IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 6 February 2024, accepted 27 February 2024, date of publication 5 March 2024, date of current version 3 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3373798

RESEARCH ARTICLE

An Ultra-Short-Term Wind Power Forecasting Model Based on EMD-EncoderForest-TCN

YU SUN[1], JUNJIE YANG[2], XIAOTIAN ZHANG[1], KAIYUAN HOU[1], JIYUN HU[1], AND GUANGZHI YAO[1]

[1]State Grid Corporation of China, Shenyang 110179, China
[2]Shenyang Institute of Computing Technology, Chinese Academy of Sciences, Shenyang 110168, China

Corresponding author: Yu Sun (sunyu2701@163.com)

ABSTRACT Accurate wind power prediction helps to stabilize the operation of the power system, improve the utilization rate of renewable energy, reduce dependence on traditional energy, and achieve sustainable energy development. An ultra short-term wind power prediction method integrating EMD-EncoderForest-TCN is proposed to address the difficulty of predicting wind power due to frequent changes in wind speed. Firstly, the time-series input data of the model is decomposed into high-frequency and low-frequency components using Empirical Mode Decomposition. Then, based on the EncoderForest model and TCN model, differential information extraction is performed on the low-frequency and high-frequency components. The EncoderForest model regularizes low-frequency information and captures trend patterns in the data. The TCN model models the high-frequency components of time series to capture complex patterns and structures in wind power. Finally, based on convolutional neural networks, the output results of each part are calculated to achieve accurate prediction of wind power. Based on the operational data of an actual wind farm, conduct a case study analysis. The results show that the proposed model can achieve accurate prediction of short-term wind power, with a prediction accuracy improvement of 2.57%.

INDEX TERMS Wind power, empirical mode decomposition, encoderforest, temporal convolutional network, ultra short term wind power prediction.

I. INTRODUCTION

Currently, the world is facing problems such as climate change, resource shortage and environmental pollution, and sustainable development is an important goal and guiding principle for realizing the global energy transition [1], [2]. The global energy transition means shifting from traditional fossil energy sources to renewable energy sources in order to reduce environmental damage and pollution and achieve sustainable development. Wind power as a kind of renewable and clean energy has become a key solution for carbon emission reduction and energy transition in various countries. It plays an important role in promoting sustainable development, coping with climate change, and ensuring energy security [3], [4], [5]. Global wind energy resources are very rich, and all countries are actively developing and utilizing wind energy to promote the development of clean energy. With the progress of technology and the growth of the industry, wind power will play an increasingly important role in the global energy system. With the vigorous development of the wind power industry, the proportion of wind power connected to the grid has been increasing [6], [7]. At the same time, wind power generation is characterized by volatility and randomness [8]. Therefore, a large number of wind power connected to the grid puts forward higher requirements for the safe and stable operation of the power system. Accurate ultra short term wind power prediction can help power grid dispatch departments better understand the output power changes of wind farms. This can more accurately arrange the power generation plan of wind farms, thereby improving the access capacity of wind power [9], [10]. Meanwhile, the output curve predicted by the wind farm can optimize the output of conventional units, achieving the goal of reducing operating costs. In addition, ultra short term wind power prediction can describe the variation pattern of wind power output, and can take measures in advance to enhance the safety, reliability, and controllability

The associate editor coordinating the review of this manuscript and approving it for publication was Ehab Elsayed Elattar.

of the system [11], [12]. In summary, accurate ultra short term wind power prediction is of great significance for improving the operational efficiency of the electricity market, reducing system operating costs, enhancing system safety, reliability, and controllability.

A large number of researchers have conducted research on ultra short term wind power prediction. There are three main methods for predicting ultra short term wind power: statistical methods, physical methods, and artificial intelligence methods [13], [14], [15]. The ultra short term wind power prediction based on statistical methods establishes statistical models based on historical wind speed and power data, such as time series analysis, regression analysis, etc. This method predicts future power output by analyzing historical data. Stathopoulos et al. [16] explored the problem of wind power prediction through numerical prediction models and statistical prediction models. The results indicate that accurate wind power prediction can be achieved under reliable local environmental data conditions. González Sopeña et al. [17] proposed a benchmark framework for wind power prediction models based on statistical data. Research has shown that modal decomposition models exhibit higher performance compared to other statistical models. Sun et al. [18] proposed a spatiotemporal wind power prediction method based on multi factor extraction. The model validates the superiority of the proposed method in statistical wind power prediction. A mathematical model is established based on meteorological conditions and the working principle of wind turbines for ultra short term wind power prediction using physical methods. This method considers factors such as wind speed, direction, and unit characteristics to predict power output. Guo et al. [19] constructed a short-term wind power prediction model considering wake effects. The article physically enhances the statistical prediction model, indicating that the accuracy of wind power prediction based on physical principles has been greatly improved. Nasery and Aziz Ezzat [20] proposed a yaw adjustment wind power curve modeling approach, which improved the accuracy of wind power simulation. Zhou et al. [21] proposed a wind power prediction method based on hybrid physical processes and machine learning. Research has shown that a hybrid model of physics and machine learning can simultaneously leverage the advantages of both physical models and machine learning methods. With the rapid development of artificial intelligence, more and more researchers are using artificial intelligence methods to predict wind power. Artificial intelligence based ultra short term wind power prediction utilizes artificial intelligence algorithms such as neural networks, support vector machines, random forests, etc. This method establishes a prediction model by learning and training a large amount of historical data. This method can better capture complex nonlinear relationships and spatiotemporal changes. Wang et al. [22] proposed a wind power prediction model based on deep learning, which integrates multiple prediction learners through a multi-layer stacked prediction model. Research has shown that the model has good generalization performance. Liu et al. [23] proposed a wind power prediction model based on sub attention and convolutional neural networks. The model integrates global and local information of wind power timing power, which can effectively improve the accuracy of the model. Abou Houran et al. [24] proposed a hybrid solar and wind power prediction method based on deep learning. This model can accurately predict wind and solar power.

In summary, most scholars predict wind power based on the fluctuation patterns of wind power and the characteristics of weather changes. Wind power has randomness and volatility, and using only wind power for prediction cannot accurately track the detailed changes in wind power [25], [26]. However, frequent changes in wind speed are the main reason for the difficulty in predicting wind power. Therefore, it is necessary to study the fluctuation patterns of wind power, and decomposing different fluctuation frequencies can improve the accuracy of wind power prediction. Meanwhile, for the first time, the EncoderForest model and TCN model are integrated to further extract information features at different frequencies. In summary, this paper proposes an ultra-short-term wind power prediction method that integrates EMD-EncoderForest-TCN. The method proposed in this paper decomposes wind power into high-frequency and low-frequency components. The differentiated prediction of high-frequency and low-frequency components can improve the accuracy of wind power prediction. The model structure is shown in Figure 1. First, the model timing input data is decomposed into high-frequency and low-frequency components by means of Empirical Modal Decomposition (EMD). Then, based on the EncoderForest model and Temporal Convolutional Network (TCN)model, differential information extraction is performed on the low-frequency and high-frequency components. The EncoderForest model regularizes low-frequency information and captures trend patterns in the data. The TCN model models the high-frequency components of time series to capture complex patterns and structures in wind power. Finally, based on convolutional neural networks, the modal functions are superimposed to achieve accurate prediction of wind power.

