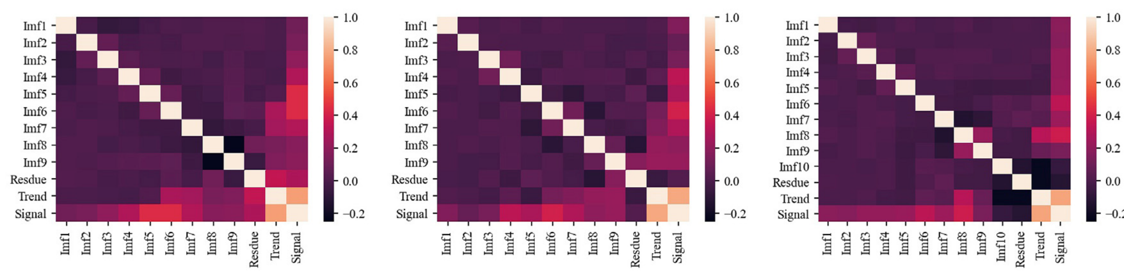
errors effectively and significantly in multi-step-ahead forecasting on various datasets, the remarkable superiority of EPT-CEEMDAN-TCN on accuracy is validated definitely.

Comparing with the single models, including ARIMA, SARIMA, SVR, BPNN, LSTM, GRU and TCN, the EPT-CEEMDAN-TCN can enhance prediction precision significantly, with the errors are almost halved. For example, the improvement percentages on RMSE of EPT-CEEMDAN-TCN compared with ARIMA, SARIMA, SVR, BPNN, LSTM, GRU and TCN are 56.98%, 47.59%, 51.41%, 57.81%, 51.01%, 50.67% and 49.06% orderly in the case of one-step-ahead forecasting in Jiangsu as is shown in Table 9. That is to say, the proposed EPT-CEEMDAN decomposition is rather powerful and necessary.

Furthermore, EPT-CEEMDAN-TCN also promotes the prediction accuracy of the three hybrid contrast models evidently, i.e. EPT-CEEMDAN-BPNN, EPT-CEEMDAN-LSTM and EPT-CEEMDAN-GRU. Specifically, the improvement percentages on RMSE of EPT-CEEMDAN-TCN compared with EPT-CEEMDAN-BPNN, EPT-



(a) CEEMDAN of wind speed in Gansu. (b) CEEMDAN of wind speed in Liaoning. (c) CEEMDAN of wind speed in Jiangsu.



(d) EPT-CEEMDAN of wind speed in Gansu. (e) EPT-CEEMDAN of wind speed in Liaoning. (f) EPT-CEEMDAN of wind speed in Jiangsu.

Fig. 6. The correlations for the components sets.

Table 4
The multi-step-ahead forecasting accuracy in Gansu.

Prediction horizon	Model	MAE	RMSE	MAPE	NRMSE	Prediction horizon	Model	MAE	RMSE	MAPE	NRMSE
1-Step	ARIMA	0.75036	0.99677	0.19740	7.66745	3-Step	ARIMA	0.97477	1.30791	0.24610	10.06081
	SARIMA	0.55813	0.77507	0.14995	5.96205		SARIMA	0.84987	1.13984	0.22330	8.76803
	SVR	0.63584	0.86547	0.17628	6.65745		SVR	0.93912	1.27634	0.24683	9.81801
	BPNN	0.59619	0.81122	0.15089	6.24012		BPNN	1.00469	1.28037	0.26799	9.84901
	LSTM	0.58013	0.79083	0.15358	6.08333		LSTM	0.89739	1.19647	0.24136	9.20362
	GRU	0.57101	0.77943	0.15640	5.99561		GRU	0.85968	1.16102	0.22978	8.93091
	TCN	0.56005	0.77494	0.15420	5.96109		TCN	0.84117	1.13651	0.21753	8.74240
	EPT-CEEMDAN-BPNN	0.37587	0.50294	0.09314	3.86879		EPT-CEEMDAN-BPNN	0.67890	0.84542	0.16118	6.50326
	EPT-CEEMDAN-LSTM	0.34791	0.45949	0.08666	3.53453		EPT-CEEMDAN-LSTM	0.53286	0.68676	0.14267	5.28279
	EPT-CEEMDAN-GRU	0.35836	0.47798	0.09569	3.67678		EPT-CEEMDAN-GRU	0.54288	0.69484	0.15445	5.34493
	EPT-CEEMDAN-TCN	**0.28890**	**0.40157**	**0.07595**	**3.08901**		EPT-CEEMDAN-TCN	**0.43800**	**0.56956**	**0.11625**	**4.38124**
2-Step	ARIMA	0.86198	1.16172	0.22132	8.93627	4-Step	ARIMA	1.09460	1.43588	0.27271	11.04521
	SARIMA	0.74627	0.98809	0.19950	7.60073		SARIMA	0.94718	1.26830	0.24697	9.75613
	SVR	0.81933	1.10615	0.21767	8.50888		SVR	1.07058	1.42277	0.27392	10.94438
	BPNN	0.86452	1.09109	0.23085	8.39298		BPNN	1.06295	1.37633	0.28317	10.58717
	LSTM	0.81430	1.08485	0.21794	8.34496		LSTM	1.01472	1.33591	0.27391	10.27620
	GRU	0.79784	1.04583	0.21999	8.04481		GRU	0.96856	1.30910	0.25017	10.06998
	TCN	0.74810	0.99327	0.20058	7.64051		TCN	0.94510	1.26588	0.24383	9.73757
	EPT-CEEMDAN-BPNN	0.63142	0.78217	0.16720	6.01666		EPT-CEEMDAN-BPNN	0.67949	0.85706	0.19267	6.59278
	EPT-CEEMDAN-LSTM	0.45798	0.58322	0.11632	4.48632		EPT-CEEMDAN-LSTM	0.56088	0.73031	0.15842	5.61779
	EPT-CEEMDAN-GRU	0.50145	0.62649	0.12058	4.81917		EPT-CEEMDAN-GRU	0.60467	0.77625	0.15207	5.97115
	EPT-CEEMDAN-TCN	**0.43056**	**0.55225**	**0.11629**	**4.24808**		EPT-CEEMDAN-TCN	**0.49176**	**0.63539**	**0.13649**	**4.88758**

CEEMDAN-LSTM and EPT-CEEMDAN-GRU are 25.86% 13.00% 18.15% individually in the case of four-step-ahead forecasting in Gansu as is shown in Table 7.

Moreover, the improvement percentages on accuracy of EPT-CEEMDAN-TCN is displayed in Fig. 8 intuitively. The deeper color is, the higher the corresponding improvement percentage is. Consequently, it is obvious that EPT-CEEMDAN-TCN can enhance the prediction accuracy to a great extent, especially in comparison with ARIMA, SARIMA, SVR, BPNN, LSTM, GRU, TCN and EPT-CEEMDAN-BPNN. In addition, as the lightest color that represents the minimum of improvement percentages is larger than zero, the superiority of EPT-CEEMDAN-TCN on accuracy is demonstrated apparently.

3.3.3. *Forecasting error analysis*

A further analysis on forecasting errors is conducted in this section. Take one-step-ahead forecasting as an example. Fig. 9 exhibits the stacked forecasting errors of the ten benchmark models and the proposed model at different time points, where the deviations of each single model at different time points can be observed directly. Accordingly, it is obvious that the hybrid models based on EPT-CEEMDAN decomposition generally possess smaller errors while EPT-CEEMDAN-TCN possess the smallest errors overall on the three datasets.

The error distributions of the various models are presented in Fig. 10, which makes it crystal clear that the errors from EPT-CEEMDAN-TCN are more concentrated around zero with a

Table 5
The multi-step-ahead forecasting accuracy in Liaoning.

Prediction horizon	Model	MAE	RMSE	MAPE	NRMSE	Prediction horizon	Model	MAE	RMSE	MAPE	NRMSE
1-Step	ARIMA	0.34028	0.42858	0.19553	8.47004	3-Step	ARIMA	0.43321	0.54912	0.24799	10.85215
	SARIMA	0.30287	0.37690	0.17388	7.44853		SARIMA	0.38086	0.48368	0.22736	9.55880
	SVR	0.34601	0.42863	0.23829	8.47091		SVR	0.41456	0.52408	0.29606	10.35724
	BPNN	0.33001	0.40359	0.17400	7.97612		BPNN	0.43516	0.54807	0.26955	10.83134
	LSTM	0.30694	0.38032	0.18941	7.51625		LSTM	0.37004	0.46587	0.23718	9.20686
	GRU	0.30358	0.37406	0.18301	7.39249		GRU	0.36852	0.46413	0.23815	9.17258
	TCN	0.30768	0.37182	0.17713	7.34829		TCN	0.35821	0.45342	0.22598	8.96090
	EPT-CEEMDAN-BPNN	0.28392	0.33864	0.17580	6.69248		EPT-CEEMDAN-BPNN	0.32530	0.40152	0.22060	7.93526
	EPT-CEEMDAN-LSTM	0.18127	0.22413	0.11207	4.42953		EPT-CEEMDAN-LSTM	0.26738	0.32382	0.17641	6.39951
	EPT-CEEMDAN-GRU	0.16936	0.20773	0.09606	4.10527		EPT-CEEMDAN-GRU	0.24721	0.29991	0.15058	5.92701
	EPT-CEEMDAN-TCN	**0.15659**	**0.19586**	**0.08896**	**3.87075**		EPT-CEEMDAN-TCN	**0.23442**	**0.28551**	**0.13423**	**5.64250**
2-Step	ARIMA	0.37898	0.48158	0.21875	9.51741	4-Step	ARIMA	0.48565	0.61647	0.27494	12.18311
	SARIMA	0.34185	0.42948	0.19991	8.48770		SARIMA	0.43724	0.55225	0.26175	10.91400
	SVR	0.36113	0.44995	0.24594	8.89226		SVR	0.47824	0.59111	0.34272	11.68200
	BPNN	0.35965	0.45118	0.21587	8.91669		BPNN	0.47081	0.59198	0.24753	11.69915
	LSTM	0.34472	0.43288	0.22639	8.55499		LSTM	0.46450	0.57594	0.31688	11.38223
	GRU	0.33862	0.42636	0.21875	8.42616		GRU	0.45456	0.56445	0.31418	11.15505
	TCN	0.33778	0.41980	0.21611	8.29646		TCN	0.41562	0.51924	0.26147	10.26164
	EPT-CEEMDAN-BPNN	0.34629	0.41896	0.18137	8.27977		EPT-CEEMDAN-BPNN	0.37920	0.47502	0.19341	9.38769
	EPT-CEEMDAN-LSTM	0.26869	0.33097	0.14945	6.54087		EPT-CEEMDAN-LSTM	0.27079	0.32975	0.16622	6.51672
	EPT-CEEMDAN-GRU	0.22383	0.26723	0.12631	5.28129		EPT-CEEMDAN-GRU	0.28275	0.34748	0.18709	6.86712
	EPT-CEEMDAN-TCN	**0.21370**	**0.25635**	**0.12290**	**5.06621**		EPT-CEEMDAN-TCN	**0.26401**	**0.32091**	**0.16370**	**6.34217**

Table 6
The multi-step-ahead forecasting accuracy in Jiangsu.

