

Journal of Engineering Science and Technology Review 16 (2) (2023) 197 - 206

Research Article

JOURNAL OF Engineering Science and Technology Review

www.jestr.org

Short-Term Wind Power Forecasting and Uncertainty Analysis Based on Hybrid Temporal Convolutional Network

Yang Jian¹, Yao Shuai¹, Chang Xuejun¹, Li Dewei¹, Gu Bo^{2,*} and Zhang Zichao³

¹Yellow River Engineering Consulting Co., Ltd. Zhengzhou 450003, China ²School of Electrical Engineering, North China University of Water Resources and Electric Power, Zhengzhou 450011, China ³Department of mechanical engineering, Texas Tech University, Texas 79409, United State

Received 20 February 2023; Accepted 11 May 2023

Abstract

The integration of large-scale wind power into power grids has made accurate short-term wind power forecasting a key technology for the safe and economical operation of power grids. A novel method based on variational mode decomposition (VMD), temporal convolutional network (TCN), and Gaussian mixture model (GMM) was proposed for accurate short-term wind power forecasting and uncertainty analysis. First, the wind speed information was decomposed into different mode components via VMD. Second, TCN was employed to capture accurately the time-series dependence of data by training and forecasting different mode component data. On this basis, GMM was used to calculate the distribution characteristics of short-term wind power forecasting errors and quantify the confidence interval of wind power forecasting for 4 h during winter is 4.69%, 3.13%, 2.48%, 1.21%, and 0.7% lower than the RMSE values of wavelet neural network, BP neural network, PSO-BP hybrid model, long short-term memory model, and TCN model, respectively. The proposed method has a certain promoting effect on improving the accuracy of short-term wind power forecasting.

Keywords: Short-term wind power forecasting, Variational mode decomposition, Temporal convolutional networks, Gaussian mixture model, Confidence interval

1. Introduction

The development and utilization of clean energy have become feasible schemes for the fight against climate change and environmental pollution. Wind energy is a type of clean energy widely developed and utilized in recent years [1]. Given the intermittence, randomness, and volatility of wind power, large-scale wind power grid connections have posed great challenges to the safe and stable operation of power grids. Accurately forecasting the wind power output has become an effective approach to ensuring the safe and stable operation of power grids and improving the consumption of wind power [2-3].

According to different forecasting time scales, wind power forecasting can be divided into medium-, long-, short-, and ultrashort-term forecasting [4-5]. Medium-term forecasting and long-term forecasting are mainly used for the maintenance plan of wind farms and the evaluation of annual power generation after the construction of a wind farm [6]. Short-term forecasting is mainly used for power grid dispatching, improving power supply quality, and promoting the participation of wind power in bidding [7]. Ultrashort-term forecasting is typically used for power grid real-time dispatching and wind turbine control [8].

The common statistical forecasting models for shortterm and ultrashort-term wind power forecasting include single-algorithm and combined-algorithm forecasting models [9], parametric and nonparametric forecasting models [10], linear and nonlinear forecasting models [11], and machine learning forecasting models [12]. Example calculations show that machine learning can be applied to short-term wind power forecasting under different terrain conditions, and the forecasting accuracy is relatively high [13]. The forecasting accuracy of the adaptive variational mode decomposition (VMD) and multiple machine learning models combined with short-term and ultra-short-term wind speed and wind power forecasting is higher than that of a single forecasting model [14]. The forecasting accuracy of a forecasting model can be improved to a certain extent by optimizing its weights and thresholds through various optimization algorithms [15-16].

The machine learning models mentioned above have achieved certain results in short-term wind power forecasting. However, they cannot mine the temporal correlation between data; thus, further improvement of their forecasting accuracy is limited [17]. Therefore, deeply mining the temporal correlation between data and constructing accurate short-term wind power forecasting models have become urgent challenges for short-term wind power forecasting.

A short-term wind power forecasting model based on VMD, temporal convolutional network (TCN), and Gaussian mixture model (GMM) was proposed in the present study. First, VMD was used to eliminate the impact of wind speed randomness on forecasting accuracy. Then, the time convolutional network (TCN) was applied to mine the temporal correlation between data deeply. Thus, accurate short-term wind power forecasting was achieved. On this basis, GMM was used to calculate the distribution characteristics of errors in short-term wind power forecasting.

2. State of the Art

The development of artificial intelligence has led to the proposal and use of new intelligent learning algorithms in wind power forecasting, further promoting the development of wind power forecasting technology [18]. Deep learning algorithms are widely studied novel methods of wind power forecasting. These algorithms include recurrent neural networks, deep belief networks, and convolutional neural networks (CNNs) [19]. Ran et al. [20] proposed a short-term wind power forecasting model based on the convolutional long short-term memory neural network (CNN-LSTM). The proposed model extracts the temporal correlation characteristics of data. However, it does not preprocess the input data or analyze the uncertainty of wind power forecasting. Zhou et al. [21] proposed a wind power interval forecasting model based on long short-term memory (LSTM) neural network. The calculation results showed that the wind power interval forecasting model can extract the temporal correlation of data and forecast the distribution interval of wind power. However, the model's accuracy in extracting the temporal correlation features must be further improved. A deep learning model for short-term wind speed and direction forecasting was proposed by [22]. The actual calculation results revealed that the forecasting accuracy of the deep learning model for wind speed and direction is higher than that of the benchmark forecasting model. However, the temporal correlation of data and the uncertainty of forecasting were not considered in the deep learning model.

