Short-term wind power forecast based on chaotic analysis and multivariate phase space reconstruction

Tianyao Ji, Jin Wang, Mengshi Li*, Qinghua Wu

School of Electric Power Engineering, South China University of Technology, Guangzhou, 510640, China

Abstract

The randomness and volatility of wind power time series, which are an external reflection of their internal chaotic dynamics, have always been important factors affecting the accuracy of wind power prediction. The chaotic characteristics of wind power data have not been studied deeply enough in existing research. Therefore, in this paper, a short-term wind power forecast method based on chaotic analysis is proposed, including chaotic time series estimation and multivariate phase space reconstruction (PSR). First, we calculate the largest Lyapunov exponent of wind power time series to measure the degree of chaos in wind power data. Then, the chaotic characteristics of several wind power time series from some neighboring wind farms are analyzed, based on which two multivariate multi-dimensional PSR models are established. Afterwards, the nearest neighbors (NNs) of the data to be predicted are selected in the phase space and delivered to the least-square support vector machine (LSSVM) model to experiment the forecasting. Wind power data collected from several adjacent wind sites are used to conduct the simulation studies, and the results have shown the effectiveness of the proposed method and its advantage over the classic LSSVM model in terms of accuracy and stability.

*The authors are with School of Electric Power Engineering, South China University of Technology, Guangzhou, 510640, China. Corresponding author: Dr. Mengshi Li, mengshili@scut.edu.cn
The research work was supported by National Natural Science Foundation of China (No. 52077081) and Guangdong Province Introduction of Innovative R&D Team Program (No. 201001N0104744201).

Keywords: Wind power forecasting, chaotic characteristics, the largest Lyapunov exponent, multivariate phase space reconstruction

1. Introduction

Global energy shortages and environmental problems are becoming more and more serious, hence, the development and application of renewable energy grows increasingly important. Wind energy is currently one of the most widely used renewable energy resources, and its proportion in energy system continues to increase. By far, the total worldwide wind capacity has reached 744 GW, which accounts for 7% of the world electricity demand [1]. Among it, the grid-connected capacity of wind power is also increasing. Under such circumstances, to ensure the stable operation of the power system and reduce wind curtailment, the effective utilization of wind power is of utmost importance. However, due to its strong randomness and volatility, it is always difficult to predict wind power generation. Therefore, researchers need to combine a variety of related factors and historical data to achieve effective wind power forecasting.

In recent years, a large number of wind power forecasting methods have been proposed. They can be grouped into three categories, including physical models [2, 3], statistical models [4, 5], and artificial intelligence (AI) models [6–9]. The AI models are now widely used and proved to be effective. They are trained using historical data to learn the relationship between the input and output data. Support vector regression (SVR) [10, 11], support vector machine (SVM), and its least-square version, least-square support vector machine (LSSVM) [12, 13] are typical AI models. LSSVM simplifies the formulation of the standard SVM, achieving better generalization performance and low computational cost [14]. In our research, we select LSSVM as the forecasting model.

Although prediction models can learn the relationship between input and output data, it is unlikely for the prediction models to fully capture the inherent characteristics and evolution dynamics of the wind power data as they are more complex. Therefore, many data analysis methods are applied to preprocess the data before forecasting, such as data decomposition algorithms [15–19], feature extraction algorithms [20–22], etc. These methods focus on decreasing the influence of the non-stationarity and strong volatility of wind power data on their prediction, but cannot provide an in-depth analysis of the intrinsic dynamics of the data. According to some research, the inherent evolution of wind power data has the characteristics of chaotic dynamics [23–25]. The changing trend of wind power time series seems to be irregular, but it is a reflection of their chaotic dynamics. Papers [26] and [27] reconstruct wind power/speed time series into higher-dimensional phase space to observe and study their chaotic characteristics. The PSR method needs two parameters: embedding dimension and time delay. To estimate the embedding dimension, [28] develops the method of false nearest neighbors (FNN), but its criterion to determine a false neighbor is somewhat subjective; Cao *et al.* [29] improves the FNN method by proposing an indicator to help distinguish between deterministic and stochastic time series. To estimate the time delay, [30] proposes the mutual information method (MI). In this paper, the Cao method and the MI method are used to calculate the two parameters respectively.

Apart from PSR, the estimation of chaotic extent is also important to analyze the chaotic characteristics of a wind power time series. The largest Lyapunov exponent (LLE) is usually used to determine whether a time series is chaotic [23, 25, 26]. However, since the LLE is closely related to the degree of divergence of adjacent orbits in the attractor (*i.e.*, the reconstructed time series), it is also a quantitative measurement of the chaotic extent of a time series [31–33]. Wolf *et al.* develops an algorithm to calculate the LLE of one-dimensional experimental time series [32]. The Wolf algorithm is improved in this paper by adjusting some parameters, and applied to analyze wind power data. In general, time series with positive LLE indicates that it is chaotic. However, the improved Wolf algorithm may cause the estimation of LLE of random signals to be positive. Thus, the improved Wolf algorithm is combined with the Cao method to analyze the chaotic characteristics of a time series, where the latter judges the chaos qualitatively and the former estimates the chaotic extent quantitatively.

Time series with a similar degree of LLE will be analyzed together as relevant variables to assist the prediction. This is an important application of chaotic analysis in our research. In this paper, the chaotic characteristics of wind power data of adjacent wind farms are estimated and among them, the data with close chaos extent (*i.e.*, similar LLE) will be analyzed in the same phase space. Then a multivariate PSR method is proposed, including two models, the high-dimensional PSR (HDPSR) model and the low-dimensional PSR (LDPSR) model. Their application scenarios vary depending on the LLEs and the embedding dimensions of wind power time series. Moreover, a calculation process to verify the chaotic characteristics of the reconstructed space is described in this paper. In addition, to improve the prediction accuracy and stability, the methods for selecting the nearest neighbors (NNs) in the phase space and the training points for forecasting models are optimized.

The main contributions of this paper are summarized as follows:

1. Detailed analysis and instructions of the calculation principle of LLE are given, explaining why and how it can be used to measure the degree of chaos. Adjustments have been made to the selection of some parameters of the calculation algorithm of LLE, which is applied to estimate the chaotic extent of wind power time series.
2. Two multivariate PSR models are proposed for wind power time series based on the LLE and PSR parameters, which are used to assist the prediction of data collected from adjacent wind farms.
3. The influence of the numbers of NNs and training points on prediction results are analyzed, and an optimal selection method for the two parameters is developed.

2. Chaotic characteristic analysis

Chaotic characteristics analysis consists of two parts: phase space reconstruction (PSR) and chaotic extent estimation. For PSR, the Cao method and the mutual information (MI) method are selected to calculate the embedding dimension and time delay, respectively; for chaotic extent estimation, the Cao method and the improved Wolf algorithm are combined to estimate whether a time series is chaotic and calculate the LLE.

2.1. Determining the embedding dimension

The estimation of the embedding dimension, m, is based on the distance change rate of NNs. According to the Cao method, as m increases, the separation between NNs will converge gradually, and the dimension at that time should be the embedding dimension [29]. The calculation procedure is described as follows.

First, a given time series $\{x_1, x_2, \cdots x_N\}$ is reconstructed into a phase space and the i_{th} reconstructed point is described as:

$$X_i(m) = (x_i, x_{i+\tau}, \cdots, x_{i+(m-1)\tau}),\ i = 1, 2, \cdots N - (m-1)\tau, \tag{1}$$

where τ is the time delay and N is the total number of the samples of the original time series. In the m-dimensional phase space, the NNs of $X_i(m)$, which is denoted as $X_{n(i,m)}(m)$, are searched. Then define

$$r(i,m) = \frac{\left\|X_i(m+1) - X_{n(i,m)}(m+1)\right\|}{\left\|X_i(m) - X_{n(i,m)}(m)\right\|},\ i = 1, 2, \cdots, N - m\tau, \tag{2}$$

where $\|\cdot\|$ denotes the measurement of Euclidean distance, $X_i(m+1) = (x_i, x_{i+\tau}, \cdots, x_{i+m\tau})$ is the i_{th} reconstructed vector with dimension $m+1$, $X_{n(i,m)}(m+1)$ is $X_{n(i,m)}(m)$ with dimension $m+1$.

The mean value of $r(i,m)$ is selected to estimate the divergence of NN:

$$E(m) = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} r(i,m) \tag{3}$$

and to investigate its variation from m to $m+1$, the following variable is defined:

$$E_1(m) = \frac{E(m+1)}{E(m)} \tag{4}$$

When $E_1(m)$ converges at a certain m_0, m_0 is the minimum embedding dimension determined.

2.2. Estimating a deterministic time series

Aside from choosing the embedding dimension, the Cao method also defines another quantity, $E_2(m)$, to distinguish deterministic signals from stochastic signals, *i.e.*, to help investigate whether a time series is chaotic. Define

$$E^*(m) = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} \left| x_{i+m\tau} - x_{n(i,m)+m\tau} \right| \tag{5}$$

and its variation from m to $m+1$

$$E_2(m) = \frac{E^*(m+1)}{E^*(m)} \tag{6}$$

where E^* is the mean separation of the NNs in the m-dimensional space. As Figure 1 shows, for stochastic time series, its data features are not related to the dimension m; hence, E_2 will be equal to one for any m. Contrarily, for deterministic time series, E_2 is certainly related to m; thus, there must exist some cases that $E_2(m) \neq 1$. Moreover, as m increases, it may converge to some extent [29].