The wind power prediction method based on the EMD-EncoderForest-TCN model exhibits higher robustness and a series of significant advantages compared to traditional prediction methods. Wind power data usually exhibits strong nonlinear and non-stationary characteristics, which makes traditional prediction methods difficult to cope with. EMD (Empirical Mode Decomposition) can decompose complex nonlinear and non-stationary signals into a series of intrinsic mode functions (IMFs) with physical significance. In addition, compared to other models with weaker aliasing phenomena, the EMD model has stronger intuitiveness and physical significance, making it easier for engineering understanding and application in the training process of machine learning. Thus, effectively processing these complex features and

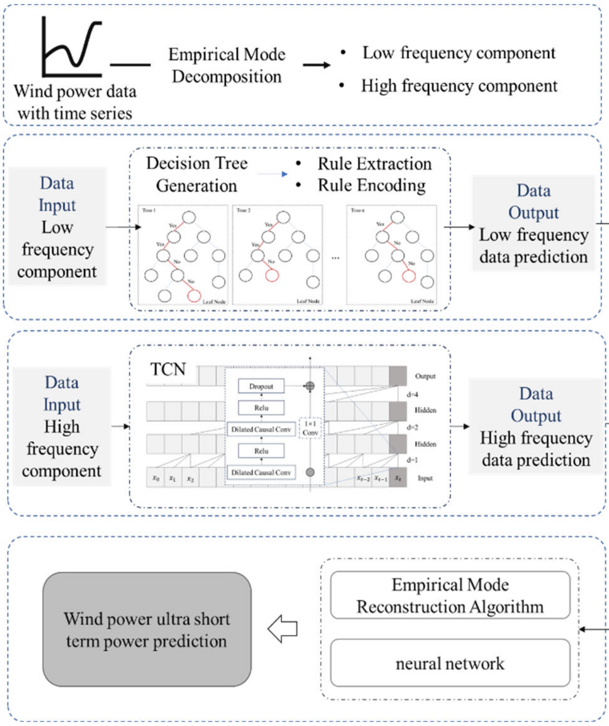


FIGURE 1. **Structural diagram of ultra short-term wind power prediction model.**

improving the robustness of the prediction model. In practical applications, wind power data is often affected by noise and outliers. EncoderForest combines the prediction results of multiple base models through ensemble learning, effectively reducing the impact of noise and outliers on the prediction results, and enhancing the robustness of the model. TCN can handle long-term dependency relationships. This enables the prediction method based on EMD-EncoderForest-TCN to adapt to the needs of long-term wind power prediction and maintain the robustness of the prediction results. By combining EMD, EncoderForest, and TCN, this method can fully utilize their respective advantages and achieve high-precision prediction of wind power. This not only contributes to the stable operation of the power system, but also provides more accurate data support for the operation, maintenance, and management of wind farms. This method can be flexibly adjusted and optimized according to actual needs. For example, predictive performance can be optimized by adjusting parameters such as the decomposition level of EMD, the number and type of base models in EncoderForest, and the network structure of TCN. In addition, this method can also be combined with other advanced algorithms and models to further expand its application scope and functionality.

II. GUIDELINES FOR MANUSCRIPT PREPARATION

To accurately evaluate the fusion EMD-EncoderForest-TCN approach to ultra-short-term wind power prediction, each component of the model is presented in this chapter. SectionII-A describes the EMD-based approach to decompose modeled wind power time series data into high-frequency and low-frequency components. SectionII-B describes the information extraction of low-frequency power in wind power based on the EncoderForest model to capture trending patterns in the data. SectionII-C describes the modeling of high-frequency components in wind power based on TCN to capture the complex changing patterns and structures in wind power. SectionII-D describes the superposition of modal functions by convolutional neural networks to ultimately achieve accurate wind power prediction.

A. TIME SERIES DECOMPOSITION METHOD FOR WIND POWER BASED ON EMD

EMD is a method used to process nonlinear, nonsmoothed signals. Its main objective is to decompose a complex signal into a series of simple oscillatory components [27], [28]. These components are called intrinsic mode functions (IMF). These IMFs represent the characteristics of different time scales in the signal. The structure and flow of EMD can be described as follows [29]:

(1) Identify all local extremes of the wind power time series, including local maxima and minima.

(2) All these local maxima and minima were fitted with a cubic spline interpolation function to form the upper and lower envelopes of the original data, respectively. The upper envelope is $X_u(t)$, and the lower envelope is $X_l(t)$.

(3) Calculate the average envelope of the upper and lower envelopes $h(t) = X(t) - m(t)$.

(4) The original data sequence is subtracted from the mean envelope $X(t)$ to obtain a new data sequence $m(t) = [X_u(t) + X_l(t)]/2$.

(5) Determine whether $h(t)$ satisfies the conditions of the IMF. IMF needs to satisfy two conditions: first, the number of extreme value points and the number of over-zero points must be equal or differ by at most one in the entire data series. The second is that at any moment, the average of the upper envelope formed by the local extreme value points and the lower envelope formed by the local extreme value points is zero.

(6) If $h(t)$ satisfies the condition of IMF, then it is the first IMF of the original data sequence, denoted as IMF1.

(7) If $h(t)$ does not satisfy the condition of IMF, then consider $h(t)$ as a new original data sequence and repeat steps 1-4 until a $h(t)$ is obtained that satisfies the condition of IMF.

(8) Subtract the first IMF from the original data sequence $X(t)$ to obtain the residual $r(t) = X(t) - IMF1$.

(9) Consider residual $r(t)$ as the new original data sequence and repeat steps 1-8 until all IMFs and the final residual term are obtained. This process will continue until the residual term becomes a monotonic function or constant.

With the above steps, the original data sequence $X(t)$ is decomposed into a series of IMFs and a residual term. These IMFs are arranged in the order of frequency from high to low, and each IMF represents an oscillating component with

Algorithm EMD

Input: Original signal; Stop criterion (residual threshold)
Output: Intrinsic Mode Function (IMF) component
if $r(t) = 0$:
  $r(t) = X(t)$
  Upper and lower envelope lines: $h(t) = X(t) - m(t)$
  Mean envelope: $m(t) = [X_u(t) + X_l(t)]/2$
  Extracting fluctuation details: $h'(t) = r(t) - m(t)$
  Update IMF components: $h(t) = h(t) - h'(t)$
  Update residual signal: $r(t) = X(t) - IMF1$
  IMFs.append($h(t)$)
return IMFs

different frequencies in the original data series. Empirical modal decomposition is a method to deal with nonlinear and nonsmoothed signals. It can decompose the wind farm output power signal into a series of quasi-single component signals based on different time scale characteristics. Each IMF component represents the oscillatory component of the wind power signal on different time scales. In this way, the dynamic characteristics of wind power can be better understood. The algorithm pseudocode is as follows.

Wind speed has fluctuations, randomness, and a certain periodicity, so wind power also has similar characteristics. The purpose of EMD decomposition is to extract different features in the signal, such as frequency, amplitude, phase, etc., to better understand and analyze the properties of the signal. The decomposed IMF components have different time scales, reflecting the local characteristics of the signal at different time scales. Therefore, the number of IMF components obtained through decomposition depends on the complexity of the signal itself and the accuracy requirements of the decomposition. Through repeated experiments, wind power is decomposed into five signals based on empirical mode decomposition, which can fully extract temporal data features while generating less noise and interference. The goal is to achieve good wind power prediction results.

B. THE LOW-FREQUENCY POWER PREDICTION OF WIND POWER BASED ON EncoderForest

EncoderForest is a tree set model based on the autoencoder model [30]. The core idea is to design an effective process that enables the forest to reconstruct the original path by using the Maximum Compatible Rule (MCR) defined by the decision path of the tree. The EncoderForest model is an effective machine learning method that can be applied to various prediction and classification tasks. EncoderForest has made improvements on the basis of random forests, introducing the idea of encoding [31]. Random forest is an ensemble learning method based on decision trees. It improves the accuracy and stability of predictions by constructing multiple decision trees and integrating their prediction results [32]. In a random forest, each decision tree is constructed through random sampling and feature selection of training data. This can increase the diversity of the model and reduce the risk of overfitting.