Prediction horizon	Model	MAE	RMSE	MAPE	NRMSE	Prediction horizon	Model	MAE	RMSE	MAPE	NRMSE
1-Step	ARIMA	0.39818	0.51980	0.29027	6.44112	3-Step	ARIMA	0.49670	0.63942	0.33435	7.92345
	SARIMA	0.32574	0.42664	0.39609	5.29331		SARIMA	0.46454	0.60320	0.78175	7.48393
	SVR	0.36148	0.46016	0.34174	5.70919		SVR	0.47061	0.61487	0.78451	7.62867
	BPNN	0.41870	0.52997	0.21891	6.56717		BPNN	0.55528	0.69628	0.40785	8.62806
	LSTM	0.35511	0.45644	0.24724	5.65605		LSTM	0.51543	0.65367	0.39224	8.10000
	GRU	0.34621	0.45328	0.22944	5.61691		GRU	0.46440	0.62107	0.33552	7.69600
	TCN	0.33825	0.43894	0.18861	5.43919		TCN	0.45576	0.60434	0.33619	7.48876
	EPT-CEEMDAN-BPNN	0.26951	0.33657	0.12004	4.17058		EPT-CEEMDAN-BPNN	0.32740	0.41192	0.20865	5.10428
	EPT-CEEMDAN-LSTM	0.19916	0.24823	0.10549	3.07591		EPT-CEEMDAN-LSTM	0.26070	0.32602	0.13913	4.03991
	EPT-CEEMDAN-GRU	0.21274	0.26546	0.09799	3.28951		EPT-CEEMDAN-GRU	0.30912	0.38894	0.18256	4.81961
	EPT-CEEMDAN-TCN	**0.17790**	**0.22361**	**0.09606**	**2.77083**		EPT-CEEMDAN-TCN	**0.25483**	**0.31743**	**0.13475**	**3.93352**
2-Step	ARIMA	0.44602	0.58380	0.32619	7.23423	4-Step	ARIMA	0.52962	0.68565	0.35423	8.49626
	SARIMA	0.40394	0.52897	0.73741	6.56296		SARIMA	0.51531	0.66953	0.83149	8.30677
	SVR	0.42099	0.54412	0.66762	6.75084		SVR	0.52298	0.68318	0.84677	8.47612
	BPNN	0.44937	0.57007	0.33142	7.06412		BPNN	0.53191	0.68667	0.39838	8.50894
	LSTM	0.42665	0.56790	0.30204	7.03723		LSTM	0.50766	0.67674	0.36210	8.38585
	GRU	0.42545	0.56861	0.31315	7.04598		GRU	0.51403	0.66905	0.39055	8.29062
	TCN	0.40297	0.52716	0.28493	6.53229		TCN	0.51048	0.66664	0.33572	8.26067
	EPT-CEEMDAN-BPNN	0.35423	0.43967	0.17721	5.44816		EPT-CEEMDAN-BPNN	0.44651	0.55103	0.28381	6.82813
	EPT-CEEMDAN-LSTM	0.27406	0.34165	0.14795	4.23358		EPT-CEEMDAN-LSTM	0.32803	0.41264	0.22263	5.11332
	EPT-CEEMDAN-GRU	0.27739	0.34532	0.15790	4.27907		EPT-CEEMDAN-GRU	0.31405	0.39014	0.20209	4.83451
	EPT-CEEMDAN-TCN	**0.22549**	**0.28258**	**0.13516**	**3.50157**		EPT-CEEMDAN-TCN	**0.27826**	**0.33971**	**0.15826**	**4.20959**

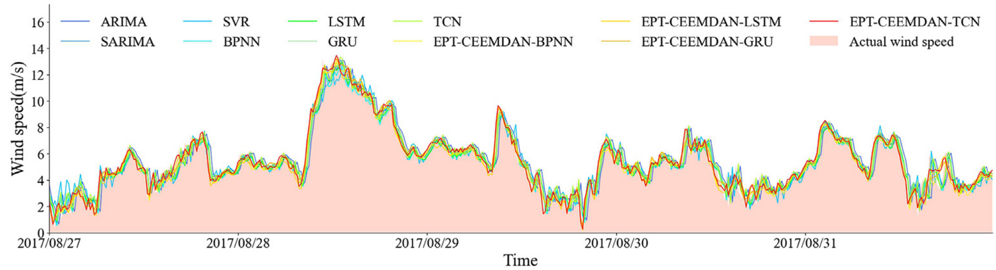
smaller variation.

Furthermore, the probability density functions of the errors are fitted under the normal distribution in Fig. 11. Evidently, the errors from hybrid models with EPT-CEEMDAN decomposition reach higher kurtoses than the errors from the other four single models where EPT-CEEMDAN-TCN achieves the highest kurtosis on each dataset. It confirms the distinction of EPT-CEEMDAN decomposition as well as the high-precision and robustness of EPT-CEEMDAN-TCN.

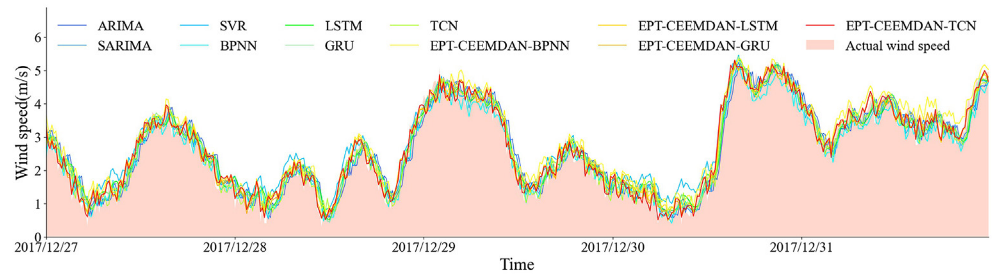
3.3.4. Stability analysis of forecasting results

Aiming to test the forecasting stability further, the variance of absolute error (VAE) of one-step-ahead (15 min ahead) to four-step-ahead (1 h ahead) forecasting results are calculated according to Eq. (26). As is listed in Table 10, the proposed EPT-CEEMDAN-TCN obtains the lowest VAE, exhibiting its stronger stability than all the other benchmark models in multi-step-ahead forecasting on the three datasets. Take the case of Liaoning for instance. The VAE of EPT-CEEMDAN-TCN is 0.01384, 0.02005, 0.02657 and 0.03328 in one-, two-, three- and four-step-ahead forecasting orderly, which are relatively lower than the corresponding VAE of the benchmark models.

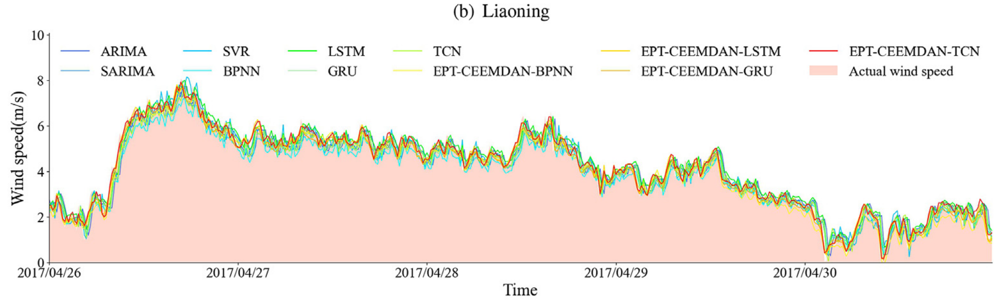
The improvement percentage on VAE of the proposed EPT-CEEMDAN-TCN compared with the other benchmark models is presented in Fig. 12. Comparing with the single models, i.e. ARIMA, SARIMA, SVR, BPNN, LSTM, GRU and TCN, the EPT-CEEMDAN-TCN reduces VAE by seventy percent approximately. It implies that the proposed EPT-CEEMDAN decomposition also plays a critical role in prediction stability. Comparing with the benchmark hybrid models, including EPT-CEEMDAN-BPNN, EPT-CEEMDAN-LSTM and EPT-CEEMDAN-GRU, the EPT-CEEMDAN-TCN deceases VAE by 4.32%–63.94%. It verifies the prediction based on the TCN is stable.



(a) Gansu



(b) Liaoning



(c) Jiangsu

Fig. 7. One-step-ahead forecasting results.

Table 7
The improvement percentages on accuracy of the proposed EPT-CEEMDAN-TCN compared with the benchmark models in Gansu.