As noted, the LLE can to some extent help estimate a chaotic time series. However, based on practical experiments, it is found that the estimation of the LLE of some stochastic time series also tends to be positive, which is not consistent with the feature of the time series (stochastic). Therefore, this paper proposes to use the Cao method to distinguish a chaotic time series and to calculate the LLE to estimate the chaotic extent.

2.3. Determining the time delay

The mutual information method (MI) takes the time delay when the mutual information function first reaches a local minimum as the optimal embedding time delay for PSR [30]. For time series $x(t) = \{x_1, x_2, \ldots, x_m, \ldots\}$, denote it as system S, *i.e.*, $S = \{s_1, s_2, \ldots, s_n\}$. When time delay is τ, denote the delayed time series $x(t+\tau) = \{x_{1+\tau}, x_{2+\tau}, \ldots, x_{m+\tau}, \ldots\}$ as system Q, *i.e.*, $Q = \{q_1, q_2, \ldots, q_n\}$. Then the mutual information is defined as

$$I(Q, S) = H(Q) + H(S) - H(S, Q) \tag{7}$$

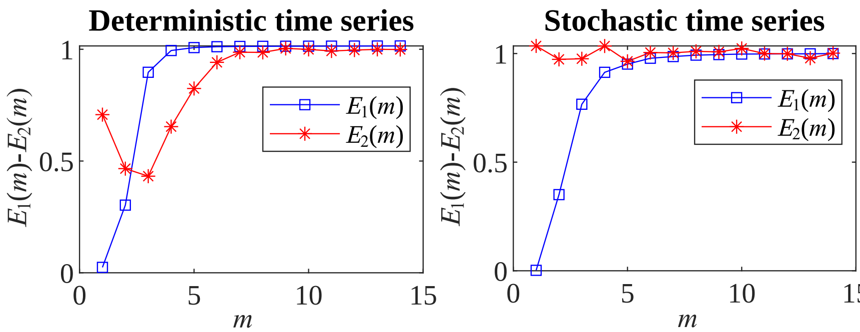

Figure 1: Plots of $E_1(m)$ and $E_2(m)$ of a deterministic and a stochastic time series.

where $H(Q)$ and $H(S)$ are the entropy of Q and S, respectively, and $H(S,Q)$ is the joint entropy of S and Q, which are defined as

$$H(S) = -\sum P_s(s_i)\log_2 P_s(s_i) \tag{8}$$

$$H(Q) = -\sum_j P_q(q_j)\log_2 P_q(q_j) \tag{9}$$

$$H(S,Q) = -\sum_{i,j} P_{sq}(s_i,q_j)\log_2 \frac{P_{sq}(s_i,q_j)}{P_s(s_i)} \tag{10}$$

where $P_s(s_i)$ is the probability that a measurement of s will yield s_i, $P_q(q_j)$ is the probability a measurement of q will yield q_j, and $P_{sq}(s_i,q_j)$ is the joint distribution probability of s_i and q_j in s and q. Thus, the calculation of mutual information $I(Q,S)$ is given as

$$I(Q,S) = \sum_i \sum_j P_{sq}(s_i,q_j)\log_2\left[\frac{P_{sq}(s_i,q_j)}{P_s(s_i)P_q(q_i)}\right] \tag{11}$$

When calculating the time delay, $I(Q,S)$ is a function of τ, denoted as $I(\tau)$. When $I(\tau)$ first reaches a local minimum, the value of τ at that time is determined as the time delay.

2.4. Calculating the largest Lyapunov exponents (LLE)

Lyapunov exponents are indicators that measure the chaotic characteristics of a time series. For a chaotic system, Lyapunov exponents are the average exponential rates of divergence or convergence of the nearby orbits in the phase space [1]. Since it costs huge calculation burden to estimate all the Lyapunov exponents of a system, which is unnecessary, only the LLE is chosen to measure the chaotic extent of a wind power time series.

In [32] Wolf proposed a general method to calculate the LLE of an experimental time series, based on which some improvements have been made in this research in the selection of some parameters to better fit the characteristics of wind power data. The improved Wolf algorithm is explained as follows.

In general, wind power time series only has one observation, but the calculation of Lyapunov exponents should be conducted in the phase space, by analyzing the chaotic attractor and the trajectory orbits of the states. Therefore, an original time series $\{x_1, x_2, \cdots, x_n\}$ is first reconstructed into a multi-dimensional phase space:

$$X = \begin{bmatrix} x_1, x_{1+\tau}, \cdots, x_{1+(m-1)\tau} \\ x_2, x_{2+\tau}, \cdots, x_{2+(m-1)\tau} \\ \cdots \\ x_N, x_{N+\tau}, \cdots, x_{N+(m-1)\tau} \end{bmatrix} \tag{12}$$

where m is the embedding dimension, τ is the time delay, $N = n - (m-1)\tau$ is the total number of state points in the phase space.

To calculate its LLE, it is needed to analyze the long-term evolution of two adjacent orbits in the reconstructed phase space. The two adjacent orbits, where two neighboring points are evolving along, should satisfy the following condition: the time interval of the two points in the original time series is at least one orbital period [32]. As long as two points have relatively small spatial distance, they can be considered as the initial states of the attractor in the phase space, although they are recorded at different moments in the original time series.

As shown in Figure 2, the estimation process can be described as follows.

Step 1: For the initial point of the reconstructed vector $p_0 = X(t_0) = \left\{x_{t_0}, x_{t_0+\tau}, \ldots, x_{t_0+(m-1)\tau}\right\}$, locate its NN in the Euclidean sense $p_0' = X(t_0')$.

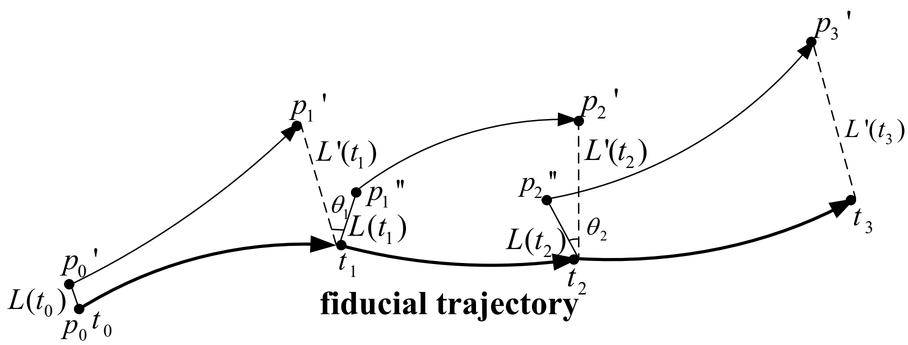

Figure 2: The procedure of calculating the LLE of an experimental time series.

The distance between the two points is defined as $L(t_0)$.

As discussed above, the two points should be at least one orbital period apart. There is no certain criterion about how to choose the orbital period of a wind power time series. Therefore, according to the principles of choosing the embedding dimension and time delay, to avoid excessive overlap or separation of the evolutionary orbits of the points, this paper proposes to choose the time interval of the first and the last coordinates of each phase point as the orbital period

$$P = (m-1)\tau \tag{13}$$

Thus, the initial point p_0 and its NN p_0' should satisfy the following condition:

$$|t_0 - t_0'| > P \tag{14}$$

Step 2: As the two points evolve one step to a later time t_1, the initial length element $L(t_0)$ becomes $L'(t_1)$, and p_0' becomes p_1'. At each moment t_k, estimate the logarithm of the change of the length element and denote it as η_k:

$$\eta_k = \log_2 \frac{L'(t_k)}{L(t_{k-1})} \tag{15}$$

To estimate the length element as accurately as possible, the evolution time $\Delta t = t_k - t_{k-1}$ should be short enough so that only a small-scale attractor structure is likely to be examined [32]. In this paper, the evolution time Δt

is set as one sampling interval, $\Delta t = 1$. This is because this research aims to predict wind power in short term and ultra-short term, which requires paying attention to the evolution of the data points at each step.

Step 3: To avoid the length element growing too large, a new NN should be chosen as time evolves. This new neighbor should satisfy two criteria: (1) its spatial distance (Euclidean distance) from the evolved fiducial point is small; (2) the angular separation between the evolved and replacement element is small. Here, the restriction is set that the separation θ should be an acute angle. For example, in Figure 2, when $L(t_0)$ evolves to time t_1, if the new length element $L(t_1)$ and the angular separation θ_1 between $L(t_1)$ and $L'(t_k)$ is small, the new NN p_1'' is an adequate replacement of the initial NN p_1'. If no adequate replacement point can be found, p_1' will be retained.

Step 4: Repeat Step 3 until the fiducial trajectory has traversed the entire data. After each evolution, calculate the change rate of the length element η_k, and the LLE is the mean value of the change rate, which is defined as

$$\lambda_1 = \frac{1}{t_S - t_0} \sum_{k=1}^{S} \log_2 \frac{L'(t_k)}{L(t_{k-1})} \tag{16}$$

where S is the number of the evolution steps.

3. Multivariate PSR and selection of NNs

Usually, when applying the PSR method to wind power time series, only single time series are analyzed. However, wind power data are related to several factors, including meteorological factors, wind power and speed data from neighboring wind farms, and so on [34]. This paper discusses the relationship between wind power time series of neighboring wind farms. Based on the characteristics analysis method described in Section 2, it is proposed to reconstruct several related time series into the same phase space and select the NNs in this phase space. The chosen points will be delivered to the forecasting model to carry out the prediction. In this research, two multivariate time series PSR methods are developed, namely HDPSR and LDPSR.

