An ultra-short-term wind speed forecasting model based on time scale recognition and dynamic adaptive modeling

An Ultra-short-term Wind Speed Forecasting Model Based on Time Scale Recognition and Dynamic Adaptive Modeling

Zhao Zhen [1,2,*], Gang Qiu [3], Shengwei Mei [1], Fei Wang [2,4,5], Xuemin Zhang [1], Rui Yin[6], Yu Li [3], Gerardo J. Osório [7], Miadreza Shafie-khah [8], João P.S. Catalão [9]

[1] State Key Lab of Power System, Department of Electrical Engineering, Tsinghua University, Beijing 100084, China
[2] Department of Electrical Engineering, North China Electric Power University, Baoding 071003, China
[3] State Grid Xinjiang Electric Power Co., Ltd, Urumqi 830018, China
[4] State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Beijing 102206, China
[5] Hebei Key Laboratory of Distributed Energy Storage and Micro-grid (North China Electric Power University), Baoding 071003, China
[6] State Grid Hebei Electric Power Co., Ltd, Shijiazhuang 050021, China
[7] Portucalense University Infante D. Henrique (UPT), 4200-072, Porto, Portugal
[8] School of Technology and Innovations, University of Vaasa, 65200 Vaasa, Finland
[9] Faculty of Engineering of the University of Porto and INESC TEC, 4200-465 Porto, Portugal
* Correspondence: georgiazhz@foxmail.com

Abstract: The forecast of wind speed is prerequisite for wind power prediction, which is one of the most effective means of promoting wind power absorption. However, when modeling for wind speed sequences with different fluctuations, most existing researches ignore the influence of time scale of wind speed fluctuation period, let alone the low compatibility between training and testing samples that severely limit the training performance of forecasting model. To improve the accuracy of wind speed and wind power forecasting, an ultra-short-term wind speed forecasting model based on time scale recognition and dynamic adaptive modeling is proposed in this paper. First, a series of wind processes are divided from the historical wind speed sequence according to the natural variation characteristics of wind speed. Second, we divide all the wind processes into two patterns based on their time scale, and an SVC model with input features extracted from meteorological data is built to identify the time scale of the current wind process. Third, for a specifically identified wind process, the complex network algorithm is applied in data screening to select high compatible training samples to train the forecast model dynamically for current input. Simulation indicates that the proposed approach presents higher accuracy than benchmark models using the same forecasting algorithms but without considering the time scale and data screening.

Keywords: Wind Speed Forecast; Wind Process; Time Scale Distribution Function; Pattern Recognition; Complex Network

1. Introduction

Wind power, solar photovoltaic power, and hydropower account for the majority of renewable energy generation [1]-[4]. Among which, wind power has developed rapidly in recent years. According to the data released by the Global Wind Energy Council (GWEC), the total installed capacity of wind power in the world has reached 591GW by 2018, nearly doubled compared with 6 years ago, and increased by 6 times compared with ten years ago [5].

However, with the increase of wind power penetration, its intermittency, randomness, and volatility bring significant challenges to the power quality of the power system [6], which

caused voltage fluctuations, frequency fluctuations, harmonics, and other issues [7-9]. It also affects the stability and reliability of the power system, which is the primary concern in the planning activities [10-12].

Wind power prediction can not only solve the above problems but also facilitate the implementation of demand response program [13,14], microgrid operation [15,16], coordinating power generation side and load side [17,18] and improving the competitiveness of new energy generators in the electricity market [19,20]. Irregular wind speed is the leading cause of fluctuations in wind power output, and its changes are affected by multiple meteorological factors and geographical locations. Wind speed forecasting is the basis of wind power prediction, whose accuracy has impacts on the efficiency of wind energy utilization. Improving the accuracy of wind speed forecasting is of great significance for power system dispatching [21-23].

According to the prediction time scale, wind speed forecasting can be classified into the ultra-short-term forecast, short-term forecast, medium-term forecast, and long-term forecast. Ultra-short-term forecasting refers to forecasts within the next four hours [24,25]. The time scale of short-term forecasting can be up to several days [26,27]. At present, the forecasting models mostly belong to statistical models.

Statistical models establish the mapping relationship between input data and output data based on historical data. Examples are Autoregressive Integrated Moving Average (ARIMA) [28], artificial neural network (ANN) [29,30], Support Vector Machines (SVM) [31,32], least-squares support vector machines (LSSVM) [33,34], and Extreme Learning Machine (ELM) [35,36]. Due to the limitations of a single model, a large amount of literature uses hybrid models combining multiple methods to improve forecasting accuracy [37-39].

For statistical models, the instability and dependence on data are an obvious limitation that needs to be addressed. Due to the multiple wind speed fluctuation patterns, the selection of training data has a significant influence on forecasting performance. Therefore, the classification of the data and the similarity screening of the training data for current model input can be adapted to reduce the redundancy of the training data and improve the forecasting accuracy. In [40], k-means is used to cluster the annual wind speeds by different time scales: one day and half a month. Based on the classified data, the SVM model is established, respectively.

The experimental results show that the forecasting accuracy is improved. In [41], the clustering similar measure function combines the Euclidean distance and Angle Cosine aiming to identify similar wind speed days to the predicting days which are treated as training samples. In [42], based on the classification of weather types, the forecasting accuracy of day-ahead photovoltaic power generation has been significantly improved.

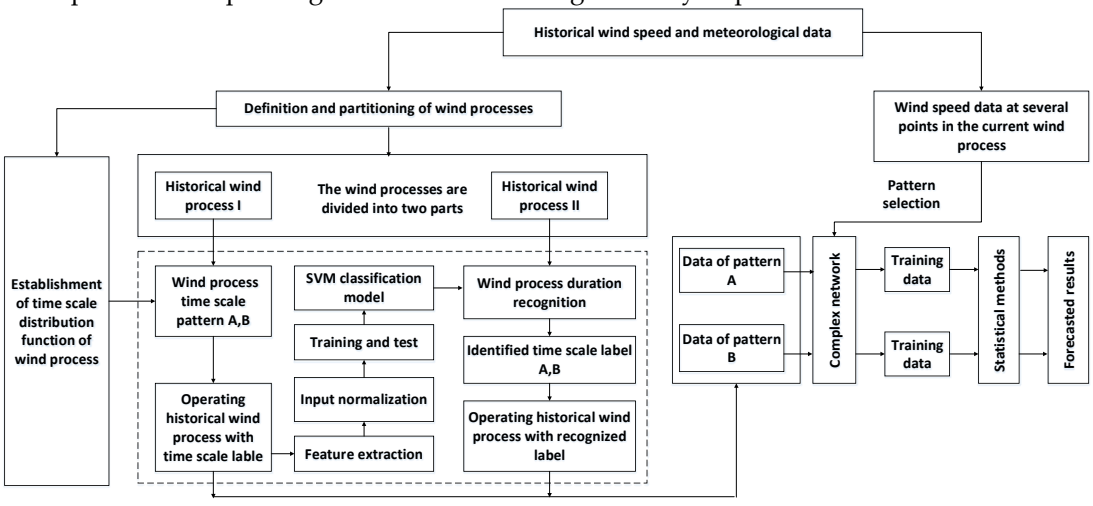


Figure 1. The detailed framework of the proposed approach.

It can be seen that wind power or wind speed forecast researches have acquired several achievements and successes. However, there are still problems and challenges to be considered.

First, the wind speed variation of wind processes with different durations usually has certain regularity, e.g. the wind speed of wind processes with short duration is likely to show a downward trend within the forecasted time, and that with long duration is likely to show a continuous fluctuation trend within the forecasted time. But most researches haven't considered the influence of wind process duration on wind speed fluctuation law.

Second, when choosing the training data for established forecast model, only the similarity between the actual forecast input and the training set input is considered, which does not guarantee the similarity of the corresponding outputs and the accuracy of mapping relationship. Therefore, in this paper, a novel ultra-short-term wind speed forecasting model is proposed to deal with the above two issues. The detailed framework of the proposed model is shown in Figure 1.

Compared with the existing researches, the contributions of this paper are summarized as follows:

- Explore the relationship between wind speed fluctuation characteristics and the time scale of current wind processes.
- An ultra-short-term wind speed forecasting approach which utilizes the time scale information of current wind process is proposed.
- The complex network is adopted to deeply mining the morphological characteristics of wind speed sequence, thus the optimal compatible training data according to current forecast input is dynamically screened out.

The structure of the paper is as follows: Section 2 introduces the definition of the wind processes and establishes the time scale distribution function of the wind processes. Section 3 describes the extracted wind speed features and the SVC algorithm. The data screening method based on the complex network is introduced in Section 4. Section 5 verifies the performance of the proposed method. Section 6 concludes this paper.