Prediction horizon	Model	MAE	RMSE	MAPE	NRMSE	Prediction horizon	Model	MAE	RMSE	MAPE	NRMSE
1-Step	ARIMA	61.50%	59.71%	61.52%	59.71%	3-Step	ARIMA	55.07%	56.45%	52.76%	56.45%
	SARIMA	48.24%	48.19%	49.35%	48.19%		SARIMA	48.46%	50.03%	47.94%	50.03%
	SVR	54.56%	53.60%	56.92%	53.60%		SVR	53.36%	55.38%	52.90%	55.38%
	BPNN	51.54%	50.50%	49.67%	50.50%		BPNN	56.40%	55.52%	56.62%	55.52%
	LSTM	50.20%	49.22%	50.55%	49.22%		LSTM	51.19%	52.40%	51.84%	52.40%
	GRU	49.41%	48.48%	51.44%	48.48%		GRU	49.05%	50.94%	49.41%	50.94%
	TCN	48.41%	48.18%	50.75%	48.18%		TCN	47.93%	49.89%	46.56%	49.89%
	EPT-CEEMDAN-BPNN	23.14%	20.16%	18.45%	20.16%		EPT-CEEMDAN-BPNN	35.48%	32.63%	27.88%	32.63%
	EPT-CEEMDAN-LSTM	16.96%	12.61%	12.36%	12.60%		EPT-CEEMDAN-LSTM	17.80%	17.07%	18.52%	17.07%
	EPT-CEEMDAN-GRU	19.38%	15.99%	20.63%	15.99%		EPT-CEEMDAN-GRU	19.32%	18.03%	24.73%	18.03%
2-Step	ARIMA	50.05%	52.46%	47.46%	52.46%	4-Step	ARIMA	55.07%	55.75%	49.95%	55.75%
	SARIMA	42.31%	44.11%	41.71%	44.11%		SARIMA	48.08%	49.90%	44.73%	49.90%
	SVR	47.45%	50.07%	46.57%	50.07%		SVR	54.07%	55.34%	50.17%	55.34%
	BPNN	50.20%	49.39%	49.62%	49.39%		BPNN	53.74%	53.83%	51.80%	53.83%
	LSTM	47.13%	49.09%	46.64%	49.09%		LSTM	51.54%	52.44%	50.17%	52.44%
	GRU	46.03%	47.19%	47.14%	47.19%		GRU	49.23%	51.46%	45.44%	51.46%
	TCN	42.45%	44.40%	42.02%	44.40%		TCN	47.97%	49.81%	44.02%	49.81%
	EPT-CEEMDAN-BPNN	31.81%	29.39%	30.45%	29.39%		EPT-CEEMDAN-BPNN	27.63%	25.86%	29.16%	25.86%
	EPT-CEEMDAN-LSTM	5.99%	5.31%	0.02%	5.31%		EPT-CEEMDAN-LSTM	12.32%	13.00%	13.84%	13.00%
	EPT-CEEMDAN-GRU	14.14%	11.85%	3.56%	11.85%		EPT-CEEMDAN-GRU	18.67%	18.15%	10.25%	18.15%

3.3.5. Significance evaluation

Diebold-Mariano (DM) test [91] is utilized to evaluate the statistical significance of the superiority of the proposed EPT-CEEMDAN-TCN model comparing with the other ten benchmark models, where the mean square error on testing samples is adopted as the loss function. Under the null hypothesis of equal forecast accuracy across models, the DM statistic is asymptotically $N(0, 1)$. Accordingly, we accept the null hypothesis at the α significance level only if the DM statistic falls in the confidence interval $\left[-Z_{\alpha/2}, Z_{\alpha/2}\right]$. On the contrary, we reject the null hypothesis when

Table 8
The improvement percentages on accuracy of the proposed EPT-CEEMDAN-TCN compared with the benchmark models in Liaoning.

Prediction horizon	Model	MAE	RMSE	MAPE	NRMSE	Prediction horizon	Model	MAE	RMSE	MAPE	NRMSE
1-Step	ARIMA	53.98%	54.30%	54.50%	54.30%	3-Step	ARIMA	45.89%	48.01%	45.87%	48.01%
	SARIMA	48.30%	48.03%	48.84%	48.03%		SARIMA	38.45%	40.97%	40.96%	40.97%
	SVR	54.75%	54.31%	62.67%	54.31%		SVR	43.45%	45.52%	54.66%	45.52%
	BPNN	52.55%	51.47%	48.87%	51.47%		BPNN	46.13%	47.91%	50.20%	47.91%
	LSTM	48.98%	48.50%	53.03%	48.50%		LSTM	36.65%	38.71%	43.40%	38.71%
	GRU	48.42%	47.64%	51.39%	47.64%		GRU	36.39%	38.49%	43.63%	38.49%
	TCN	49.11%	47.32%	49.77%	47.32%		TCN	34.56%	37.03%	40.60%	37.03%
	EPT-CEEMDAN-BPNN	44.85%	42.16%	49.40%	42.16%		EPT-CEEMDAN-BPNN	27.94%	28.89%	39.15%	28.89%
	EPT-CEEMDAN-LSTM	13.62%	12.61%	20.62%	12.61%		EPT-CEEMDAN-LSTM	12.33%	11.83%	23.91%	11.83%
	EPT-CEEMDAN-GRU	7.54%	5.71%	7.39%	5.71%		EPT-CEEMDAN-GRU	5.18%	4.80%	10.86%	4.80%
2-Step	ARIMA	43.61%	46.77%	43.81%	46.77%	4-Step	ARIMA	45.64%	47.94%	40.46%	47.94%
	SARIMA	37.49%	40.31%	38.52%	40.31%		SARIMA	39.62%	41.89%	37.46%	41.89%
	SVR	40.83%	43.03%	50.03%	43.03%		SVR	44.79%	45.71%	52.23%	45.71%
	BPNN	40.58%	43.18%	43.06%	43.18%		BPNN	43.92%	45.79%	33.87%	45.79%
	LSTM	38.01%	40.78%	45.71%	40.78%		LSTM	43.16%	44.28%	48.34%	44.28%
	GRU	36.89%	39.88%	43.82%	39.88%		GRU	41.92%	43.15%	47.90%	43.15%
	TCN	36.73%	38.94%	43.13%	38.94%		TCN	36.48%	38.20%	37.39%	38.20%
	EPT-CEEMDAN-BPNN	38.29%	38.81%	32.23%	38.81%		EPT-CEEMDAN-BPNN	30.38%	32.44%	15.36%	32.44%
	EPT-CEEMDAN-LSTM	20.47%	22.55%	17.76%	22.55%		EPT-CEEMDAN-LSTM	2.50%	2.68%	1.51%	2.68%
	EPT-CEEMDAN-GRU	4.53%	4.07%	2.70%	4.07%		EPT-CEEMDAN-GRU	6.63%	7.64%	12.50%	7.64%

Table 9
The improvement percentages on accuracy of the proposed EPT-CEEMDAN-TCN compared with the benchmark models in Jiangsu.

Prediction horizon	Model	MAE	RMSE	MAPE	NRMSE	Prediction horizon	Model	MAE	RMSE	MAPE	NRMSE
1-Step	ARIMA	55.32%	56.98%	66.91%	56.98%	3-Step	ARIMA	48.70%	50.36%	59.70%	50.36%
	SARIMA	45.39%	47.59%	75.75%	47.65%		SARIMA	45.14%	47.38%	82.76%	47.44%
	SVR	50.79%	51.41%	71.89%	51.47%		SVR	45.85%	48.37%	82.82%	48.44%
	BPNN	57.51%	57.81%	56.12%	57.81%		BPNN	54.11%	54.41%	66.96%	54.41%
	LSTM	49.90%	51.01%	61.15%	51.01%		LSTM	50.56%	51.44%	65.65%	51.44%
	GRU	48.61%	50.67%	58.13%	50.67%		GRU	45.13%	48.89%	59.84%	48.89%
	TCN	47.41%	49.06%	49.07%	49.06%		TCN	44.09%	47.47%	59.92%	47.47%
	EPT-CEEMDAN-BPNN	33.99%	33.56%	19.98%	33.56%		EPT-CEEMDAN-BPNN	22.17%	22.94%	35.42%	22.94%
	EPT-CEEMDAN-LSTM	10.67%	9.92%	8.94%	9.92%		EPT-CEEMDAN-LSTM	2.25%	2.63%	3.15%	2.63%
	EPT-CEEMDAN-GRU	16.38%	15.77%	1.97%	15.77%		EPT-CEEMDAN-GRU	17.56%	18.39%	26.19%	18.39%
2-Step	ARIMA	49.44%	51.60%	58.56%	51.60%	4-Step	ARIMA	47.46%	50.45%	55.32%	50.45%
	SARIMA	44.18%	46.58%	81.67%	46.65%		SARIMA	46.00%	49.26%	80.97%	49.32%
	SVR	46.44%	48.07%	79.75%	48.13%		SVR	46.79%	50.27%	81.31%	50.34%
	BPNN	49.82%	50.43%	59.22%	50.43%		BPNN	47.69%	50.53%	60.27%	50.53%
	LSTM	47.15%	50.24%	55.25%	50.24%		LSTM	45.19%	49.80%	56.29%	49.80%
	GRU	47.00%	50.30%	56.84%	50.30%		GRU	45.87%	49.22%	59.48%	49.22%
	TCN	44.04%	46.40%	52.56%	46.40%		TCN	45.49%	49.04%	52.86%	49.04%
	EPT-CEEMDAN-BPNN	36.34%	35.73%	23.73%	35.73%		EPT-CEEMDAN-BPNN	37.68%	38.35%	44.24%	38.35%
	EPT-CEEMDAN-LSTM	17.72%	17.29%	8.64%	17.29%		EPT-CEEMDAN-LSTM	15.17%	17.67%	28.91%	17.67%
	EPT-CEEMDAN-GRU	18.71%	18.17%	14.40%	18.17%		EPT-CEEMDAN-GRU	11.40%	12.93%	21.69%	12.93%

the DM statistic falls outside the confidence interval, holding that the accuracy of the proposed model is significantly different from the benchmark model.

As is listed in Table 11, the absolute value of DM statistic is generally larger than the critical value of the 1% significance level as 2.58. More specifically, the DM statistics among the single models, such as ARIMA, SARIMA, SVR, BPNN, LSTM, GRU and TCN, range from −12.74 to −8.33, which are far less than −2.58. It proves the significant improvement of the proposed EPT-CEEMDAN-TCN model comparing with the single models. Among the hybrid benchmark models based on EPT-CEEMDAN decomposition, the DM statistic is significant in most cases. The superiority of EPT-CEEMDAN-TCN model still holds.

3.4. Analysis on the hyper parameters

To analysis the effect of the hyper parameters in TCN, we apply it on dataset 2 with different parameter sets, which vary in input dimension, dropout rate, hidden size, and filter size.

Table 12 presents the one-step-ahead forecasting error (RMSE) on training set, validation set and testing set. It implies that the forecasting accuracy of TCN changes with parameter sets. As for input dimension, we set it as 12, 16, 20 respectively with the other parameters unchanged. When 16 is selected, TCN realizes the minimum RMSE on the three sets. Likewise, we set filter size as 1, 2 and 3, hidden size as 32, 64 and 128, dropout rate as 0.00, 0.05 and 0.10. Among these various parameter sets, the selected one always achieves the most accurate forecasting.