3.1. High-dimensional phase space reconstruction (HDPSR)

Given M experimental time series, $\{X_1, X_2, \ldots, X_i, \ldots, X_M\}$, where $X_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,j}, x_{i,n})\,, i = 1, 2, \ldots, M$. Define

$$\begin{array}{c}\left(\begin{array}{c} x_{1,1}, x_{1,2}, \cdots, x_{1,j}, \cdots, x_{1,n} \\ x_{2,1}, x_{2,2}, \cdots, x_{2,j}, \cdots, x_{2,n} \\ M \\ x_{M,1}, x_{M,2}, \cdots, x_{M,j}, \cdots, x_{M,n} \end{array}\right) \\ \downarrow, \quad \downarrow, \quad \wedge, \quad \downarrow, \quad \wedge, \quad \downarrow \\ Y_1, \quad Y_2, \; \cdots, \; Y_j, \; \cdots, \; Y_n \end{array} \tag{17}$$

Thus, a multivariate time series $Y = (Y_1, Y_2, \ldots, Y_n)$ is obtained. Similar to univariate time series, PSR is carried out on Y to obtain the reconstructed point

$$\begin{aligned} \boldsymbol{V}_j &= \left(Y_j, Y_{j+\tau}, \ldots, Y_{j+(m-1)\tau}\right) \qquad (18) \\ &= \left(\begin{array}{c} x_{1,j}, x_{1,j+\tau_1}, \ldots, x_{1,j+(m_1-1)\tau_1} \\ x_{2,j}, x_{2,j+\tau_2}, \ldots, x_{2,j+(m_2-1)\tau_2} \\ \cdots, \\ x_{M,j}, x_{M,j+\tau_M}, \ldots, x_{M,j+(m_M-1)\tau_M} \end{array}\right) \quad (j = 1, 2, \ldots, N) \end{aligned}$$

where m_i and $\tau_i (i = 1, 2, \ldots, M)$ are the embedding dimension and time delay of the i_{th} time series X_i; $N = \min\left(n - (m_i - 1)\,\tau_i\right)$ is the total number of phase points in the reconstructed space. Stretch $\boldsymbol{V}_j, j = 1, 2, \ldots, N$ into a phase point with m coordinates:

$$\begin{aligned} \boldsymbol{V}_j = & \left(x_{1,j}, x_{1,j+\tau_1}, \ldots, x_{1,j+(m_1-1)\tau_1},\right. \\ & x_{2,j}, x_{2,j+\tau_2}, \ldots, x_{2,j+(m_2-1)\tau_2}, \\ & \cdots, \\ & \left. x_{M,j}, x_{M,j+\tau_M}, \ldots, x_{M,j+(m_M-1)\tau_M}\right) \\ = & \left(v_{j,1}, v_{j,2}, \cdots, v_{j,m}\right) \qquad (j = 1, 2, \ldots, N) \end{aligned} \tag{19}$$

where $m = \sum_{i=1}^{M} m_i$ is the dimension of the reconstructed phase space. Therefore, the reconstructed vector is $\boldsymbol{V} = (\boldsymbol{V_1}, \boldsymbol{V_2}, \ldots, \boldsymbol{V_N})$, $\boldsymbol{V_j}$ represents the j_{th}

state of the reconstructed attractor.

Comparing with the vector reconstructed from single time series, $\boldsymbol{V}$ not only contains the information of the data to be analyzed and predicted, but also involves the information of the related variables. Therefore, each state point $\boldsymbol{V}_j$ can be considered as a microsystem which reflects the state of the data and related factors at evolution time t_i. As time evolves, the evolution trajectory reveals the variation of the system state. Figure 3 shows the process of the reconstruction. The example contains three wind power time series, $W1$, $W2$, and $W3$, whose embedding dimensions are equal to three, and they are reconstructed into a nine-dimensional phase space, where $v_1 - v_9$ represent the coordinates of the reconstructed vectors. Take point $\boldsymbol{V_j}$ as an example. $v_{j,q}(q = 1, 2, \ldots, 9)$ is the value of its q_{th} coordinate, and q just represents the coordinate, not the index of this value in its original time series. Each point $\boldsymbol{V}_j(j = 1, 2, \ldots, N)$ is defined by nine coordinates, and they contain the information of three time series.

Further, the selection of the NNs in the space is discussed. For a state point $\boldsymbol{V_p}$, its k NNs are selected based on Euclidean distance. In this way, several points whose states of wind power data and corresponding factors are similar can be obtained. Then these points are delivered to the forecasting model to predict the following state. Note that the prediction performance is strongly related to the number of NNs, k, and the determination of k will be discussed in detail in Section 4.3.

3.2. Low-dimensional phase space reconstruction (LDPSR)

As discussed above, the dimension of the high-dimensional reconstructed attractor is the sum of the dimensions of all the time series involved; thus, the high-dimensional PSR method has no strict requirement on the dimensions of each time series involved. However, for low-dimensional PSR, the dimensions of all the time series are expected to be equal, so that they can be reconstructed into a phase space with this dimension. The algorithm of this method is similar to the PSR of univariate time series, where the critical difference is that this

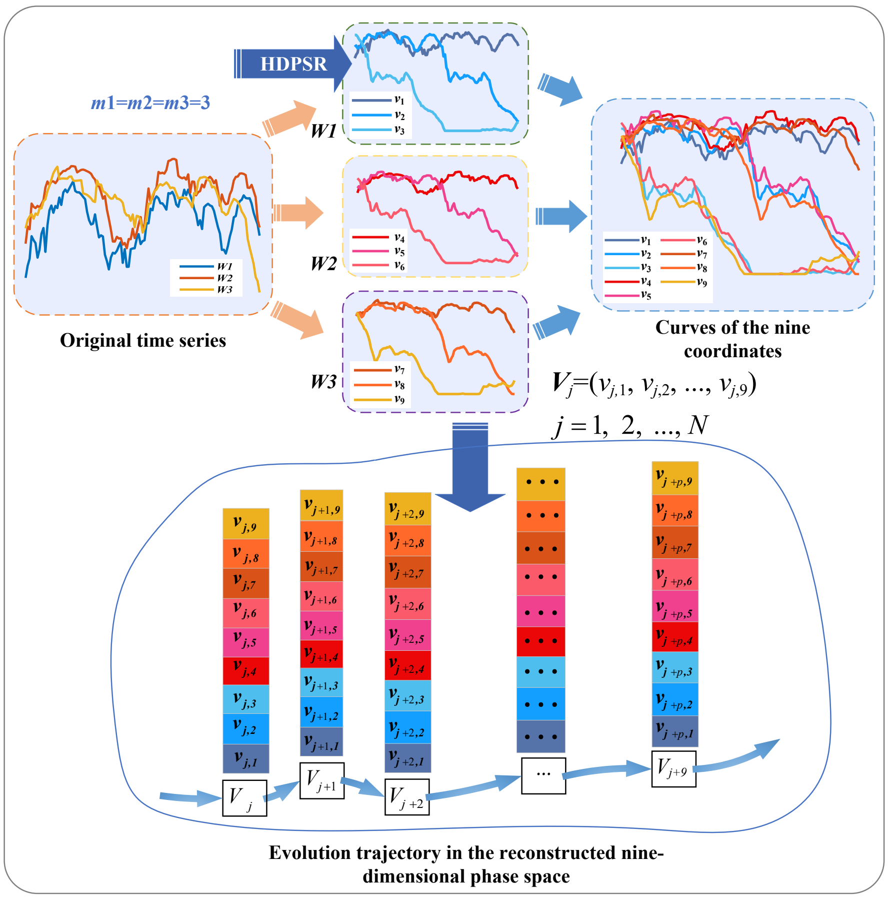

Figure 3: The process of high-dimensional phase space reconstruction (HDPSR)

method aims to expand the phase space and obtain more similar state points, which can help to evaluate the evolution of the data. The process can be described as follows.

For M experimental time series, $\{X_1, X_2, \cdots, X_i, \cdots, X_M\}$, $X_i = (x_{i,1}, x_{i,2}, \cdots, x_{i,j}, x_{i,n})$, $i = 1, 2, \cdots, M$, estimate their chaos-related indicators, including the LLE, embedding dimension, and time delay. Since each reconstructed point represents a state of the phase space, the time series involved should satisfy the following two conditions: (1) the magnitude of their LLE is close, which indicates that their chaotic extent is similar; (2) their embedding dimensions are the same.

Then these M time series are reconstructed into phase space, respectively:

$$\begin{aligned} \boldsymbol{V}_i &= (\boldsymbol{V}_{i,1}, \boldsymbol{V}_{i,2}, \ldots, \boldsymbol{V}_{i,N_i}) \\ &= \begin{pmatrix} x_{i,1}, x_{i,1+\tau_i}, \cdots, x_{i,1+(m-1)\tau_i} \\ x_{i,2}, x_{i,2+\tau_i}, \cdots, x_{i,2+(m-1)\tau_i} \\ \cdots \\ x_{i,N_i}, x_{i,N_i+\tau_i}, \cdots, x_{i,N_i+(m-1)\tau_i} \end{pmatrix} i = 1, 2, \cdots, M \end{aligned} \tag{20}$$

where $\boldsymbol{V}_i$ is the reconstructed attractor of the i_{th} time series, τ_i and m are its time delay and embedding dimension, respectively, $\boldsymbol{V}_{i,j} = (x_{i,j}, x_{i,j+\tau_i}, \cdots, x_{i,j+(m-1)\tau_i})$, $j = 1, 2, \cdots, N_i$ is the j_{th} reconstructed point of $\boldsymbol{V}_i$, and $N_i = n - (m-1)\tau_i$ is the total number of the reconstructed phase points. Since these time series have different time delays, they will have different numbers of reconstructed points. The embedding dimensions of all the M time series are equal to m; in other words, they are reconstructed into a phase space with dimension m. Although these phase points are from different original time series, they have similar chaotic characteristics and evolution trends, and thus, they can provide more state information about this attractor.

