

## Research Article

# A Hybrid Method for Ultrashort-Term Wind Power Prediction considering Meteorological Features and Seasonal Information

Baobin Zhou, Che Liu, Jianjing Li, Bo Sun , and Jun Yang 

School of Control Science and Engineering, Shandong University, Jinan, Shandong 250000, China

Correspondence should be addressed to Bo Sun; [sunbo@sdu.edu.cn](mailto:sunbo@sdu.edu.cn)

Received 12 March 2020; Revised 24 July 2020; Accepted 20 August 2020; Published 28 September 2020

Academic Editor: Fausto Arpino

Copyright © 2020 Baobin Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

High-precision wind power prediction is important for the planning, economics, and security maintenance of a power grid. Meteorological features and seasonal information are strongly related to wind power prediction. This paper proposes a hybrid method for ultrashort-term wind power prediction considering meteorological features (wind direction, wind speed, temperature, atmospheric pressure, and humidity) and seasonal information. The wind power data are decomposed into stationary subsequences using the ensemble empirical mode decomposition (EEMD). The principal component analysis (PCA) is used to reduce the redundant meteorological features and the algorithm complexity. With the stationary subsequences and extracted meteorological features data as inputs, the long short-term memory (LSTM) network is used to complete the wind power prediction. Finally, the seasonal autoregressive integrated moving average (SARIMA) is innovatively used to fit seasonal features (quarterly and monthly) of wind power and reconstruct the prediction results of LSTM. The proposed method is used to predict 15-minute wind power. In this study, three datasets were collected from a windfarm in Laizhou to validate the prediction performance of the proposed method. The experimental results showed that the prediction accuracy was significantly improved when meteorological features were considered and further improved with seasonal correction.

## 1. Introduction

With the sharp decline in fossil energy reserves and increasing prominence of environmental problems, wind power has seen rapid development as a renewable and clean energy source around the world [1]. At the end of 2016, the global wind energy cumulative capacity reached 486.74 GW, with a new installation for 54.6 GW [2]. With the increasing proportion of installed wind capacity, the impact of the volatility and randomness of wind power on the power grid has become obvious, and the subsequent abandonment of wind has become more serious [3]. To suppress the influence of randomness, ensure the safe operation of the power grid, and improve the rate of wind power consumption, accurate prediction of wind power is important.

According to the time scale, the wind power forecasting model is divided into four categories: long-term, medium-term, short-term, and ultrashort-term prediction. The long-

term prediction is usually in years and mainly serviced for the formulation of wind farm plan and annual power generation plan. The medium-term prediction is usually in weeks or months and generally used to arrange maintenance plans. The short-term prediction is generally in days and generally used to reduce the discarded wind power, optimize the maintenance scheduling, and optimize the generation scheduling [4]. So the accuracy of short-term prediction is particularly important for planning of power grid.

Ultrashort-term wind power prediction refers to forecasting at intervals of 5, 10, or 15 minutes. Such prediction results are mainly applied to increasing wind power consumption, reducing wind turbine regulation pressures, and ensuring the safe operation of power grids [5]. The ultrashort-term prediction is theoretically the most demanding forecasting method; it is mainly used to control daily operations of wind farm units [6, 7]. In addition, it is considered essential for maximizing the potential revenues of operators due to

price fluctuations. This is because future price in the renewable energy market is determined via a bidding mechanism between wind power operators and renewable market operators; the potential revenue for a wind power operator is directly affected by their ability to predict when and how much electricity will be generated in the future [8].

Currently, there are three basic types of wind power prediction methods: physical, statistical, and hybrid. Some research focused on building physical methods using mechanism analysis and equations. For ensuring the best efficiency of power outputs, the authors in [9, 10], respectively, established turbine models to discuss the influence of different hub heights and configurations and the effects of inflow turbulence intensity and turbine arrangements. But due to the fluctuations and uncertainty of meteorological change, it is difficult to obtain the satisfactory forecasting results only through a single mechanism prediction model. Statistical methods, such as the autoregressive moving average model, are based on pattern recognition, parameter estimation, and model checking to construct historical data and predicted values [11, 12]. However, statistical models cannot dig deep into data features, which lead to large prediction errors. Hybrid methods have two main constructions. The first construction focuses on data preprocessing, in which a decomposition algorithm is used to decompose the original wind power sequence into several sublayers to weaken the volatility. Each sublayer is predicted separately by the prediction model, and the results of all sublayers are then recombined to obtain the final predicted values [13, 14]. The second construction usually uses several different prediction models to predict the wind power separately and then recombines the results to obtain the final predicted values [15, 16]. Because hybrid methods have a better predictive performance, research with this approach has peaked in recent years. Hong et al. [17] proposed a hybrid method based on empirical mode decomposition (EMD) and an artificial neural network (ANN). The EMD was used to decompose the wind and wind power to obtain the stationary sub-sequences, and the ANN was used to predict the wind power. Yang et al. [14] proposed a hybrid method using the improved EMD to decompose the wind power. The ANN and SVM were used to establish prediction model according to different frequencies of sub-sequences. In recent years, long short-term memory (LSTM) networks have been widely used because of their unique memory and forgetting mode; they can flexibly adapt to timing fluctuations in wind power and have better prediction accuracies than conventional methods. Liu et al. [18] proposed a hybrid method based on a discrete wavelet transform (DWT) and the LSTM. The DWT is used to decompose the nonstationary wind power into components which have more stationarity, and each component is dug by an independent LSTM. Yu et al. [19] proposed a hybrid method using improved LSTM network to predict the wind power and further optimized the prediction results with clustering analysis. Han et al. [20] proposed a hybrid method using improved variational mode decomposition (VMD) to decompose wind power, and using improved LSTM to get the prediction.

These studies achieved many optional models for wind power prediction, and the forecasting accuracy was improved by applied the hybrid methods. However, most of the current studies only used wind power data, which do not make full use of meteorological information such as wind speed, wind direction, and temperature. In addition, the prediction model fails to fully extract seasonal information which will greatly improve the prediction accuracy if these messages are considered.

Based on previous work, a hybrid method based on meteorological features and seasonal information is proposed. This paper uses LSTM-SARIMA hybrid model to achieve a 15-minute prediction. The innovations are as follows:

- (i) The meteorological features were regarded as the input variables, and the principal component analysis (PCA) was used to reduce the redundant components that these excessive influencing variables may cause higher computational complex.
- (ii) With the stationary subsequences and extracted meteorological features data as inputs, the LSTM was used to predict the wind power, taking advantage of seasonal autoregressive integrated moving average (SARIMA) method that can extract the seasonality characteristics (including quarterly and monthly) of wind power [21].

The rest of this paper is organized as follows. Section 2 describes the process of data preprocessing, including the decomposition of the wind power and the extraction of the related meteorological features. Section 3 introduces the prediction algorithms used in this paper. Section 4 introduces the case study. The method was applied to data from a windfarm in Laizhou for verification, and the prediction accuracy was compared with those of the four other methods. Finally, Section 5 summarizes this study and presents plans for future work.