3.5. Comparison of decomposition methods

To compare the prediction performance with different decomposition methods, different forecasting techniques and decomposition methods are combined for forecasting the wind speed on dataset 1, the results of which are listed in Table 13. The widely used VMD [68], EMD [69], CEEMDAN [77] and newly developed WSTD [31], VMD-SSA [46] are adopted as the contrast decomposition methods herein.

Following conclusions can be drawn from sufficient analyses. Firstly, when combining with the EPT-CEEMDAN decomposition,

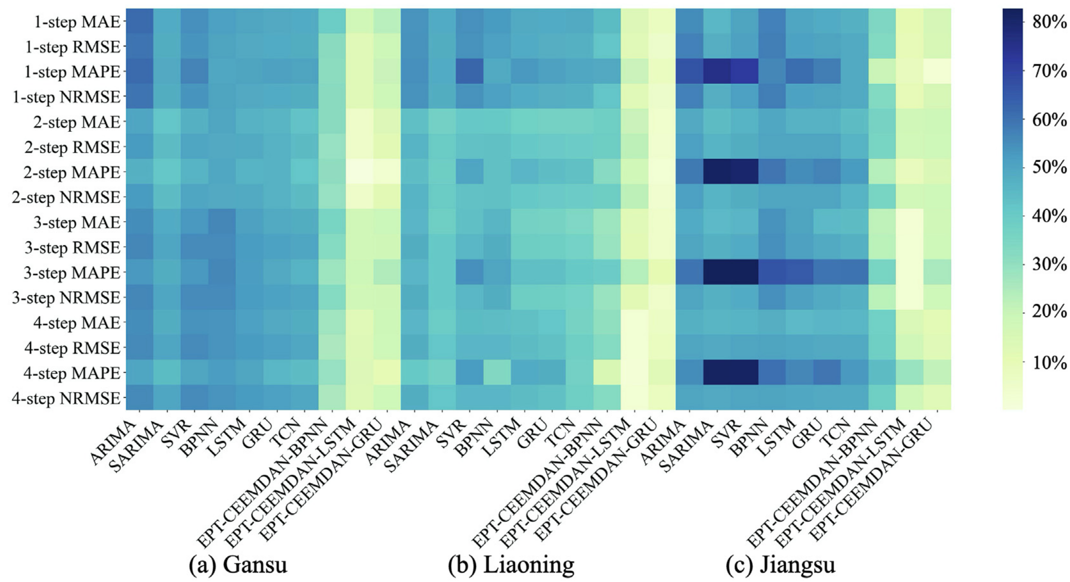


Fig. 8. The improvement percentages on accuracy of the proposed EPT-CEEMDAN-TCN compared with the benchmark models.

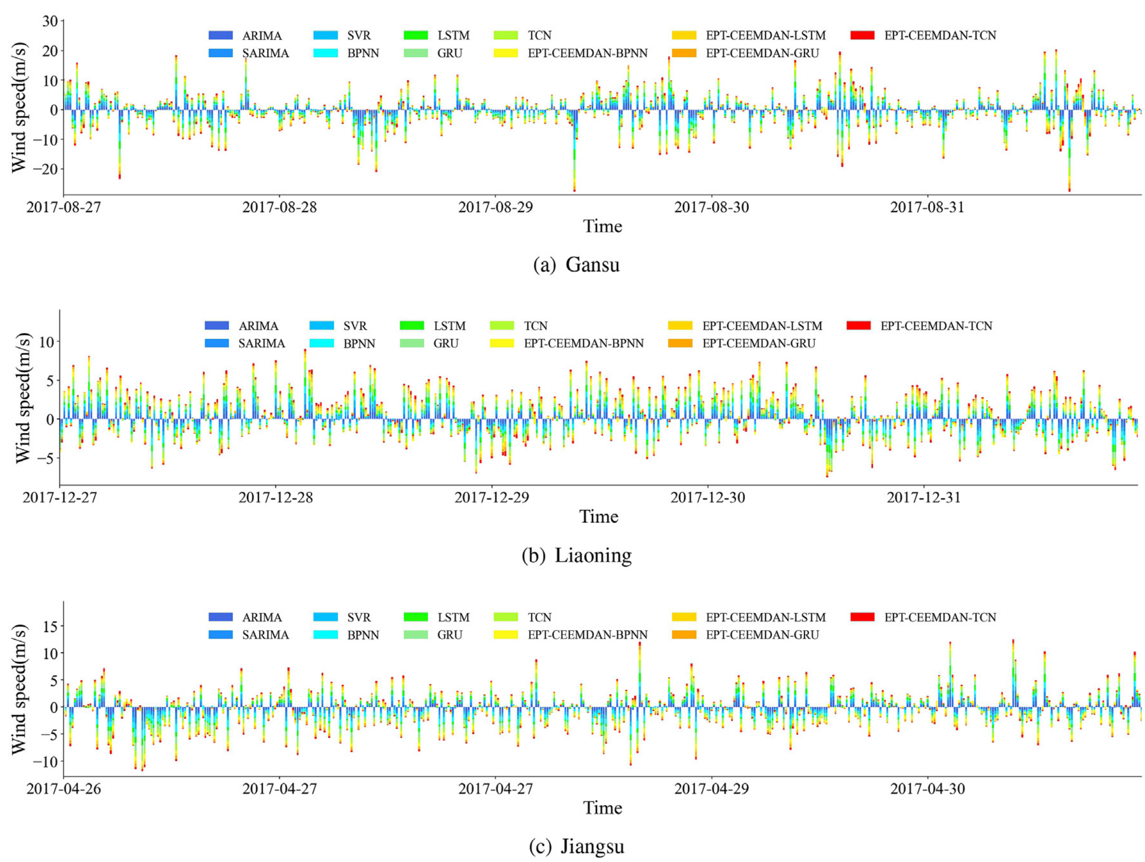
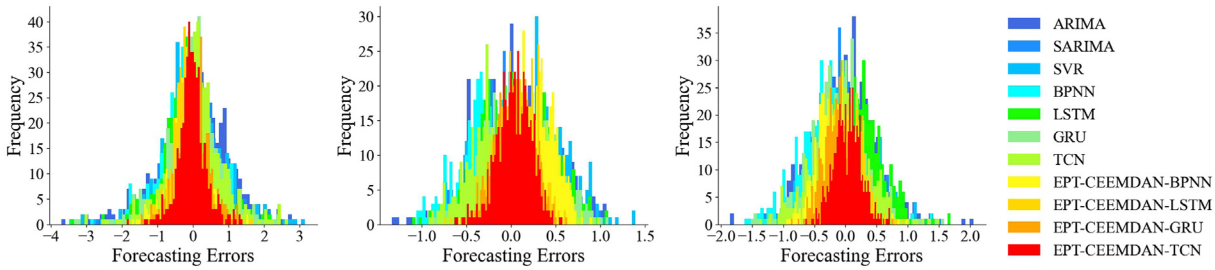


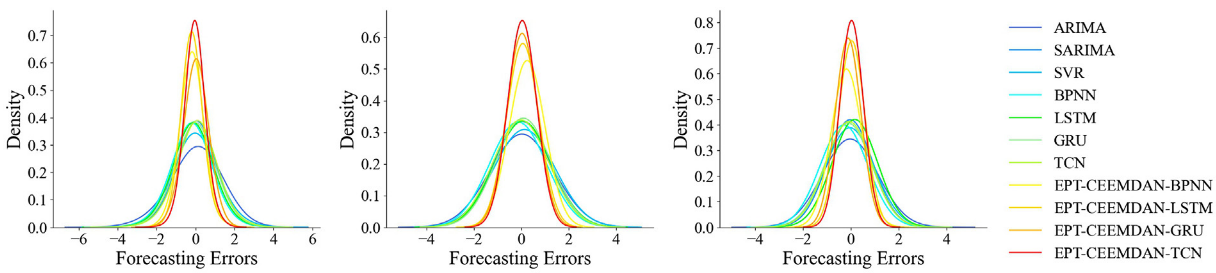
Fig. 9. One-step-ahead forecasting errors.

each forecasting technique, except ARIMA model, reaches the minimum of forecasting root mean square errors, i.e. 0.65575 m/s, 0.50294 m/s, 0.45949 m/s, 0.47798 m/s and 0.40157 m/s. It verifies the superiority of the proposed decomposition method. Secondly, cooperating with the more accurate forecasting technique, the decomposition methods perform more effectively. Specifically, compared with the traditional methods (ARIMA and SVR), the deep learning models (BPNN, LSTM, GRU and TCN) can realize greater accuracy improvements through decomposition. For example, EPT-CEEMDAN decomposition reduces the prediction RMSE of TCN from 0.77494 m/s to 0.40157 m/s by 48.18%, but it only reduces the RMSE of ARIMA from 0.99677 m/s to 0.73334 m/s by 26.43%.



(a) Gansu (b) Liaoning (c) Jiangsu

Fig. 10. The distributions of one-step-ahead forecasting errors.



(a) Gansu (b) Liaoning (c) Jiangsu

Fig. 11. The fitting probability density functions of one-step-ahead forecasting errors.

Table 10
The multi-step-ahead forecasting stability results (VAE) in Gansu, Liaoning and Jiangsu.