Therefore, the related reconstructed vectors can be integrated into an identical space $\boldsymbol{V}$, defined as

$$\boldsymbol{V} = \{\boldsymbol{V}_1, \boldsymbol{V}_2, \cdots, \boldsymbol{V}_M\} \tag{21}$$

where the number of phase points in $\boldsymbol{V}$ is $m = \sum_{j=1}^{M} m_j$. In this way, an expanded phase space is obtained.

The process of LDPSR is displayed in Figure 4. Same as Figure 3, three wind power time series $W1$, $W2$, and $W3$ are reconstructed into a multi-dimensional phase space. After LDPSR, they are reconstructed to their attractors $\boldsymbol{V}_1$, $\boldsymbol{V}_2$ and $\boldsymbol{V}_3$, respectively, where $v_{j,1}-v_{j,3}$ (j=1,2,3) represent the coordinates of their reconstructed vectors, j is the index of the time series variable. Integrate the three vectors and obtain an expanded phase space $\boldsymbol{V}$, the phase points in which are defined as $\boldsymbol{V}_{j,i} = (v_{j,i,1}, v_{j,i,2}, v_{j,i,3})$, $j = 1, 2, 3$, $i = 1, 2, \cdots, N_j$. Different from HDPSR, reconstructed attractors in $\boldsymbol{V}$ in LDPSR are independent with one another, they have their own evolution trajectories, although their chaotic characteristics are similar.

Next, we discuss how to select the nearest neighbors (NNs) in such an expanded space. For a phase point $\boldsymbol{V}_{j,p}\,(j = 1, 2, \cdots, M; p = 1, 2, \cdots, N_j)$, look for its k nearest points based on the Euclidean distance criterion in the integrated phase space $\boldsymbol{V}$. These chosen points not only include points from $\boldsymbol{V}_j$ where $\boldsymbol{V}_{j,p}$ is in, but also involve some from other reconstructed vectors $\boldsymbol{V}_i\,(i \neq j)$. Compared to the selection of NNs in univariate phase space, the advantage of this method is: under the condition that the number of NNs is equal, this method can gain more points whose states are similar to $\boldsymbol{V}_{j,p}$. This is because in a reconstructed univariate phase space, as the spatial distance of two points increases, the difference of their states will become larger. Therefore, it can only provide a limited number of points whose state similarity with $\boldsymbol{V}_{j,p}$ is in a certain range; whereas, the integrated multivariate phase space $\boldsymbol{V}$, reconstructed from several related time series, can provide more points that meet the requirement on state similarity.

3.3. Validation of the multivariate reconstructed phase space

In our research, multiple variables are reconstructed into the same phase space, and the state of each point in the phase space will be determined by these variables. Further, the evolution of these phase points will help implement the

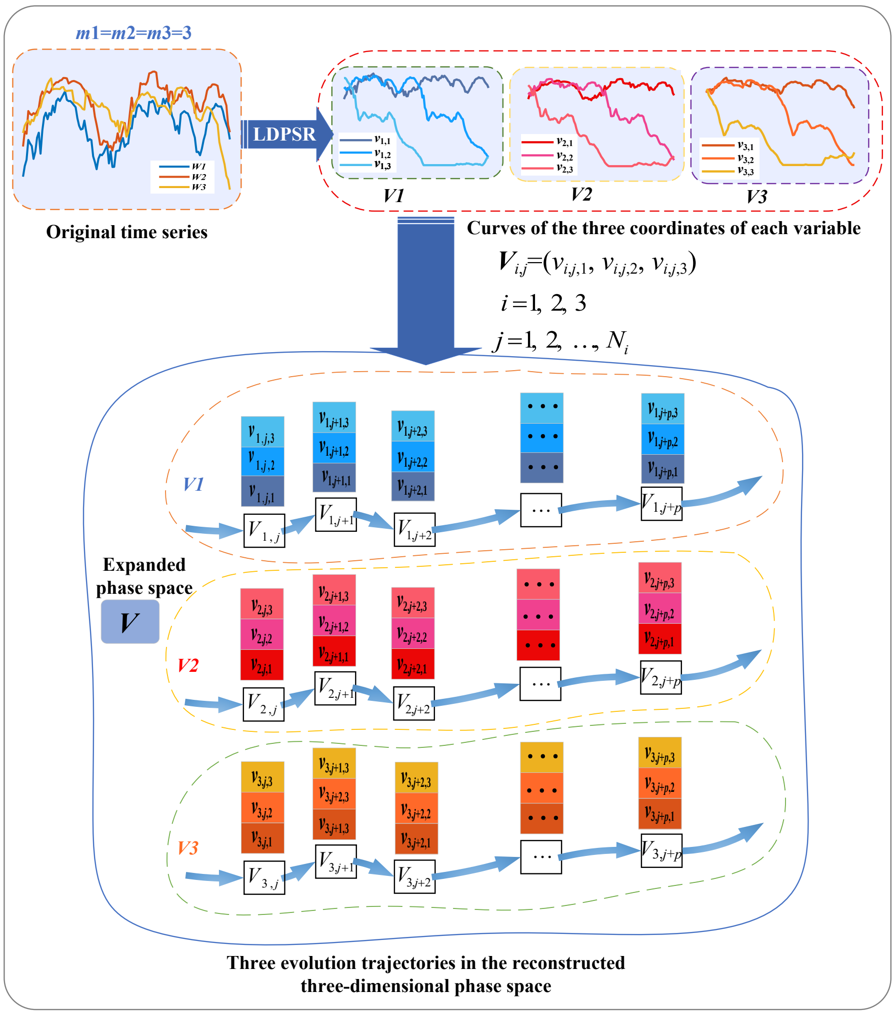

Figure 4: The process of low-dimensional phase space reconstruction (LDPSR).

prediction. Therefore, it is necessary to verify that the chaotic characteristics of the reconstructed phase space are similar to those of the original signals. During the validation process, the LLE is selected to estimate the chaotic extent of the reconstructed vectors. Here some adjustments to the LLE calculation algorithm are made based on the reconstructed phase space.

In the low-dimensional reconstructed phase space, each reconstructed vector is considered as an evolution trajectory. To calculate the LLE, the separation of adjacent trajectories should be analyzed. The calculation method is explained as follows.

First, for each reconstructed vector $\boldsymbol{V}_i$ $(i = 1, 2, \cdots, M)$, look for the nearest neighbor of its first state point $\boldsymbol{V}_{i,1}$ in the other reconstructed vectors, identify the adjacent trajectory $\boldsymbol{V}_j$ $(j \neq i)$ where the nearest neighbor point is located, and calculate the LLE based on (15) and (16). Note that the difference between this validation algorithm and the original LLE estimation algorithm is that, the pair of nearest neighbors come from different reconstructed vectors. Therefore, the time interval restriction defined in (14) is not needed any more. The calculation result is denoted as λ_i, $i = 1, 2, \cdots, M$.

Take the mean value of the calculation results as the LLE of the reconstructed phase space, which is defined as

$$\lambda = \frac{1}{M} \sum_{j=1}^{M} \lambda_j \tag{22}$$

The magnitude of λ reflects the chaotic extent of the reconstructed phase space.

In the high-dimensional reconstructed phase space, the calculation of LLE is consistent with that described in (13)-(16), but the parameters need to be chosen properly. For the embedding dimension, it is selected as the summary of the dimensions of each variable involved, *i.e.*, $m = \sum_{j=1}^{M} m_j$; and for the time delay, to avoid possible trajectory crossover problems, select $\tau = \max(\tau_i)$, $i = 1, 2, \cdots, M$.

3.4. The proposed model

Based on the algorithms discussed above, a short-term multivariate wind power time series forecasting model is proposed, as is shown in Figure 5. This model consists of three main modules:

1. Chaotic characteristics analysis. This procedure aims to select the factors that have strong relationships with the data to be analyzed and forecasted. The Cao and mutual information methods are used to calculate the embedding dimension and time delay of the time series, and the Cao and Wolf methods are used to analyze whether a time series is chaotic and to estimate its LLE to assess its chaotic extent. Time series that have similar chaotic characteristics (especially LLE) are considered as related data and will be grouped. According to their embedding dimension, these groups will be classified into two categories: time series with equal dimensions and time series with different dimensions.
2. Multivariate time series phase space reconstruction and model validation. For the groups of time series that have different embedding dimensions, implement HDPSR on them and reconstruct them into high-dimensional phase space. And for time series that have equal embedding dimensions, implement LDPSR on them and reconstruct them into low-dimensional but expanded phase space.
3. Nearest neighbor selection and prediction. Based on the principle of minimum error, a proper value of the number of training points, as well as the number of nearest neighbors are chosen. Finally, the chosen input points are delivered into a well-constructed forecasting model, such as LSSVM, to implement the forecasting.