2. Definition and time scale distribution function of wind process

2.1. Definition and division of wind process

The formation of wind is the result of horizontal airflow, which is caused by solar radiant heat. The magnitude of the wind is the wind speed [43]. Because the surface wind is affected by solar radiation, geostrophic force, ocean and topography, the spatial and temporal distribution of wind speed is more complicated [44], and the variation law of wind speed is weak, unlike most daily irradiances which are from minimum values begin to increase, increases to the maximum value, and then decreases, and finally decreases to the minimum value, and the variation law is obvious, which is conducive to speculating the variation trend of the future. Inspired by the variation law of daily irradiance, this paper finds the internal law from the wind speed sequence itself and finds that the wind speed contains many processes of wind rising and falling. The variation trend of each process is similar to that of daily irradiance.

The change of wind speed reflects a very complex atmospheric physical process, to identify the wind fluctuation pattern, wind process division is very necessary as it can transform the original study of wind speed into the study of wind process sequence [45, 46]. At present, the definition of wind process can be mainly divided into three categories. The first is based on mathematical methods, such as references [47] using extreme points to define, that is, the wind speed "wave" between two adjacent minimum points is a wind process. And in [48], the swinging door algorithm based on mathematical methods is used. The second is based on wind speed time series with a fixed duration. Reference [49] proposes that the wind speed time series data are separated into different segments with each representing a period of 24-hours. Similarly, reference [50] selects wind speed series based on a 10min sliding window from historical data as

wind processes. The third is based on physical meaning. The research in [51] considers the physical significance of wind speed changing. The wind processes are divided and modeled based on the correlation between wind mast data and numerical weather prediction data.

In many papers, the purpose of wind process division and pattern recognition is to apply in wind speed and power prediction. Establishing different prediction models according to different pattern recognition results to improve the prediction accuracy, such as studies in [49-51]. In addition, reference [48] uses the divided wind processes to identify wind speed ramp events.

Defining and dividing the wind processes according to the time axis, it is necessary to consider the time scale parameters. Typical time scale parameters include two types: zero scale parameter and extreme point scale parameter. For any sequence, the zero-scale parameter refers to the span of two adjacent zeros in time, and the extreme-point scale parameter refers to the span of two adjacent extreme points in time. For the wind speed sequence, there may be no zero-crossing points. Therefore, in this paper, wind processes are extracted from wind speed sequence by using the extreme point scale parameter. From a minimum value below the threshold (which is derived from local years of historical wind speed data statistics) to another adjacent minimum value below the threshold, the change between these two minimum values is defined as a wind process (Figure. 2). The formula of the specific mathematical model of wind process is given as:

$$S(w_j)=\begin{cases} w_j \in \{w\}, j=1,\cdots,n \\ w_j \in [\varepsilon, w_{\max}] \\ w_1, w_n \in [0,\varepsilon] \\ w_1, w_n \in \{w_{\min}\} \end{cases} \tag{1}$$

where $\{w\}$ is the wind speed sequence, $\{w_{\min}\}$ is the local minimum sequence of wind speed, $w_{\max}$ is the maximum of wind speed, ε is the threshold.

2.2. Wind process time scale distribution

Due to the influence of meteorological factors, the duration of each wind processes is different. Wind processes with different durations have different variation trends. Figure 3 shows the fluctuations of two wind processes with different durations. The one on the left is a small scale wind process with a duration of about 90 minutes, and the one on the right is a large scale wind process with a duration of about 3 hours.

The part circled by the red line in the two figures represents the variation of wind speed from 30 minutes after the beginning of the wind process to the next hour. It can be seen that the variation of wind speed in the small scale wind process presents a downward trend, while the variation of wind speed in the large scale wind process presents a continuous fluctuation and upward trend, indicating that the variation trends of wind speed are different in different durations. Therefore, we need to study the time scale distribution of the wind processes which is used to determine appropriate classification criteria.

By classifying wind processes according to the duration scale, wind processes with similar variation trend can be classified into one category, which can reduce the redundancy of training data for forecasting and improve the accuracy of mapping relationship and the forecasting.

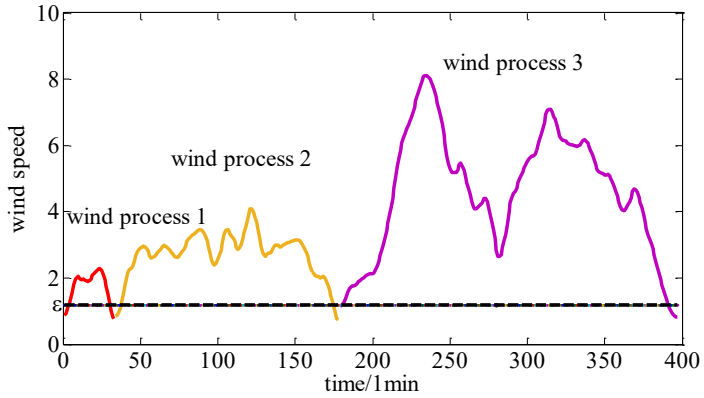


Figure 2. Wind process partitioning results.

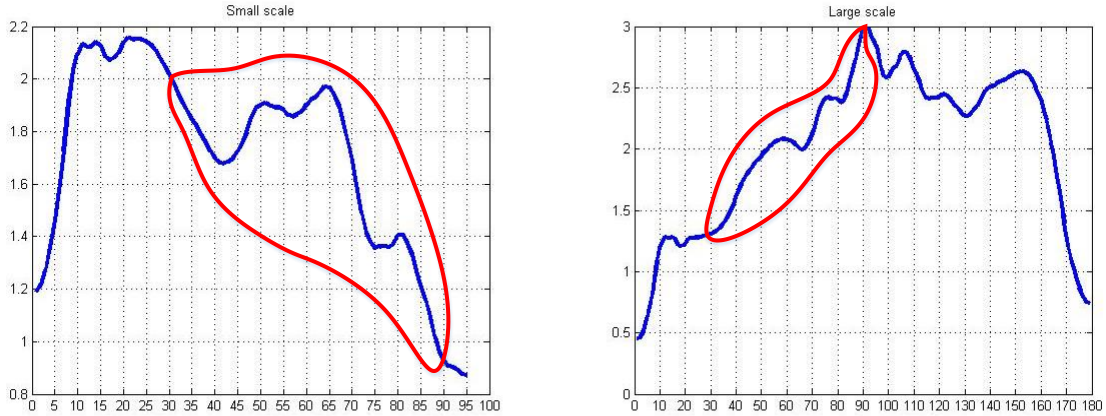


Figure 3. Wind processes of different scales.

The wind speed data of the SXF meteorological station (Sioux Falls, South Dakota, USA), the DRA meteorological station (Desert Rock, Nevada, USA) and the PSU meteorological station (Penn State, Pennsylvania, USA) are selected to analyze the time scale distribution of the wind processes. The time resolution is 1 minute.

Based on the definition in section II-*A*, all the wind processes are extracted from the wind speed sequence of the three meteorological stations, and the time scale distribution diagrams of the SXF, DRA and PSU meteorological stations with a time interval of 15 minutes are shown in Figure. 4(a), Figure. 4(b), and Figure. 4(c), respectively.

From Figure. 4, it can be seen that the wind process time scale distributions of three meteorological stations are similar. The largest number of wind processes occurs within the range of 30~45 minutes for all the three meteorological stations, and when the time scale is less than 45 minutes, the number of wind processes of the three stations increases with the increase of the scale.

When the time scale is greater than 45 minutes, the number of wind processes decreases with the increase of the scale. Through further observation, it is found that the wind processes distributions of the three stations are all similar to the Gamma distribution, which is a standard asymmetric distribution with probability density function shown below:

$$f(\mathrm{x}) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \tag{2}$$

where α is the shape parameter, β is the scale parameter.

The expressions of mean value and variance of Gamma distribution are shown in equation (3) and equation (4) respectively

$$\mu = \alpha\beta \tag{3}$$

$$\sigma^2 = \alpha\beta^2 \tag{4}$$

where μ represents mean value, σ^2 represents variance.