A day-ahead wind power forecasting model based on CNNs was proposed, and the model parameters were optimized by [23]. The forecasting results showed that the method demonstrates good forecasting performance. However, the model does not consider the temporal correlation between data, and the forecasting uncertainty was not analyzed. Aslam et al. [24] proposed a multistep advanced wind power forecasting model based on the dual attention mechanism. They also optimized the hyperparameters of the model. The study results demonstrated that the forecasting effect of the model is good. However, the model does not preprocess the data or analyze the forecasting uncertainty. Maryam et al. [25] compared and analyzed the performances of LSTM, gated recurrent unit (GRU), CNN, and CNN-LSTM models in the accuracy of wind power forecasting. The forecasting results revealed that the forecasting accuracy of GRU is better than the forecasting accuracies of the three other models. However, data preprocessing methods and forecasting uncertainty were not studied.

Optimizing the hyperparameters of a forecasting model can effectively improve forecasting accuracy [26]. Seyed et al. [27] utilized evolutionary search optimizers to optimize the parameters of deep CNNs, thereby improving model forecasting performance. However, they did not preprocess the data or analyze the uncertainty of wind power forecasting. Ahmed et al. [28] applied a heap optimizer to optimize the parameters of the LSTM model. This approach improves the forecasting accuracy of wind power to some extent. However, the uncertainty of wind power forecasting was not analyzed. A short-term wind power forecasting based on the wavelet-ARIMA model was presented by [29]. The calculation results demonstrated that this model has good forecasting performance. However, the temporal correlation between data was not mined in the model. Dominik et al. [30] proposed a neural unfolding analysis model based on time-series forecasting. This model has good performance in wind power forecasting. However, the data preprocessing and forecasting uncertainty were not considered. A novel interval forecasting method was proposed, and its effectiveness was verified by [31]. However, the data preprocessing and temporal correlation between data were not considered in the forecasting method. Bo et al. [32] considered the temporal correlation between data and constructed a nonlinear mapping network model for wind power forecasting. The effectiveness of the model was verified through examples. However, the related work of data preprocessing and forecasting uncertainty was not considered.

Combining the advantages of multiple forecasting algorithms for achieving wind power forecasting can effectively overcome the shortcomings of a single forecasting algorithm [33-34]. The combination forecasting methods of multiple forecasting models mainly include parameter optimization methods and ensemble forecasting methods [26]. In terms of parameter optimization, Hossain et al. [35] combined a convolutional layer, a gated recursive unit layer, and a fully connected neural network to form a hybrid deep neural network model. Example calculations demonstrated that the forecasting accuracy of this hybrid deep neural network model is higher than that of other forecasting models. However, the uncertainty of wind power forecasting was not analyzed. Wang et al. [36-39] used data decomposition methods, such as wavelet transform, empirical mode decomposition, integrated empirical mode decomposition, and empirical wavelet transform, to decompose wind speed data into high- and low-frequency signals. Then, they used deep learning algorithms to forecast low- and high-frequency signal sequences, thereby effectively improving the accuracy of wind power forecasting. However, they did not analyze data preprocessing and forecasting uncertainty. A short-term wind speed forecasting model combining the automatic encoder of the CNN with LSTM units was presented by [40-41]. The calculation results demonstrated that the short-term wind speed forecasting model can accurately forecast shortterm wind speed. However, this model lacks forecasting uncertainty analysis.

The above scholars conducted in-depth studies on the forecasting methods of wind speed and wind power. They also made certain progress in temporal-related forecasting models, data preprocessing, and uncertainty analysis. comprehensive exploration of However а data preprocessing, temporal-related forecasting models and uncertainty analysis is lacking. Therefore, the authors of the present study proposed a short-term wind power forecasting and uncertainty analysis method based on the VMD-TCN-GMM. First, the proposed method applied the VMD to decompose wind speed information into different modal components, thereby eliminating the impact of wind speed randomness on forecasting accuracy. Second, the TCN model was used to explore deeply the temporal correlation between data, further improving the accuracy of short-term wind power forecasting. On this basis, the GMM was used to calculate accurately the distribution characteristics of short-term wind power forecasting errors. A confidence interval for wind power forecasting was constructed based on this calculation.

The remainder of this study is organized as follows. Section 3 details the calculation principles of VMD, TCN model, and GMM. The evaluation indicators of forecasting performance and the construction process of the forecasting model are also discussed. The effectiveness and superiority of the VMD-TCN model proposed in this study are verified in Section 4 by comparing the forecasting accuracies of different forecasting models. Moreover, the uncertainty of wind power forecasting is analyzed using GMMs. Section 5 summarizes this article and provides relevant conclusions.

3. Methodology

3.1 Variational mode decomposition (VMD)

VMD is a nonstationary signal decomposition method proposed by Dragomiretskiy et al. [42]. VMD realizes signal decomposition by introducing variational constraints, which effectively overcome the problems of mode aliasing and endpoint effects in traditional empirical mode decomposition methods.

When VMD is used to decompose a signal f(t), it is employed to obtain K intrinsic mode functions $u_k(t)$ and minimize the sum of the bandwidths of these K modes. The process for constructing VMD is as follows:

Step 1: Define a band-limited intrinsic mode function with limited bandwidth.

$$u_k(t) = A_k(t)\cos(\varphi_k(t)) \tag{1}$$

Where $A_k(t)$ is the instantaneous amplitude, and $\varphi_k(t)$ is the phase.

Step 2: Perform Hilbert transform on $u_k(t)$ to construct the corresponding analytical signal. The corresponding unilateral spectrum of each mode is obtained, as shown in Eq. (2).

$$u_k = [\delta(t) + \frac{j}{\pi t}] * u_k(t)$$
⁽²⁾

where $\delta(t)$ is the pulse function, $j=\sqrt{-1}$, and u_k is the i^{th} variational mode component.