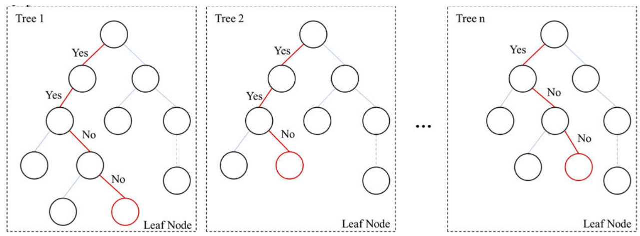


FIGURE 2. Generator based on EncoderForest.

In EncoderForest, each decision tree is treated as an encoder used to transform input data into an easy to process encoding. These codes contain the main features and information of the input data, providing better input for subsequent models.

Specifically, the algorithmic process of EncoderForest is as follows:

Constructing Encoder: Based on Random Forest, multiple decision trees are constructed by self-sampling and feature random selection, each of which encodes the input data and generates a coding vector. These encoding vectors constitute the output of the Encoder.

Training Decoder: Using the generated coding vectors for prediction, various machine learning algorithms, such as neural networks, support vector machines, etc., are utilized to learn and predict the coding vectors.The goal of Decoder is to learn the mapping relationship from coding vectors to the target variables, so as to achieve prediction of the target variables. The model generator is shown in Figure. 2.

Perform prediction: new input data is fed into the Encoder to generate the corresponding encoding vectors. These coding vectors are then input into the trained Decoder to get the final prediction results. The final model prediction result is shown in Equation 1.

$$F(x) = \sum f_i(x) \tag{1}$$

where $F(x)$ is the prediction of the random forest and $f_i(x)$ is the prediction of the ith decision tree.

C. HIGH-FREQUENCY POWER PREDICTION FOR WIND POWER BASED ON TCN

TCN is a neural network for processing sequence data [33], [34]. It is based on the convolutional idea of Convolutional Neural Networks (CNNs), but performs convolutional operations only in the temporal dimension of the sequence data, rather than in the spatial dimension. This maintains the temporal information of the sequence data and makes the output dependent only on the current and previous inputs without being affected by future inputs. TCN effectively improves the range of sensibility of the traditional CNN model by introducing dilated convolution, causal convolution, and residual networks [35]. Thus the model is able to enhance the breadth of the processed data and is a convolutional algorithm that performs well in time series prediction. Among them, causal convolution is able to avoid the problem of data leakage caused by future data being recognized by the model in the process of processing data. Whereas, dilation convolution expands the horizon interval of time series data by stacking fewer network layers, which improves the feature mining capability of the input data. At the same time, TCN also uses the residual structure of ResNet to replace the convolutional layer, thereby training a deeper network, as shown in Equation 2. The residual module includes two layers of extended convolution, two layers of weight normalization, two layers of activation function ReLU, and two layers of Dropout. Unlike ResNet, standard ResNet adds input directly to output. In TCN, due to the possibility of inconsistent input and output dimensions, 1∗1 convolution is first introduced to change the input feature dimension, and then added to the output.

$$H(T) = \sum_{i=0}^{n-1} f(i) \cdot x_{T-d \cdot i} \tag{2}$$

$$o = Activation(x + F(x)) \tag{3}$$

where, T is the time series element, f is the filter, n is the filter size, d is the inflation factor, $x_{T-d \cdot i}$ is the historical data, *Activation* is the output data, and 6 is the residual calculation.

The advantages of the TCN model are as follows: (1) Due to the adoption of a convolutional neural network structure, TCN can perform convolution operations in parallel, improving the efficiency of the model. This enables TCN to have better performance when processing long sequence data. (2) The receptive field in TCN can be adjusted by parameters such as layer number, dilation factor, and filter size. This allows the model to capture information at different time scales more flexibly. This flexibility is crucial when dealing with sequence data with multiple time scales. (3) Due to the fact that the gradient of TCN is not in the temporal direction, but in the depth direction of the network. Therefore, when the input length is longer, the gradient in TCN becomes more stable. This helps to solve the problem of vanishing or exploding gradients, allowing the model to train more effectively. (4) Due to the fact that TCN has only one filter per layer, its memory usage is relatively low. This makes TCN more feasible in processing large-scale sequence data.

Through the dilated convolution structure of TCN, the effective window size of the convolution module will increase exponentially with the number of layers. After the model construction is completed, use the training dataset to train the TCN model. Improve the accuracy and stability of predictions by optimizing the parameters and structure of the model. The model achieves wider data feature mining with fewer layers, which can effectively solve the problem of large input data scale in wind power high-frequency power prediction.

D. WIND POWER PREDICTION BASED ON CONVOLUTIONAL NEURAL NETWORKS

By using convolutional neural networks for data convolution calculation, the prediction results of different modules are stacked to achieve accurate prediction of wind power. Convolutional neural networks (CNN) are common algorithms

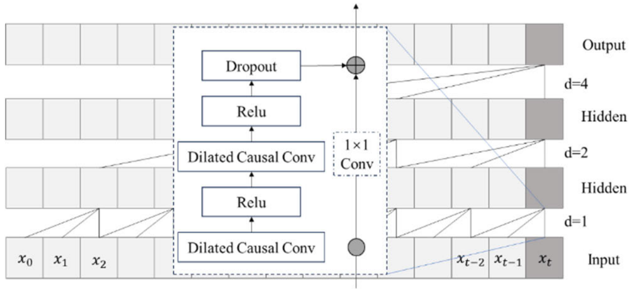


FIGURE 3. TCN model structure diagram.

in deep learning and widely used in image processing [36]. The CNN model structure is shown in Figure 2. The CNN model preprocesses the original image data and normalizes the image amplitude. Currently, in load forecasting, scholars treat load data as different grayscale points and preprocess them. The CNN model consists of two parts: convolutional layer and pooling layer. In the convolutional layer, the feature matrix of the input data is generated through convolutional kernel calculation, and the formula for calculating the feature matrix is shown in equation 5. Train using gradient descent and backpropagation algorithms, and the gradient formulas for convolutional and pooling layers are shown in equation 6 [37].

$$W' = \frac{(W + 2q - m)}{s} + 1 \tag{4}$$

$$\frac{\partial E}{\partial W_{ij}} = \frac{\partial E}{\partial V_{ij}} \frac{\partial V_{ij}}{\partial W_{ij}} = \delta_{ij} \frac{\partial V_{ij}}{\partial W_{ij}} \tag{5}$$

where, W is the feature weight matrix, V is the input matrix, m is the size of the convolution kernel, and q is the number of zero padding layers.

CNN models share convolutional kernels and can capture high-dimensional data features. At the same time, there is no need for manual feature selection, and linear and nonlinear relationships between different variables are explored through weight feature matrices. By using CNN convolutional layers to calculate the low-frequency and high-frequency prediction components of wind power, accurate prediction of ultra short term wind power can be achieved. The pseudocode based on the EMD-EncoderForest-TCN model is shown below.