Prediction horizon	ARIMA	SARIMA	SVR	BPNN	LSTM	GRU	TCN	EPT-CEEM DAN-BPNN	EPT-CEEM DAN-LSTM	EPT-CEEM DAN-GRU	EPT-CEEM DAN-TCN
Gansu											
1-step-ahead	0.43051	0.28983	0.34475	0.30263	0.28887	0.28145	0.28688	0.11167	0.09009	0.10004	**0.07780**
2-step-ahead	0.60657	0.42029	0.55227	0.44308	0.51380	0.45721	0.42692	0.21309	0.13040	0.14104	**0.11960**
3-step-ahead	0.76045	0.57817	0.74710	0.62996	0.62623	0.60892	0.58410	0.25384	0.18771	0.18808	**0.13255**
4-step-ahead	0.86360	0.71292	0.87814	0.76443	0.75500	0.77563	0.70924	0.27599	0.21876	0.23694	**0.16189**
Liaoning											
1-step-ahead	0.06789	0.05043	0.06400	0.05398	0.05043	0.04776	0.04358	0.03407	0.01738	0.01447	**0.01384**
2-step-ahead	0.08829	0.06773	0.07204	0.07422	0.06856	0.06712	0.06214	0.05560	0.03735	0.02131	**0.02005**
3-step-ahead	0.11386	0.08908	0.10279	0.11101	0.08010	0.07961	0.07728	0.05540	0.03336	0.02883	**0.02657**
4-step-ahead	0.14417	0.11404	0.12070	0.12877	0.11595	0.11197	0.09687	0.08185	0.03540	0.04079	**0.03328**
Jiangsu											
1-step-ahead	0.11164	0.07607	0.08108	0.10556	0.08223	0.08561	0.07826	0.04064	0.02195	0.02521	**0.01835**
2-step-ahead	0.14189	0.11689	0.11883	0.12305	0.14049	0.14231	0.11551	0.06783	0.04162	0.04230	**0.02900**
3-step-ahead	0.16216	0.14837	0.15660	0.17648	0.16162	0.15476	0.15751	0.06248	0.03832	0.05041	**0.03583**
4-step-ahead	0.189613	0.18310033	0.19322	0.18859	0.20025	0.18929	0.18381	0.10427	0.06267	0.06676	**0.03798**

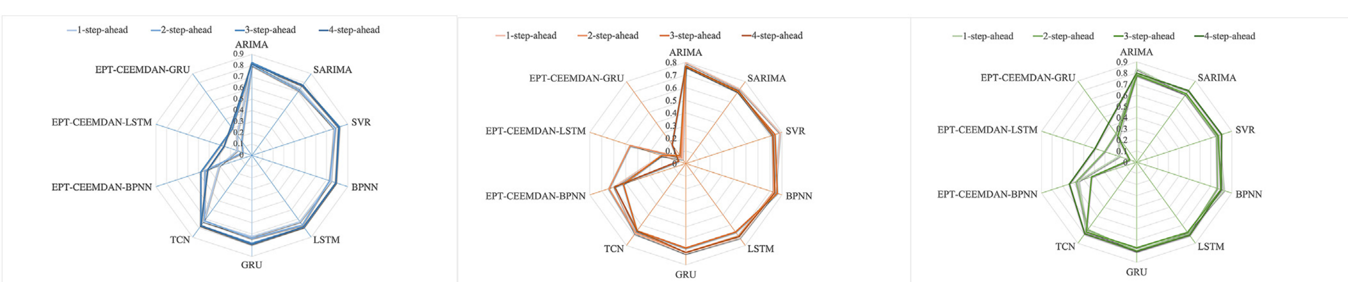


Fig. 12. The improvement percentages on VAE of the proposed EPT-CEEMDAN-TCN compared with the benchmark models.

Thirdly, based on the same decomposition method, TCN generally achieves the best prediction accuracy, further contributing to the accuracy improvement. Last but not the least, among the 42 combinations of 6 forecasting models and 7 decomposition techniques, the proposed EPT-CEEMDAN-TCN model accomplishes the minimum of forecasting root mean square errors, i.e. 0.40157 m/s,

Table 11
The results of Diebold-Mariano test.

Prediction horizon	ARIMA	SARIMA	SVR	BPNN	LSTM	GRU	TCN	EPT-CEEM DAN-BPNN	EPT-CEEM DAN-LSTM	EPT-CEEM DAN-GRU
Gansu										
1-step-ahead	−10.9***	−8.33***	−9.29***	−8.95***	−8.66***	−8.65***	−8.45***	−6.22***	−6.07***	−4.83***
2-step-ahead	−9.81***	−9.54***	−9.31***	−11.93***	−9.11***	−10.06***	−9.21***	−10.29***	−1.78*	−3.92***
3-step-ahead	−10.36***	−9.73***	−9.74***	−12.25***	−9.62***	−9.4***	−9.15***	−10.09***	−6.25***	−6.16***
4-step-ahead	−10.55***	−9.75***	−10.07***	−11.71***	−10.22***	−9.01***	−9.12***	−8.7***	−4.91***	−6.06***
Liaoning										
1-step-ahead	−12.47***	−11.64***	−12.74***	−12.66***	−11.74***	−11.92***	−12.66***	−14.1***	−5.8***	−2.63***
2-step-ahead	−10.39***	−10.39***	−11.28***	−9.94***	−10.06***	−9.9***	−10.2***	−11.23***	−6.77***	−1.55
3-step-ahead	−11.07***	−10.26***	−11.38***	−12.36***	−9.84***	−10.04***	−9.26***	−9.65***	−4.44***	−2.23**
4-step-ahead	−10.56***	−10.23***	−12.61***	−10.95***	−11.54***	−11.93***	−9.85***	−8.08***	−0.95	−3.12***
Jiangsu										
1-step-ahead	−10.6***	−10.53***	−12.01***	−12.71***	−11.15***	−10.49***	−11.39***	−9.93***	−4.17***	−5.42***
2-step-ahead	−10.34***	−9.18***	−9.24***	−10.83***	−9.54***	−9.42***	−8.97***	−9.54***	−4.99***	−7.64***
3-step-ahead	−11.06***	−9.98***	−9.72***	−12.23***	−11***	−9.54***	−9.67***	−6.86***	−1.32	−7.67***
4-step-ahead	−10.04***	−10.64***	−9.65***	−10.38***	−8.92***	−9.23***	−10.25***	−12.02***	−6.7***	−8.67***

*** is the 1% significance level.
** is the 5% significance level.
* is the 10% significance level.

Table 12
The forecasting accuracy of TCN with the different parameter sets.

Interested parameter	Parameter sets				RMSE (m/s)		
	Input dimension	Filter size	Hidden size	Dropout	Training	Validation	Testing
(The selected set)	**16**	**2**	**64**	**0.05**	**0.34746**	**0.36446**	**0.37182**
Input dimension	12	2	64	0.05	0.37866	0.38431	0.38979
	20	2	64	0.05	0.37056	0.39128	0.44916
Filter size	16	1	64	0.05	0.40870	0.41338	0.41276
	16	3	64	0.05	0.39486	0.40281	0.40431
Hidden size	16	2	32	0.05	0.36160	0.38175	0.39708
	16	2	128	0.05	0.34999	0.36688	0.39477
Dropout	16	2	64	0.00	0.34786	0.36515	0.37934
	16	2	64	0.10	0.35495	0.36459	0.40388

Table 13
The comparison on the combinations of different forecasting techniques and decomposition methods in RMSE (m/s).

Decomposition method	ARIMA	SVR	BPNN	LSTM	GRU	TCN
Non-decomposition	0.99677	0.85180	0.81122	0.79083	0.77943	0.77494
WSTD	0.80608	0.88297	0.71328	0.71590	0.68382	0.73238
VMD	**0.71914**	0.77627	0.62559	0.62877	0.62417*	0.60346
VMD-SSA	0.72625	0.71218	0.57107	0.59579**	0.59159	0.54632
EMD	0.73449	0.70997	0.55308	0.50348	0.51003	0.49142
CEEMDAN	0.73370	0.68248	0.53853	0.49041	0.50207	0.45153
EPT-CEEMDAN	0.73334	**0.65575**	**0.50294**	**0.45949**	**0.47798**	**0.40157**

* The WSTD-GRU model is proposed for wind speed forecasting by Peng et al. [31].
** The VMD-SSA-LSTM model is proposed for wind speed forecasting by Rodrigues Moreno et al. [46].

unveiling the best prediction performance.

3.6. Real-time forecasting strategy

In the practical application, only the data up to the current time t is ‘known’ and can be decomposed at time t. In this case, the decomposition process should be implemented successively with the new arrival of data in real-time forecasting, which can moderate the nonstationarity of original series, but may amplify the end effect and the fluctuation in each subseries, and yield illusive components, leading to a loss of effectiveness [80,84]. On this account, Wang et al. [71] proposed an approximated forecasting method based on EMD, where the approximated time series generated by removing the highest frequency portion IMF_1 from the original series [70] is forecasted by a single prediction model. Whereas the IMF_1 varies with the new obtained data in real-time decomposition, it only accounts for a negligibly small proportion of the original series and will not produce a significant change of the approximated series. Accordingly, the real-time wind speed and solar irradiation forecasting cases indicated that this method transcended the existing EMD-based prediction algorithms and the non-decomposition based model. Following this approximated forecasting strategy, we construct the real-time EPT-CEEMDAN-TCN and CEEMDAN-TCN forecasting models as illustrated in Fig. 13.

To evaluate the real-time forecasting performance of the proposed EPT-CEEMDAN-TCN model, we apply it on the short-term wind speed data collected from Gansu in March 2017, along with the TCN model and the real-time CEEMDAN-TCN model for comparison. As is displayed Fig. 14 and Table 14, the whole wind speed dataset is divided into three parts in consistency with Section 3.1.1.

Table 15 and Fig. 15 exhibit the real-time forecasting performance on the testing set, where EPT-CEEMDAN-TCN still outperforms TCN and CEEMDAN-TCN in both accuracy and stability, achieving the minimum of forecasting errors (MAE, MSE, MAPE and

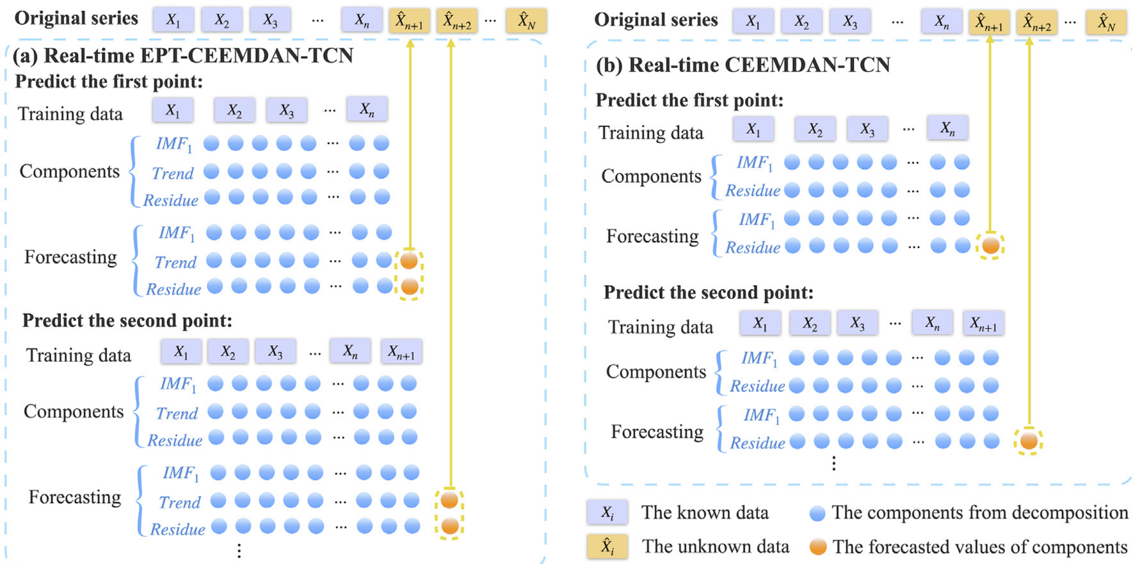


Fig. 13. Schematic illustration of the real-time EPT-CEEMDAN-TCN and CEEMDAN-TCN forecasting models.