Step 3: Add an exponential term $e^{-j\omega_k t}$ to each variational mode component u_k to adjust its corresponding center frequency ω_k and modulate the spectrum of each mode to the corresponding fundamental frequency band, as shown in Eq. (3).

$$u_{k} = [(\delta(t) + \frac{j}{\pi t}) * u_{k}(t)] * e^{-j\omega_{k}t}$$
(3)

Step 4: Demodulate the signal in Eq. (3) by Gaussian smoothing (i.e., the square root of the L^2 norm gradient) and estimate the bandwidth of each mode signal to obtain the variational constraint equation shown in Eq. (4).

$$\begin{cases} \min_{\{u_k, w_k\}} \{\sum_{k=1}^{K} \left\| \partial_t [(\delta(t) + \frac{j}{\pi t}) * u_k(t)] e^{-j\omega_k t} \right\|_2^2 \} \\ \text{s.t.} \quad \sum_{k=1}^{K} u_k = f \end{cases}$$

$$\tag{4}$$

where *K* is the number of components of the variational mode, ω_k is the center frequency of the k^{th} variational mode component, *t* is the time, and ∂_t is the partial derivative of time.

3.2 Temporal convolutional network (TCN)

3.2.1 Principle of TCN

TCN is a time-series model based on a CNN. Unlike CNN, which is mainly used for image or text feature extraction, TCN is mainly used for forecasting time-series data. When the input time-series data of the TCN model are $X = \{x_1, x_2, x_3, \dots, x_t\}$, the model output can be $Y = \{y_1, y_2, y_3, \dots, y_t\}$ with equal sequence length or the feature $H = \{h_1, h_2, h_3, \dots, h_t\}$ of the middle layer.

All outputs satisfy the causal constraints. The current output y_t is only related to $\{x_1, x_2, x_3,, x_t\}$ and is unrelated to the "future" input $\{x_{t+1}, x_{t+2}, x_{t+3},, x_{t+T}\}$. This scenario aligns with the application in real society, where only the historical data of the scenario are known, and the future state of the scenario is forecasted.

The mapping relationship between the input and output of the TCN model is given by Eq. (5).

$$Y(y_1, y_2, y_3, \dots, y_t) = f(x_1, x_2, x_3, \dots, x_t)$$
(5)

where the function f represents the mapping function of the TCN model. The input of each sequence X corresponds to the corresponding output Y by learning and adjusting the parameters of the TCN model.

TCN consists of three main parts: causal, dilated, and residual connection modules.

3.2.2 Causal convolution

In the design and implementation of a TCN, the following criteria must be observed: (1) The length of the network's output time-series data is equal to that of the input time-series data. (2) The output of the current time is only related to the input of the current time and the input of historical time. It is unrelated to the input of future time. When the input time-series information is $X = \{x_1, x_2, x_3, ..., x_T\}$ and the filter is $F = \{f_1, f_2, f_3, ..., f_k\}$, the causal convolution at x_t can be calculated according to Eq. (6).

$$(F * X)(x_t) = \sum_{k=1}^{K} f_k \cdot x_{t-K+k}$$
(6)

where K is the filter size, and t-K+k is the historical timepoint.

The structure of the causal convolution is illustrated in Fig. 1. As shown in Fig. 1, the value of time t for the next layer depends only on the value of time t and before time t for the previous layers. Compared with traditional CNN, causal convolution does not consider future data. Causal convolution is a one-way structure; thus, only the previous cause can have the latter effect, which is a strict time constraint model. The number of hidden layers also increases as historical data increases.



3.2.3 Dilated convolution

A deep network structure or a large convolution kernel is required when the traced historical time-series data are long. Thus, the model training time and the demand for computing resources increase. Therefore, the TCN introduces a dilated convolution based on causal convolution.

Dilated convolution increases the range of the receptive field by injecting holes into the receptive field of an ordinary convolution. When the input time-series information is $X = \{x_1, x_2, x_3, \dots, x_T\}$ and the filter is $F = \{f_1, f_2, f_3, \dots, f_k\}$, the dilated convolution at x_t can be calculated according to Eq. (7).

$$(F *_{d} X)(x_{t}) = \sum_{k=1}^{K} f_{k} \cdot x_{t-(K-k)d}$$
(7)

where *d* is the dilated factor, *K* is the filter size, and $t-(K-k) \cdot d$ is the historical time-point.

d is usually set to the exponential form of $2(1,2,4,8,...,2^i)$ to avoid the grid effect when dilated convolution is calculated. When *d* is one, every point of the input is sampled. When *d* is 2, one point is sampled at every two input points. Thus, the higher the level is, the greater the value of *d* is. Dilated convolution causes the size of the effective window to grow exponentially with the number of layers, and the convolutional network can obtain a large receptive field with few layers.

The structure of dilated convolution is shown in Fig. 2. The first layer is a causal convolution, and the length of the time-series data after convolution is equal to that of the timeseries data input by the model. Owing to an increase in the number of dilated convolutional layers, the current state of the hidden layer is exponentially related to the length of the historical input time-series data.



3.2.4 Residual connection block

Gradient disappearance and gradient explosion problems tend to occur during the training process of deep learning networks. Thus, the ability of the network to converge to the optimal solution is affected. The TCN model introduces a residual connection block to solve the problems of gradient disappearance and gradient explosion in deep learning networks. This model realizes the weighted fusion of the input x of the residual connection block into the output of the residual connection block. The residual connection block is calculated using Eq. (8).

$$O = \operatorname{Activation}(x + F(x)) \tag{8}$$

where x is the input of the residual connection block, F(x) is the output of the convolution calculation of the residual connection block, and Activation() is the activation function.