## 2. Data Preprocessing

The architectural design of this prediction method is shown in Figure 1.

The details of how the proposed method decomposes the original wind power sequence and considers the coupling relationship between the wind power and meteorological features are presented here.

*2.1. Ensemble Empirical Mode Decomposition of Wind Power.* EMD is a direct, empirical, and adaptive data processing method designed specifically for nonlinear and non-stationary data. Unlike traditional methods, such as Fourier decomposition, that use fixed bases, EMD directly decomposes nonlinear and non-stationary time series into multiple sublayers (multiple IMF layers and a Res term) by the Hilbert transform without any base functions or filter functions. Because of this unique advantage, EMD is widely used in complex system analyses [22, 23] and has been demonstrated to be advantageous for time series prediction.

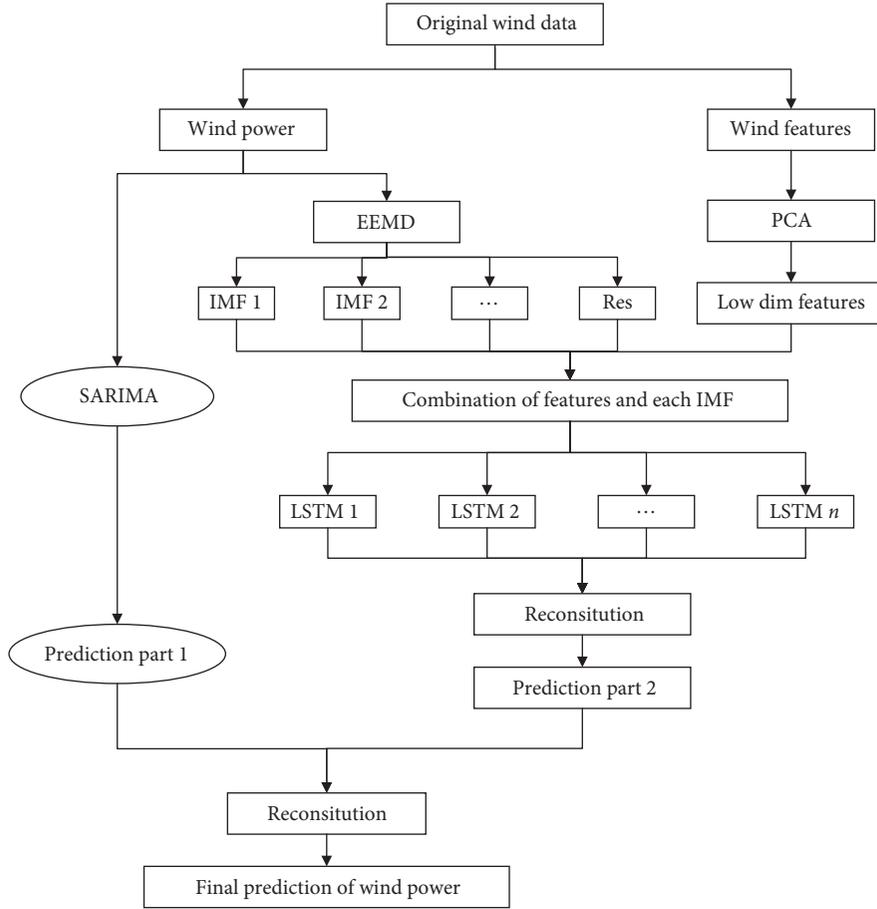


FIGURE 1: Overall process of the wind power prediction method.

EEMD was developed as an extended version of EMD to overcome the modal confusion problem [24] by adding white noise to the original time series multiple times; then, EMD is performed on the modified time series to obtain IMF sublayers, and the overall average of the corresponding IMF sublayers is considered as the final decomposition result. The main steps of EEMD are as follows.

Add white noise  $w^i(t) - N(0, \sigma^2)$  to the original wind power series  $p(t)$  to convert it to a new time series:

$$p^i(t) = p(t) + w^i(t), \quad i = 1, 2, \dots, m. \quad (1)$$

Decompose the time series  $p^i(t)$  into  $n$  IMF sublayers  $e_j^i(t)$  ( $j = 1, 2, \dots, n$ ) and a Res term  $r^i(t)$  whose frequency is close to zero and cannot be extracted further:

$$p^i(t) = \sum_{j=1}^n e_j^i(t) + r^i(t), \quad (2)$$

where  $e_j^i(t)$  is the  $j$ th IMF sublayer of the  $i$ th wind power sequence.

Repeat the above operation several times, in which the corresponding IMF sublayers could be obtained by adding different white noise to the original time series.

Calculate the average of all IMF sublayers obtained in  $m$  decompositions as the final IMF sublayer:

$$e_j(t) = \frac{1}{m} \sum_{i=1}^m e_j^i(t). \quad (3)$$

Once EEMD is complete, the original time series can be represented as a linear combination of IMF sublayers and the Res term:

$$P(t) = \sum_{j=1}^n e_j(t) + r(t), \quad (4)$$

where  $e_j(t)$  is the average of the  $j$ th IMF extracted during  $m$  decompositions,  $r(t)$  is the final residual, and  $n$  is the number of IMF sublayers.

**2.2. Meteorological Feature Extraction Based on Principal Component Analysis.** PCA is a multivariate statistical analysis technique that combines data compression and feature extraction [25]. PCA data inputs increase the overall accuracy by preserving the actually useful information. Also, by reducing the number of input data, the model training and prediction times can be significantly reduced. PCA has

thus been effectively used in wind power prediction [26, 27]. The idea of PCA is to construct new variables from linear combinations of the original and select the variables that reflect the original as much as possible. PCA maps  $n$ -dimensional features to  $k$  dimensions ( $k < n$ ), which results in a new set of orthogonal features called principal components (PC), rather than simply removing the remaining  $n - k$  dimensional features from the  $n$ -dimensional. The PCA algorithm calculates the eigenvalues of the input variable matrix, finds the variance corresponding to each input variable, and then determines the PC according to the cumulative value [28]. The main steps are as follows.

Center all meteorological features (including wind direction, wind speed, temperature, atmospheric pressure, and humidity): find the average of each feature  $f_j$  ( $j = 1, 2, \dots, 5$ ) and then subtract the mean from each feature  $f_j$ :

$$f_{ij}^* = \frac{f_{ij} - \bar{f}_j}{\sqrt{\text{var}(f_j)}}, \quad i = 1, 2, \dots, n, \quad (5)$$

$$\text{var}(f_j) = \frac{1}{n-1} \sum_{i=1}^n (f_{ij} - \bar{f}_j)^2, \quad (6)$$

where  $f_{ij}$  is the  $i$ th value of feature  $j$ ,  $\bar{f}_j$  is the average of feature  $j$ ,  $n$  is the number of values of feature  $j$ , and  $p$  is the number of features.