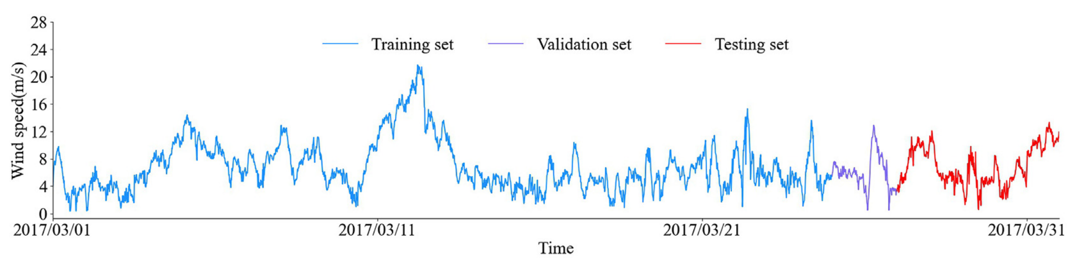


Fig. 14. The wind speed time series.

Table 14
The descriptions of the dataset.

Dataset	Location	Samples	Time interval	# of samples
4	Gansu	All	March 1st - March 31st	2976
		Training	March 1st - March 24th	2304
		Validation	March 25th - March 26th	192
		Testing	March 27th - March 31st	480

Table 15
The real-time forecasting results.

Model	MAE	RMSE	MAPE	NRMSE	VAE
TCN	0.68285	0.88677	0.11963	6.98244	0.32008
CEEMDAN-TCN	0.63219	0.86774	0.11950	6.83258	0.35330
EPT-CEEMDAN-TCN	**0.59888**	**0.82218**	**0.11676**	**6.47387**	**0.31732**

NMAPE) and the variance of absolute error (VAE) among the three models. It verifies the effectiveness and superiority of the EPT-CEEMDAN decomposition in the practical application.

4. Discussion

Wind speed is affected by many factors such as temperature, humidity, air pressure, and surface obstacles, resulting in its nonstationary and nonlinear characteristics. Single models often fail to predict the complex series well. However, the power systems requires high-precision wind speed forecasting, because the higher-precision prediction can save more power system operating costs. Therefore, it is worthwhile to use the more complex models to achieve higher precision predictions.

4.1. The computational efficiency

Compared with other prediction models, the proposed method has higher computational efficiency, for: (i) the decomposition structure can take advantage of the distributed storage and the parallel computing technology to achieve much higher efficiency, for the prediction of each component is independent from each other; (ii) the proposed EPT-CEEMDAN decomposition combines EPT and CEEMDAN, where the CEEMDAN requires much less trials than EEMD; (iii) the convolutions in TCN can be conducted in parallel due to the same filter in each layer, while the predictions in RNNs must be computed in sequence; (iv) TCN possesses the more stable gradients and easy to train, for its backpropagation path is distinct from the temporal direction of the sequence, avoiding the problem of exploding/vanishing gradients in RNNs [32].

4.2. The computational complex

Compared with non-decomposition forecasting models, the proposed model increases the computational complexity within an acceptable range. The EPT-CEEMDAN-TCN model has to construct and train TCNs on each component respectively to obtain the prediction values of subseries, multiplying the computational complexity. For real-time forecasting, all the decomposition-based

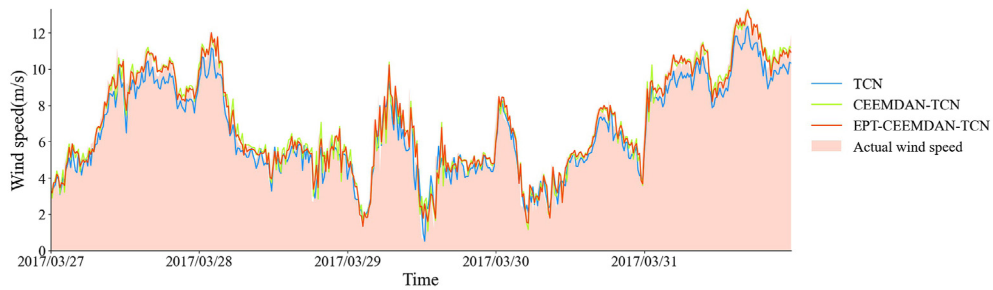


Fig. 15. Real-time forecasting results.

models will generate higher computational complexity than single models, because the time series should be decomposed the step-by-step, which needs more computational resources [80]. Fortunately, the higher-precision prediction will slash the power system operating costs. Considering that, the extra computational complexity is negligible.

5. Conclusion

Recent years have witnessed the vigorous growth of wind power industry. However, considering the intermittency and complexity of the wind, it is challenging to obtain the precise and robust prediction of the wind speed series, which is crucial for the reliable wind power generation. On this account, this paper proposes a novel approach coupling a hybrid decomposition method and the temporal convolutional networks for a more accurate and robust wind speed prediction. To begin with, the ensemble patch transform (EPT) is employed on the wind speed series to extract the trend component of the original series, with the volatility separated out simultaneously. Furthermore, the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) effectively decomposes the volatility into fluctuation components with different frequency characteristics. Subsequently, the temporal convolutional networks (TCN) forecast the trend and fluctuation components individually. Ultimately, the wind speed prediction values are deduced through the reconstruction of the forecasting results. To test the performance of the proposed EPT-CEEMDAN-TCN model, the proposed model and several benchmark models are implemented on various wind speed datasets from three wind farms across China.

The results from the contrast experiments between CEEMDAN and EPT-CEEMDAN decomposition illustrate that the common CEEMDAN method decomposes a sequence into several smooth intrinsic mode functions and a residue overlooking the essential trend of the original sequence, while the proposed EPT-CEEMDAN decomposition has the capability of capturing the vital daily trend of the wind speed series exactly and decomposing the nonlinear volatility completely. Furthermore, the multi-step-ahead wind speed forecasting results on the various datasets attest the remarkable effectiveness of the proposed EPT-CEEMDAN decomposition, as the composite models based on EPT-CEEMDAN decomposition cut the prediction error by nearly a half. Additionally, the comparison experiment results of different decomposition methods confirm the superiority EPT-CEEMDAN decomposition over the widely used and newly developed decomposition methods.

Moreover, among the single models, the temporal convolutional network (TCN) generally outperforms the statistical models (ARIMA, SARIMA), traditional machine learning algorithm (SVR), and other deep learning models (BPNN, LSTM, GRU) in multi-step-ahead wind speed forecasting. In the meanwhile, the proposed EPT-CEEMDAN-TCN model demonstrates the significant superiority on accuracy and stability, for it consistently reaches a more precise and stable prediction than ARIMA, SARIMA, SVR, BPNN, LSTM, GRU, TCN and the hybrid models, such as EPT-CEEMDAN-BPNN, EPT-CEEMDAN-LSTM, and EPT-CEEMDAN-GRU. Even in real-time forecasting, the effectiveness of the proposed EPT-CEEMDAN-TCN model still holds.

In addition, the proposed method has higher computational efficiency with the acceptable computational complexity, especially under a parallel or distributed computation environment. It is beneficial and competitive in the big data era. Furthermore, there is a potential application of the decomposition-prediction framework on the very short-term and large-scale wind speed prediction. For example, Chen et al. [23] successfully applied the framework based on EEMD-GA-LSTM method to the short-term (5 min) wind speed prediction with more than 100,000 records of historical wind speed throughout a year. Besides, the proposed EPT-CEEMDAN-TCN model can also apply to any other nonstationary and nonlinear time series prediction, such as wind power, electricity consumption and solar radiation intensity.

Credit author statement

Dan Li: Conceptualization, Methodology, Software, Validation, Writing. Fuxin Jiang: Software, Validation, and Editing. Min Chen: Data curation, Conceptualization, Methodology, and Supervision. Tao Qian: Methodology and Editing.

Declaration of competing interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Acknowledgments

This research is supported by the National Natural Science Foundation of China, China (Grant No. 11690014, 11731015) and the Science and Technology Development Fund, Macau SAR (File No. 0123/2018/A3).

A. Appendix

The specific parameter settings of these models are listed in Table 1, including the determination approaches.

Table 1
The parameters of the neural networks.

Model	Parameters	Determination approach	Values
ARIMA	AR	Partial Autocorrelation Function	[0,10]
	I	Augmented Dickey–Fuller Test	0 or 1
	MA	Autocorrelation Function	[0,10]
SARIMA	Seasonal period	Preset	96
	AR	Partial Autocorrelation Function	[0,10]
	I	Augmented Dickey–Fuller Test	0 or 1
	MA	Autocorrelation Function	[0,10]
SVR	Regularization coefficient c	Grid search	[1,1000]
	Kernel parameter g	Grid search	$[2^{-10}, 2^{10}]$
BPNN	Input dimension	Trial and error approach	16
	Number of hidden layer nodes	Trial and error approach	64
	Output dimension	Preset	1
	Dropout	Trial and error approach	0.05
	Initial learning rate	Trial and error approach	0.001
	Batch size	Trial and error approach	96
	Maximum of epochs	Preset	200
LSTM	Input dimension	Trial and error approach	16
	Number of hidden layer nodes	Trial and error approach	64
	Output dimension	Preset	1
	Dropout	Trial and error approach	0.05
	Initial learning rate	Trial and error approach	0.001
	Batch size	Trial and error approach	96
	Maximum of epochs	Preset	200
GRU	Input dimension	Trial and error approach	16
	Number of hidden layer nodes	Trial and error approach	64
	Output dimension	Preset	1
	Dropout	Trial and error approach	0.05
	Initial learning rate	Trial and error approach	0.001
	Batch size	Trial and error approach	96
	Maximum of epochs	Preset	200
TCN	Input dimension	Trial and error approach	16
	Number of hidden layer nodes	Trial and error approach	64
	Filter size in each convolutional layer	Trial and error approach	2
	Dilations	Trial and error approach	[1,2,4,8]
	Output dimension	Preset	1
	Dropout	Trial and error approach	0.05
	Initial learning rate	Trial and error approach	0.001
	Batch size	Trial and error approach	96
	Maximum of epochs	Preset	200

References

[1] Dong F, Li W. Research on the coupling coordination degree of "upstream-midstream-downstream" of China's wind power industry chain. J Clean Prod 2021;283:124633. https://doi.org/10.1016/j.jclepro.2020.124633.