The internal structure of the residual connection block is illustrated in Fig. 3. As shown in Fig. 3, a residual connection block is composed of two layers: dilated causal convolution network and the nonlinear function ReLU. The weighted-norm and dropout layers are added to each layer of the dilated causal convolution network. In the residual connection block, the dilated causal convolution mainly completes the calculation of causal and dilated convolutions. The nonlinear ReLU function enables the TCN model to obtain nonlinear expression capabilities. The weighted-norm layer can effectively prevent gradient explosion in the network. The dropout layer can effectively prevent the overfitting of the TCN model.



Fig. 3. Residual block structure

In Fig. 3, the residual connection block fuses its input data x weighted into its output data. A 1×1 convolution operation on the input data x is usually necessary to prevent the dimension of the input data x from being inconsistent with that of the output data during the calculation.

The overall structure of the TCN model is shown in Fig. 4. As depicted in Fig. 4, the TCN comprises multiple residual blocks for tracing the depth of historical data and further improving the accuracy of wind power forecasting.



Fig. 4. Topology diagram of the TCN model

3.3 Gaussian mixture model (GMM)

GMM is a linear combination of a certain number of Gaussian probability density functions to approximate the probability density distribution of the sample set. It is also associated with high fitting accuracy and fast calculation speed. The probability density functions of GMM are shown in Eqs. (9) to (11).

$$P(x_i) = \sum_{k=1}^{K} \pi_k \varphi(x_i, \ \mu_k, \ \Sigma_k)$$
⁽⁹⁾

$$\sum_{k=1}^{K} \pi_k = 1 \tag{10}$$

$$\varphi(x_{i},\mu_{k},\Sigma_{k}) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_{k}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x_{i}-\mu_{k})^{T} \sum_{k}^{-1}(x_{i}-\mu_{k})\right)$$
(11)

where $P(x_i)$ represents the probability density function of the GMM, x_i represents the *i*th sample in the dataset, π_k is the weight of the k^{th} Gaussian probability density function, μ_k and \sum_k are the mean and covariance matrices of the k^{th} Gaussian probability density function, respectively, and *D* is the dimension of the sample space.

Eqs. (9), (10), and (11) contain unknown parameters $(\pi_k, \mu_k \text{ and } \Sigma_k)$ that must be solved through the sample set. The commonly used method of solving unknown parameters $(\pi_k, \mu_k \text{ and } \Sigma_k)$ is the expectation maximization (EM) algorithm. The EM algorithm is an iterative algorithm used for the maximum likelihood estimation of probability model parameters with hidden variables. Each iteration of the EM algorithm consists of two steps: Step E is used to calculate the expectation, and Step M is used for maximization. The specific calculation process is as follows.

Step E: Calculate the possibility that each feature x_i is derived from submodel *k* according to the current parameters.

$$\gamma_{ik} = \frac{\pi_k \varphi(x_i \mid \mu_k, \Sigma_k)}{\sum_{k=1}^{K} \pi_k \varphi(x_i \mid \mu_k, \Sigma_k)}, i = 1, 2, \dots, N; k = 1, 2, \dots, K$$
(12)

where N is the number of samples, and γ_{ik} represents the probability that the *i*th sample belongs to the *k*th Gaussian probability density function.

Step M: Calculate the model parameters of a new round of iterations.

$$\mu_{k} = \frac{\sum_{i=1}^{N} (\gamma_{ik} x_{i})}{\sum_{i=1}^{N} (\gamma_{ik})}, \ k = 1, \ 2, \ \cdots, \ K$$
(13)

$$\Sigma_{k} = \frac{\sum_{i=1}^{N} \gamma_{ik} (x_{i} - \mu_{k}) (x_{i} - \mu_{k})^{T}}{\sum_{i=1}^{N} \gamma_{ik}}, \ k = 1, \ 2, \ \cdots, K$$
(14)

$$\pi_{k} = \frac{\sum_{i=1}^{N} \gamma_{ik}}{N}, \ k = 1, \ 2, \ \cdots, \ K$$
(15)

Steps E and M are repeated until they converge $\|\theta_{i+1} - \theta_i\| < \varepsilon$; ε is a small positive real number). At this point, the GMM parameters are calculated.

3.4 Forecasting performance evaluation indicators and forecasting model construction

3.4.1 Forecasting performance evaluation indicators

Forecasting performance evaluation indicators are used to evaluate the performance of wind power forecasting models. The present study uses mean absolute error (MAE), root mean square error (RMSE), confidence interval coverage rate, and width to evaluate the forecasting performance of wind power forecasting models. MAE is the average of absolute error, which truly reflects the size of the forecasting error. The calculation formula for MAE is shown in Eqs. (16) and (17).

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| P_{\text{true, }i} - P_{\text{fore, }i} \right|$$
(16)

$$P_{\text{MAE}} = \frac{\frac{1}{N} \sum_{i=1}^{N} |P_{\text{true}, i} - P_{\text{fore}, i}|}{P_{\text{cap}}}$$
(17)

where N represents the data point for wind power forecasting, $P_{\rm ture}$ represents the actual value of wind power, $P_{\rm fore}$ is the forecasting value of wind power, and $P_{\rm cap}$ is the total installed capacity of the wind farm. $P_{\rm MAE}$ is the ratio of MAE to the total installed capacity, it is usually in percentage form in practical applications.

RMSE is the root mean square error, and the calculation formulas are shown in Eqs. (18) and (19).

RMSE =
$$\sqrt{\frac{1}{N} \sum_{i=1}^{N} (P_{\text{true},i} - P_{\text{fore},i})^2}$$
 (18)

$$P_{\rm RMSE} = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^{N} (P_{\rm true, i} - P_{\rm fore, i})^2}}{P_{\rm cap}}$$
(19)

where P_{RMSE} is the ratio of RMSE to the total installed capacity; it is usually in percentage form in practical applications.