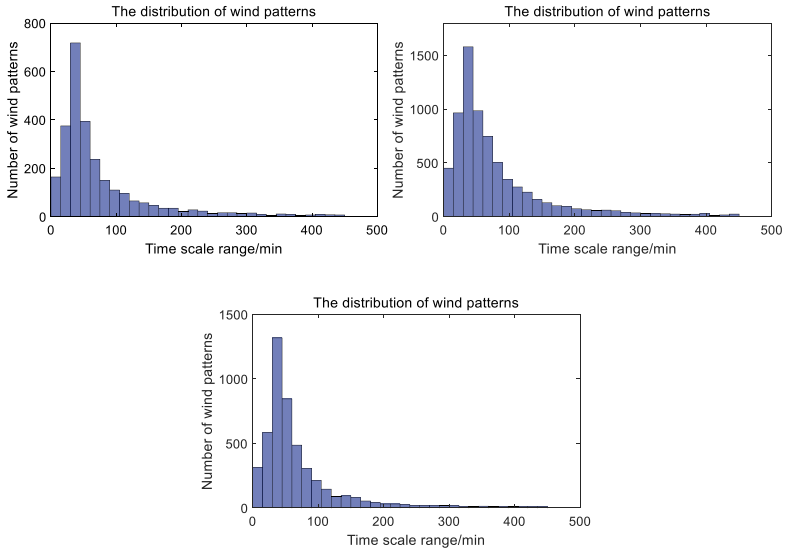
The k-s test is used to test whether a single sample obeys a pre-assumed specific distribution, and the quantitative test problem is transformed into a hypothesis testing problem [52]. That is:

$$\text{Null hypothesis } H_0 : F(x) = F_n(x)$$

$$\text{Opposite hypothesis } H_1 : F(x) \neq F_n(x)$$

where: $F_n(x)$ is pre-assumed particular distribution; $F(x)$ is the cumulative probability (frequency) function of the sample, and n is the sample size. Take a significant level α=0.05, $D_{n,\alpha}$ is the k-s statistical critical value of whether the null hypothesis is accepted under the confidence interval, D is the actual statistical value of k-s, P is the probability for accepting the null hypothesis. If $D < D_{n,\alpha}$, accept H_0, that is H=0, which implies that the time scale distribution of wind process obeys Gamma distribution. Conversely, reject H_0, that is H=1.

In this paper, it is first assumed that the time scale distribution of wind process obeys the Gamma distribution, and then use the k-s test to verify whether the hypothesis is true. If the hypothesis is true, we use the curve fitting method to obtain three new gamma curves approximating the real values.



(a)SXF (b)DRA (c)PSU

Figure 4. The wind process time scale distribution of three meteorological stations.

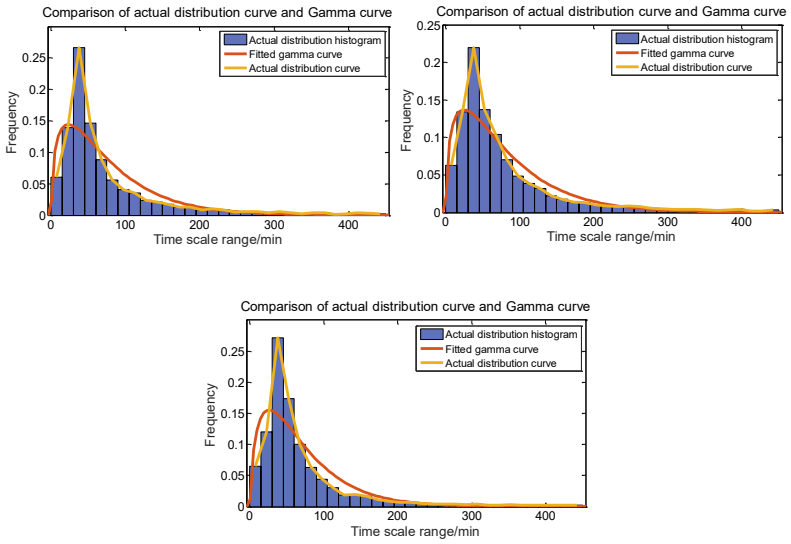
Table 1 is the K-S test results of time scale distribution of wind process. From Table 1, the D values of the three meteorological stations are 0.20, 0.21, and 0.22, respectively, all less than the corresponding values of $D_{n,a}$. Therefore, all the H values of the three meteorological stations are 0, which quantitatively verifies that the time scale distributions of the three meteorological stations can be effectively modeled as the Gamma distribution. Then, the gamma curves of the three meteorological stations are fitted by curve fitting, and the result is shown in Figure. 5.

The solid yellow lines in Figure. 5(a), Figure. 5(b), and Figure. 5(c) are the fitted gamma curves of SXF weather station, DRA weather station, and PSU weather station, respectively, and the parameters of three fitting gamma curves are shown in Table 2. The red dotted lines in Figure. 5(a), Figure. 5(b), and Figure. 5(c) represent the corresponding actual curves. It could be seen that the trend of these two kinds of curves is basically the same.

According to the parameters of the fitted gamma curves in Table 2, the cumulative probability distribution diagrams of SXF, DRA, and PSU weather stations are obtained, as shown in Figure. 6(a), Figure. 6(b), and Figure. 6(c), respectively.

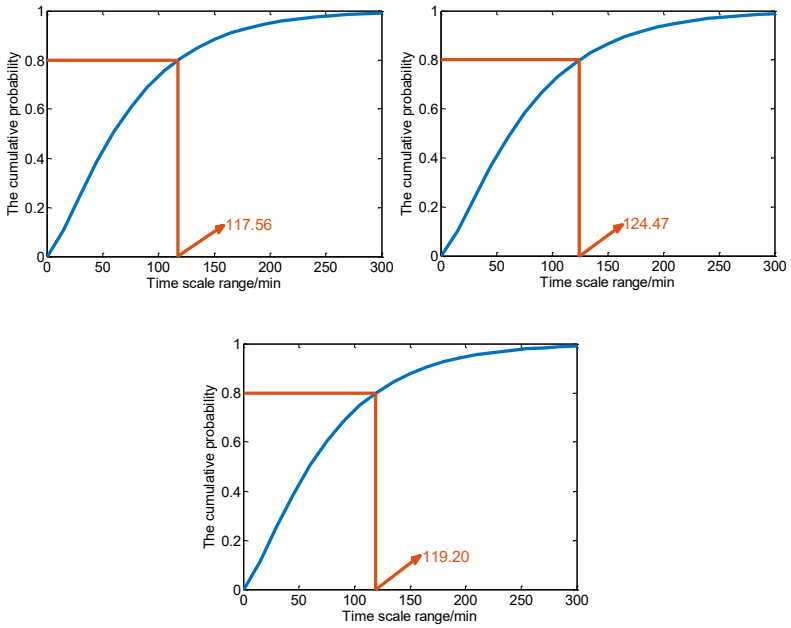
When the cumulative probability of the three weather stations reaches 0.8, the corresponding time scales are respectively 117.56, 124.47 and 119.20, all located in around 120 minutes, which indicates that the duration of wind processes is mostly less than 120 minutes, and only a few (about 20%) are more than 120 minutes.

Here, considering the general duration of wind processes, this paper divides wind processes into two patterns with a boundary of 120 minutes. One is the common short-duration wind pattern whose duration is less than 120 minutes, labeled as A; the other is the uncommon long-duration wind pattern whose duration is more than 120 minutes, labeled as B.



(a)SXF (b)DRA (c)PSU

Figure 5. The wind process time scale distribution function of three meteorological stations.



(a)SXF (b)DRA (c)PSU

Figure 6. The Cumulative probability distribution of three meteorological stations.

Table 1. K-S test results.

Parameter	H	P	$D_{n,a}$	D
SXF	0	0.16	0.242	0.20
DRA	0	0.07	0.224	0.21
PSU	0	0.10	0.250	0.22

Table 2. The parameters of the fitted gamma curves.

Name	α	β	μ	σ^2
SXF	1.45	52.53	76.17	4001.13
DRA	1.44	55.62	80.09	4454.76
PSU	1.40	54.62	76.47	4176.68

3. Wind pattern identification

In this section, it is applied the feature extraction process and the support vector classification (SVC) model to realize the time scale identification of wind patterns.

3.1. Feature extraction

According to the time scale distribution of wind processes and considering the equilibrium of samples, this paper divides the wind processes into two patterns based on their durations. The time scale of one category is less than 2 hours, labeled as A, and the time scale of the other is greater than 2 hours, labeled as B. For different classification cases, there is no universal solution for feature extraction.

Most researches generally calculate and analyze the statistical features of data itself of different categories and extract statistical features with large differences for pattern recognition [53-55]. In addition to the calculation and analysis of the statistical features of the wind speed of different categories, this paper also considers the influence of historical meteorological factors on the duration of the wind process. Therefore, five features describing wind speed

variation and nine features describing meteorological changes are extracted for recognition in this paper.

3.1.1. Maximum of wind speed

The maximum wind speed is different for wind processes with different durations. For wind processes in category A, the duration is short and the maximum wind speed is generally small. However, for wind processes in category B, the duration is long and the maximum wind speed is generally large. Feature one is the maximum of several points at the beginning of wind processes which is expressed as $v_{\max}$, defined in formula (5):

$$F1 = v_{\max} = \max_{i=1,\cdots,k} \{v_i\} \tag{5}$$

where v_i represents the ith value of the current wind process.