Calculate the covariance matrix of features:

$$\text{Cov}(f) = \begin{pmatrix} r_{11} & \cdots & r_{1p} \\ \vdots & \ddots & \vdots \\ r_{p1} & \cdots & r_{pp} \end{pmatrix}, \quad (7)$$

$$r_{ij} = \frac{1}{n-1} \sum_{i=1}^n (f_{ti}^* - \bar{f}_{ti}^*), \quad ij = 1, 2, \dots, p. \quad (8)$$

Calculate the eigenvalues of the covariance matrix and corresponding eigenvectors. Calculate the contribution rate of all eigenvectors and select the appropriate PC that was regarded as the extracted meteorological features. A higher contribution rate indicates that the original feature contains stronger information. The selection of a PC is mainly determined by the cumulative contribution rate of eigenvectors. In other words, the cumulative contribution rate usually needs to be more than 85% to ensure that the PC can contain most of the information of the original feature:

$$\text{con}_i = \frac{r_{ij}}{\sum_{i=1}^p r_i}, \quad i = 1, 2, \dots, p. \quad (9)$$

### 3. Prediction Algorithms

**3.1. Seasonal Auxiliary Prediction Based on the SARIMA Model.** As an extension of the autoregressive integrated moving average (ARIMA) model, SARIMA considers the periodicity of the time series. The predicted results can be reconstructed using the results of the LSTM network to

correct the wind power values. When a wind power sequence contains seasonal and non-seasonal information, the typical ARIMA model cannot capture the behavior of the seasonal portion, which results in a wrong choice of model parameters [29]. The general formula for SARIMA ( $p, d, q$ ) ( $P, D, Q$ )<sub>s</sub> [30] is given by

$$\phi(L)(1-L)^d \phi(L)(1-L^s)^D P_t = a + \theta(L)\theta(L)e_t, \quad (10)$$

where  $P_t$  is the wind power sequence,  $e_t$  is white noise,  $L$  is the lag operator,  $D$  is the seasonal difference order, and  $d$  is the regular difference order, which represents the number of differences that smooth the wind power sequence,  $a$  is the regression model parameter, and  $s$  is the seasonality order that represents the seasonal change cycle of wind power.

Equations (11) and (12) are autoregressive and moving average polynomials, respectively, and represent the dependence of the future time series on past values and errors:

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p, \quad (11)$$

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q. \quad (12)$$

Similarly, equations (13) and (14) represent the seasonal autoregressive and seasonal moving average polynomials, respectively. Adding these polynomials to the ARIMA equation can help capture seasonal variations in time series:

$$\phi(L) = 1 - \phi_{p1} L - \phi_{p2} L^{p^2} - \dots - \phi_p L^p, \quad (13)$$

$$\theta(L) = 1 - \theta_{q1} L - \theta_{q2} L^{q^2} - \dots - \theta_q L^Q. \quad (14)$$

**3.2. Wind Power Prediction Based on the LSTM Networks.** As an improved model of the recurrent neural network (RNN), the LSTM network has not only the recursiveness of RNN but also a unique memory and forgetting mode that can flexibly adapt to the timing characteristics of wind power. The LSTM network solves the gradient disappearance and gradient explosion problems of RNN during training and can fully utilize the time dependence of historical wind power data. The LSTM network consists of an input layer, output layer, and hidden layer. Compared to the traditional RNN, the hidden layer is no longer an ordinary neural unit but an LSTM unit with a unique memory mode. The structure of the memory cell is shown in Figure 2.

Each LSTM unit has a cell [31] whose state at time  $t$  is denoted as  $c_t$ ; this can also be regarded as a memory unit which will input in the same cell the next time. The reading and modification of the memory cell are accomplished by controlling the input gate, forgetting gate, and output gate. Specifically, the workflow of the LSTM unit is as follows: at each given time, the LSTM unit receives the input state of combined feature vectors:

$$x_t = [\text{IMF}_t, \text{PC}_1, \text{PC}_2, \dots, \text{PC}_k], \quad (15)$$

where  $x_t$  is the  $t$ th combined feature vector,  $\text{IMF}_t$  is the  $t$ th wind power sub-sequence of different frequency scales, and  $(\text{PC}_1, \text{PC}_2, \dots, \text{PC}_k)$  is  $k$  meteorological feature vectors.

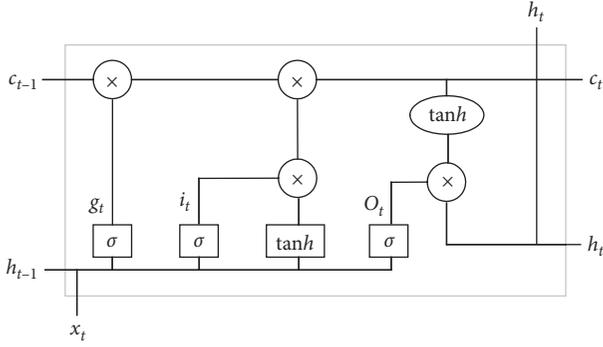


FIGURE 2: Structural unit of LSTM.

In addition, the LSTM unit also receives the hidden state of the previous time  $h_{t-1}$  and the internal information  $c_{t-1}$  as the state of the last cell. After receiving the information, the logic function of a gate determines whether the input is active. Then, the input is transformed by a nonlinear function and superimposed with the state of the memory cell processed by the forgetting gate to form a new memory cell state  $c_t$ . Finally, the state  $c_t$  forms the wind power prediction  $h_t$  by operation of the nonlinear function and dynamic control of the output gate. The variables are calculated as follows:

$$i_t = \sigma(w_{xi}x_t + w_{hi}h_{t-1} + w_{ci}c_{t-1} + b_i), \quad (16)$$

$$g_t = \sigma(w_{xg}x_t + w_{hg}h_{t-1} + w_{cg}c_{t-1} + b_g), \quad (17)$$

$$c_t = g_t c_{t-1} + i_t \tanh(w_{xc}x_t + w_{hc}h_{t-1} + b_c), \quad (18)$$

$$o_t = \sigma(w_{xo}x_t + w_{ho}h_{t-1} + w_{co}c_t + b_o), \quad (19)$$

$$h_t = o_t \tanh(c_t). \quad (20)$$

where  $w_{xc}$ ,  $w_{xi}$ ,  $w_{xg}$ ,  $w_{xo}$  are the weight matrices connected to the input signal  $x_t$ ;  $w_{hc}$ ,  $w_{hi}$ ,  $w_{hg}$ ,  $w_{ho}$  are the weight matrices connected to the input of the hidden layer;  $h_{t-1}$ ,  $w_{ci}$ ,  $w_{cg}$ ,  $w_{co}$  are the weight matrices connected to the internal information; and  $c_{t-1}$ ,  $b_i$ ,  $b_g$ ,  $b_c$ ,  $b_o$  are the offset vectors.  $\sigma$  is the activation function, which is usually a tanh or sigmoid function.