The confidence interval coverage rate δ_n describes the

situation where the confidence interval covers the true value of wind power. The larger the coverage rate is, the more accurately the confidence interval can reflect the actual value distribution of wind power. The calculation formula for the coverage rate is shown in Eq. (20).

$$\delta_p = \frac{1}{N} \times \sum_{i=1}^{N} \partial_i$$
(20)

where N is the total number of samples, and ∂_i is the coverage factor. When the actual value of wind power falls within the confidence interval, $\partial_i = 1$; otherwise, $\partial_i = 0$.

The interval width $\Delta P'$ is an indicator for measuring the effectiveness of wind power forecasting. Under the premise of ensuring coverage rate, the smaller the interval width is, the better the forecasting effect is. The calculation formula for the interval width is shown in Eq. (21).

$$\Delta P' = \frac{\sum_{i=1}^{N} \Delta P_i}{N}$$
(21)

where ΔP_i is the difference between the upper and lower limits of the confidence interval where the i^{th} power value is located.

3.4.2 Forecasting model construction

A short-term wind power forecasting model based on the VMD-TCN-GMM was constructed according to the VMD principle, TCN model principle, GMM, and confidence interval calculation method introduced in the previous sections. The specific calculation process for the model is as follows:

(1) Normalize input data, such as wind speed, wind direction, pressure, temperature, and power, to meet the input data requirements of the forecasting model.

(2) Determine the number of mode components in the VMD and calculate the forecasting effect for different numbers of mode components. When the number of mode components is four, the forecasting effect of the forecasting model is the best. Therefore, the number of mode components in the VMD is four in this study.

(3) Decompose the normalized wind speed data by VMD and change the decomposed wind speed data from the original n data to $4 \times n$ matrix data.

(4) Divide the decomposed wind speed data and the wind direction, pressure, temperature, and power data into training and testing sample sets.

(5) Use the training sample set to train the constructed TCN model until the convergence condition is achieved.

(6) Input the testing sample set into the trained TCN model and inversely normalize the forecasting results of the TCN model to obtain the forecasting value.

(7) Use GMM to calculate the distribution characteristics of wind power forecasting error and obtain the probability density distribution of wind power forecasting error.

(8) Calculate the confidence interval of wind power forecasting according to the probability density distribution characteristics of wind power forecasting error.

The calculation process of the entire forecasting model is shown in Fig. 5.

4. Result Analysis and Discussion

4.1 Sample set division

The data in this study were obtained from a wind farm with 90 wind turbines in northern China. The data collection period for the wind power forecasting occurred from January 1, 2010 to August 31, 2011. The data were divided into training and testing sample sets to complete the model training and testing verification. The data of the training sample set included the annual data obtained from January 1, 2010 to December 31, 2010, with 25,384 data points. The data of the testing sample set were obtained from January 1, 2011 to August 31, 2011. Two datasets in different seasons from the testing sample set were selected to evaluate the model and verify that the proposed forecasting model demonstrated a good forecasting effect in different seasons and different forecasting time scales. These two datasets included the 4 and 24 h data from February 12, 2011 in winter and August 1, 2011 in summer.



Fig. 5. Calculation process of the forecasting model

4.2 Performance comparison of the forecasting models

Figs. 6(a) and 6(b) show the wind power forecasting results for 4 h on February 12 (winter) and August 1 (summer), respectively. The forecasting models include the wavelet neural network (WNN), BP neural network, PSO-BP hybrid model, LSTM model, TCN model, and VMD-TCN model. As shown in Fig. 6, the similarity between the forecasted value of the VMD-TCN model and the actual value of wind power was high. This finding indicated that the forecasting effect of the VMD-TCN model was better than the forecasting effects of the other models. The following RMSE values were obtained for the forecasting results of each model during the 4 h forecasting on August 1: 1.03%, VMD-TCN; 1.30%, TCN; 3.36%, LSTM; 7.58%, WNN; 4.65%, BP; 4.35%, PSO-BP. The MAE values of the forecasting results of each model are as follows: 0.37%, VMD-TCN; 1.05%, TCN; 2.72%, LSTM; 7.04%, WNN; 3.89%, BP; 4.02%, PSO-BP. Based on the calculation results, the RMSE and MAE values of the VMD-TCN model were less than those of the other models. This finding proved that the forecasting effect of the VMD-TCN model was better than the forecasting effects of the other models.



Fig. 6. 4 h wind power forecasting for different forecasting models. (a) 4 h wind power forecasting in February. (b) 4 h wind power forecasting in August

Figs. 7(a) and 7(b) show the forecasting results of wind power for 24 h on February 12 (winter) and August 1 (summer), respectively. As depicted in Fig. 7, the similarity between the forecasted value of the VMD-TCN model and the real value of wind power was high in the 24 h wind power forecasting. This finding indicated the high forecasting accuracy of the VMD-TCN model. The VMD-TCN model demonstrated a good forecasting effect because VMD effectively eliminated the influence of wind speed randomness on forecasting accuracy. Furthermore, the TCN model could deeply mine the temporal correlation between data, thereby improving the forecasting accuracy of VMD-TCN further. During the 24 h forecasting using VMD-TCN on August 1 (summer), the following RMSE values for the forecasting results of each model were obtained: 3.39%, VMD-TCN; 3.59%, TCN; 4.62%, LSTM; 9.59%, WNN; 5.21%, BP; 4.99%, PSO-BP. The MAE values of the forecasting results of each model are as follows: 2.66%, VMD-TCN; 2.92%, TCN; 3.70%, LSTM; 7.67%, WNN; 3.80%, BP; 3.78%, PSO-BP. The calculation results showed that the RMSE and MAE values for the forecasting error of the VMD-TCN model remained the lowest. This finding further indicated that the forecasting effect of the VMD-TCN model was better than the forecasting effects of the other models.