[2] World Wind Energy Association. Worldwide wind capacity reaches 744 gigawatts - an unprecedented 93 gigawatts added in 2020. 2021. https://wwindea.org/worldwide-wind-capacity-reaches-744-gigawatts/. [Accessed 24 March 2021]. Accessed.

[3] Wang J, Du P, Niu T, Yang W. A novel hybrid system based on a new proposed algorithm—multi-objective whale optimization algorithm for wind speed forecasting. Appl Energy 2017;208:344–60. https://doi.org/10.1016/j.apenergy.2017.10.031.

[4] Wang L, Li J. Estimation of extreme wind speed in SCS and NWP by a non-stationary model. Theoretical and Applied Mechanics Letters 2016;6(3):131–8. https://doi.org/10.1016/j.taml.2016.04.001.

[5] Zhao J, Guo ZH, Su ZY, Zhao ZY, Xiao X, Liu F. An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed. Appl Energy 2016;162:808–26. https://doi.org/10.1016/j.apenergy.2015.10.145.

[6] Yang J, Astitha M, Monache LD, Alessandrini S. An analog technique to improve storm wind speed prediction using a dual NWP model approach. Mon Weather Rev 01 Dec. 2018;146(12):4057–77. https://doi.org/10.1175/MWR-D-17-0198.1.

[7] Zhao J, Guo Z, Guo Y, Lin W, Zhu W. A self-organizing forecast of day-ahead wind speed: selective ensemble strategy based on numerical weather predictions. Energy 2021;218:119509. https://doi.org/10.1016/j.energy.2020.119509.

[8] Karakuş O, Kuruo§lu EE, Altınkaya MA. One-day ahead wind speed/power prediction based on polynomial autoregressive model. IET Renew Power Gener 2017;11(11):1430–9. https://doi.org/10.1049/iet-rpg.2016.0972.

[9] Kavasseri RG, Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. Renew Energy 2009;34(5):1388–93. https://doi.org/10.1016/j.renene.2008.09.006.

[10] Erdem E, Shi J. ARMA based approaches for forecasting the tuple of wind speed and direction. Appl Energy 2011;88(4):1405–14. https://doi.org/10.1016/j.apenergy.2010.10.031.

[11] Liu X, Lin Z, Feng Z. Short-term offshore wind speed forecast by seasonal ARIMA - a comparison against GRU and LSTM. Energy 2021;227:120492. https://doi.org/10.1016/j.energy.2021.120492.

[12] Shen Z, Ritter M. Forecasting volatility of wind power production. Appl Energy 2016;176:295–308. https://doi.org/10.1016/j.apenergy.2016.05.071.

[13] 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. https://doi.org/10.1016/j.renene.2015.07.004.

[14] Chen J, Zeng GQ, Zhou W, Du W, Lu KD. Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization. Energy Convers Manag 2018;165:681–95. https://doi.org/10.1016/j.enconman.2018.03.098.

[15] Liu M, Cao Z, Zhang J, Wang L, Huang C, Luo X. Short-term wind speed forecasting based on the Jaya-SVM model. Int J Electr Power Energy Syst 2020;121:106056. https://doi.org/10.1016/j.ijepes.2020.106056.

[16] 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 Base Syst 2014;56:226–39. https://doi.org/10.1016/j.knosys.2013.11.015.

[17] 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. https://doi.org/10.1016/j.renene.2016.03.103.

[18] Yang Z, Wang J. A hybrid forecasting approach applied in wind speed forecasting based on a data processing strategy and an optimized artificial intelligence algorithm. Energy 2018;160:87–100. https://doi.org/10.1016/j.energy.2018.07.005.

[19] Duan J, Zuo H, Bai Y, Duan J, Chang M, Chen B. Short-term wind speed forecasting using recurrent neural networks with error correction. Energy 2021;217:119397. https://doi.org/10.1016/j.energy.2020.119397.

[20] Ergen T, Kozat SS. Efficient online learning algorithms based on LSTM neural networks. IEEE Transactions on Neural Networks and Learning Systems 2018;29(8):3772–83. https://doi.org/10.1109/TNNLS.2017.2741598.