According to (18), the current cell state  $c_t$  is composed of the current input feature  $x_t$  and the previous cell state  $c_{t-1}$  in which there is the memory of previous input features that is controlled by forgetting gate  $g_t$ . When predicting the subsequent time point  $c_{t+1}$ , the information contained in  $c_{t-1}$  is not abandoned but is input into the next cell. As the sequence grows, all the features are used as important information for every subsequent wind power prediction.

Finally, the prediction results of the LSTM network are corrected by the results of the SARIMA model. The final wind power predicted values can be expressed as

$$\tilde{p}_t = \omega p_t + (1 - \omega)h_t \quad (0 \leq \omega \leq 1), \quad (21)$$

where  $p_t$  is the prediction results of SARIMA and  $\omega$  is the ratio by which the SARIMA model modifies the predicted values of the LSTM network, which takes the minimum error between the predicted value and the real value as the standard.

## 4. Case Study

**4.1. Evaluation Indices.** In this study, several indices were used to evaluate the proposed wind power prediction method: the mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE). MAE reflects the true state of the predicted value error; MAPE measures the deviation between the predicted and true values; and RMSE represents the sample standard deviation between the predicted and true values, which is sensitive to larger errors. The specific equations are given below:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Y_{\text{pre}}(t) - Y_{\text{act}}(t)|, \quad (22)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_{\text{act}}(t) - Y_{\text{pre}}(t)}{Y_{\text{act}}(t)} \right|, \quad (23)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_{\text{pre}}(t) - Y_{\text{act}}(t))^2}, \quad (24)$$

where  $n$  is the size of the test data and  $Y_{\text{pre}}(t)$  and  $Y_{\text{act}}(t)$  represent the predicted and true values of wind power, respectively. To compare the prediction performances of relevant models, the percentage improvements of MAE, MAPE, and RMSE were used as evaluation indices:

$$P_{\text{MAE}} = \frac{\text{MAE}_1 - \text{MAE}_2}{\text{MAE}_1} \times 100\%, \quad (25)$$

$$P_{\text{MAPE}} = \frac{\text{MAPE}_1 - \text{MAPE}_2}{\text{MAPE}_1} \times 100\%, \quad (26)$$

$$P_{\text{RMSE}} = \frac{\text{RMSE}_1 - \text{RMSE}_2}{\text{RMSE}_1} \times 100\%. \quad (27)$$

**4.2. Data Description.** Laizhou is located in the northeast of Shandong Province (37.0215°N, 119.4546°N), which borders the ocean and is rich in wind power resources. This study used three time series datasets of wind power collected at 15-minute intervals from a Laizhou wind farm. The wind turbine siting layout is fixed. The datasets contain a total of 35040 samples (from 15 January 2017 to 27 February 2018). As shown in Table 1, the detail dates for prediction model are presented. The wind farm has a total installed capacity of 100 MW. The information of wind speed, pressure, temperature, and relative humidity is directly acquired through the system of Supervisory Control and Data Acquisition (SCADA), which could realize real-time collection of local

TABLE 1: Dates for forecasting model in all seasons.

	Season	Training set	Validation set	Test set
1	Summer	1 June to 10 Aug. 2017	11 Aug. to 20 Aug. 2017	21 Aug. to 31 Aug. 2017
2	Winter	1 Dec. 2017 to 7 Feb. 2018	8 Feb. to 17 Feb. 2018	18 Feb. to 27 Feb. 2018
3	Spring and Autumn	15 Feb. to 10 May 2017	11 May to 20 May 2017	21 May to 31 May 2017
		1 Sept. to 31 Oct. 2017	1 Nov. to 10 Nov. 2017	11 Nov. to 20 Nov. 2017

meteorological information. There are many kinds of sensors in the system, which can collect many kinds of data such as meteorological environment and running state. The data has six features, as shown in Figure 3: wind direction, wind speed, temperature, atmospheric pressure, humidity, and wind power. Wind power shows high volatility and seasonality. The unit of wind speed is meters per second; the wind direction ranges from 0 to 360°, the temperature unit is Celsius, and the humidity ranges from 0 to 100 percent. During the modelling process, the first 80% of each dataset was used for training; the next 10% was used for cross-validation; and the last 10% was used for the test.

### 4.3. Experiments

**4.3.1. Parameter Settings and Decomposition Results Using EEMD.** As shown in Figure 3, the original wind power sequence had strong instability with no obvious rule to follow. It is difficult to obtain reliable results if the prediction model is directly established for analysis. In order to decrease the volatility of the sequence, EEMD was used to decompose the wind power. The standard deviation of the additional white noise was set to 0.2, and the EMD execution frequency was set to 100 times to make the Res item small enough. The algorithm produced 10 IMF sublayers and a Res term at different time scales. As shown in Figure 4, the stability of the IMF sublayers was significantly improved. The 11 subsequences generated after decomposition were used for the next LSTM modelling analysis.

**4.3.2. Extracting Results Using PCA.** The feature properties of wind power include five vectors: wind direction, wind speed, temperature, atmospheric pressure, and humidity. In order to analyze the correlation between the features and wind power, the Person correlation coefficient between wind power and each feature was tested. The calculation formula is shown as follows:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}, \quad (28)$$

where  $r$  represents the Person correlation coefficient,  $X_i$  and  $Y_i$  represent the features, respectively, and  $\bar{X}$  and  $\bar{Y}$  represent the mean value of respective samples.

As shown in Table 2, the wind speed is the most relevant to wind power, and the following is wind direction. It can be concluded that the wind turbine output is directly affected by the wind speed and wind direction. The temperature and humidity have a low correlation with wind power but have a

certain correlation with wind speed. This is because temperature and humidity can indirectly reflect wind speed. In summary, this article considered all these five meteorological features for experiments.

In order to reduce the redundant components of meteorological characteristics and reduce the computational complex, PCA was used to extract features, and the components with a large contribution rate were kept. Table 3 presents the contribution rates of the extracted PC. The third and fourth PC had very small contribution rates and contained negligible information of the original features. Therefore, only the first and second PC were considered in this experiment.

### 4.3.3. Parameter and Hyperparameter Settings of LSTM.

The IMF sublayers and Res term were combined with the principal components extracted by PCA to obtain 11 combined feature vectors  $X_t$  (IMF<sub>it</sub>, PC<sub>1</sub>, PC<sub>2</sub>), which were inputted into the LSTM network for training. For the experimental design of the training process, the Keras deep learning framework was selected. The sigmoid function was used as the activation function, and the model was optimized with the Adam optimization algorithm. To prevent overfitting, we set a dropout of 0.8 after cross-validation. The number of training samples per unit of time (batch size) was set to 96 that equals the amount of data in a day. The learning rate affects convergence effect and speed of the model, and we set it to 0.02 after experiments. In terms of the LSTM layers and neurons of dense, the following values were considered: {layers: 1, 2, 3, 4}, {neurons of dense: 1(70), 1(80), 2(70, 80), 2(80, 80)}. As shown in Table 4, RMSE is calculated to evaluate the forecasting performance using training set and validation set. As the number of layers increases, the training model will have a better effect, even reaching 100% prediction accuracy. However, what follows is that the model is overfitted, and the prediction effect is seriously reduced when the model is put on the test data. Above all, increasing the number of hidden layers may not be beneficial to forecasting performance. Also, based on their effects on the prediction and training speed, the number of LSTM layers was set to 2 with a dense layer that has 80 neurons.