Fig. 7. 24 h wind power forecasting for different forecasting models. (a) 24 h wind power forecasting in February. (b) 24 h wind power forecasting in August

Table 1 demonstrates the RMSE and MAE values for the forecasting results of each model under different forecasting time scales and climatic conditions. As depicted in Fig. 6, Fig. 7, and Table 1, the RMSE and MAE values for the forecasting error of the VMD-TCN model were lower than those of the other models under different climatic conditions and forecasting time scales. Such a finding proved that the forecasting accuracy of the VMD-TCN model was higher than that of the other models.

Table. 1. RMSE and MAE values of the forecasting models at different time scales and climatic conditions

Month	Forecast time	Models	RMSE	MAE
Feb.	4 h	VMD-TCN	1.50%	1.22%
		LSTM	2.71%	2.20%
		TCN	2.20%	1.48%
		PSO-BP	3.98%	3.76%
		BP	4.63%	4.42%
		WNN	6.19%	6.05%
	24 h	VMD-TCN	4.45%	3.08%
		LSTM	5.93%	4.33%
		TCN	5.37%	4.14%
		PSO-BP	7.26%	6.54%
		BP	7.38%	6.66%

Yang Jian, Yao Shuai, Chang Xuejun, Li Dewei, Gu Bo and Zhang Zichao/ Journal of Engineering Science and Technology Review 16 (2) (2023) 197 - 206

		WNN	10.34%	9.69%
Aug.	4 h	VMD-TCN	1.03%	0.37%
		LSTM	3.36%	2.72%
		TCN	1.30%	1.05%
		PSO-BP	4.35%	4.02%
		BP	4.65%	3.89%
		WNN	7.59%	7.04%
	24 h	VMD-TCN	3.39%	2.66%
		LSTM	4.62%	3.70%
		TCN	3.59%	2.92%
		PSO-BP	4.99%	3.78%
		BP	5.21%	3.80%
		WNN	9.59%	7.67%

4.3 Confidence interval of wind power forecasting

4.3.1 Probability density estimation of wind power forecasting errors

The probability density distribution of the wind power forecasting errors must be determined first to calculate the confidence interval of wind power forecasting. In this study, GMM was used to determine the probability density distribution of the wind power forecasting error shown in Fig. 8. Fig. 8 reveals the probability density distribution of the 24 h wind power forecasting error. The blue box in Fig. 8 represents the frequency histogram, the red solid line represents the probability density distribution of the wind power forecasting errors calculated by GMM, and the red dotted line represents the probability density distribution of the wind power forecasting errors obtained by the single Gaussian model. As shown in Fig. 8, the characteristics of the wind power forecasting errors described by the density distribution of the wind power probability forecasting errors obtained by GMM were more accurate than those described by the probability density distribution of the wind power forecasting errors obtained by a single Gaussian model.



Fig. 8. Probability density distribution of the 24 h wind power forecasting errors. (a) Probability density distribution of the 24 h wind

power forecasting errors in February. (b) Probability density distribution of the 24 h wind power forecasting errors in August

4.3.2 Confidence interval of wind power forecasting

After the probability density distribution of wind power forecasting error was obtained, the confidence intervals of wind power forecasting values at different confidence levels could be calculated. Figs. 9 and 10 demonstrate the distribution of the confidence intervals of the VMD-TCN model at 97.5%, 95%, 90%, and 85% confidence levels when the forecasting time scales are 4 and 24 h, respectively. As shown in Figs. 9 and 10, the real value of wind power was within the range of the confidence interval. This finding proved that the GMM could be reasonably used to determine the probability density distribution of wind power forecasting errors. The width of the confidence level. This scenario aligns with the calculation principle of the confidence interval.



Fig. 9. Confidence intervals of the 4 h wind power forecasting using the VMD-TCN model. (a) Confidence interval of the 4 h wind power forecasting in February. (b) Confidence interval of the 4 h wind power forecasting in August

Table 2 demonstrates the coverage rate of the confidence interval of wind power forecasting based on the VMD-TCN model under different climatic conditions and forecasting time scales. The coverage rate of the confidence interval was higher than the confidence level. Thus, the confidence interval calculation method based on GMM could accurately describe the distribution range of the actual output power of the wind farm.



Fig. 10. Confidence intervals of the 24 h wind power forecasting using the VMD-TCN model. (a) Confidence interval of the 24 h wind power forecasting in February. (b) Confidence interval of the 24 h wind power forecasting in August

Month	Confidence level	24 h	4 h
	97.5%	97.93%	100%
Eab	95%	95.87%	100%
reb.	90%	90.72%	94.12%
	85%	85.54%	88.24%
	97.5%	97.94%	100%
4.110	95%	95.88%	100%
Aug.	90%	91.76%	94.12%
	85%	90.72%	88.24%

5 Conclusions

VMD, time convolutional network (TCN), and GMM were combined to construct a short-term wind power forecasting model and improve further the accuracy of short-term wind power forecasting. The calculation results indicate the following:

(1) VMD can effectively overcome the impact of wind speed randomness on model forecasting accuracy, thereby improving the accuracy of short-term wind power forecasting.