4. Case study

4.1. Experiment setting

In this section, we select wind power data from several neighboring wind farms in Michigan to validate the proposed model and apply it to short-term

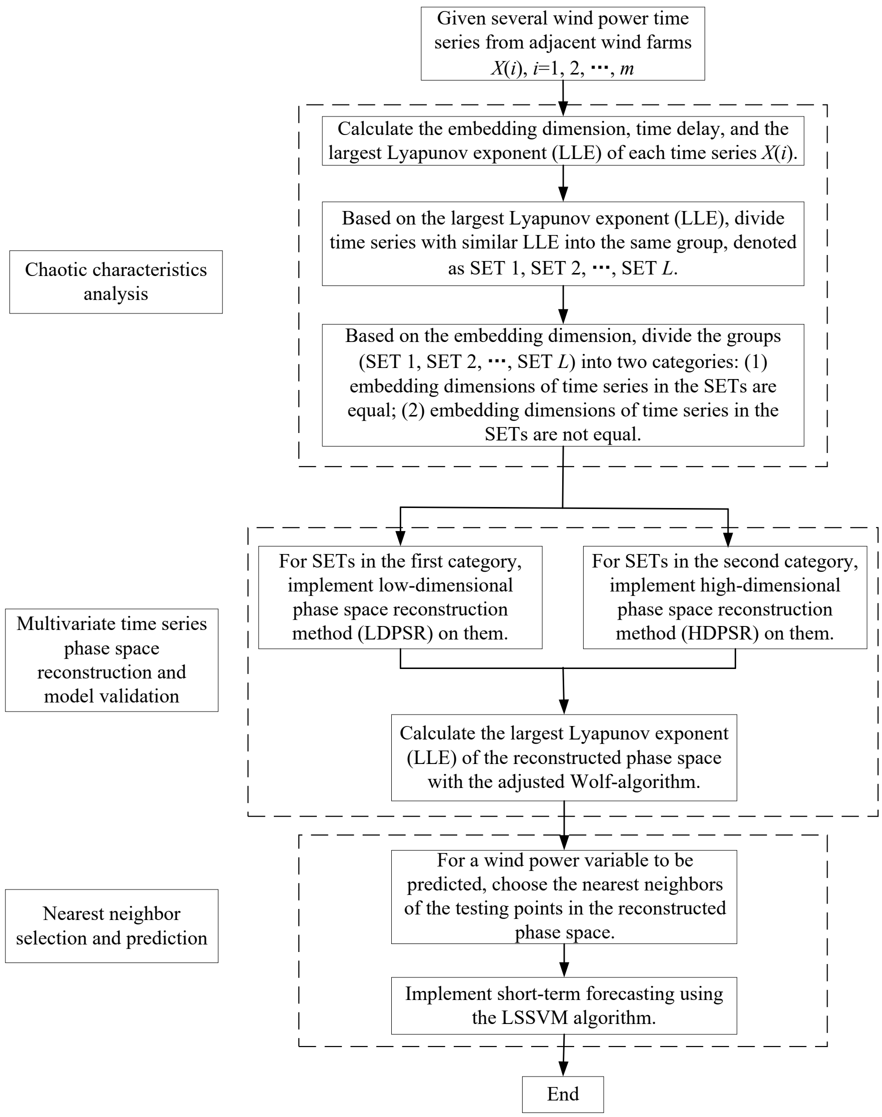

Figure 5: Flowchart of the proposed model.

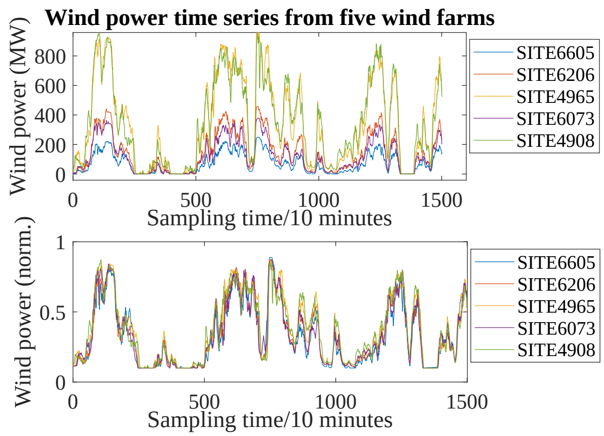

Figure 6: Wind power time series from five wind farms. The upper one shows the original wind power time series, and the lower one shows the normalized ones.

forecasting. Affected by geographical environment factors, wind sites that are close to one another may have some similar features in wind power data. Five wind sites are chosen to carry out the experiments, which are Site 6605, Site 6206, Site 4965, Site 6073 and Site 4908. Their locations are shown in Table 1. For each wind site, wind power data for a whole year are available, and the data in January, 2005 are used to carry out the validation. The data sampling interval is 10 minutes. Data curves of the original and the normalized time series are shown in Figure 6, where the normalized range is $[0.1, 0.9]$. We can see that the dynamic trends of these five wind power curves are similar, which meets the basic requirements of multivariate phase space reconstruction.

In our research, the least-square support vector machine (LSSVM) model is selected as the forecasting model, and it will be considered as a benchmark to evaluate the prediction performance of the proposed LSSVM-LDPSR model and LSSVM-HDPSR model. We will carry out 1-step ahead (10 minutes), 3-step ahead (30 minutes), and 6-step ahead (1 hour) forecast on the data. The forecasting performance is measured by $MAPE$ and $NRMSE$, which assess the forecasting accuracy and model stability, respectively. They are defined as

Table 1: Locations of the five wind sites

Site	Latitude	Longitude
6605	43°3′N	82°67′W
6206	43°39′N	82°75′W
4965	43°42′N	82°63′W
6073	43°38′N	82°86′W
4908	43°62′N	82°7′W

follows:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}|\frac{y_i - \hat{y}_i}{y_i}| \times 100 \tag{23}$$

$$NRMSE = \frac{1}{Y}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \times 100 \tag{24}$$

where n is the number of testing points; y_i and $\hat{y}_i$ are the prediction value and actual value, respectively; Y is the installed capacity of the wind farm.

4.2. Characteristics analysis

The chaotic characteristics of the time series from each wind site are tested and the results are shown in Table 2. It shows that the LLE of Site 6206, Site 4965 and Site 4908 are similar, within the range of 0.15 and 0.165, and that of Site 6605 and Site 6073 are similar, within the range of 0.075 and 0.09. Therefore, based on the LLE and PSR parameters, the time series are divided into two groups: (1) Site 6206, Site 4965, and Site 4908; (2) Site 6605 and Site 6073. For the first group, they have similar LLE, similar time delay (τ) and equal embedding dimension (m), so both LDPSR and HDPSR can be implemented on them. For the second group, they can only be reconstructed into high-dimensional phase space.

Based on the above analysis, three datasets are selected to carry out the prediction: (1) two-dimensional wind power time series from Site 6206 and Site 4965 (SET 1); (2) three-dimensional wind power time series from Site 6206, Site 4965, and Site 4908 (SET 2); (3) two-dimensional wind power time series from

Table 2: LLE and PSR parameters of wind power time series

Site	LLE	Embedding dimension (m)	Time delay (τ)
6605	0.0786	5	39
6206	**0.1648**	**3**	**56**
4965	**0.1572**	**3**	**57**
6073	0.0894	4	37
4908	**0.1537**	**3**	**55**

Site 6605 and Site 6073 (SET 3). Before prediction, we will discuss the selection of the training data for the forecasting model and nearest neighbors for the data to be predicted in the following subsection.

4.3. The determination of the numbers of training data points and nearest neighbors (NNs)

For a forecasting model, the number of training points has a great influence on the effect of model training, and the nearest neighbors selected will affect the data regression and fitting, thus directly influence the prediction performance. Denote the number of training points as L and the number of nearest neighbors as k. In our experiments, we have found that the prediction evaluation indicators ($MAPE$ and $NRMSE$) show different change trends with the variation of k and L. Moreover, $MAPE$ and $NRMSE$ show a certain trend of change with the variation of the combinations of k and L.

Wind power data from Site 6206 are selected as testing example, and we discuss the performance of the LSSVM-HDPSR model. Conditions of LSSVM, LSSVM-LDPSR models and data from other wind sites are similar. Three groups of testing have been experimented, analyzing the influence of k, L and the combination of k and L (denoted as kL) on $MAPE$ and $NRMSE$ in 1-step ahead (10 minutes), 3-step ahead (30 minutes), 6-step ahead (1 hour), and 12-step ahead (2 hours) prediction, respectively. The results are shown in Figure

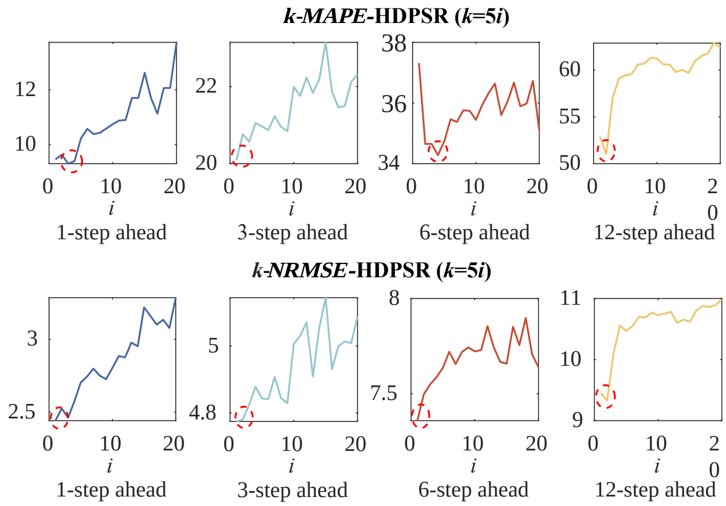

Figure 7: The variation of $MAPE$ and $NRMSE$ with the increase of k in different steps ahead prediction.

7 $\sim$ Figure 9.

To avoid excessive calculations, not the results of all the values of the parameters are tested. Instead, the time series of k and L are defined as $k = 5i\,(i = 1, 2, \cdots, 20)$ with an interval of 5 and $L = 100+10i\,(i = 1, 2, \cdots, 45)$ with an interval of 10. In each figure, the minimum value of the curves is circled with a red ellipse. In Figure 7, the horizontal axis is i ($k = 5i$). It shows that the prediction achieves better performance when k is small, and as k increases, the mean trend of $MAPE$ and $NRMSE$ increases generally. This indicates that there's no need to choose too many NNs when experimenting short-term forecasting. Keeping the number of k within the range of $[10, 15]$ is proper.