[21] Duan J, Wang P, Ma W, Tian X, Fang S, Cheng Y, et al. Short-term wind power forecasting using the hybrid model of improved variational mode decomposition and correntropy long short -term memory neural network. Energy 2021;214:118980. https://doi.org/10.1016/j.energy.2020.118980.
[22] Marndi A, Patra G, Gouda KC. Short-term forecasting of wind speed using time division ensemble of hierarchical deep neural networks. Bulletin of Atmospheric Science and Technology 2020;1(1):91–108. https://doi.org/10.1007/s42865-020-00009-2.
[23] Chen Y, Dong Z, Wang Y, Su J, Han Z, Zhou D, et al. Short-term wind speed predicting framework based on EEMD-GA-LSTM method under large scaled wind history. Energy Convers Manag 2021;227:113559. https://doi.org/10.1016/j.enconman.2020.113559.
[24] Yıldırım DC, Toroslu IH, Fiore U. Forecasting directional movement of Forex data using LSTM with technical and macroeconomic indicators. Financial Innovation 2021;7(1):1–36. https://doi.org/10.1186/s40854-020-00220-2.
[25] Chimmula VKR, Zhang L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos, Solit Fractals 2020;135:109864. https://doi.org/10.1016/j.chaos.2020.109864.
[26] Choi H, Cho K, Bengio Y. Context-dependent word representation for neural machine translation. Comput Speech Lang 2017;45:149–60. https://doi.org/10.1016/j.csl.2017.01.007.
[27] Liu H, Mi X, Li Y, Duan Z, Xu Y. Smart wind speed deep learning based multi-step forecasting model using singular spectrum analysis, convolutional gated recurrent unit network and support vector regression. Renew Energy 2019;143:842–54. https://doi.org/10.1016/j.renene.2019.05.039.
[28] Xiang L, Li J, Hu A, Zhang Y. Deterministic and probabilistic multi-step forecasting for short-term wind speed based on secondary decomposition and a deep learning method. Energy Convers Manag 2020;220:113098. https://doi.org/10.1016/j.enconman.2020.113098.
[29] Zhang G, Liu D. Causal convolutional gated recurrent unit network with multiple decomposition methods for short-term wind speed forecasting. Energy Convers Manag 2020;226:113500. https://doi.org/10.1016/j.enconman.2020.113500.
[30] Wei D, Wang J, Niu X, Li Z. Wind speed forecasting system based on gated recurrent units and convolutional spiking neural networks. Appl Energy 2021;292:116842. https://doi.org/10.1016/j.apenergy.2021.116842.
[31] Peng Z, Peng S, Fu L, Lu B, Tang J, Wang K, et al. A novel deep learning ensemble model with data denoising for short-term wind speed forecasting. Energy Convers Manag 2020;207:112524. https://doi.org/10.1016/j.enconman.2020.112524.
[32] Bai S, Kolter JZ, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. 2018. arXiv:1803.01271.
[33] Zhao W, Gao Y, Ji T, Wan X, Ye F, Bai G. Deep temporal convolutional networks for short-term traffic flow forecasting. IEEE Access 2019;7:114496–507. https://doi.org/10.1109/ACCESS.2019.2935504.
[34] Zhang K, Liu Z, Zheng L. Short-term prediction of passenger demand in multi-zone level: temporal convolutional neural network with multi-task learning. IEEE Trans Intell Transport Syst 2020;21(4):1480–90. https://doi.org/10.1109/TITS.2019.2909571.
[35] Yating G, Wu W, Qiongbin L, Fenghuang C, Qinqin C. Fault diagnosis for power converters based on optimized temporal convolutional network. IEEE Transactions on Instrumentation and Measurement 2021;70:1–10. https://doi.org/10.1109/TIM.2020.3021110.
[36] Gan Z, Li C, Zhou J, Tang G. Temporal convolutional networks interval prediction model for wind speed forecasting. Elec Power Syst Res 2021;191:106865. https://doi.org/10.1016/j.epsr.2020.106865.
[37] Büyükşahin Ümit Çavuş, Ertekin Şeyda. Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. Neurocomputing 2019;361:151–63. https://doi.org/10.1016/j.neucom.2019.05.099.
[38] Zhang J, Wei Y, Tan Z. An adaptive hybrid model for short term wind speed forecasting. Energy 2020;190:115615. https://doi.org/10.1016/j.energy.2019.06.132.
[39] Liu H, Yu C, Wu H, Duan Z, Yan G. A new hybrid ensemble deep reinforcement learning model for wind speed short term forecasting. Energy 2020;202:117794. https://doi.org/10.1016/j.energy.2020.117794.
[40] Duan J, Zuo H, Bai Y, Duan J, Chang M, Chen B. Short-term wind speed forecasting using recurrent neural networks with error correction. Energy 2021;217:119397. https://doi.org/10.1016/j.energy.2020.119397.
[41] Jiang P, Liu Z, Niu X, Zhang L. A combined forecasting system based on statistical method, artificial neural networks, and deep learning methods for short-term wind speed forecasting. Energy 2021;217:119361. https://doi.org/10.1016/j.energy.2020.119361.
[42] 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. https://doi.org/10.1016/j.renene.2013.08.011.
[43] Liu H, wei Mi X, fei Li Y. 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 Manag 2018;156:498–514. https://doi.org/10.1016/j.enconman.2017.11.053.
[44] Aasim, Singh SN, Mohapatra A. Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting. Renew Energy 2019;136:758–68. https://doi.org/10.1016/j.renene.2019.01.031.
[45] Tang G, Xue X, Saeed A, Hu X. Short-term wind speed interval prediction based on ensemble GRU model. IEEE Transactions on Sustainable Energy 2019:1. https://doi.org/10.1109/TSTE.2019.2926147.
[46] Rodrigues Moreno S, Gomes da Silva R, Cocco Mariani V, dos Santos Coelho L. Multi-step wind speed forecasting based on hybrid multi-stage decomposition model and long short-term memory neural network. Energy Convers Manag 2020;213:112869. https://doi.org/10.1016/j.enconman.2020.112869.
[47] Hu H, Wang L, Tao R. Wind speed forecasting based on variational mode decomposition and improved echo state network. Renew Energy 2021;164:729–51. https://doi.org/10.1016/j.renene.2020.09.109.
[48] Zhang Y, Yan B, Aasma M. A novel deep learning framework: prediction and analysis of financial time series using CEEMD and LSTM. Expert Syst Appl 2020;159:113609. https://doi.org/10.1016/j.eswa.2020.113609.
[49] Liu MD, Ding L, Bai YL. Application of hybrid model based on empirical mode decomposition, novel recurrent neural networks and the ARIMA to wind speed prediction. Energy Convers Manag 2021;233:113917. https://doi.org/10.1016/j.enconman.2021.113917.
[50] Chen XJ, Zhao J, Jia XZ, Li ZL. Multi-step wind speed forecast based on sample clustering and an optimized hybrid system. Renew Energy 2021;165:595–611. https://doi.org/10.1016/j.renene.2020.11.038.
[51] Pei S, Qin H, Zhang Z, Yao L, Wang Y, Wang C, et al. Wind speed prediction method based on empirical wavelet transform and new Cell update long short-term memory network. Energy Convers Manag 2019;196:779–92. https://doi.org/10.1016/j.enconman.2019.06.041.
[52] Zhang Y, Zhao Y, Kong C, Chen B. A new prediction method based on VMD-PRBF-ARMA-E model considering wind speed characteristic. Energy Convers Manag 2020;203:112254. https://doi.org/10.1016/j.enconman.2019.112254.
[53] Bracewell R. The fourier transform and its applications. second ed. Tokyo: McGraw-Hill Kogakusha, Ltd.; 1978.
[54] Yi Y, Johnson JT, Wang X. On the estimation of wind speed diurnal cycles using simulated measurements of CYGNSS and ASCAT. Geosci Rem Sens Lett IEEE 2019;16(2):168–72. https://doi.org/10.1109/LGRS.2018.2872354.
[55] Aly HH. A novel deep learning intelligent clustered hybrid models for wind speed and power forecasting. Energy 2020;213:118773. https://doi.org/10.1016/j.energy.2020.118773.
[56] Xia G, Draxl C, Raghavendra A, Lundquist JK. Validating simulated mountain wave impacts on hub-height wind speed using sodar observations. Renew Energy 2021;163:2220–30. https://doi.org/10.1016/j.renene.2020.10.127.
[57] Hess-Nielsen N, Wickerhauser M. Wavelets and time-frequency analysis. Proc IEEE 1996;84(4):523–40. https://doi.org/10.1109/5.488698.
[58] 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. https://doi.org/10.1016/j.renene.2013.08.011.
[59] Catalão JPS, Pousinho HMI, Mendes VMF. Hybrid Wavelet-PSO-ANFIS approach for short-term wind power forecasting in Portugal. IEEE Transactions on Sustainable Energy 2011;2(1):50–9. https://doi.org/10.1109/TSTE.2010.2076359.
[60] 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. https://doi.org/10.1016/j.renene.2016.02.054.
[61] Gokhale M, Khanduja DK, et al. Time domain signal analysis using wavelet packet decomposition approach. Int J Commun Netw Syst Sci 2010;3(3):321.
[62] Liu H, qi Tian H, fu Pan D, fei Li Y. Forecasting models for wind speed using wavelet, wavelet packet, time series and artificial neural networks. Appl Energy 2013;107:191–208. https://doi.org/10.1016/j.apenergy.2013.02.002.
[63] 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. https://doi.org/10.1016/j.apenergy.2015.01.038.
[64] Meng A, Ge J, Yin H, Chen S. Wind speed forecasting based on wavelet packet decomposition and artificial neural networks trained by crisscross optimization algorithm. Energy Convers Manag 2016;114:75–88. https://doi.org/10.1016/j.enconman.2016.02.013.
[65] Gilles J. Empirical wavelet transform. IEEE Trans Signal Process 2013;61(16):3999–4010. https://doi.org/10.1109/TSP.2013.2265222.
[66] Hu J, Wang J, Ma K. A hybrid technique for short-term wind speed prediction. Energy 2015;81:563–74. https://doi.org/10.1016/j.energy.2014.12.074.
[67] Golyandina N, Zhigljavsky A. Singular spectrum analysis for time series, ISBN 978-3-642-34913-3. https://doi.org/10.1007/978-3-642-34913-3.
[68] Dragomiretskiy K, Zosso D. Variational mode decomposition. IEEE Trans Signal Process 2014;62(3):531–44. https://doi.org/10.1109/TSP.2013.2288675.
[69] Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London Series A: Mathematical, Physical and Engineering Sciences 1998;454(1971):903–95. https://doi.org/10.1098/rspa.1998.0193.
[70] Guo Z, Zhao W, Lu H, Wang J. Multi-step forecasting for wind speed using a modified emd-based artificial neural network model. Renew Energy 2012;37(1):241–9. https://doi.org/10.1016/j.renene.2011.06.023.
[71] Wang Y, Wu L. On practical challenges of decomposition-based hybrid forecasting algorithms for wind speed and solar irradiation. Energy 2016;112:208–20. https://doi.org/10.1016/j.energy.2016.06.075.
[72] Fu W, Wang K, Tan J, Zhang K. A composite framework coupling multiple feature selection, compound prediction models and novel hybrid swarm optimizer-based synchronization optimization strategy for multi-step ahead short-term wind speed forecasting. Energy Convers Manag 2020;205:112461. https://doi.org/10.1016/j.enconman.2019.112461.
[73] Wu Z, Huang N. Ensemble empirical mode decomposition: a noise-assisted

data analysis method. Adv Adapt Data Anal 2009;1:1−41. https://doi.org/10.1142/S1793536909000047.
[74] Sun N, Zhou J, Chen L, Jia B, Tayyab M, Peng T. An adaptive dynamic short-term wind speed forecasting model using secondary decomposition and an improved regularized extreme learning machine. Energy 2018;165:939−57. https://doi.org/10.1016/j.energy.2018.09.180.
[75] Yeh JR, Shieh JS, Huang N. Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method. Adv Adapt Data Anal 2010;2:135−56. https://doi.org/10.1142/S1793536910000422.
[76] Wu C, Wang J, Chen X, Du P, Yang W. A novel hybrid system based on multi-objective optimization for wind speed forecasting. Renew Energy 2020;146: 149−65. https://doi.org/10.1016/j.renene.2019.04.157.
[77] Torres ME, Colominas MA, Schlotthauer G, Flandrin P. A complete ensemble empirical mode decomposition with adaptive noise. IEEE international conference on acoustics, speech and signal processing. 2011. p. 4144−7. https://doi.org/10.1109/ICASSP.2011.5947265.
[78] 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 Manag 2017;136:439−51. https://doi.org/10.1016/j.enconman.2017.01.022.
[79] Qu Z, Mao W, Zhang K, Zhang W, Li Z. Multi-step wind speed forecasting based on a hybrid decomposition technique and an improved back-propagation neural network. Renew Energy 2019;133:919−29. https://doi.org/10.1016/j.renene.2018.10.043.
[80] Qian Z, Pei Y, Zareipour H, Chen N. A review and discussion of decomposition-based hybrid models for wind energy forecasting applications. Appl Energy 2019;235:939−53. https://doi.org/10.1016/j.apenergy.2018.10.080.
[81] Alkhayat G, Mehmood R. A review and taxonomy of wind and solar energy forecasting methods based on deep learning. Energy and AI 2021;4:100060. https://doi.org/10.1016/j.egyai.2021.100060.
[82] Jun W, Yuyan L, Lingyu T, Peng G. A new weighted CEEMDAN-based prediction model: an experimental investigation of decomposition and non-decomposition approaches. Knowl Base Syst 2018;160:188−99. https://doi.org/10.1016/j.knosys.2018.06.033.
[83] Kim D, Choi G, Oh HS. Ensemble patch transformation: a flexible framework for decomposition and filtering of signal. EURASIP J Appl Signal Process 2020;2020(1):1−27. https://doi.org/10.1186/s13634-020-00690-7.
[84] Jiang Y, Huang G, Peng X, Li Y, Yang Q. A novel wind speed prediction method: hybrid of correlation-aided DWT, LSSVM and GARCH. J Wind Eng Ind Aerod 2018;174:28−38. https://doi.org/10.1016/j.jweia.2017.12.019.
[85] Rilling G, Flandrin P, Goncalves P, et al. On empirical mode decomposition and its algorithms. In: IEEE-EURASIP workshop on nonlinear signal and image processing, vol. 3. Citeseer; 2003. p. 8−11. https://hal.inria.fr/inria-00570628.
[86] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: Proceedings of the 32nd international conference on machine learning, vol. 37. Lille, France; 2015. p. 448−56.
[87] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 2014;15(56):1929−58. http://jmlr.org/papers/v15/srivastava14a.html.
[88] Wang J, Li Y. Multi-step ahead wind speed prediction based on optimal feature extraction, long short term memory neural network and error correction strategy. Appl Energy 2018;230:429−43. https://doi.org/10.1016/j.apenergy.2018.08.114.
[89] Gemperline PJ, Long JR, Gregoriou VG. Nonlinear multivariate calibration using principal components regression and artificial neural networks. Anal Chem 1991;63(20):2313−23. https://doi.org/10.1021/ac00020a022.
[90] Diederik P. Kingma and Jimmy Ba. Adam: a method for stochastic optimization. 2017. arXiv:1412.6980.
[91] Francis D, Roberto M. Comparing predictive accuracy. J Bus Econ Stat 2002;20: 134−44. https://doi.org/10.1080/07350015.1995.10524599.