(2) Under different climate conditions and forecasting time scales, the short-term wind power forecasting accuracy of the VMD-TCN model is higher than the short-term wind power forecasting accuracies of WNN, BP neural network, PSO-BP hybrid model, LSTM model, and TCN model. This finding proves the feasibility and superiority of the shortterm wind power forecasting method based on the VMD-TCN model proposed in the study.

(3) GMM can accurately describe the distribution characteristics of short-term wind power forecasting errors. The confidence interval coverage rate constructed is greater than the confidence level under different climate conditions and forecasting time scales.

This study combines the advantages of VMD, TCN, and GMM to construct a short-term wind power forecasting model. This model effectively overcomes the impact of wind speed randomness on the accuracy of wind power forecasting. Moreover, it deeply mines the temporal correlation between data, thereby improving the forecasting accuracy of short-term wind power effectively. However, the model constructed in this study cannot effectively extract the spatial distribution characteristics of data. Therefore, in future research, constructing a short-term wind power forecasting model that can deeply explore the spatial and temporal distribution characteristics of data will be a key breakthrough technology for wind power forecasting.

Acknowledgments

This work was supported by the National Key Research and Development Program of China [grant number 2019YFE0104800].

This is an Open Access article distributed under the terms of the Creative Commons Attribution License.



References

- He, Y., Li. H., Wang, S., Yao, X. "Uncertainty analysis of wind power probability density forecasting based on cubic spline interpolation and support vector quantile regression". *Neurocomputing*, 430, 2021, pp. 121–137.
- Tang, X. Z., Gu, N. W., Huang, X. Q., Peng, R.T. "Progress on Short Term Wind Power Forecasting Technology". *Journal of Mechanical Engineering*, 58(12), 2022, pp. 213–236.
- Sun, R. F., Zhang, T., He, Q., Xu, H. X. "Review on Key Technologies and Applications in Wind Power Forecasting". *High Voltage Engineering*, 47(04), 2021, pp. 1129–1143.
- Chen, Y. L., Hu, X., Zhang, L. X. "A review of Ultra-short-term Forecasting of Wind Power Based on Data Decompositionforecasting Technology Combination Model". *Energy Reports*, 8, 2022, pp. 14200–14219.
- Ajay, K. B. "Sizing and Forecasting Techniques in Photovoltaic-wind Based Hybrid Renewable Energy System: A Review". *Journal of Cleaner Production*, 369, 2022, pp. 133376.

- Duan, Z., Liu, H., Li, Y., Nikolaos, N. "Time-variant Post-processing Method for Long-term Numerical Wind Speed Forecasts Based on Multi-region Recurrent Graph Network". *Energy*, 259, 2022, pp. 125021.
- Tawn, R., Browell, J. "A Review of Very Short-term Wind and Solar Power Forecasting". *Renewable and Sustainable Energy Reviews*, 153, 2022, pp. 111758.
- Wang, F., Chen, P., Zhen, Z., et al. "Dynamic Spatio-temporal Correlation and Hierarchical Directed Graph Structure Based Ultrashort-term Wind Farm Cluster Power Forecasting Method". *Applied Energy*, 323, 2022, pp. 119579.
- Qian, Z., Pei, Y., Cao, L. X., et al. "Review of Wind Power Forecasting Method". *High Voltage Engineering*, 42(04), 2016, pp. 1047–1060.

Yang Jian, Yao Shuai, Chang Xuejun, Li Dewei, Gu Bo and Zhang Zichao/ Journal of Engineering Science and Technology Review 16 (2) (2023) 197 - 206

- Chi, Y. N., Liang, W., Zhang, Z. K., et al. "An Overview on Key Technologies Regarding Power Transmission and Grid Integration of Large Scale Offshore Wind Power". *Proceedings of the CSEE*, 36(14), 2016, pp. 3758–3771.
- Ye, L., Lu, P., Zhao, Y. N., et al. "Review of Model Predictive Control for Power System With Large-scale Wind Power Gridconnected". *Proceedings of the CSEE*, 41(18), 2021, pp. 6181– 6198.
- Yang, B., Zhong, L., Wang, J., et al. "State-of-the-art One-stop Handbook on Wind Forecasting Technologies: An Overview of Classifications, Methodologies, and Analysis". *Journal of Cleaner Production*, 283(6), 2020, pp. 124628.
- Demolli, H., Dokuz, A. S., Ecemis, A., et al. "Wind Power Forecasting Based on Daily Wind Speed Data Using Machine Learning Algorithms". *Energy Conversion and Management*, 198, 2019, pp. 111823.
- Tian, Z. "Modes Decomposition Forecasting Approach for Ultrashort-term Wind Speed". *Applied Soft Computing*, 105(11), 2021, pp. 107303.
- Wang, J., Yang, Z. "Ultra-short-term Wind Speed Forecasting Using an Optimized Artificial Intelligence Algorithm". *Renewable Energy*, 171(5), 2021, pp. 1418–1435.
- Nie, Y., Liang, N., Wang, J. Z. "Ultra-short-term Wind-speed Biforecasting System Via Artificial Intelligence and a Doubleforecasting Scheme". *Applied Energy*, 301, 2021, pp. 117452.
- Tatiane, C., Paulo, A. C., Paulo, C. M., Luis, M. "Ridge regression ensemble of machine learning models applied to solar and wind forecasting in Brazil and Spain". *Applied Energy*, 314, 2022, pp. 118936.
- Zhu, Q. F., Li, J. T., Qiao, J., et al. "Application and Prospect of AI Technology in Renewable Energy Forecasting". *Proceedings of the CSEE*, 2022, 1-23. https://doi.org/10.13334/j.0258-8013.pcsee.213114.
- Sheraz, A., Herodotos, H., Syed, M. M., et al. "A Survey on Deep Learning Methods for Power Load and Renewable Energy Forecasting in Smart Microgrids". *Renewable and Sustainable Energy Reviews*, 144, 2021, pp. 110992.
 Li, R., Ma, T., Zhang, X., et al. "Short-term Wind Power Prediction
- Li, R., Ma, T., Zhang, X., et al. "Short-term Wind Power Prediction Based on Convolutional Long-short-term Memory Neural Networks". ACTA Engergiae Solaris Sinica, 42(06), 2021, pp. 304– 311.
- Zhou, M., Wang, B., Guo, S. D., et al. "Multi-objective Prediction Intervals for Wind Power Forecast Based on Deep Neural Networks". *Information Sciences*, 550, 2021, pp. 207–220.
- Tang, Z., Zhao, G., Ouyang, T. "Two-phase Deep Learning Model for Short-term Wind Direction Forecasting". *Renewable Energy*, 173(72), 2021, pp. 114002.
- Hong, Y. Y., Satriani, T. "Day-ahead Spatiotemporal Wind Speed Forecasting Using Robust Design-based Deep Learning Neural Network". *Energy*, 209, 2020, pp. 118441.
- Aslam, M., Kim, J. S., Jung, J. "Multi-step Ahead Wind Power Forecasting Based on Dual-attention Mechanism". *Energy Reports*, 9, 2023, pp. 239–251.
- Maryam, Y., Narjes, A., Meisam, F., et al. "Deep Learning-based Multistep Ahead Wind Speed and Power Generation Forecasting Using Direct Method". *Energy Conversion and Management*, 281, 2023, pp. 116760.
- 26. Wan, C., Song, Y. H. "Theories, Methodologies and Applications of Probabilistic Forecasting for Power Systems with Renewable Energy Sources". *Automation of Electric Power Systems*, 45(01), 2021, pp. 2–16.