The change trend of $MAPE$ and $NRMSE$ with the variation of L is shown in Figure 8. The horizontal axis in this figure is also $i\,(L = 100 + 10i)$. The curves indicate that in 1-step and 3-step ahead forecasting, the prediction obtains better performance when i is larger than 30 (L>400). And in 6-step ahead forecasting, both $MAPE$ and $NRMSE$ obtain the smallest value when $i = 5\,(L = 140)$, while in 12-step ahead forecasting they both get the best results

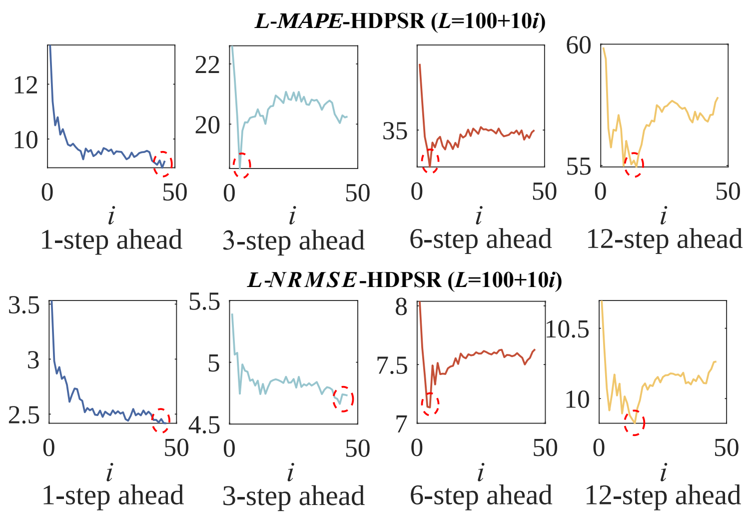

Figure 8: The variation of $MAPE$ and $NRMSE$ with the increase of L in different steps ahead prediction.

385 when $i = 14\,(L = 230)$. This implies that forecasts of different time scales have different requirements for the number of training points (L).

Next, we discuss how the evaluation indicators vary under different combinations of k and L. For k, keep it within the range of [10,15], and L is still defined as $L = 100 + 10i$, $(i = 0, 1, 2, \cdots, 45)$ thus can obtain 276 pairs of $[k, L]$,
390 $\{[10, 100], \cdots, [10, 550], [11, 100], \cdots, [11, 550], \cdots, [15, 550]\}$. The horizontal axis in this figure is set as $p\,(p = 1, 2, \cdots, 276)$, each point on the horizontal axis represents a pair of $[k, L]$ Testing results are displayed in Figure 9. The variation of both $MAPE$ and $NRMSE$ appears periodic, and the sudden changes in the curves appear at the time when the value of k changes. When k remains
395 the same, with the increase of L, $MAPE$ and $NRMSE$ change in a similar trend, leading to the periodic curves. Figure 9 shows that different values of k in this range do not result in obvious differences in the performance of $MAPE$ or $NRMSE$. Therefore, it is not necessary to further subdivide the scale of $[k, L]$. In the following forecasting analysis, we first determine a proper scale of
400 k and L, respectively, then utilize the principle of the minimum error to select

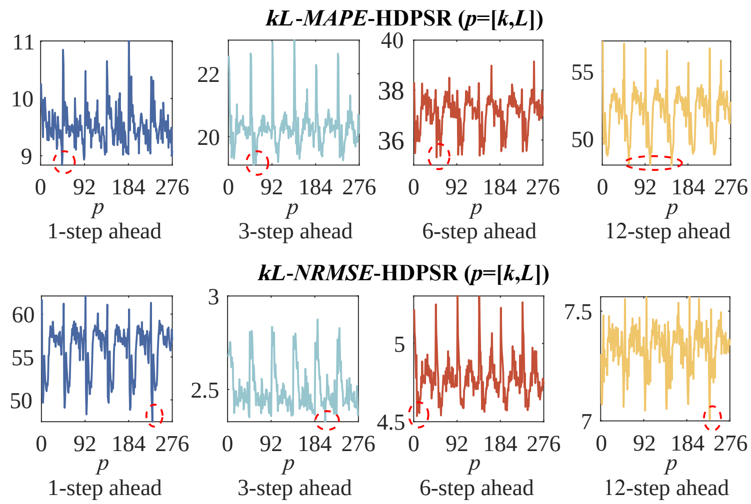

Figure 9: The variation of $MAPE$ and $NRMSE$ with the increase of kL in different steps ahead prediction.

the optimal combination of k and L.

4.4. Forecasting of two-dimensional wind power data from Site 6206 and Site 4965 (SET 1)

According to Table 2, these two time series have equal embedding dimensions $m = 3$. Both LDPSR and HDPSR are implemented to test and compare their performance, thus they are reconstructed into a three-dimensional phase space and a six-dimensional phase space, respectively. After reconstruction, the optimal combinations of k (the number of the nearest neighbors of each point to be predicted) and L (the number of training data for the forecasting model) for each time series are selected using the principle of minimum error described in Section 4.3. Next, the selected input samples (the nearest neighbors and training data) are delivered to the forecasting model LSSVM to implement the prediction. The forecasting performance is shown in Table 3. Moreover, to visually observe the prediction effect of the models, the results of $MAPE$ and $NRMSE$ are displayed in histograms in Figure 10.

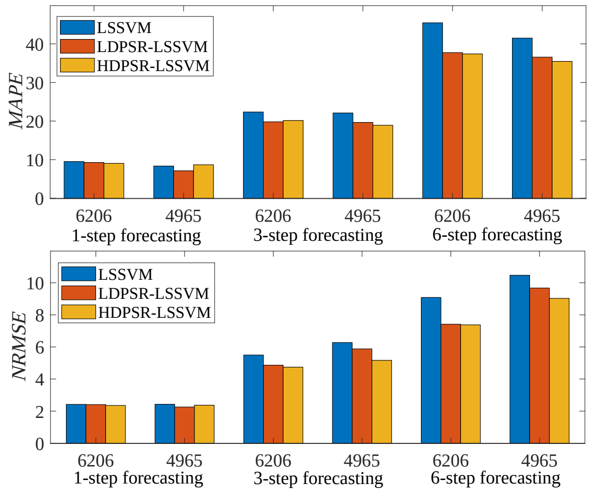

Figure 10: Prediction performance of different models on SET 1.

In Table 3, the smallest value of $MAPE$ and $NRMSE$ is bold, indicating the best forecasting performance. As the results show, both LDPSR-LSSVM and HDPSR-LSSVM can achieve better performance than LSSVM in general. In 1-step ahead forecasting, their improvement effect is not obvious. Nonetheless, as the look-ahead step increases, the advantages of multivariate PSR gradually show. For $MAPE$, in 3-step ahead forecast, the LDPSR-LSSVM and HDPSR-LSSVM models improve 11.39% and 9.84% on Site 6206, respectively; and improve 11.06% and 14.39% on Site 4965. In 6-step ahead forecasting, the improvements increase to 16.99% and 17.73% on Site 6206, and 11.87% and 14.52% on Site 4965, respectively. As for $NRMSE$, they both achieve better performance than LSSVM for all the three scales of forecasting, which indicates that these models can maintain good forecasting stability when reconstructed into an expanded phase space or into higher-dimensional space.

Comparing the performance of LDPSR-LSSVM and HDPSR-LSSVM models, it is found that HDPSR-LSSVM can achieve higher prediction accuracy ($MAPE$) and better prediction stability ($NRMSE$) than LDPSR-LSSVM. Especially in terms of stability, the advantages of HDPSR are more obvious: in the three-step and six-step prediction, HDPSR-LSSVM can achieve better stability than LDPSR-LSSVM. In this case, the low-dimensional phase space is three-dimensional, while the high-dimensional phase space is six-dimensional. In the high-dimensional phase space, each phase point contains the reconstruction information of two variables, while phase points in the low-dimensional phase space only contain the information of one variable, but the number of phase points in this space is more. Therefore, it can be inferred that a single phase point containing the information of all variables is more helpful to improve the stability of the prediction model than multiple phase points with the information of individual variable, especially when the prediction step is long (e.g. longer than three steps), the requirements for model stability will be higher. What's more, HDPSR-LSSVM has better stability effect under such circumstances, it is also proved that the chaotic system composed of phase points containing all variables can better track the dynamic evolution law of the points to be predicted and realize the prediction effect at a longer forecasting scale. This is the condition with two variables. Next, data from three variables will be tested to see how the two models perform as the number of variables increases (and, correspondingly, the reconstructed phase space dimension increases).

4.5. Forecasting of three-dimensional wind power data from Site 6206, Site 4965, and Site 4908 (SET 2)

Based on the above analysis on SET 1, time series from Site 4908 is added to form a three-dimensional model. As Table 2 shows, these three time series have similar LLE (0.1648, 0.1572, 0.1537) and time delay (56, 57, 55), and their embedding dimensions are equal ($m = 3$). Implement LDPSR and HDPSR on them, and they are reconstructed into a three-dimensional and a nine-dimensional phase space, respectively. Similarly, select the nearest neighbors and training data for the points to be predicted in the reconstructed phase space based on the principle of minimum error described in Section 4.3, and deliver them to the forecasting model LSSVM. Their prediction performance is displayed in Table 4 and Figure 11.