3.1.2. Average wind speed

In some cases, the last feature is invalid. For example, the wind process lasts for a short time, but its maximum value is large. Therefore, in order to further improve the accuracy of recognition, feature two is extracted, i.e. the average of several points at the beginning of wind processes which is expressed as v_{ave}, defined in formula (6). For wind processes with short duration, the average value is generally low, while for wind processes with long duration, the average value is high.

$$F2 = v_{ave} = \sum_{i=1}^{k} v_i \Big/ k \tag{6}$$

3.1.3. The variance of wind speed

The wind speed fluctuation is different for wind processes with different durations. The variance can reflect the fluctuation level to a certain extent. For the wind process in category A, the duration is short, the wind speed fluctuation is small, and the variance is small. However, for the wind process in category B, the duration is long, the wind speed fluctuates greatly, and the variance is large. Feature three is the variance of several points at the beginning of wind processes which is expressed as v_s, defined in formula (7):

$$F3 = v_s = \frac{1}{k} \sum_{i=1}^{k} (v_i - v_{ave})^2 \tag{7}$$

3.1.4. The ramp rate of wind speed

For a wind process with short duration, the time from the beginning to the first local maximum wind speed is shorter and the ramp rate is higher, while for a wind process with long duration, the time from the beginning to the first local maximum wind speed is longer and the ramp rate is lower. Feature four is the ramp rate of several points at the beginning of wind processes which is denoted by v_r, defined in formula (8):

$$F4 = v_r = (v_{\max} - v_1) \Big/ (t_{\max} - t_1) \tag{8}$$

where $t_{\max}$ represents the time when the maximum wind speed occurs and t_1 represents the initial time instant.

3.1.5. Tail variation of wind speed

Within the same and relatively long time window, the wind process with short duration, from the maximum wind speed to the end of the time window, the wind speed value drops more, and the absolute value is bigger. While the wind process with long duration, from the maximum wind speed to the end of the time window, the wind speed value drops less, the absolute value is smaller. Feature five is the tail variation of several points at the beginning of wind processes, which is denoted by v_{re}, defined in formula (9):

$$F5 = v_{re} = {v_k - v_{\max}} \Big/ {t_k - t_{\max}} \tag{9}$$

3.1.6. Features of meteorological change

Considering the influence of temperature, humidity, and pressure on the time scale of wind processes, this paper extracts the statistical features of these three meteorological factors as feature six (F6) to describe historical meteorological variation. The average of historical temperature is denoted by T_{ave}, defined in formula (10). The variation trend of historical temperature is denoted by T_r, defined in formula (11). The variance of historical temperature is denoted by T_s, defined in formula (12).

$$T_{ave} = {\sum_{j=1}^{n} T_j} \Big/ {n} \tag{10}$$

$$T_r = \sum_{j=2}^{2} (T_j - T_{j-1}) \tag{11}$$

$$T_s = \frac{1}{n} \sum_{j=1}^{n} (T_j - T_{ave})^2 \tag{12}$$

The average of historical humidity is denoted by H_{ave}, defined in formula (13). The variation trend of historical humidity is denoted by H_r, defined in formula (14). The variance of historical humidity is denoted by H_s, defined in formula (15).

$$H_{ave} = {\sum_{j=1}^{n} H_j} \Big/ {n} \tag{13}$$

$$H_r = \sum_{j=2}^{2} (H_j - H_{j-1}) \tag{14}$$

$$H_s = \frac{1}{n} \sum_{j=1}^{n} (H_j - H_{ave})^2 \tag{15}$$

The average of historical air pressure is denoted by P_{ave}, defined in formula (16). The variation trend of historical air pressure is denoted by P_r, defined in formula (17). The variance of historical air pressure is denoted by P_s, defined in formula (18):

$$P_{ave} = {\sum_{j=1}^{n} P_j} \Big/ {n} \tag{16}$$

$$P_r = \sum_{j=2}^{2} (P_j - P_{j-1}) \tag{17}$$

$$P_s = \frac{1}{n} \sum_{j=1}^{n} (P_j - P_{ave})^2 \tag{18}$$

where n is the number of samples, T_j, H_j, P_j represent the temperature, humidity, and air pressure of the jth point respectively.

3.2. Support vector classification (SVC)

Given its excellent performance in many applications, SVC is adopted in this paper [56,57]. The traditional SVC is a two-classifier that divides objects into two categories that result in the minimum of generalization errors. SVC consists of two types of models: linear separable SVC, and nonlinear separable SVC. Figure. 7 illustrates linear separable SVC, where $wx + b = 0$ is a hyperplane, x is the input, and w, b are two parameters of the hyperplane.

SVC could be converted into the corresponding dual form by Lagrangian function. Suppose that $\alpha^* = \left[a_1^*, a_2^*, \cdots, a_m^*\right]$ is the solution to the dual form, then, the separation functions of linear separable SVC and nonlinear separable SVC can be expressed by formula (19) and (20), respectively.

$$g(x) = \operatorname{sgn}\left[\sum_{i=1}^{m} a_i^* y_i (xx_i) + b^*\right] \tag{19}$$

$$g(x) = \operatorname{sgn}\left[\sum_{i=1}^{m} a_i^* y_i K(x_i, x) + b^*\right] \tag{20}$$

where y_i is the output of training data, m is the dimension of the input vector and b^* is the optimal value.

Then, using the extracted features as model input, and the actual wind pattern labels as output, the SVC based wind pattern identification model can be built and trained.

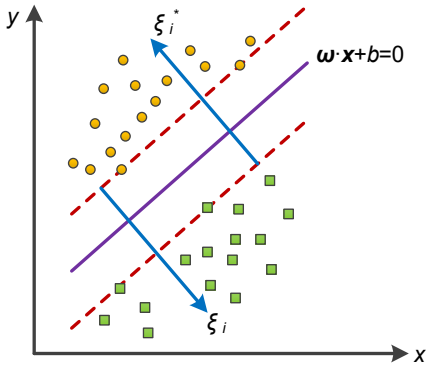


Figure 7. The schematic diagram of SVC.

3.3. Evaluation index of classification model

M is defined in equation (21) and includes all results produced by a classification model.

$$M=\begin{bmatrix} m_{11} & \cdots & m_{1n} \\ \cdots & \cdots & \cdots \\ m_{n1} & \cdots & m_{n1} \end{bmatrix} \quad i,j=1,\cdots,n \tag{21}$$

where m_{ij} means that the actual label of the object is i but the identified label is j, n is the number of total categories.

The product's accuracy (PA), the user's accuracy (UA) and the overall accuracy (OA) are three kinds evaluation indexes commonly used to evaluate the performance of classification model, defined in formula (22) to (24).

$$PA_i=\frac{m_{ii}}{\sum_{j=1}^{n} m_{ij}} \quad i=1,\cdots,n \tag{22}$$

$$UA_i=\frac{m_{ii}}{\sum_{j=1}^{n} m_{ji}} \quad i=1,\cdots,n \tag{23}$$

$$OA=\frac{\sum_{i=1}^{n} m_{ii}}{\sum_{j=1}^{n}\sum_{i=1}^{n} m_{ji}} \quad i=1,\cdots,n \tag{24}$$

where PA_i, UA_i and OA_i indicate respectively PA, UA and OA of i category.

4. Data screening method based on complex network

Based on the identification of the duration of the current wind process, the screening processing of the training data is further carried out. The selection of training data and the establishment of its mapping relationship determine the accuracy of the forecasting. Assuming that the screened training data are completely consistent with the test data, the mapping relationship established will be very accurate and the forecasting accuracy will be very high.

However, in the actual prediction, the case where the training data are the same as the test data does not exist, but the training data matching with the test data can be selected as much as possible by the data screening method to improve the forecasting accuracy.

Commonly used data screening methods include Euclidean distance (ED), Pearson correlation coefficient (PCC), Cosine similarity (CS) and so on, but all of them have defects. Taking Euclidean distance as an example, there are two curves A and B in Figure. 8, it can be seen that the variation trends of curves A and B are the same, but the Euclidean distance is large.

Suppose A is the input of the test set, B is the output of the test set, and the output of the training set screened by A based on Euclidean distance may be significantly different from B, resulting in inaccurate mapping relationship and poor forecasting accuracy. Similarly, the same problems exist in other data screening methods.

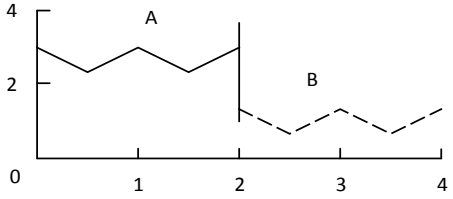


Figure 8. An example of data screening.

In recent years, time series analysis methods based on complex networks have been widely used. Complex network (CN) is an abstract description of a complex system, which can map nonlinear time series to network and mine the correlation between data by analyzing its topology.