- 27. Seyed, M. J. J., Sajad, A., Mahdi, K., et al. "An Advanced Shortterm Wind Power Forecasting Framework Based on the Optimized Deep Neural Network Models". *International Journal of Electrical Power & Energy Systems*, 141, 2022, pp. 108143.
- Ahmed, A. E., Mohammed, A. A. A., Laith, A., et al. "HBO-LSTM: Optimized Long Short Term Memory with Heap-based Optimizer for Wind Power Forecasting". *Energy Conversion and Management*, 268, 2022, pp. 116022.
- Akshita, G., Arun, K., Boopathi, K. "Intraday Wind Power Forecasting Employing Feedback Mechanism". *Electric Power* Systems Research, 201, 2021, pp. 107518.
- Dominik, P., Michael, G., Hans, A. "A Novel Approach to Multihorizon Wind Power Forecasting Based on Deep Neural Architecture". *Renewable Energy*, 178, 2021, pp. 494–505.
- Daniel, V. P., Mario, J. R., Peder, B., et al. "Assessing Stacked Physics-informed Machine Learning Models for Colocated Windsolar Power Forecasting". *Sustainable Energy, Grids and Networks*, 32, 2022, pp. 100943.
- Bo, Y., Xiaohui, Y., Fei, T. "Improved Nonlinear Napping Network for Wind Power Forecasting in Renewable Energy Power System Dispatch". *Energy Reports*, 8, 2022, pp. 124–133.
- 33. Yun, W., Runmin, Z., Fang, L., et al. "A Review of Wind Speed and Wind Power Forecasting with Deep Neural Networks". *Applied Energy*, 304, 2021, pp. 117766.
- Ahmadi, M., Khashei, M. "Current Status of Hybrid Structures in Wind Forecasting". *Engineering Applications of Artificial Intelligence*, 99, 2021, pp. 104133.
- Hossain, M. A., Ripon, K. C., Sondoss, E., et al. "Very Short-term Forecasting of Wind Power Generation Using Hybrid Deep Learning Model". *Journal of Cleaner Production*, 296, 2021, pp. 126564.
- Wang, R., Ran, F., Lu, J. "Wind Power Prediction Based on Run Discriminant Method and VMD Residual Correction". *Journal of Hunan University* (Natural Sciences), 49(08), 2022, pp. 128–137.
- 37. Zhao, L. Y., Liu, Y. B., Shen, X. D. "Short-term Wind Power Prediction Model Based on CEEMDAN and an Improved Time Convolutional Network". *Power System Protection and Control*, 50(01), 2022, pp. 42–50.
- 38. Yang, L., Zhong, J.H., Zhang, Y., Bai, S.C., Li, G.S., Yang, Y., Zhang, J., "An Improving Faster-RCNN With Multi-Attention ResNet for Small Target Detection in Intelligent Autonomous Transport With 6G". *IEEE Transactions on Intelligent Transportation Systems*, 2022, doi: 10.1109/TITS.2022.3193909.
- Sheng, S. Q., Jin, H., Liu, C. R. "Short-term and Mid-short-term Wind Power Forecasting Based on VMD-WSGRU". *Power System Technology*, 46(03), 2022, pp. 897-904.
- Zhang, S., Chen, Y., Xiao, J. H., et al. "Hybrid Wind Speed Forecasting Model Based on Multivariate Data Secondary Decomposition Approach and Deep Learning Algorithm with Attention Mechanism". *Renewable Energy*, 174, 2021, pp. 688– 704.
- Li, Y., Qin, H.M., Su, H.W., "Research on security risk assessment of SCADA system based on the cloud model and combination weighting". *CAAI Transactions on Intelligent Systems*, 17(5),2022, pp. 969-979.
- Dragomiretskiy, K., Zosso, D. "Variational Mode Decomposition". *IEEE Transactions on Signal Processing*, 62(3), 2014, pp. 531– 543.