III. RESULTS AND DISCUSSIONS

A. EVALUATION INDICATORS

This paper adopts Normalized Mean Absolute Error (NMAE), Normalized Root Mean Square Error (NRMSE), and Coefficient of Determination (R2) as metrics for evaluating wind power forecasts. NMAE measures average prediction errors, NRMSE emphasizes larger errors by squaring them, and R2 quantifies the fit between predicted and actual values. The calculation methods for these metrics are outlined below:

$$NMAE = \frac{1}{n} \sum\nolimits_{t=1}^{n} \left| \frac{\hat{y_t} - y_t}{Cap} \right| \tag{6}$$

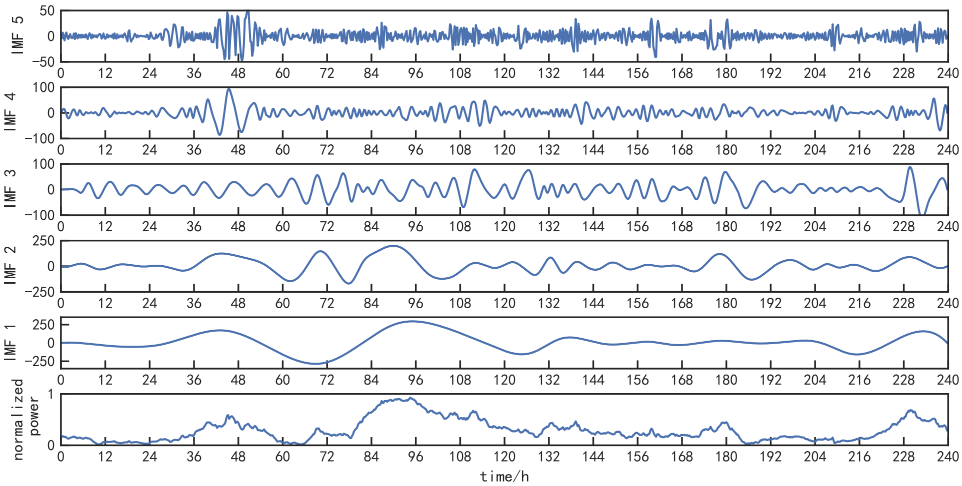


FIGURE 4. TCN model structure diagram.

Algorithm EMD-EncoderForest-TCN

```
def wind_power_prediction (wind_data, emd_component_
num,encoder_forest_params, tcn_params):
        emd = EMD()
        imfs = emd.emd(wind_data)
        high_frequency_component = imfs[0:i]
        low_frequency_component = np.sum(imfs[i:], axis=0)
        # Using EncoderForest model for low-frequency
components
        encoder_forest = EncoderForest(**encoder
        _forest_params) low_frequency_prediction =
encoder_forest.predict(low_frequency_component)
        # Using TCN model for high-frequency components
        tcn = TCN(**tcn_params)
        high_frequency_prediction=
tcn.predict(high_frequency_component)
        # Combination prediction results
        final_prediction = low_frequency_prediction +
high_frequency_prediction
            return final_prediction
```

$$NRMSE = \frac{1}{n} \sum\nolimits_{t=1}^{n} \left(\frac{\hat{y}_t - y_t}{Cap} \right)^2 \tag{7}$$

$$R2 = 1 - \frac{\sum\limits_{t=1}^{T} \left(\hat{y}_t - y_t\right)^2}{\sum\limits_{t=1}^{T} \left(\hat{y}_t - \bar{y}\right)^2} \tag{8}$$

where $\hat{y}_t$ is the actual power, y_t is the predicted value, Cap is the rated capacity of the wind farm, $\bar{y}$ is the average value of power, and n is the number of sampling points.

B. EVALUATION INDICATORS\

Data quality is a critical foundation for ensuring neural networks achieve their intended performance. The proposed method requires a coverage of over one year, aligning with the common demands of current deep learning approaches. The dataset utilized originates from the historical measurements of a wind farm in China, spanning from January 2021 to December 2022. To address issues of missing and anomalous data, we initially calculate the upper and lower quartiles of the historical data distribution to identify outliers, subsequently replacing both outliers and missing values. The replacement strategy employs a moving average technique. It's important to note that models trained on local data are only effective locally. Deploying prediction systems in other wind farms necessitates training models with data from those specific locations.

The deep learning models involved in this article are all developed on the pytorch (version 1.7.1) framework through the python (version 3.7.9) language.

Utilizing the aforementioned dataset, we configured the parameters including optimizer, learning rate, model depth, and other relevant settings as detailed in Table 1. Within the table, the TCN is constructed by stacking three neural network layers, with the internal cell connection rules adhering to standard default settings. The hyperparameters values, including Optimizer, Learning Rate, Epoch, and Batch Size listed in the table, were all determined through the application of a Bayesian stochastic optimization plugin.

TABLE 1. Parameter value.

Parameters	Values/Types
Optimizer	Adam
Learning_rate	0.0001
Epoch	350
Batch_size	64
TCN	hidden layers=3; nb_filters=64; kernel_size=2
Encoder Forest	n_bins=200; n_estimators=12

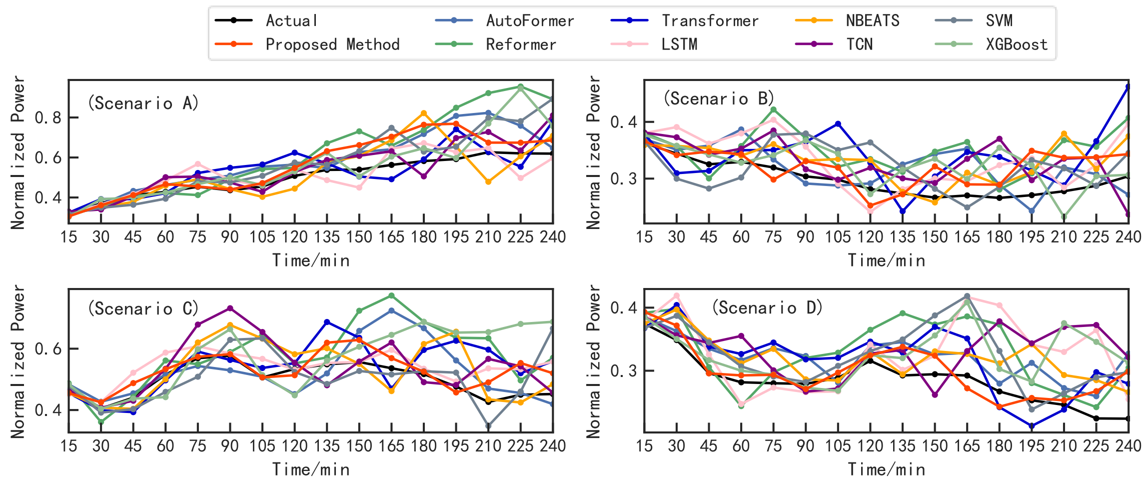


FIGURE 5. Wind power forecasting results of different methods. The forecasting error of each method increases with the extension of the lead time and is influenced by the uncertainty level of wind speed fluctuations, resulting in significant errors at certain moments. However, compared to the control model, the proposed method yields results that are closer to the actual values, with smaller error magnitudes, demonstrating higher precision.

TABLE 2. Allocation plan of each modal component and its prediction accuracy.

	EncoderForest	TCN	NMAE/%	NRMSE/%
Plan 1	IMF1	IMF2-5	5.61	7.97
Plan 2	IMF1-2	IMF3-5	5.30	7.59
Plan 3	IMF1-3	IMF4-5	6.42	8.27
Plan 4	IMF1-4	IMF5	6.46	8.35

C. EMD DECOMPOSITION RESULTS FOR WIND POWER SERIES

Based on the original data, this article selects 240h data points at intervals as a group. One group of data is taken as an example. The EMD method is used to decompose the original waveform signal into the classification of each mode. The original power sequence of one selected group and its first five modal components are shown in Figure 4. Each component reflects different characteristics of wind power output. The IMF4 and IMF5 components mainly manifest themselves as high-frequency fluctuations, and their amplitudes change significantly with time. IMF1∼IMF3 shows a smoother fluctuation trend. The information contained in different modal components is not consistent, and a targeted feature extraction network model should be established.