In this paper, in order to mine the fluctuation characteristics of wind speed series and select the training data with the same fluctuation as the data to be predicted, the phase space coarse granulation algorithm is adopted to map the wind speed into the corresponding fluctuation network [58]. Suppose there is a time series $x_t, t = 1, 2, 3, \cdots N$, its fluctuation sequence is shown in equation (25).

$$f(t) = x(t) - x(t-1) \tag{25}$$

Further symbolizing $f(t)$, the formula is (26):

$$ls_i = \begin{cases} r, & f(t) > 0 \\ s, & f(t) = 0 \\ d, & f(t) < 0 \end{cases} \tag{26}$$

where, r indicates the rising state, s indicates the steady state, d indicates the falling state.

Through the above process, time series $x_t, t = 1, 2, 3, \cdots N$ is transformed into a symbol series $LS = \{ls_1, ls_2, ls_3, \cdots\}$. The symbol sequence is segmented by a time window of length l, and each segment represents the fluctuation characteristics of the network node. Then the network nodes are connected in order of time, and the weight of the edge is the number of conversions. In this paper, the length of the time window is set to the first k minutes of the current wind process.

Based on the establishment of wind speed fluctuation network, the training data with similar fluctuation and morphological characteristics as the current wind process can be screened out. In order to evaluate the performance of four data screening methods: Complex network, Euclidean distance, Pearson correlation coefficient, and Cosine similarity, Relative matching rate is defined here.

Each data screening method screens out the training data respectively according to the input and output of test set, and the ratio of the number of matching samples in the selected training data to the total number of matching samples selected by the four methods is the Relative matching rate (RMR). The formula is as shown in equation (27). The higher the RMR means that compared with other methods, the more samples matching the test set selected by this method.

$$RMR_i = \frac{len(inter(N_{ik}, N_{ij}))}{\sum_{i=1}^{4} len(inter(N_{ik}, N_{ij}))} (i = 1, 2, 3, 4) \tag{27}$$

where, $len(\cdot)$ represents the function finding the number, $inter(\cdot)$ represents the function finding the intersection. N_{ik}, N_{ij} are the training samples screened by ith method according to the input and output of the test data separately.

5. Case study and simulation

5.1. Data source

In this paper, meteorological data from DRA meteorological station (Desert Rock, Nevada, USA) are employed for the case study. The dataset consists of three years' data from January 1st, 2011 to December 31th, 2013 with 1 observation value every minute. All the simulations are implemented on MATLAB (version R2014a).

5.2. Evaluation index

Three metrics are used to evaluate the forecasting accuracy of the proposed method: the mean absolute percentage error (MAPE), the root mean square error (RMSE), and the mean absolute error (MAE). Their definitions are as follows:

$$e_{MAPE} = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{y(t)^{*} - \mathrm{y}(t)}{y(t)}\right| \times 100\% \tag{29}$$

$$e_{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y(t)^{*} - \mathrm{y}(t))^{2}} \tag{30}$$

$$e_{MAE} = \frac{1}{N}\sum_{t=1}^{N}\left|y(t)^{*} - \mathrm{y}(t)\right| \tag{31}$$

where N is the number of test samples, $y(t)^{*}$ and $y(t)$ represent the predicted and actual wind speeds at time t, respectively.

5.3. Simulation design

First, the recognition accuracy of the time scale of wind processes needs to be tested. Obviously, the recognition result of the time scale of wind processes is affected by input data's time range (denoted as k) and different input feature dimensions (IFD). The k value is small, the information contained is small, and the recognition accuracy is poor, which is not conducive to subsequent research. The k value is large, the information contained is large, but this needs to consider the classification boundary (120 min), so the k value should not be too small or too large.

The input feature dimension is high, and there may be redundancy, resulting in poor recognition accuracy. Here, different recognition models are constructed to analyze the influence of k and IFD. Second, the data is screened out and the performance of the data screening method is evaluated. The test set (including input and output) is selected from the historical data.

Taking the complex network as an example, the complex network is used to extract the morphological characteristics of the input of test set, and the training data whose input has the similar morphological characteristics are screened from the historical data by the extracted characteristics, which is the data screening process.

When performing the evaluation, the fluctuation characteristics of the output of test sets need to be extracted on the basis of the previous process, and the training data whose output has the same fluctuation characteristics are selected from the historical data according to the extracted characteristics.

The number of intersecting samples in these two training data sets, namely this method selects the number of samples matching the test set. For the other three methods mentioned in the paper: Euclidean distance, Pearson coefficient, and Cosine similarity, the simulation process is the same. When evaluating, the same amount of training data is selected according to the input and output of the test set respectively, and then the number of intersecting samples is calculated. Finally, the relative matching ratio is calculated according to the formula (22) to evaluate the four methods.

Third, the forecasting of wind speed is performed. According to the recognition results of the time scale and the data screening results, the optimal compatible training data used for prediction in the proposed model is obtained, and then corresponding mapping relationships are established based on SVM to forecast the future wind speed. In order to verify the effectiveness of the proposed model, the other three prediction models are established.

In the first model, only the identification of wind patterns is performed, and the historical wind processes with the same label are selected as the training data for prediction according to the identified label of the time scale of the current wind process. In the second model, only CN-based data screening is performed, and the historical data with the same fluctuation characteristics are selected as the training data for prediction according to the fluctuation characteristics of the current wind process in the first k minutes.

In the third model, historical data is directly used as training data for prediction without any processing. The forecasting process of these three models is the same as the proposed model. Finally, we evaluate the performance of these four models by calculating evaluation indexes and then analyze the advantages, disadvantages, and applicable scope of the proposed model.

Fourth, the performance of the proposed model is further verified. The data screening method in the proposed model is respectively changed into Euclidean distance, Pearson correlation coefficient, and Cosine similarity, and the rest simulation process remains unchanged to forecast the future wind speed. Finally, evaluation indexes are calculated to evaluate the performance.

Due to the different test purposes of different simulations, the corresponding benchmark selection is also different. For time-scale pattern recognition simulation, the benchmark models are those with different feature groups. For the data screening, the benchmark methods are commonly used methods include Euclidean distance, Pearson correlation coefficient, and Cosine similarity. For the testing of the complete proposed forecasting model, we choose the three prediction models as benchmarks: model only have identification without data screening, model only have data screening without identification, and model have no identification and no data screening.

5.4. Results of time-scale pattern recognition

According to the cumulative probability distribution of wind process time scales, the wind processes are divided into two patterns. The time scale of one pattern is shorter than 2 hours, which is denoted as A, and the other pattern is longer than 2 hours, which is denoted as B. Table 3 shows the time scale classification results of the three-year historical wind processes of the DRA weather station. In Table 3, it can be seen that pattern A contains 5878 samples and pattern B contains 2176 samples, which indicates the time scale of history wind patterns mostly is less than 2 hours.

In order to analyze the influence caused by input data's time range (denoted as k) and different input feature dimensions (IFD) on classification accuracy, different classification models are constructed and tested. For input data's time range k, the higher the value of k, the more information it contains, the better corresponding classification results should be, but considering the classification boundary (120 min), k value should not be too large.

Here, only k equaling to 30min, 45min, 60min, and 75min is discussed. For IFD, four groups models denoted by 6F to 3F which means IFD vary from 6 to 3 are discussed. In 6F group, there is only one model with all features defined in section III. In group 5F to 3F, there are too many models with the same dimension but different input feature combinations.

Table 3. Results of time scale classification.

Type	Pattern A	Pattern B	Total
Number of wind patterns	5878	2176	8054

Percentage/%	72.98	27.02	100

In order to analyze the influence caused by different input features (IF) on classification accuracy, this paper select two representative models' results (the best OA and the worst OA) from each group for display.

Table 4 shows the classification results corresponding to different k when all defined features as input. From Table 4, it can be seen that UA1, PA2, UA2, and OA all show an upward tendency with the increase of k, and the maximum values are 78.72, 93.66, 89.26, and 86.77 respectively. But PA1's overall tendency is downward with the increase of k, and the maximum value is 90.74. The results above revealed the classification performance is relevant to input data's time range, the higher the value of k, the greater the OA. Considering the best classification accuracy, it was considered k to be 75min.

Next, the influence of different input feature dimensions on classification accuracy with k equaling to 75 is discussed. The classification results of all representative models in group 6F to 3F are shown in Table 5 In group 6F, there is only one model, i.e. IF=[F1, F2, F3, F4, F5, F6]. In group 5F, the best model (IF=[F1, F2, F3, F4, F6]), and the worst model (IF=[F1, F3, F4, F5, F6]) of 5-D models whose input features are included in group 6F are selected as the representative models of this group.

In group 4F, the best model (IF=[F1, F2, F4, F5]), and the worst model (IF=[F1, F3, F5, F6]) of 4-D models whose input features are included in group 6F are selected as the representative models of this group. In group 3F, the best model (IF=[F1, F2, F5]), and the worst model (IF=[F3, F4, F5]) of 3-D models whose input features are included in group 6F is selected as the representative models of this group.