**4.3.4. Parameter Settings of SARIMA.** At the same time, the SARIMA ( $p, d, q$ ) ( $P, D, Q$ )<sub>s</sub> model was built according to the information of the wind power sequence. The Akaike information criterion (AIC) was used to evaluate and compare models with different parameters [32]. AIC also considers the overall model complexity when measuring how a model fits the data. A model suitable for wind power data would have a lower AIC score than others. In

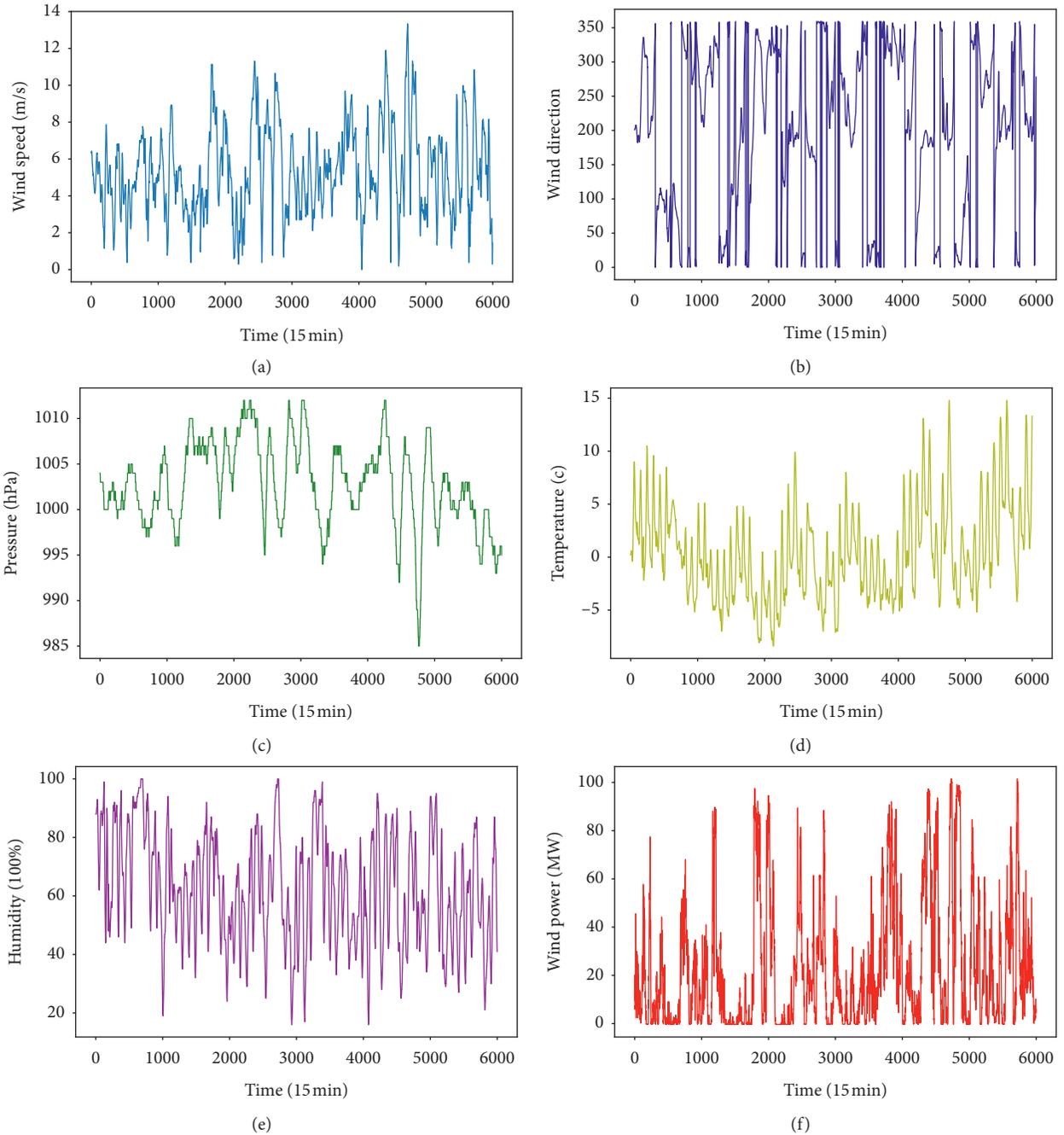


FIGURE 3: Wind power sequence and meteorological features.

applications,  $d$  rarely exceeds 1, and  $P$  and  $Q$  are usually no more than 3 [33]. In this study, the SARIMA function of the statsmodels framework was used to test the effects of different model parameters on AIC, and the results are presented in Table 5. Based on the minimum AIC score, SARIMA (3, 0, 1) (3, 1, 1)<sub>12</sub> was chosen as the wind power prediction model.

**4.4. Comparative Analysis.** To test prediction ability of the proposed three-stage wind power prediction method, it was compared with four other methods: single LSTM, Gated Recurrent Unit (GRU), EEMD-LSTM, and PCA-EEMD-

LSTM. GRU is a sequence prediction network. Compared with the LSTM, it only contains update gate and reset gate, which also has a good prediction effect while simplifying the structure [34]. As the existing paper has demonstrated that the LSTM model is superior to traditional models in prediction effect, the traditional methods are not considered in this experiment. The three Laizhou wind farm datasets were used in this comparison. This study took the predicted wind power data for the following day as example, which contains 96 data points. The prediction interval is 15 minutes. To better explain the performance of the methods, we used the determination coefficient ( $R^2$ ) to test the degree of