Based on the above modal decomposition results, we designed multiple modal component allocation schemes based on the differential extraction capabilities of EncoderForest and TCN for low-frequency fluctuation information and high-frequency information. The input modal component information of EncoderForest gradually increases from IMF1 to IMF1-4, and the input information of TCN gradually decreases from IMF2-5 to IMF5. The prediction accuracy obtained by different plans is shown in Table 2. It can be seen from the table that plan 2 has the highest accuracy, and this type of allocation plan will be used as fixed parameters of the model in the future.

D. COMPARISON WITH TRADITIONAL METHODS

To assess the superiority of the proposed method over conventional power prediction models, this section compares it with several benchmark models including AutoFormer, Reformer, Transformer, LSTM, NBEATS, TCN, SVM, and XGBoost. The essential training and modeling parameters for these models are documented in Table 9, located in the Appendix. The forecasting outcomes for four different types of power fluctuation scenarios are depicted in Figure 5. In the figure, the actual power fluctuations are represented in black, while the predictions made by the proposed method are shown in red.

The forecasting errors of all methods tend to increase as the prediction lead time extends, and they are affected by the degree of uncertainty in wind speed fluctuations, leading to significant errors at certain moments. Among the four scenarios, both the wind speed increase and decrease phases show a clear pattern of escalating errors, while scenarios with minor wind speed changes are more prone to random errors. However, compared to the benchmark models, the proposed method's predictions are closer to the actual values, with smaller error spikes, indicating superior precision. A comprehensive comparison reveals that the proposed method exhibits a stronger capability to accurately match actual power, demonstrating higher overall accuracy.

To further compare the advantages of the proposed methods, we use NRMSE, NMAE, and R2 as indicators to calculate the prediction errors of various methods at different time steps. The resulting error histograms are shown in Fig. 6 - Fig. 8. As can be seen from the figure, the proposed method achieves the lowest prediction error in each prediction step. Among the models, Reformer, LSTM, and SVM showed higher errors. The Transformer's underperformance is likely

TABLE 3. Forecasting error for various methods at the fourth hour.

	Proposed Method	AutoFormer	Reformer	Transformer	LSTM	NBEATS	TCN	SVM	XGBoost	Average improvement
NMAE/%	5.30	9.22	11.87	8.23	6.06	7.61	7.25	5.92	6.17	**2.49**
NRMSE/%	7.59	9.22	11.87	11.93	8.81	10.69	10.12	8.47	10.81	**2.65**
R2/%	89.97	86.33	80.94	84.30	86.98	83.74	84.36	85.75	83.52	**5.48**

TABLE 4. Ablation protocol.

Method	EMD	EncoderForest	TCN
Proposed method	√	IMF1-2	IMF3-5
Comparison method 1 (CM1)	-	Original	Original
Comparison method 2 (CM2)	-	Original	-
Comparison method 3 (CM3)	-	-	Original
Comparison method 4 (CM4)	√	IMF1-5	IMF1-5
Comparison method 5 (CM5)	√	IMF1-5	-
Comparison method 6 (CM6)	√	-	IMF1-5

TABLE 5. Prediction error of the 4th-hour for different ablation experimental test plans.

	Proposed Method	CM1	CM2	CM3	CM4	CM5	CM6	Average improvement
NMAE/%	5.30	6.29	7.14	7.37	6.25	5.98	5.67	**1.15**
NRMSE/%	7.59	9.12	10.31	10.59	8.93	8.66	8.11	**1.69**
R2/%	89.97	85.5	84.13	84.56	87.82	87.66	87.27	**3.81**

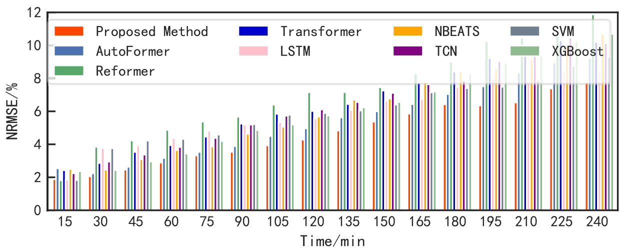


FIGURE 6. Stepwise NMAE for various models.

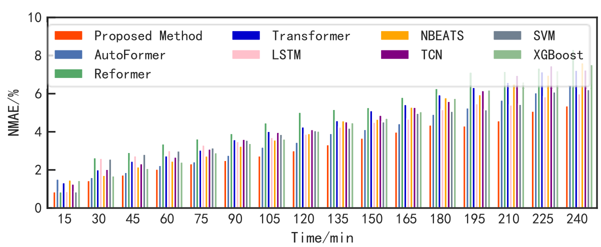


FIGURE 7. Stepwise NRMSE for various models.

due to its focus on long sequences, which doesn't align well with the short temporal spans analyzed in this study.

Given the wind power industry's focus on the accuracy of ultra-short-term power forecasting at the 16th step (the fourth hour), which serves as a standard for assessing wind farms' grid integration, we have documented the forecasting error of each method at this specific hour as a case in point.

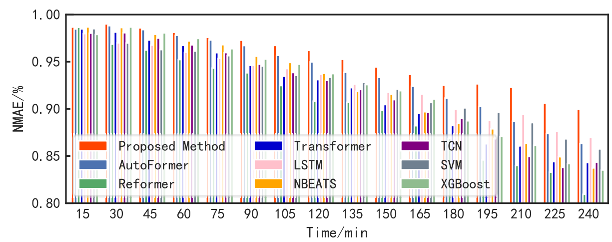


FIGURE 8. Stepwise R2 for various models.

From the data, it can be concluded that the proposed method can reduce the NMAE value by an average of 2.49% and the NRMSE value by an average of 2.65%. The accuracy, as measured by the R2 value, can be improved by an average of 5.48%. This superior performance is primarily attributed to the innovative integration of EMD, EncoderForest, and TCN techniques, which effectively decompose and analyze wind speed variations, capturing complex patterns more accurately than traditional models.

E. ABLATION EXPERIMENT

In order to analyze the effectiveness of EMD decomposition and TCN and EncoderForest in the proposed method, this section designs an ablation experiment as shown in Table 4. '√' in the table indicates that each frequency domain component obtained by EMD decomposition is used as data input, and '-' indicates that the EMD method is not used

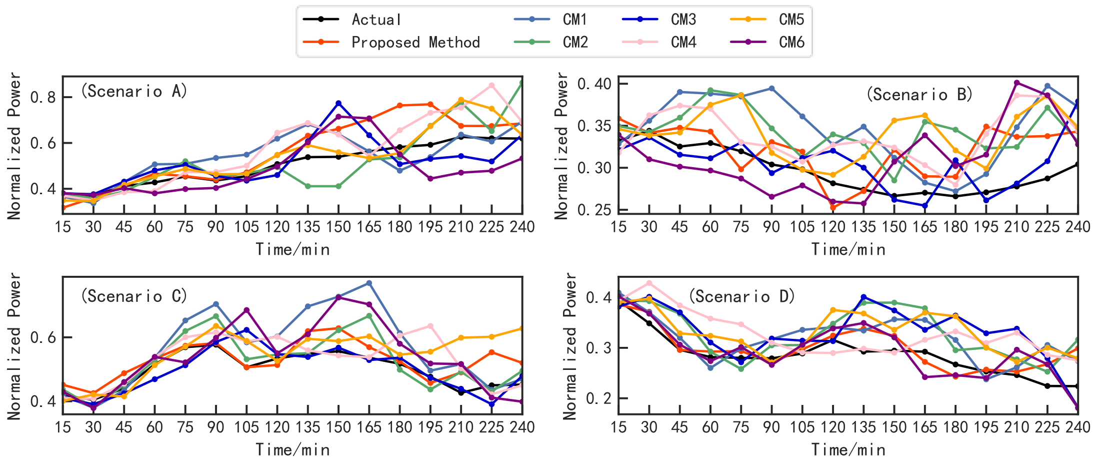


FIGURE 9. **Power prediction results of different ablation experimental test plans. Overall, predictions without frequency domain decomposition (CM1, CM2) show larger deviations, both in terms of random errors and errors at the fourth hour. This is most evident in scenario B. Additionally, methods using EncoderForest and TCN to parallel process the decomposition results exhibit smaller bias and less frequent large random fluctuations, ranking just below the proposed method among the comparison plans.**

and the original data is input. The part of the table titled "IMF/Original" is used to indicate whether the information input to the module is the modal component decomposed by EMD or the original power data. Mainly designed the following types of test plans:

- Test whether the EMD decomposition scheme is effective. The experiment includes CM1-CM6, where CM1-CM3 is a method without EMD, and CM4-CM6 is a method using EMD.
- Test the adaptability of the EncoderForest module to low-frequency components. The experiment includes all original information as input information, mainly CM1 and CM2. The input information is all modal component information, mainly CM4 and CM5.
- Test the adaptability of the TCN module to high-frequency components. The experiment includes all original information as input information, mainly CM1 and CM3. The input information is all modal component information, mainly CM4 and CM6.

Based on the described approach, the power prediction results for four scenarios of output changes are illustrated in Figure 8. In the figure, the red curve represents the prediction results of the proposed method, while the curves in other colors correspond to the results from various ablation experimental plans. Overall, predictions not involving frequency domain decomposition (CM1, CM2) exhibit larger deviations, evident in both error randomness and error amplitude, at the fourth hour, particularly in scenario B. Moreover, the method that uses EncoderForest and TCN to process decomposition results in parallel shows smaller bias and fewer occurrences of significant random fluctuations, ranking just below the proposed method among the comparison plans. A comprehensive comparison of curve fitting across various prediction scenarios demonstrates that the proposed method outperforms others, confirming its advantages.

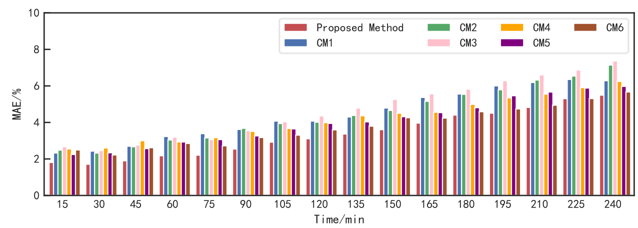


FIGURE 10. **Stepwise NMAE for different ablation experimental test protocols.**

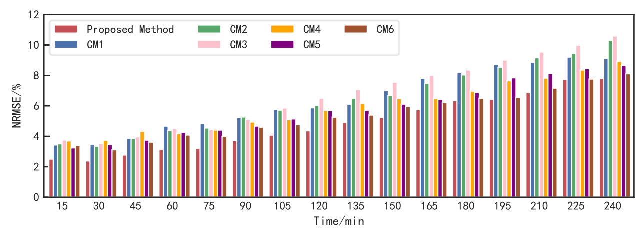


FIGURE 11. **Stepwise NRMSE for different ablation experimental test protocols.**

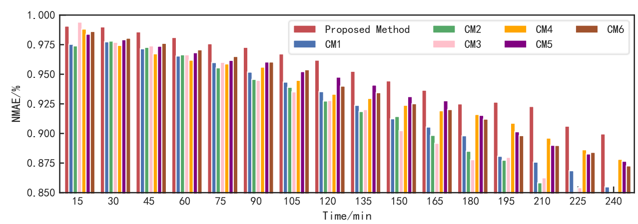


FIGURE 12. **Stepwise R2 for different ablation experimental test protocols.**

TABLE 6. The 4th-Hour forecasting accuracy metrics of SCS1.

	NMAE/%	NRMSE/%	R2/%
Proposed Method	5.16	8.06	91.57
AutoFormer	6.51	8.85	90.51
Reformer	8.15	12.94	80.94
Transformer	7.66	10.71	79.24
LSTM	6.48	9.04	93.95
NBEATS	7.15	10.37	79.55
TCN	7.03	10.12	86.05
SVM	6.03	10.1	83.18
XGBoost	8.21	10.47	87.7
Average improvement	1.99	2.27	6.43

TABLE 7. The 4th-hour forecasting accuracy metrics of SCS2.

	NMAE/%	NRMSE/%	R2/%
Proposed Method	5.35	8.4	88.06
AutoFormer	5.5	9.85	82.91
Reformer	7.75	12.53	79.18
Transformer	7.02	11.58	83.02
LSTM	6.94	9.93	80.96
NBEATS	7.21	9.02	78.71
TCN	6.27	9.38	87.56
SVM	6.69	11.3	86.05
XGBoost	7.33	8.92	85.59
Average improvement	1.49	1.91	5.06

TABLE 8. Percent improvement in accuracy of the proposed method.

Method of comparison	NMAE/%	NRMSE/%	R2/%
Traditional methods	2.49	2.65	5.48
Without EMD	1.63	2.41	5.24
Use only low-frequency components	1.12	1.67	3.69
Use only high-frequency component	1.09	1.59	3.68

We still use NMAE, NRMSE, and R2 as evaluation indicators for the prediction accuracy of each step size, and the results are shown in Fig.10, Fig.12 respectively. As can be seen from the figure, in general, the prediction errors of various methods gradually increase with the prediction step length. Among them, the errors of CM1-CM3 that do not use the EMD method increase significantly, and the worst performance is CM3 that uses TCN alone. Among these three types of methods, the best is CM1 which uses TCN and EncoderForest. Among the CM4-CM6 that use the modal component data obtained by EMD, the best performance is the CM6 that uses TCN as the encoder alone. The effects of CM4 and CM5 are similar.

Further analysis of the 16th step prediction errors from various ablation experimental plans is presented in Table 5. The table reveals that methods incorporating Empirical Mode Decomposition (EMD) exhibit lower prediction errors compared to those without EMD (CM1-CM3), with an average reduction of 1.63% in NMAE error, 2.41% in RMSE error, and an average improvement of 5.24% in R2. When compared to methods that do not utilize EncoderForest for extracting low-frequency modal information (CM1, CM2, CM4, CM5), the proposed method achieves a reduction of 1.12% in NMAE error, 1.67% in NRMSE error, and an average improvement of 3.69% in R2. Additionally, against methods lacking TCN for high-frequency component extraction (CM1, CM3, CM4, CM6), the proposed method shows a reduction of 1.09% in NMAE error, 1.59% in NRMSE error, and an improvement of 3.68% in R2. These findings underscore the superior prediction accuracy of the proposed method, highlighting the significant contribution of the EMD method, followed by the EncoderForest module, and finally the TCN module in enhancing forecasting precision.

F. VALIDATING EFFECTIVENESS ON OTHER WIND FARMS

To validate the effectiveness of the proposed method on data from wind farms in other locations, we conducted additional tests using two years of operational data from wind farms in Northwest China (the first supplementary case study, SCS1) and Northeast China (the second supplementary case study, SCS2). The data underwent the same integrity checks and cleaning processes as described previously and were compared against the same control models mentioned earlier. The predictive accuracy of the test sets is detailed in Tables 6 and 7.