The classification model of group 6F gets better results where OA is 86.77%, the classification model of group 5F obtains the promising results where the lowest OA is 81%, and the results of the classification model of group 4F are general, with the lowest OA equaling to 76%. In group 3F, there are cases where a certain category cannot be identified. Such as model [F3, F4, F5], whose PA1= UA1=0, means the model cannot identify A at all.

From Table 5, it can be seen that the lowest OA in group 6F, 5F, 4F, and 3F are 86.77%, 81%, 76%, and 74.67%. The results indicate the OA in each group shows a downward tendency with the decrease of input feature dimension, which represents the classification accuracy is related to IFD. The larger the IFD is, and the higher the accuracy is.

Table 4. The classification results corresponding to different k.

6F Accuracy (%)					
Time range	A		B		OA
	PA1	UA1	PA2	UA2	
30min	90.74	73.13	49.30	77.78	74.30
45min	76.54	72.94	77.00	80.21	76.80
60min	62.70	78.72	92.31	84.51	83.07
75min	65.96	77.50	93.66	89.26	86.77

Table 5. The classification results of four groups with different IFD and IF.

k=75min Accuracy (%)						
Group	Feature combinations	A		B		OA
		PA1	UA1	PA2	UA2	
6F	[F1,F2,F3,F4,F5,F6]	65.96	77.50	93.66	89.26	86.77
5F	[F1,F2,F3,F4,F6]	45.83	75.86	95.39	84.80	83.50
	[F2,F3,F4,F5,F6]	51.85	70.00	91.78	83.75	81.00
4F	[F1,F2,F4,F5]	28.26	76.47	97.40	81.97	81.50
	[F1,F3,F5,F6]	28.57	66.67	94.44	77.27	76.00
3F	[F1,F2,F5]	33.09	62.75	93.38	80.56	78.19

	[F3,F4,F5]	0	0	100	74.67	74.67

From Table 5, it can also be seen that the higher IFD models are not necessarily superior to lower IFD models in all cases. For example, the OAs of model [F2, F3, F4, F5, F6] in 5F and model [F1, F2, F4, F5] in 4F are 81% and 81.5%, which represents the model [F1, F2, F4, F5] in 4F is better than the model [F2, F3, F4, F5, F6] in 5F. The discussion and analysis above revealed that the performance of the classification model is related to input data's time range, input feature dimensions, and input feature.

5.5. Evaluate the performance of data screening methods

According to the RMR defined in section V, the performance of four data screening methods is evaluated. A sample matching the input and output means that the sample is selected according to the input of the test data, and the sample is also selected according to the output of the test data, which has the same or similar characteristics with the test data. Different data screening methods select samples according to different features.

For example, Complex networks focus on wave and morphological characteristics, Euclidean distance focuses on spatial distance, Pearson correlation coefficient focuses on correlation, and cosine similarity focuses on the shape. Different features are concerned, the samples selected according to the input of test data are different, and the number of samples matching the input and output is also different, which is mainly due to different data features concerned by different data screening methods.

Table 6 is the result of RMR of data screening method based on Complex network, Euclidean distance, Pearson correlation coefficient, and Cosine similarity. It can be seen from Table 5 that the data screening method based on complex network has the highest RMR, indicating that compared with other traditional methods, this method can deeply mine the morphological characteristics of wind speed sequence, thus screening and obtaining the optimal compatible training data, which is helpful to establish more accurate mapping relationship.

5.6. Performance of proposed forecasting method

The time resolution of the original data is 1 minute, which needs to be compressed to 15 minutes for the prediction, namely, four wind speed values in one hour. It was used SVR to establish four prediction models to evaluate the effectiveness of the proposed model.

The four prediction models respectively are: both wind pattern identification and CN based data screening (Proposed method), only identification without data screening (Only classification), only data screening without identification (Only CN based data screening), and no identification and no data screening (Direct prediction).

For the prediction model including identification, the training input is the first 75-minute wind speed of the same label wind processes, and the output is the next 1-hour wind speed. For the prediction model without classification, the training input is the first 75-minute wind speed, and the output is the next 1-hour wind speed. When performing the forecasting, the test input is the first 75-minute data of the current wind process.

Figure. 10 shows the wind speed forecasting results generated by four prediction models based on SVM. It shows that the forecasting results generated by the proposed model are closer to the actual wind speed than that of the other three models. In other words, the proposed model has a better performance than the other models, implying the effectiveness of the proposed model.

According to the formulas defined in V, RMSE, MAE, and MAPE evaluation indexes of forecasting results based on four prediction models are calculated under different prediction horizons, as shown in Table 7.

Table 6. Matching rate results.

Method	CN	ED	PCC	CS
RMR	32.77%	18.20%	22.32%	26.71%

From Table 7, it can be seen that the three evaluation indexes show an upward trend with the forecasting horizon increases, indicating forecasting accuracy gradually decreases, and the 15-minute scale accuracy is highest, 60-minute scale accuracy is worst. By comparing the prediction model containing only classification and the prediction model containing only data screening, it is found that the accuracy of the former is better than that of the latter, verifying the effectiveness of classifying the wind processes according to time scale.

Among the four prediction models, the evaluation indexes of the proposed model are the smallest, the evaluation indexes of the direct prediction model are the largest, and the evaluation indexes of the other two models are between them, which proves the excellent performance of the proposed model.

The comparison of MAPE among the four prediction models for four different time scales is shown in Figure. 11. It can be observed that as the time scales increases, the value of MAPE gradually increases, and the accuracy decreases. Only at the 30-minute scale, the MAPE of the proposed model is slightly higher than that of the prediction model containing the only classification. At other time scales, the MAPE of the proposed model is always lower than that of the other three models, which verifies the superiority of the proposed model.

In order to further analyze the advantages, disadvantages and applicable scope of the proposed model, different scenarios are set up for simulation, including small scale wind process, large scale wind process, correct scale classification, and wrong scale classification. The simulation results of various scenarios are shown in Table 8.

From Table 8, it can be seen that no matter for small scale wind process or large scale wind process, when the scale recognition is accurate, data with similar variations to the wind speed to be predicted can be selected from the historical data as the training set, which reduces the redundancy and improves the forecasting accuracy, and the forecasting results are the best.

On the contrary, when the scale recognition is wrong, data with dissimilar variations to the trend of wind speed to be predicted is selected from the historical data as the training set, resulting in incorrect mapping relationship establishment and the worst forecasting results. When without classification, the training set is redundancy, and its forecasting results are between the best and the worst.

When the scale recognition of small scale wind process is correct, MAPE, RMSE, and MAE increase by 7.2%, 0.2%, and 7.1%, respectively, compared with the non-classification. Except for RMSE, the other two indexes have a great improvement. When the scale recognition of large scale wind process is correct, MAPE, RMSE, and MAE increase by 3.0%, 5.5% and 3.0% respectively compared with the non-classification. The results above revealed compared with large scale wind process, this model is more suitable for small scale wind process, and the key to improving the forecasting accuracy of this model is to accurately recognize the time scale of wind process.

In order to further verify the effectiveness of the data screening method based on the complex network in the proposed model and its influence on forecasting accuracy. Here, the data screening method in the proposed model is sequentially changed into Euclidean distance, Pearson correlation coefficient and Cosine similarity to forecast the future wind speed, and the evaluation indexes of the forecasting results are shown in Table 9.

It can be seen from the Table 9, when using complex network to screen the data, the evaluation indexes are the lowest and the forecasting accuracy is the highest, when using Euclidean distance to screen the data, the evaluation indexes are the highest and the forecasting accuracy is the worst, and the evaluation indexes of the other two data screening methods are between them, which verifies the effectiveness of the data screened by complex network.

Combining the results of Table 5 and Table 7, it is found that the higher the RMR of the data screening method, the lower the corresponding forecasting indexes, indicating that the RMR can reflect the forecasting accuracy to some extent.