Table 3: Forecast performance of the proposed models on SET 1 (%)

Prediction scale (m)	Evaluation indicator	Site	LSSVM	LDPSR-LSSVM	HDPSR-LSSVM
1-Step	$MAPE$	6206	9.5062	9.2749	**9.0412**
		4965	8.3438	**7.1292**	8.6736
	$NRMSE$	6206	2.4184	2.4065	**2.3516**
		4965	2.4265	**2.2576**	2.3696
3-Step	$MAPE$	6206	22.3460	**19.8002**	20.1473
		4965	22.0936	19.6505	**18.9147**
	$NRMSE$	6206	5.4966	4.8640	**4.7414**
		4965	6.2732	5.8781	**5.1645**
6-Step	$MAPE$	6206	45.4564	37.7355	**37.3949**
		4965	41.4899	36.5637	**35.4673**
	$NRMSE$	6206	9.0853	7.4156	**7.3762**
		4965	10.4678	9.6729	**9.0283**

The bold number indicates the best performance.

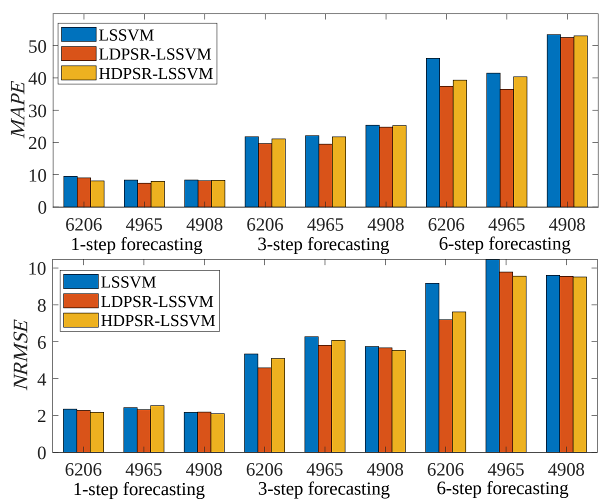

Figure 11: Prediction performance of different models on SET 2.

Table 4: Forecast performance of the proposed models on SET 2 (%)

Prediction scale (m)	Evaluation indicator	Site	LSSVM	LDPSR-LSSVM	HDPSR-LSSVM
1-Step	$MAPE$	6206	9.5069	9.0172	**8.0754**
		4965	8.3438	**7.3778**	7.9567
		4908	8.3596	**8.1314**	8.2458
	$NRMSE$	6206	2.3459	2.2761	**2.1696**
		4965	2.4265	**2.3169**	2.5326
		4908	2.1680	2.1857	**2.0963**
3-Step	$MAPE$	6206	21.7646	**19.6587**	21.0980
		4965	22.0936	**19.4812**	21.7329
		4908	25.3540	**24.7523**	25.2351
	$NRMSE$	6206	5.3367	**4.5840**	5.0914
		4965	6.2732	**5.8115**	6.0772
		4908	5.7404	5.6695	**5.5320**
6-Step	$MAPE$	6206	46.0630	**37.4277**	39.3157
		4965	41.4899	**36.4896**	40.3506
		4908	53.4156	**52.5466**	53.0351
	$NRMSE$	6206	9.1750	**7.1945**	7.6181
		4965	10.4678	9.7823	**9.5590**
		4908	9.6052	9.5509	**9.5140**

The bold number indicates the best performance.

Like the two-dimensional model, the LDPSR-LSSVM and HDPSR-LSSVM can achieve smaller $MAPE$ than LSSVM, and their advantages are more obvious in 3-step ahead and 6-step ahead forecast. It is noticed that for data from Site 6206 and Site 4965, the three-dimensional model can obtain relatively good performance in $MAPE$, especially for the 1-step ahead forecast, but as $NRMSE$ shows, the stability of this model in 1-step ahead forecast is a little weaker than the two-dimensional model, which indicates that the ultra-short-term prediction stability may be influenced by the high dimension or expanded space with multiple state variables. However, in 3-step and 6-step ahead forecast, LDPSR-LSSVM and HDPSR-LSSVM still show their advantages in $NRMSE$.

Further, the performance of the two-dimensional and three-dimensional models is compared. For the two-dimensional model, LDPSR-LSSVM and HDPSR-LSSVM have similar performances in $MAPE$ and $NRMSE$, and the HDPSR slightly defeats LDPSR; whereas, for the three-dimensional model, the overall performance of LDPSR is better than HDPSR, which means it can achieve better prediction accuracy, and maintain acceptable stability at the same time. This indicates that when analyzing multiple variables, if their total dimensions are relatively small (in our research, for example, no more than nine), then each state point in the high-dimensional space can provide more data information, thus can obtain satisfactory prediction results. But if their embedding dimensions are too high, the state points contain too much information, so the difference between them and the variable to be analyzed will increase, which will in turn affect the prediction performance. In this condition, it is suitable to choose the low-dimension phase space reconstruction method.

4.6. Forecasting of two-dimensional wind power data from Site 6605 and Site 6073 (SET 3)

The embedding dimensions of these two variables are higher than the other three variables and are not equal, so we implement HDPSR on them, then they are reconstructed into a nine-dimensional phase space. Table 5 and Figure

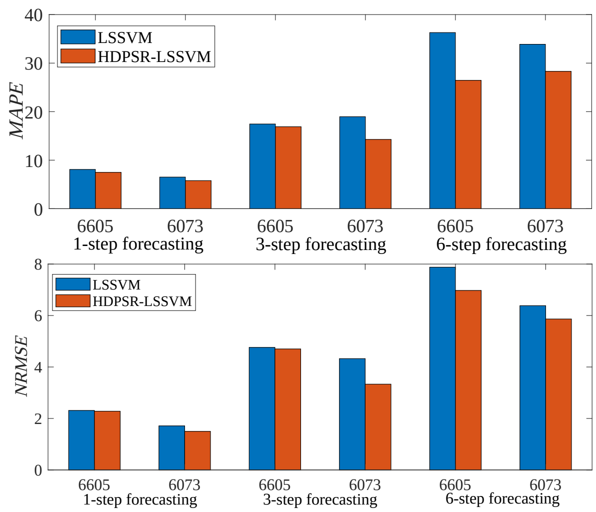

Figure 12: Prediction performance of different models on SET 3.

12 show the prediction results. As the figures show, HDPSR-LSSVM achieves better performance in prediction accuracy and stability than LSSVM in 1-step ahead, 3-step ahead and 6-step ahead forecast. In 1- step ahead forecast, $MAPE$ for Site 6605 and Site 6073 are improved by 7.39% and 11.21%; in 3-step ahead forecast, the improvements are 3.21% and 24.70%; in 6-step ahead forecast, these improvements reach 27.11% and 16.45%. And the $NRMSE$ also maintains a good level. As the look-ahead step increases, advantages of multivariate PSR gradually show. That is because as the prediction scale gets longer, the forecasting model needs to resolve more data information, which is exactly what the multivariate PSR model can provide.

4.7. Discussion

The simulation results show that the proposed multivariate phase space spatial reconstruction method can achieve satisfactory effect in analyzing wind power time series from adjacent wind farms. Based on the chaotic characteristics of several variables, we reconstruct them into one phase space as a highly connected system, then select training points in this system to implement the prediction. What's more, we give a chaotic analysis of the reconstructed phase

Table 5: Forecast performance of the proposed models on SET 3 (%)

Prediction scale (m)	Evaluation indicator	Site	LSSVM	HDPSR-LSSVM
1-Step	$MAPE$	6605	8.0872	**7.4896**
		6073	6.5133	**5.7834**
	$NRMSE$	6605	2.3114	**2.2810**
		6073	1.7128	**1.4986**
3-Step	$MAPE$	6605	17.4517	**16.8918**
		6073	18.9569	**14.2752**
	$NRMSE$	6605	4.7614	**4.7017**
		6073	4.3248	**3.3328**
6-Step	$MAPE$	6605	36.2936	**26.4528**
		6073	33.8811	**28.3063**
	$NRMSE$	6605	7.8778	**6.9750**
		6073	6.3828	**5.8650**

The bold number indicates the best performance.

space, and the calculation results of LLE are displayed in Table 6. As we can see, the chaotic extent of reconstructed phase space is lower compared to the original individual variables, which is possible, since the interaction between multiple variables may lead to this result. However, the decrease of LLEs is not too obvious, they are still in the same level of magnitude, indicating that our reconstruction method is effective. Moreover, each variable is predicted based on the reconstructed multivariate space, and the results prove to be effective. Therefore, we believe that for wind power time series from adjacent wind farms which have similar chaotic dynamic evolutions, they can be analyzed as an integrated system, and the multivariate information provided by this system can be utilized to support the prediction, achieving the improvement of prediction performance.

Table 6: LLE of the reconstructed LDPSR and HDPSR of three SETs

LLE	SET 1	SET 2	SET 3
LDPSR	0.1236	0.1150	–
HDPSR	0.1126	0.0962	0.0466

5. Conclusion

In this paper, two multivariate time series phase space reconstruction methods based on chaos theory are proposed to analyze the chaotic characteristics of wind power time series from adjacent wind farms and to help implement the prediction. Our research includes two main parts: (1) chaos signal recognition and chaos extent estimation. We utilized the Cao method to distinguish between deterministic signals and random signals. And we improved Wolf's algorithm for calculating the largest Lyapunov exponent to estimate the chaotic extent of wind power time series; (2) based on the chaos analysis, we carried out the multivariate low-dimensional and high-dimensional phase space reconstruction methods (LDPSR and HDPSR) and predicted wind power time series from adjacent wind farms.