The results demonstrate that the proposed method also exhibits high performance on additional datasets. In the first supplementary case study, the method achieved an average reduction of 1.99% in NMAE error, 2.27% in NRMSE error, and an average increase of 6.43% in the R2 accuracy metric. In the second supplementary case study, the method saw an average reduction of 1.49% in NMAE error, 1.91% in NRMSE error, and an average increase of 5.06% in the R2 accuracy metric. The performance improvements in different regions highlight our model's adaptability and reliability in forecasting wind power, demonstrating its broad applicability across varied environmental conditions.

IV. CONCLUSION

Enhancing the precision of ultra-short-term wind power forecasting is essential for ensuring the stability of the power grid and boosting the revenues of wind farms connected to it. To address the challenges posed by the multi-scale stochastic variations in wind speed, which complicate wind power prediction, we have developed an integrated short-term wind power prediction method that combines EMD, EncoderForest, and TCN. This methodology underwent evaluation using authentic operational data, demonstrating enhancements across key evaluation metrics as detailed in Table 8. The principal findings include:

- Modal decomposition is a crucial strategy for enhancing ultra-short-term wind power prediction accuracy.

By differentially extracting information from various frequency modes, our method leverages the EncoderForest module for low-frequency trend data and the TCN module for high-frequency details, significantly boosting accuracy. The results show a reduction in NMAE error by an average of 2.49% and in NRMSE error by 2.65%, with an average increase of 5.48% in the R2 accuracy metric.

- Regarding the impact of model components, the EMD method stands out, reducing NMAE error by 1.63%, RMSE error by 2.41%, and enhancing R2 by 5.24% on average. The EncoderForest module, focusing on low-frequency information, follows closely, with a reduction in NMAE error by 1.12%, NRMSE error by 1.67%, and a 3.69% average increase in R2, underscoring the critical role of trend information in ultra-short-term prediction. The TCN module's extraction of high-frequency details also contributes, reducing NMAE error by 1.09%, NRMSE error by 1.59%, and improving R2 by 3.68%.

It can be seen from the existing research conclusions that the current challenge of ultra-short-term power prediction mainly lies in improving the prediction accuracy in step 16. The power prediction trend information and fluctuation details should be further jointly considered, and ultra-short-term numerical weather forecasting should be introduced, which will effectively increase the prediction accuracy of the 4th hour time node. This will be one of the focuses of our future research work.

V. DISCUSSION

The accurate prediction of wind power based on EMD-EncoderForest-TCN has various potential applications. Firstly, the wind power prediction results can provide accurate information for power system dispatchers about the future generation of wind farms. This helps dispatchers to develop more reasonable power generation plans, optimize the allocation of power resources, and ensure the stable operation of the power system. Secondly, accurate wind power prediction results can provide important basis for the management of energy storage systems. Energy storage systems can provide supplementary electricity when wind power output is insufficient to balance power supply and demand. Finally, wind power forecasting results can provide decision support for trading in the electricity market. Buyers and sellers in the electricity market can evaluate future electricity supply and demand based on wind power forecasting results and develop reasonable trading strategies. In summary, the wind power prediction results based on the EMD-EncoderForest-TCN model have broad application prospects in power system scheduling and optimization, energy storage system management, wind farm operation and optimization, electricity market trading, and renewable energy policy formulation. These applications can not only improve the efficiency and reliability of the wind power industry, but also promote the integration and utilization of renewable energy, and promote the transformation and upgrading of the energy structure.

VI. APPENDIX

TABLE 9. Compare model parameter settings.

Model	Learning Rate	Batch Size	Model Structure Parameters
AutoFormer	0.001	32	Encoder layers=3, Decoder layers=3, Attention heads=4
Reformer	0.001	32	Hash bucket size=64, Chunk length=64, Attention heads=8
Transformer	0.0001	64	Encoder layers=6, Decoder layers=3, Attention heads=4
LSTM	0.001	64	Hidden units=128, Layers=2, Dropout rate=0.2
NBEATS	0.001	128	Stacks=3, Hidden units=512
TCN	0.001	32	Kernel size=2, Dilations=[1, 2, 4], Layers=6
SVM	N/A	N/A	Kernel type='rbf', C (regularization)=1, Gamma='scale'
XGBoost	0.001	N/A	Max depth=6, N estimators=120

REFERENCES

[1] A. Rozhkov, "Harnessing European policies for energy planning in illinois: Overcoming barriers and transitioning to a climate-neutral society," *Sustain. Cities Soc.*, vol. 98, Nov. 2023, Art. no. 104803.

[2] Y. Feng, J. Zhang, Y. Geng, S. Jin, Z. Zhu, and Z. Liang, "Explaining and modeling the reduction effect of low-carbon energy transition on energy intensity: Empirical evidence from global data," *Energy*, vol. 281, Oct. 2023, Art. no. 128276.

[3] J. Shi, L. Ma, C. Li, N. Liu, and J. Zhang, "A comprehensive review of standards for distributed energy resource grid-integration and microgrid," *Renew. Sustain. Energy Rev.*, vol. 170, Dec. 2022, Art. no. 112957.

[4] Y. Chen and H. Lin, "Overview of the development of offshore wind power generation in China," *Sustain. Energy Technol. Assessments*, vol. 53, Oct. 2022, Art. no. 102766.

[5] S. Zhang, J. Wei, X. Chen, and Y. Zhao, "China in global wind power development: Role, status and impact," *Renew. Sustain. Energy Rev.*, vol. 127, Jul. 2020, Art. no. 109881.

[6] A. Bamooeifard, "Future studies in Iran development plans for wind power, a system dynamics modeling approach," *Renew. Energy*, vol. 162, pp. 1054–1064, Dec. 2020.

[7] C. Hitaj, "Wind power development in the United States," *J. Environ. Econ. Manage.*, vol. 65, no. 3, pp. 394–410, May 2013.

[8] Y. Wang, X. Shao, C. Liu, G. Cai, L. Kou, and Z. Wu, "Analysis of wind farm output characteristics based on descriptive statistical analysis and envelope domain," *Energy*, vol. 170, pp. 580–591, Mar. 2019.

[9] J. Shi, N. Liu, Y. Huang, and L. Ma, "An edge computing-oriented net power forecasting for PV-assisted charging station: Model complexity and forecasting accuracy trade-off," *Appl. Energy*, vol. 310, Mar. 2022, Art. no. 118456.

[10] C. Ai, S. He, H. Hu, X. Fan, and W. Wang, "Chaotic time series wind power interval prediction based on quadratic decomposition and intelligent optimization algorithm," *Chaos, Solitons Fractals*, vol. 177, Dec. 2023, Art. no. 114222.

[11] Y. Han, X. Tong, S. Shi, F. Li, and Y. Deng, "Ultra-short-term wind power interval prediction based on hybrid temporal inception convolutional network model," *Electr. Power Syst. Res.*, vol. 217, Apr. 2023, Art. no. 109159.

[12] L. Xiang, J. Liu, X. Yang, A. Hu, and H. Su, "Ultra-short term wind power prediction applying a novel model named SATCN-LSTM," *Energy Convers. Manage.*, vol. 252, Jan. 2022, Art. no. 115036.

[13] L. Li, Y. Li, B. Zhou, Q. Wu, X. Shen, H. Liu, and Z. Gong, "An adaptive time-resolution method for ultra-short-term wind power prediction," *Int. J. Electr. Power Energy Syst.*, vol. 118, Jun. 2020, Art. no. 105814.

[14] R. Fang, Y. Wang, R. Shang, Y. Liang, L. Wang, and C. Peng, "The ultra-short term power prediction of wind farm considering operational condition of wind turbines," *Int. J. Hydrogen Energy*, vol. 41, no. 35, pp. 15733–15739, Sep. 2016.