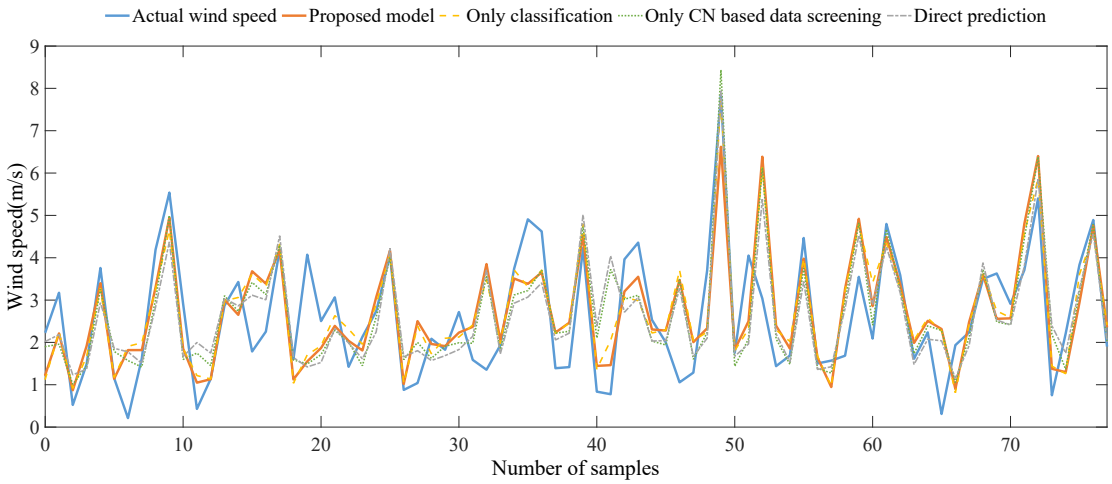


Figure 10. The classification forecasting and direct forecasting results of three algorithms.

Table 7. Evaluation indexes of forecasting results based on four models.

Evaluation index	Forecast ahead	Model			
		Proposed model	Only classification	Only CN based data screening	Direct prediction
MAPE	15min	6.32%	6.30%	6.82%	6.86%
	30min	7.48%	7.68%	7.92%	8.49%
	45min	8.96%	9.05%	8.94%	9.44%
	60min	10.87%	10.96%	10.88%	10.91%
RMSE	15min	0.8195	0.8156	0.8722	0.8683
	30min	0.9769	0.9867	1.0217	1.0846
	45min	1.1880	1.2168	1.2079	1.2666
	60min	1.4057	1.4321	1.4154	1.4266
MAE	15min	0.6340	0.6321	0.6844	0.6887
	30min	0.7510	0.7712	0.7950	0.8519
	45min	0.8992	0.9090	0.8975	0.9477
	60min	1.0916	1.0999	1.0921	1.0953

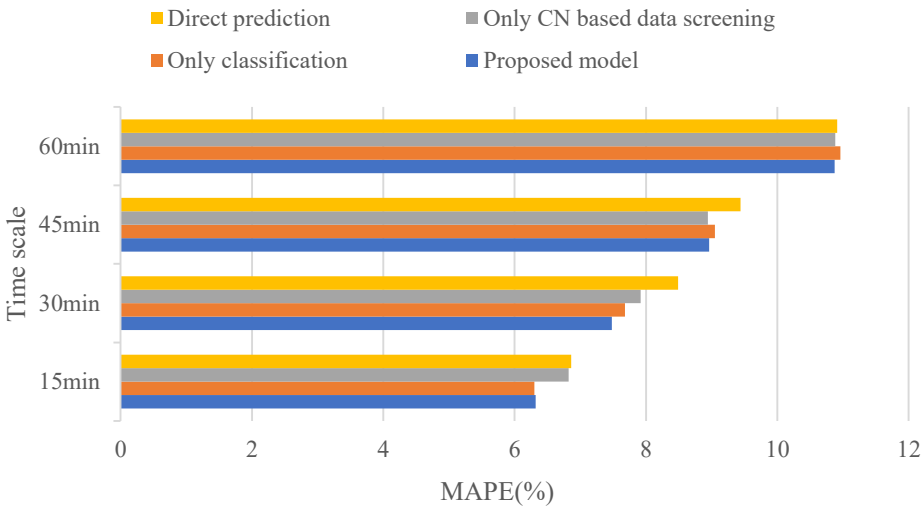


Figure 11. MAPE comparison among the four prediction models.

Table 8. The simulation results of different scenarios.

Evaluation index	Small scale wind process			Large scale wind process		
	Correct classification	Wrong classification	Without classification	Correct classification	Wrong classification	Without classification
MAPE	16.36%	20.91%	17.62%	9.35%	16.60%	9.64%
RMSE	0.7930	0.9280	0.7946	1.2499	2.2847	1.3231
MAE	0.6079	0.7770	0.6546	0.9576	1.7004	0.9876

Table 9. Evaluation indexes of forecasting results based on different data screening methods.

Evaluation index	Method			
	Classification+ CN	Classification + ED	Classification + PCC	Classification + CS
MAPE	8.41%	8.81%	8.75%	8.67%
RMSE	1.1195	1.2195	1.1770	1.1501
MAE	0.8440	0.8840	0.8786	0.8708

6. Conclusion

This paper proposed an ultra-short-term wind speed forecasting model based on time scale recognition and dynamic adaptive modeling to improve the accuracy of wind speed and wind power forecasting. The proposed approach consists of four main steps: definition and division of wind processes, the establishment of time scale distribution functions of wind processes, the establishment of pattern recognition model based on feature extraction, and data screening based on complex network.

Through wind process classification based on the time scale distribution information, the natural variation law of wind speed can be better analyzed. Then combined with the applied data screening method based on the complex network which can select more highly compatible training samples that matching with test data, a more accurate wind speed forecasting model is achieved. A case study using historical data from DRA meteorological station shows that the proposed model has a better forecasting accuracy than the benchmark model in this paper.

Funding: This work is supported by National Natural Science Foundation of China (52007092), and the Fundamental Research Funds for the Central Universities (2019MS084).

References

1. Wang F, Zhen Z, Liu C. Mi Z, Hodge B.-M, Shafie-khah M, Catalão J.P.S. Image phase shift invariance based cloud motion displacement vector calculation method for ultra-short-term solar PV power forecasting. Energy Convers. Manag.2018; 157:123–135.
2. Wu Y.K, Chang,S.M, Mandal P. Grid-Connected Wind Power Plants: A Survey on the Integration Requirements in Modern Grid Codes. In Proceedings of the IEEE Transactions on Industry Applications. Institute of Electrical and Electronics Engine-ers Inc., 2019;Vol. 55, pp.5584–5593.
3. Wu Y.K, Tan W.S, Huang S.R, Chiang Y.S, Chiu C.P, Su C.L. Impact of Generation Flexibility on the Operating Costs of the Taiwan Power System under a High Penetration of Renewable Power. In Proceedings of the IEEE Transactions on Industry Applications; Institute of Electrical and Electronics Engineers Inc. 2020;Vol.56, pp.2348–2359.
4. Zhen Z, Pang S, Wang F, Li K, Li Z, Ren H. Shafie-khah M, Catalao J.P.S. Pattern Classification and PSO Optimal Weights Based Sky Images Cloud Motion Speed Calculation Method for Solar PV Power Forecasting. IEEE Trans. Ind. Appl. 2019;1–1.
5. Global Wind Energy Council. Global Wind Report 2018. Brussels, 2019.
6. Pan Y, Liu F, Chen L, Wang J, Qiu F, Shen C, Mei S. Towards the Robust Small-Signal Stability Region of Power Systems under Perturbations Such as Uncertain and Volatile Wind Generation. IEEE Trans. Power Syst. 2018.
7. Chen H, Zhang J, Tao Y, Tan F. Asymmetric GARCH type models for asymmetric volatility characteristics analysis and wind power forecasting. Protection and Control of Modern Power Systems.2019;4(4):356-366.
8. Li J, Wang S, Ye L, Fang J. A coordinated dispatch method with pumped-storage and battery-storage for compensating the variation of wind power. Protection and Control of Modern Power Systems. 2018;3(1):21-34.
9. Wu Y.K, Zeng J.J, Lu G.L, Chau S.W, Chiang Y.C. Development of an Equivalent Wind Farm Model for Frequency Regulation. IEEE Trans. Ind. Appl. 2020;56;2360–2374.
10. Pabitra K.G, Abhik B. Stability enhancement of wind energy integrated hybrid system with the help of static synchronous compensator and symbiosis organisms search algorithm. Protection and Control of Modern Power Systems.2020;5(2):138-150.
11. Zhao L, Zhou Y, Matsuo I.B.M, Korkua S.K, Lee W.J. The Design of a Remote Online Holistic Monitoring System for a Wind Turbine. IEEE Trans. Ind. Appl.2020;56:–21.
12. Wang P, Zhang Z, Huang Q, Lee W.J. Wind Farm Dynamic Equivalent Modeling Method for Power System Probabilistic Stability Assessment. IEEE Trans. Ind. Appl.2020;56:2273–2280.
13. Chen Q, Wang F, Hodge B.M, Zhang J, Li Z, Shafie-Khah M, Catalao J.P.S. Dynamic Price Vector Formation Model-Based Automatic Demand Response Strategy for PV-Assisted EV Charging Stations. IEEE Trans. Smart Grid 2017.
14. Wang F, Xu H, Xu T, Li K, Shafie-khah M, Catalão J.P.S. The values of market-based demand response on improving power system reliability under extreme circumstances. Appl. Energy 2017; 193:220–231.
15. Wang F, Zhou L, Ren H, Liu X. Search improvement process-chaotic optimization-particle swarm optimization-elite retention strategy and improved combined cooling-heating-power strategy based. Energies 2017.
16. Lai J, Lu X, Wang F, Dehghanian P, Tang R. Broadcast gossip algorithms for distributed peer-to-peer control in AC microgrids. IEEE Trans. Ind. Appl.2019.
17. Wang F, Li K, Duić N, Mi Z, Hodge B.-M, Shafie-khah M, Catalão J.P.S. Association rule mining based quantitative analysis approach of household characteristics impacts on residential electricity consumption patterns. Energy Convers. Manag.2018;171:839–854.
18. Xuan Z, Gao X, Li K, Wang F, Ge X, Hou Y. PV-Load Decoupling Based Demand Response Baseline Load Estimation Approach for Residential Customer with Distributed PV System. IEEE Trans.Ind.Appl.2020.
19. Lu X, Li K, Xu H, Wang F, Zhou Z, Zhang Y. Fundamentals and business model for resource aggregator of demand response in electricity markets. Energy 2020.
20. Li K, Wang F, Mi Z, Fotuhi-Firuzabad M, Duić N, Wang T. Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation. Appl. Energy 2019;253, 113595.
21. Lin Y, Yang M, Wan C, Wang J, Song Y. A Multi-Model Combination Approach for Probabilistic Wind Power Forecasting. IEEE Trans. Sustain. Energy 2019.
22. Du P, Wang J, Guo Z, Yang W. Research and application of a novel hybrid forecasting system based on multi-objective optimization for wind speed forecasting. Energy Convers. Manag.2017.
23. Yu C, Li Y, Zhang M. Comparative study on three new hybrid models using Elman Neural Network and Empirical Mode Decomposition based technologies improved by Singular Spectrum Analysis for hour-ahead wind speed forecasting. Energy Convers. Manag.2017.