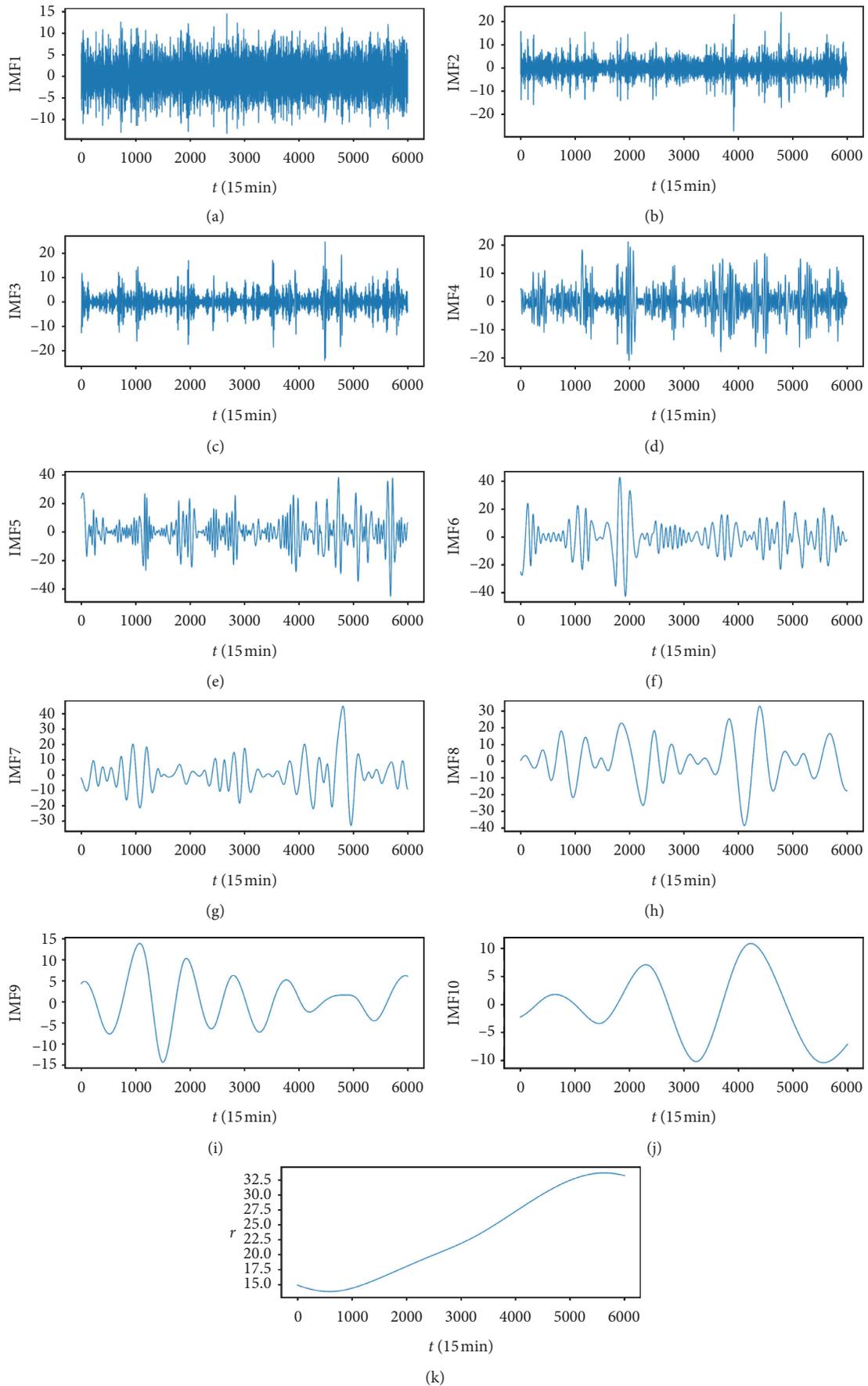


FIGURE 4: Decomposition results for the wind power sequence.

TABLE 2: Correlation coefficient between meteorological features and wind power.

Correlation	Wind power	Wind speed	Wind direction	Atmospheric pressure	Temperature	Humidity
Wind power	1.00	0.62	0.29	0.21	0.07	0.01
Wind speed	0.62	1.00	0.19	0.22	0.26	0.30

TABLE 3: Contribution rate of each PC.

Principal component	First PC	Second PC	Third PC	Fourth PC
Contribution rate	0.9687	0.0278	0.0023	0.0006

TABLE 4: Comparison of different model structures in RMSE.

Layers of LSTM	Neurons of dense	RMSE of training set	RMSE of test set
1	1 (70)	3.76	4.02
2	1 (80)	3.25	3.38
3	2 (70, 80)	3.09	4.17
4	2 (80, 80)	2.88	5.32

TABLE 5: AIC of different model parameters.

Parameters	AIC
(0, 0, 1) (2, 0, 0) <sub>12</sub>	46635.9589
(0, 0, 0) (3, 0, 0) <sub>12</sub>	46346.3836
(1, 1, 0) (2, 0, 0) <sub>12</sub>	31578.0029
(1, 1, 0) (3, 0, 0) <sub>12</sub>	31498.5598
(1, 1, 1) (3, 1, 0) <sub>12</sub>	32659.8558
(2, 0, 1) (1, 1, 0) <sub>12</sub>	33692.8103
(3, 0, 1) (2, 1, 0) <sub>12</sub>	32929.4933
(3, 0, 1) (3, 1, 1) <sub>12</sub>	28932.1469

TABLE 6: Results of  $R^2$  with the cross-validation dataset.

Indices	$R^2$
GRU	0.9748
LSTM	0.9805
EEMD-LSTM	0.9807
PCA-EEMD-LSTM	0.9955
PROPOSED METHOD	0.9961

TABLE 7: Results of  $R^2$  with the test dataset.

Indices	$R^2$
GRU	0.9724
LSTM	0.9790
EEMD-LSTM	0.9819
PCA-EEMD-LSTM	0.9921
PROPOSED METHOD	0.9933

association between the predicted and true values.  $R^2$  is defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{pre} - y_{act})^2}{\sum_{i=1}^n (y_{act} - \bar{y})^2}, \quad (29)$$

where  $y_{act}$  represents the original values,  $y_{pre}$  represents the predicted values, and  $\bar{y}$  is the average of the original values. The closer  $R^2$  is to 1, the smaller the deviation is between the predicted and true values [35]. Table 6 provides the results of  $R^2$  for the cross-validation datasets.  $R^2$  is close to 1, which shows that the value predicted by this model is very close to the true value; the method performed well during the cross-validation. Table 7 provides the results of  $R^2$  with the test dataset; here too, it can be seen that the predicted value is close to the true value.

Figures 5–7 show the prediction results for all methods for datasets 1–3. A one-day forecast curve was drawn with a total of 96 data points. Table 8 provides the prediction errors for the different methods. Table 9 presents the percentage improvement in the evaluation indices in the case of the proposed method, compared to the others.

The main purpose of this experiment was to compare the proposed method with others, thereby demonstrating its effectiveness. The minimum error value is taken as the evaluation standard, and the proposed method performed the best at wind power prediction among the four methods,

with MAPE, RMSE, and MAE values of 4.78, 3.38, and 2.86, respectively. Taking dataset 1 as example, detailed results are presented as follows:

- (i) The PCA-EEMD-LSTM combination method has a significantly higher prediction accuracy than the EEMD-LSTM combination method. Tables 7 and 8 indicate that the PCA-EEMD-LSTM combination method has MAPE, RMSE, and MAE values of 5.33, 3.4, and 3.09, respectively. Compared to the EEMD-LSTM combination method, the percentage improvements in MAPE, RMSE, and MAE were 33.87%, 18.83%, and 13.93%, respectively. Figure 5 shows the prediction results for these two methods. The experimental results show that the meteorological information is effective in improving the accuracy, and the PCA can have a satisfactory effect on reducing the redundant components and get the components with high correlation with wind power, which improves the prediction accuracy and computing speed.