The results in the experiments show that: (a) the multivariate phase space reconstruction methods proposed can analyze and help predict the wind power time series from adjacent wind power sites well; (b) the proposed LDPSR and HDPSR models can ensure that the chaotic extent of reconstructed vectors remains the same level as that of the involved original time series; (c) the improved Wolf algorithm can effectively estimate the largest Lyapunov exponent of wind power time series and can be used to verify the chaotic characteristics of the reconstructed vectors. These results indicate that we have successfully predicted multivariate wind power time series using the chaotic dynamics analysis method. In the future, we will further study the chaotic dynamics of wind power data. In addition, we will discuss the application of the proposed multivariate analysis methods in other types of data, including wind speed data and other meteorological data that affect wind power generation.

References

[1] WWEA, Worldwide wind capacity reaches 744 gigawatts – an unprecedented 93 gigawatts added in 2020, http://wwindea.org/worldwide-wind-capacity-reaches-744-gigawatts/ (2021).

[2] S. Al-Yahyai, Y. Charabi, A. Gastli, Review of the use of numerical weather prediction (nwp) models for wind energy assessment, Renewable and Sustainable Energy Reviews 14 (9) (2010) 3192–3198.

[3] S. S. Soman, H. Zareipour, O. Malik, P. Mandal, A review of wind power and wind speed forecasting methods with different time horizons, in: North American Power Symposium 2010, 2010, pp. 1–8. doi:10.1109/NAPS.2010.5619586.

[4] K. Yunus, T. Thiringer, P. Chen, Arima-based frequency-decomposed modeling of wind speed time series, IEEE Transactions on Power Systems 31 (4) (2016) 2546–2556.

[5] E. Erdem, J. Shi, Arma based approaches for forecasting the tuple of wind speed and direction, Applied Energy 88 (4) (2011) 1405–1414.

[6] W. Fu, K. Wang, C. Li, J. Tan, Multi-step short-term wind speed forecasting approach based on multi-scale dominant ingredient chaotic analysis, improved hybrid gwo-sca optimization and elm, Energy Conversion and Management 187 (2019) 356–377.

[7] S. Salcedo-Sanz, Ángel M. Pérez-Bellido, E. G. Ortiz-García, A. Portilla-Figueras, L. Prieto, D. Paredes, Hybridizing the fifth generation mesoscale model with artificial neural networks for short-term wind speed prediction, Renewable Energy 34 (6) (2009) 1451–1457.

[8] R. Ak, O. Fink, E. Zio, Two machine learning approaches for short-term wind speed time-series prediction, IEEE Transactions on Neural Networks and Learning Systems 27 (8) (2016) 1734–1747. doi:10.1109/TNNLS.2015.2418739.

[9] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, K. P. Wong, Probabilistic forecasting of wind power generation using extreme learning machine, IEEE Transactions on Power Systems 29 (3) (2014) 1033–1044. doi:10.1109/TPWRS.2013.2287871.

[10] Q. Hu, S. Zhang, M. Yu, Z. Xie, Short-term wind speed or power forecasting with heteroscedastic support vector regression, IEEE Transactions on Sustainable Energy 7 (1) (2016) 241–249. doi:10.1109/TSTE.2015.2480245.

[11] Y. Ren, P. N. Suganthan, N. Srikanth, A novel empirical mode decomposition with support vector regression for wind speed forecasting, IEEE Transactions on Neural Networks and Learning Systems 27 (8) (2016) 1793–1798. doi:10.1109/TNNLS.2014.2351391.

[12] M. Mohandes, T. Halawani, S. Rehman, A. A. Hussain, Support vector machines for wind speed prediction, Renewable Energy 29 (6) (2004) 939–947. doi:10.1016/j.renene.2003.11.009.

[13] G. Zhang, H. Liu, J. Zhang, Y. Yan, L. Zhang, C. Wu, X. Hua, Y. Wang, Wind power prediction based on variational mode decomposition multi-frequency combinations, Journal of Modern Power Systems and Clean Energy 7 (2) (2019) 281–288. doi:10.1007/s40565-018-0471-8.

[14] L. Yu, H. Chen, S. Wang, K. K. Lai, Evolving least squares support vector machines for stock market trend mining, IEEE Transactions on Evolutionary Computation 13 (1) (2009) 87–102. doi:10.1109/TEVC.2008.928176.

[15] J. L. Wu, T. Y. Ji, M. S. Li, P. Z. Wu, Q. H. Wu, Multistep wind power forecast using mean trend detector and mathematical morphology-based local predictor, IEEE Transactions on Sustainable Energy 6 (4) (2015) 1216–1223. doi:10.1109/TSTE.2015.2424856.

[16] D. Hong, T. Ji, M. Li, Q. Wu, Ultra-short-term forecast of wind speed and wind power based on morphological high frequency filter and double similarity search algorithm, International Journal of Electrical Power &

Energy Systems 104 (2019) 868–879. doi:10.1016/j.ijepes.2018.07.061.

[17] W. Sun, Y. Wang, Short-term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space reconstruction, sample entropy and improved back-propagation neural network, Energy Conversion and Management 157 (2018) 1–12. doi:10.1016/j.enconman.2017.11.067.

[18] C. Zhang, H. Wei, J. Zhao, T. Liu, T. Zhu, K. Zhang, Short-term wind speed forecasting using empirical mode decomposition and feature selection, Renewable Energy 96 (2016) 727–737. doi:10.1016/j.renene.2016.05.023.

[19] H. Liu, X. Mi, Y. Li, Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, lstm network and elm, Energy Conversion and Management 159 (2018) 54–64. doi:10.1016/j.enconman.2018.01.010.

[20] J. Wang, Y. Li, Multi-step ahead wind speed prediction based on optimal feature extraction, long short term memory neural network and error correction strategy, Applied Energy 230 (2018) 429–443. doi:10.1016/j.apenergy.2018.08.114.

[21] Q. Wang, K. K. Lai, D. Niu, Q. Zhang, A multivariate wind power forecasting model based on ls-svm, in: 2012 Fifth International Joint Conference on Computational Sciences and Optimization, 2012, pp. 127–131. doi:10.1109/CSO.2012.35.

[22] M. Alexiadis, P. Dokopoulos, H. Sahsamanoglou, Wind speed and power forecasting based on spatial correlation models, IEEE Transactions on Energy Conversion 14 (3) (1999) 836–842. doi:10.1109/60.790962.

[23] S. Balkissoon, N. Fox, A. Lupo, S. E. Haupt, Y. Charles Li, P. Market, S. Walsh, Determining chaotic characteristics and forecasting tall tower

wind speeds in missouri using empirical dynamical modeling (edm), Renewable Energy 170 (2021) 1292–1307. doi:10.1016/j.renene.2021.01.108.

[24] Z. Shu, P. Chan, Q. Li, Y. He, B. Yan, Investigation of chaotic features of surface wind speeds using recurrence analysis, Journal of Wind Engineering and Industrial Aerodynamics 210 (2021) 104550. doi:10.1016/j.jweia.2021.104550.

[25] Z. Liang, J. Liang, L. Zhang, C. Wang, Z. Yun, X. Zhang, Analysis of multi-scale chaotic characteristics of wind power based on hilbert–huang transform and hurst analysis, Applied Energy 159 (2015) 51–61. doi:10.1016/j.apenergy.2015.08.111.

[26] T. Ouyang, H. Huang, Y. He, Z. Tang, Chaotic wind power time series prediction via switching data-driven modes, Renewable Energy 145 (2020) 270–281. doi:10.1016/j.renene.2019.06.047.

[27] Z. Guo, D. Chi, J. Wu, W. Zhang, A new wind speed forecasting strategy based on the chaotic time series modelling technique and the apriori algorithm, Energy Conversion and Management 84 (2014) 140–151. doi:10.1016/j.enconman.2014.04.028.

[28] M. B. Kennel, R. Brown, H. D. I. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction, Phys. Rev. A 45 (1992) 3403–3411. doi:10.1103/PhysRevA.45.3403.

[29] L. Cao, Practical method for determining the minimum embedding dimension of a scalar time series, Physica D: Nonlinear Phenomena 110 (1) (1997) 43–50. doi:10.1016/S0167-2789(97)00118-8.

[30] A. M. Fraser, H. L. Swinney, Independent coordinates for strange attractors from mutual information, Phys. Rev. A 33 (1986) 1134–1140. doi:10.1103/PhysRevA.33.1134.
URL https://link.aps.org/doi/10.1103/PhysRevA.33.1134

[31] S. Sato, M. Sano, Y. Sawada, Practical methods of measuring the generalized dimension and the largest lyapunov exponent in high dimensional chaotic systems, Progress of Theoretical Physics 77 (1) (1987) 1–5. doi:10.1143/PTP.77.1.

[32] A. Wolf, J. B. Swift, H. L. Swinney, J. A. Vastano, Determining lyapunov exponents from a time series, Physica D: Nonlinear Phenomena 16 (3) (1985) 285–317. doi:10.1016/0167-2789(85)90011-9.

[33] M. T. Rosenstein, J. J. Collins, C. J. De Luca, A practical method for calculating largest lyapunov exponents from small data sets, Physica D: Nonlinear Phenomena 65 (1) (1993) 117–134. doi:10.1016/0167-2789(93)90009-P.

[34] M. Li, Z. Zhang, T. Ji, Q. H. Wu, Ultra-short term wind speed prediction using mathematical morphology decomposition and long short-term memory, CSEE Journal of Power and Energy Systems 6 (4) (2020) 890–900. doi:10.17775/CSEEJPES.2019.02070.