24. Ma J, Yang M, Lin, Y. Ultra-short-term probabilistic wind turbine power forecast based on empirical dynamic modeling. IEEE Trans. Sustain. Energy 2020.
25. Ma J, Yang M, Han X, Li Z. Ultra-short-term wind generation forecast based on multivariate empirical dynamic modeling. In Proceedings of the 2017 IEEE Industry Applications Society Annual Meeting, IAS 2017; 2017.
26. Kavasseri R.G, Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. Renew. Energy 2009.
27. Wang F, Xuan Z, Zhen Z, Li K, Wang T, Shi M.A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework. Energy Convers. Manag.2020,
28. Shukur O.B, Lee M.H. Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA. Renew. Energy 2015.
29. Wang F, Mi Z, Su S, Zhao H. Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters. Energies 2012.
30. Yujing Sun, Fei Wang, Zhao Zhen, Zengqiang Mi, Chun Liu, Bo Wang, Jing Lu. Research on short-term module temperature prediction model based on BP neural network for photovoltaic power forecasting. In Proceedings of the 2015 IEEE Power & Energy Society General Meeting; IEEE, 2015; pp. 1–5.
31. Wang F, Zhen Z, Wang B, Mi Z. Comparative Study on KNN and SVM Based Weather Classification Models for Day Ahead Short Term Solar PV Power Forecasting. Appl. Sci.2017, 8,28.
32. Liang Z, Liang J, Wang C, Dong X, Miao X. Short-term wind power combined forecasting based on error forecast correction. Energy Convers. Manage. 2016.
33. Zhou J, Shi J, Li G. Fine tuning support vector machines for short-term wind speed forecasting. Energy Convers Manage.2011.
34. Jiang Y, Huang G. Short-term wind speed prediction: Hybrid of ensemble empirical mode decomposition, feature selection and error correction. Energy Convers. Manag.2017.
35. Liu H, Tian H.Q, Li Y.F. Four wind speed multi-step forecasting models using extreme learning machines and signal decomposing algorithms. Energy Convers. Manag.2015.
36. Salcedo-Sanz S, Pastor-Sánchez A, Prieto L, Blanco-Aguilera A, García-Herrera R. Feature selection in wind speed prediction systems based on a hybrid coral reefs optimization - Extreme learning machine approach. Energy Convers Manage.2014.
37. Peng H, Liu F, Yang X. A hybrid strategy of short term wind power prediction. Renew. Energy 2013.
38. He Q, Wang J, Lu H. A hybrid system for short-term wind speed forecasting. Appl. Energy 2018.
39. Peng T, Zhou J, Zhang C, Zheng Y. Multi-step ahead wind speed forecasting using a hybrid model based on two-stage decomposition technique and AdaBoost-extreme learning machine. Energy Convers. Manag.2017.
40. Ding Z, Yang P, Yang X, Zhang Z. Wind power prediction based on sequential time clustering using SVM. In Proceedings of the 2011 International Conference on Electrical and Control Engineering, ICECE 2011 - Proceedings; 2011.
41. Sun G, Jiang C, Cheng P, Liu Y, Wang X, Fu Y, He Y. Short-term wind power forecasts by a synthetical similar time series data mining method. Renew. Energy 2018.
42. Wang F, Zhang Z, Liu C, Yu Y, Pang S, Duić N, Shafie-khah M, Catalão J.P.S. Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting. Energy Convers. Manag.2019;181,443–462.
43. Taboada F.G, Stock C.A, Griffies S.M, Dunne J, John J.G, Small R.J, Tsujino H. Surface winds from atmospheric reanalysis lead to contrasting oceanic forcing and coastal upwelling patterns. Ocean Model. 2019.
44. Morales-Acuña E, Torres C.R, Linero-Cueto J.R. Surface wind characteristics over Baja California Peninsula during summer. Reg. Stud. Mar. Sci. 2019.
45. Yu Y, Han X, Yang M, Yang J. Probabilistic Prediction of Regional Wind Power Based on Spatiotemporal Quantile Regression. IEEE Trans Ind Appl 2020;56. https://doi.org/10.1109/TIA.2020.2992945.
46. Ma J, Yang M, Lin Y. Ultra-short-term probabilistic wind turbine power forecast based on empirical dynamic modeling. IEEE Trans Sustain Energy 2020;11:906–15. https://doi.org/10.1109/TSTE.2019.2912270.
47. Zhao J, Guo Y, Xiao X, Wang J, Chi D, Guo Z. Multi-step wind speed and power forecasts based on a WRF simulation and an optimized association method. Appl Energy 2017;197. https://doi.org/10.1016/j.apenergy.2017.04.017.
48. Li H, Wang Y, Zhang X, Fu G. Evaluation Method of Wind Power Consumption Capacity Based on Multi-Fractal Theory. Front Energy Res 2021;9. https://doi.org/10.3389/fenrg.2021.634551.
49. Munshi AA. Clustering of Wind Power Patterns Based on Partitional and Swarm Algorithms. IEEE Access 2020;8. https://doi.org/10.1109/ACCESS.2020.3001437.
50. Calif R, Emilion R, Soubdhan T. Classification of wind speed distributions using a mixture of Dirichlet distributions. Renew Energy 2011;36. https://doi.org/10.1016/j.renene.2011.03.024.
51. Buhan S, Özkazanç Y, Çadirci I. Wind Pattern Recognition and Reference Wind Mast Data Correlations With NWP for Improved Wind-Electric Power Forecasts. IEEE Trans Ind Informatics 2016;12. https://doi.org/10.1109/TII.2016.2543004.
52. Frey J. An exact Kolmogorov–Smirnov test for whether two finite populations are the same. Stat. Probab. Lett. 2016.
53. Wang F, Zhen Z, Mi Z, Sun H, Su S, Yang G. Solar irradiance feature extraction and support vector machines based weather status pattern recognition model for short-term photovoltaic power forecasting. Energy Build. 2015,86:427–438.

54. Wang F, Li K, Wang X, Jiang L, Ren J, Mi Z, Shafie-khah M, Catalão J. A Distributed PV System Capacity Estimation Approach Based on Support Vector Machine with Customer Net Load Curve Features. Energies 2018, 11, 1750.
55. Liang Y, Niu D, Hong W.C. Short term load forecasting based on feature extraction and improved general regression neural network model. Energy 2019.
56. Wang F, Li K, Zhou L, Ren H, Contreras J, Shafie-khah M, Catalão J.P.S. Daily pattern prediction based classification modeling approach for day-ahead electricity price forecasting. Int. J. Electric Power Energy Syst.2019,105.
57. Han Q, Meng F, Hu T, Chu F. Non-parametric hybrid models for wind speed forecasting. Energy Convers Manage. 2017.
58. An H, Gao X, Fang W, Ding Y, Zhong W. Research on patterns in the fluctuation of the co-movement between crude oil futures and spot prices: A complex network approach. Appl. Energy 2014.