[15] Y. Zhang, J. Han, G. Pan, Y. Xu, and F. Wang, ''A multi-stage predicting methodology based on data decomposition and error correction for ultra-short-term wind energy prediction,'' *J. Cleaner Prod.*, vol. 292, Apr. 2021, Art. no. 125981.

[16] C. Stathopoulos, A. Kaperoni, G. Galanis, and G. Kallos, ''Wind power prediction based on numerical and statistical models,'' *J. Wind Eng. Ind. Aerodynamics*, vol. 112, pp. 25–38, Jan. 2013.

[17] J. M. G. Sopeña, V. Pakrashi, and B. Ghosh, ''A benchmarking framework for performance evaluation of statistical wind power forecasting models,'' *Sustain. Energy Technol. Assessments*, vol. 57, Jun. 2023, Art. no. 103246.

[18] S. Sun, Z. Du, K. Jin, H. Li, and S. Wang, ''Spatiotemporal wind power forecasting approach based on multi-factor extraction method and an indirect strategy,'' *Appl. Energy*, vol. 350, Nov. 2023, Art. no. 121749.

[19] N.-Z. Guo, K.-Z. Shi, B. Li, L.-W. Qi, H.-H. Wu, Z.-L. Zhang, and J.-Z. Xu, ''A physics-inspired neural network model for short-term wind power prediction considering wake effects,'' *Energy*, vol. 261, Dec. 2022, Art. no. 125208.

[20] P. Nasery and A. Aziz Ezzat, ''Yaw-adjusted wind power curve modeling: A local regression approach,'' *Renew. Energy*, vol. 202, pp. 1368–1376, Jan. 2023.

[21] H. Zhou, Y. Qiu, Y. Feng, and J. Liu, ''Power prediction of wind turbine in the wake using hybrid physical process and machine learning models,'' *Renew. Energy*, vol. 198, pp. 568–586, Oct. 2022.

[22] H. Wang, Z. Tan, Y. Liang, F. Li, Z. Zhang, and L. Ju, ''A novel multi-layer stacking ensemble wind power prediction model under tensorflow deep learning framework considering feature enhancement and data hierarchy processing,'' *Energy*, vol. 286, Jan. 2024, Art. no. 129409.

[23] C.-L. Liu, T.-Y. Chang, J.-S. Yang, and K.-B. Huang, ''A deep learning sequence model based on self-attention and convolution for wind power prediction,'' *Renew. Energy*, vol. 219, Dec. 2023, Art. no. 119399.

[24] M. A. Houran, S. M. S. Bukhari, M. H. Zafar, M. Mansoor, and W. Chen, ''COA-CNN-LSTM: Coati optimization algorithm-based hybrid deep learning model for PV/wind power forecasting in smart grid applications,'' *Appl. Energy*, vol. 349, Nov. 2023, Art. no. 121638.

[25] Y. Chen, X. Hu, and L. Zhang, ''A review of ultra-short-term forecasting of wind power based on data decomposition-forecasting technology combination model,'' *Energy Rep.*, vol. 8, pp. 14200–14219, Nov. 2022.

[26] Z. Tian and H. Chen, ''A novel decomposition-ensemble prediction model for ultra-short-term wind speed,'' *Energy Convers. Manage.*, vol. 248, Nov. 2021, Art. no. 114775.

[27] M. Lazhari and A. Sadhu, ''Decentralized modal identification of structures using an adaptive empirical mode decomposition method,'' *J. Sound Vibrat.*, vol. 447, pp. 20–41, May 2019.

[28] X. Zhang, X. Du, and J. Brownjohn, ''Frequency modulated empirical mode decomposition method for the identification of instantaneous modal parameters of aeroelastic systems,'' *J. Wind Eng. Ind. Aerodynamics*, vol. 101, pp. 43–52, Feb. 2012.

[29] W. Guan, L. Dong, A. Zhang, and Y. Cai, ''Output-only modal identification with recursive dynamic mode decomposition for time-varying systems,'' *Measurement*, vol. 224, Jan. 2024, Art. no. 113852.

[30] J. Feng and Z.-H. Zhou, ''AutoEncoder by forest,'' in *Proc. 32nd AAAI Conf. Artif. Intell. (AAAI)*, New Orleans, LA, USA, 2018, pp. 1–7.

[31] Z. Sun, G. Wang, P. Li, H. Wang, M. Zhang, and X. Liang, ''An improved random forest based on the classification accuracy and correlation measurement of decision trees,'' *Expert Syst. Appl.*, vol. 237, Mar. 2024, Art. no. 121549.

[32] N. E. I. Karabadji, A. A. Korba, A. Assi, H. Seridi, S. Aridhi, and W. Dhifli, ''Accuracy and diversity-aware multi-objective approach for random forest construction,'' *Expert Syst. Appl.*, vol. 225, Sep. 2023, Art. no. 120138.

[33] J. Zhu, L. Su, and Y. Li, ''Wind power forecasting based on new hybrid model with TCN residual modification,'' *Energy AI*, vol. 10, Nov. 2022, Art. no. 100199.

[34] T. Limouni, R. Yaagoubi, K. Bouziane, K. Guissi, and E. H. Baali, ''Accurate one step and multistep forecasting of very short-term PV power using LSTM-TCN model,'' *Renew. Energy*, vol. 205, pp. 1010–1024, Mar. 2023.

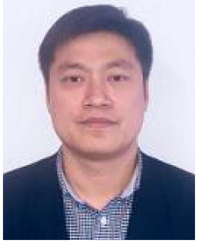
[35] C. Hu, Y. Zhao, H. Jiang, M. Jiang, F. You, and Q. Liu, ''Prediction of ultra-short-term wind power based on CEEMDAN-LSTM-TCN,'' *Energy Rep.*, vol. 8, pp. 483–492, Nov. 2022.

[36] V. B. Gowda, M. T. Gopalakrishna, J. Megha, and S. Mohankumar, ''Foreground segmentation network using transposed convolutional neural networks and up sampling for multiscale feature encoding,'' *Neural Netw.*, vol. 170, pp. 167–175, Feb. 2024.

[37] A. Babalhavaeji, M. Radmanesh, M. Jalili, and S. A. Gonzalez, ''Photovoltaic generation forecasting using convolutional and recurrent neural networks,'' *Energy Rep.*, vol. 9, pp. 119–123, Nov. 2023.

YU SUN was born in Gaizhou, Liaoning, China, in 1979. He received the Ph.D. degree from Xi'an Jiaotong University, China. He is currently with State Grid Corporation of China. His research interests include electric power system analysis and new energy generation technology.

JUNJIE YANG was born in Linfen, Shanxi, China, in 1995. He received the bachelor's degree from Xi'an University of Technology, China. He is currently with Shenyang Institute of Computing Technology, Chinese Academy of Sciences. His research interest includes new energy science and engineering.



XIAOTIAN ZHANG was born in Gaizhou, Liaoning, China, in 1978. He received the master's degree from Northeast Electric Power University, China. He is currently with State Grid Corporation of China. His research interests include power dispatching and power market trading.

KAIYUAN HOU was born in Changchun, Jilin, China, in 1973. He received the Ph.D. degree from Tsinghua University, China. He is currently with State Grid Corporation of China. His research interests include power system security and stability are divided into control and power grid dispatching operation management.

JIYUN HU was born in Jinzhou, Liaoning, China, in 1995. She received the master's degree from Shanghai Jiaotong University, China. She is currently with State Grid Corporation of China. Her research interest includes power system operation optimization.

GUANGZHI YAO was born in Huichun, Jilin, China, in 1993. He received the master's degree from Huazhong University of Science and Technology, China. He is currently with State Grid Corporation of China. His research interest includes power grid regulation and control.