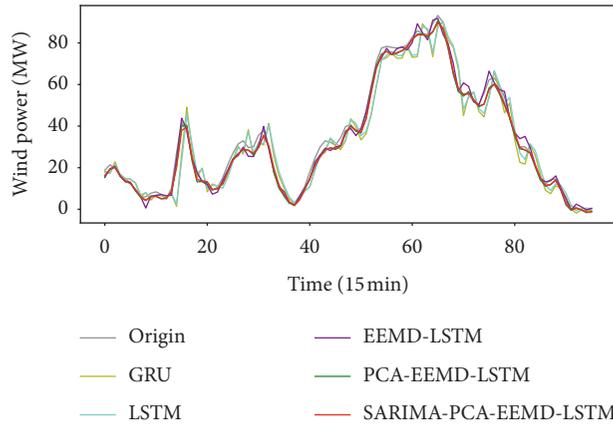


FIGURE 5: Prediction results for all methods for dataset 1.

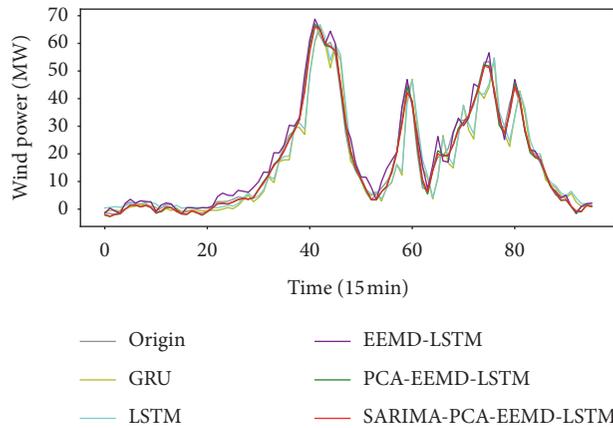


FIGURE 6: Prediction results for all methods for dataset 2.

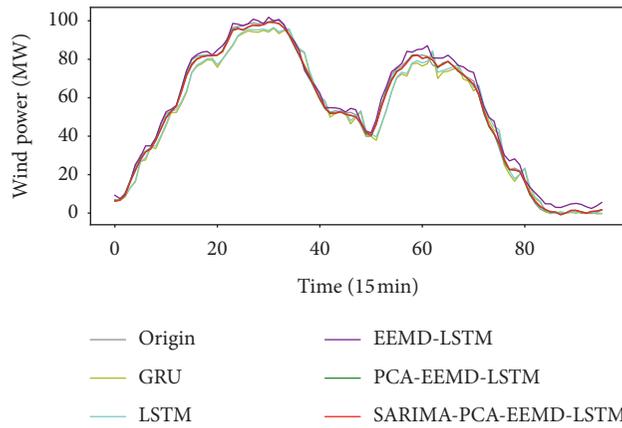


FIGURE 7: Prediction results for all methods for dataset 3.

TABLE 8: Prediction error assessment results for the different methods.

	Indices	MAPE	RMSE	MAE
Meteorological features uninvolved	GRU	13.59	5.84	3.76
	LSTM	11.74	5.67	3.75
	EEMD-LSTM	8.06	4.30	3.59
Meteorological features involved	PCA-EEMD-LSTM	5.33	3.49	3.09
	PROPOSED METHOD	4.78	3.38	2.86

TABLE 9: Percentage improvement in the proposed method compared to others.

	Indices	MAPE (%)	RMSE (%)	MAE (%)
Meteorological features uninvolved	GRU	64.82	42.12	23.93
	LSTM	59.28	40.38	23.73
	EEMD-LSTM	40.69	21.39	20.33
Meteorological features involved	PCA-EEMD-LSTM	10.31	3.15	7.44

(ii) The proposed method has a higher prediction accuracy than the PCA-EEMD-LSTM combination method. Tables 7 and 8 indicate that the proposed method has MAPE, RMSE, and MAE values of 4.78, 3.38, and 2.86, respectively. Compared to the PCA-EEMD-LSTM combination method, the percentage improvements in MAPE, RMSE, and MAE were 10.31%, 3.15%, and 7.44%, respectively. Figure 5 shows the prediction results for these two methods. The experimental results show that the SARIMA model can track the seasonality and periodicity of the wind power sequence, which improves the results predicted by the LSTM network.

## 5. Conclusions

In this study, a hybrid ultrashort-term wind power prediction method was developed. Based on the prediction using historical wind power data, the meteorological information regarded as the features of wind power is helpful to improve prediction accuracy. The PCA can be used to reduce the redundant meteorological features and only get the components with high correlation with wind power. In addition, the SARIMA can effectively extract seasonal information and be regarded as the auxiliary predictor of LSTM. The hybrid prediction model was applied to three datasets collected from a wind farm in Laizhou to verify prediction ability, and the experimental results showed that this method got great prediction accuracy. In the future, the proposed method can be applied to real-time prediction systems. Although it takes more than 3 hours to build the models on a general computer, the time required for real-time prediction using the built models is extremely short, which fully meets ultrashort-term requirements. Moreover, the future research should comprise the following:

- (i) The EEMD algorithm can be applied not only to the wind power sequence but also to meteorological features.
- (ii) The proposed method shows satisfactory results for a 15-minute prediction. We will verify predictions at 30-minute, 1-hour, or even longer intervals in the future.

## Nomenclature

$P_t$ : The original wind power series  
 $e_t$ : The sublayer of the wind power  
 $r_t$ : The redundant item  
 $f_t$ : The meteorological feature  
 $\bar{f}_t$ : The average of each feature  
 $r_{ij}$ : Eigenvalues of covariance matrix  
 $con_i$ : Contribution rate of eigenvectors

$L$ : The lag operator  
 $D$ : The seasonal difference order  
 $d$ : The regular difference order  
 $a$ : The regression model parameter  
 $s$ : The seasonality order  
 $x_i$ : The combined feature vector  
 $h_t$ : Predicted values of wind power by LSTM  
 $c_t$ : The state of the LSTM cell  
 $b$ : The offset vectors  
 $\omega$ : The ratio of final wind power  
 $Y_{pre}$ : Predicted values of wind power  
 $Y_{act}$ : True values of wind power.

## Abbreviations

EEMD: Ensemble empirical mode decomposition  
IMF: Intrinsic mode function  
PCA: Principal component analysis  
PC: Principal component  
LSTM: Long short-term memory network  
SARIMA: Seasonal autoregressive integrated moving average.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

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

## Acknowledgments

This research has been supported by the National Natural Science Foundation of China (grant nos. 61821004, 61733010, and 61803231), Department of Science and Technology of Shandong Province (grant no. 2019JZZY010901), Natural Science Foundation of Shandong Province (grant no. ZR2019ZD09), and the Young Scholars Program of Shandong University (grant no. 2016WLJH29).

## References

- [1] Y. Liu, S. Han, and Y. Hu, "Review on short-term wind power prediction," *Modern Electric Power*, vol. 5, pp. 6–11, 2017.
- [2] B. K. Sahu, "Wind energy developments and policies in China: a short review," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1393–1405, 2018.
- [3] G. Ren, J. Liu, and J. Wanet al., "Overview of wind power intermittency: impacts, measurements, and mitigation solutions," *Applied Energy*, vol. 204, pp. 47–65, 2017.

- [4] Y. Guo, P. Wang, S. Zhai, D. Hou, S. Wang, and Y. Zhou, "Ultra short-term probability prediction of wind power based on LSTM network and condition normal distribution," *Wind Energy*, vol. 23, no. 1, pp. 63–76, 2020.
- [5] A. M. Foley, P. G. Leahy, and A. Marvuglia, "Current methods and advances in forecasting of wind power generation," *Renewable Energy*, vol. 37, no. 1, pp. 1–8, 2012.
- [6] A. McKeogh and M. Uzunoglu, "A review of combined approaches for prediction of short-term wind speed and power," *Renewable and Sustainable Energy Reviews*, vol. 34, pp. 243–254, 2014.
- [7] A. Kusiak, Z. Zhang, and A. Verma, "Prediction, operations, and condition monitoring in wind energy," *Energy*, vol. 60, pp. 1–12, 2013.
- [8] L. Ma, S. Luan, C. Jiang et al., "A review on the forecasting of wind speed and generated power," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 4, pp. 915–920, 2009.
- [9] Y. T. Wu, C. Y. Lin, and T. J. Chang, "Effects of inflow turbulence intensity and turbine arrangements on the power generation efficiency of large wind farms," *Wind Energy*, vol. 23, no. 7, pp. 1640–1655, 2020.
- [10] Y.-T. Wu, T.-L. Liao, C.-K. Chen, C.-Y. Lin, and P.-W. Chen, "Power output efficiency in large wind farms with different hub heights and configurations," *Renewable Energy*, vol. 132, pp. 941–949, 2019.
- [11] L. Dong, L. Wang, S. F. Khahro, and X. Liao, "Wind power day-ahead prediction with cluster analysis of NWP," *Renewable and Sustainable Energy Reviews*, vol. 60, pp. 1206–1212, 2016.
- [12] M. Cui, D. Ke, and D. Gan, "Statistical scenarios forecasting method for wind power ramp events using modified neural networks," *Journal of Modern Power Systems and Clean Energy*, vol. 3, no. 3, pp. 371–380, 2015.
- [13] A. Sun, Q. Tan, X. Yuan et al., "Short-term wind speed prediction using EEMD-LSSVM model," *Advances in Meteorology*, vol. 2017, Article ID 6856139, 2017.
- [14] M. Yang, X. Chen, and J. Du, "Ultra-short-term multistep wind power prediction based on improved EMD and reconstruction method using run-length Analysis," *IEEE Access*, vol. 6, pp. 31908–31917, 2018.
- [15] F. U. Cui, "A new hybrid approach for wind speed prediction using fast block least mean square algorithm and artificial neural network," *Mathematical Problems in Engineering*, vol. 2016, Article ID 8395751, 2016.
- [16] Z. Sun, H. Sun, and J. Zhang, "Multistep wind speed and wind power prediction based on a predictive deep belief network and an optimized random forest," *Mathematical Problems in Engineering*, vol. 2018, Article ID 6231745, 2018.
- [17] Y.-Y. Hong, T.-H. Yu, and C.-Y. Liu, "Hour-ahead wind speed and power forecasting using empirical mode decomposition," *Energies*, vol. 6, no. 12, pp. 6137–6152, 2013.
- [18] Y. Cui, L. Guan, C. Hou et al., "Wind power short-term prediction based on LSTM and discrete wavelet transform," *Applied Sciences*, vol. 9, pp. 9–15, 2019.
- [19] R. Yu, J. Gao, M. Yu et al., "LSTM-EFG for wind power forecasting based on sequential correlation features," *Future Generation Computer Systems*, vol. 93, pp. 33–42, 2019.
- [20] L. Lu, H. Jing, R. Zhang et al., "Wind power forecast based on improved Long Short Term Memory network," *Energy*, vol. 189, Article ID 116300, 2019.
- [21] D. B. Alencar, C. M. Affonso, and R. C. L. Oliveira, "Hybrid approach combining SARIMA and neural networks for multistep ahead wind speed forecasting in Brazil," *IEEE Access*, vol. 6, pp. 55986–55994, 2018.
- [22] S.-W. Filho, "A hybrid model of EMD and multiple-kernel RVR algorithm for wind speed prediction," *International Journal of Electrical Power & Energy Systems*, vol. 78, pp. 910–915, 2016.
- [23] X. An, D. Jiang, and M. Zhao, "Short-term prediction of wind power using EMD and chaotic theory," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 2, pp. 1036–1042, 2012.
- [24] C. Liu, H. Zhang, and W. Fan, "A new chaotic time series hybrid prediction method of wind power based on EEMD-SE and full-parameters continued fraction," *Energy*, vol. 138, pp. 977–990, 2017.
- [25] H. Ma and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433–459, 2010.
- [26] C. Skittides and W.-G. Früh, "Wind forecasting using principal component analysis," *Renewable Energy*, vol. 69, pp. 365–374, 2014.
- [27] M. Khan, T. Liu, and F. Ullah, "A new hybrid approach to forecast wind power for large scale wind turbine data using deep learning with TensorFlow framework and principal component analysis," *Energies*, vol. 12, no. 12, p. 2229, 2019.
- [28] A. Bosisio, A. Berizzi, D.-D. Le, and G. Giannuzzi, "Improving DTR assessment by means of PCA applied to wind data," *Electric Power Systems Research*, vol. 172, pp. 193–200, 2019.
- [29] G. Bassi, C. Dikaiakos, and A. Dramountanis, "Analysis and modeling for short- to medium-term load forecasting using a hybrid manifold learning principal component model and Comparison with classical statistical models (SARIMAX, exponential smoothing) and artificial intelligence models (ANN, SVM): the case of Greek electricity market," *Energies*, vol. 9, no. 8, p. 635, 2016.
- [30] V. Papaioannou and N. M. Pindoriya, "A SARIMA-RVFL hybrid model assisted by wavelet decomposition for very short-term solar PV power generation forecast," *Renewable Energy*, vol. 140, pp. 124–139, 2019.
- [31] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [32] J. Wang, W. Yang, and P. Du, "A novel hybrid forecasting system of wind speed based on a newly developed multi-objective sine cosine algorithm," *Energy Conversion and Management*, vol. 163, pp. 134–150, 2018.
- [33] T. Niu and R. Lahdelma, "Evaluation of a multiple linear regression model and SARIMA model in forecasting heat demand for district heating system," *Applied Energy*, vol. 179, pp. 544–552, 2016.
- [34] R. Wang, C. Li, W. Fu et al., "Deep learning method based on gated recurrent unit and variational mode decomposition for short-term wind power interval prediction," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 17, 2016.
- [35] J. Cao, Z. Li, and J. Li, "Financial time series forecasting model based on CEEMDAN and LSTM," *Physica A: Statistical Mechanics and Its Applications*, vol. 519, pp. 127–139, 2019.