Direct Multistep Wind Speed Forecasting Using LSTM Neural Network Combining EEMD and Fuzzy Entropy

Qiong Qin [1], **Xu Lai [1,*] and Jin Zou [2]**

[1] State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, Hubei, China; qinqiong@whu.edu.cn

[2] Electric Power Research Institute, China Southern Power Grid, Guangzhou 510080, Guangdong, China; zoujin@csg.cn

* Correspondence: laixu@whu.edu.cn; Tel.: +86-134-7613-6201

Abstract: Accurate wind speed forecasting is of great significance for a reliable and secure power generation system. In order to improve forecasting accuracy, this paper introduces the LSTM neural network and proposes a wind speed statistical forecasting method based on the EEMD-FuzzyEn-LSTMNN model. Moreover, the MIC is used to analyze the autocorrelation of wind speed series, and the predictable time of wind speed statistical forecasting method for direct multistep forecasting is taken as four hours. In the EEMD-FuzzyEn-LSTMNN model, the original wind speed series is firstly decomposed into a series of components by using EEMD. Then, the FuzzyEn is used to calculate the complexity of each component, and the components with similar FuzzyEn values are classified into one group. Finally, the LSTMNN model is used to forecast each subsequence after classification. The forecasting result of the original wind speed series is obtained by aggregating the forecasting result of each subsequence. Three forecasting cases under different terrain conditions were selected to validate the proposed model, and the BPNN model, the SVM model and the LSTMNN model were used for comparison. The experimental results show that the forecasting accuracy of the EEMD-FuzzyEn-LSTMNN model is much higher than that of the other three models.

Keywords: wind speed forecasting; long short-term memory; neural network; ensemble empirical mode decomposition; fuzzy entropy; maximal information coefficient; autocorrelation

1. Introduction

With the large consumption of fossil energy and the impact of global climate change, increasing attention has turned to renewable energy. As a kind of clean and renewable energy, wind power has been widely valued and promoted in the world. According to the World Wind Energy Association statistics, the total cumulative installed capacity from wind power around the world was approximately 539 GW by the end of 2017. However, with the expansion of wind power generation, its problems have gradually become more prominent. Due to the randomness, fluctuation and intermittence nature of wind speed, the wind power output has great uncertainty [1,2]. Therefore, large-scale wind power integration will bring great hidden dangers to the safe and stable operation of the power system. In order to solve the above problems, it is necessary to effectively improve the controllability and predictability of wind power. Accurate wind power forecasting can provide an important basis for power dispatching and improve the utilization efficiency of wind energy resources. Since wind power has a direct relationship with wind speed, wind power forecasting can be achieved based on wind speed forecasting.

The accuracy of wind speed forecasting is not only dependent on the forecasting method but also on the forecasting time horizon. In general, there are two main categories for wind speed forecasting based on the forecasting time horizon, namely, short-term forecast (time horizon of minutes, hours, and days) and long-term forecast (time horizon of days, weeks and months) [3]. The short-term forecast is important for the operation of wind turbines so that dynamic controls can be accomplished to increase the energy conversion efficiency and reduce the risk of overloading. The long-term forecast can provide important references for site location and planning of wind farms.

In recent years, researchers have done many studies on the theory and methods of wind speed forecasting. According to the type of input variables of the forecasting model, the wind speed forecasting methods can be divided into physical forecasting methods, statistical forecasting methods and hybrid forecasting methods [4]. The input variables of the physical models are meteorological data from numerical weather prediction (NWP) and terrain feature data. Many physical models have been introduced, and the most popular three are Prediktor [5], LocalPred [6] and Previento [7]. Generally, physical methods have advantages in long-term forecast, while statistical methods have better performance in short-term forecast [1].

The input variables of the statistical models are long-term historical wind speed data. The statistical models achieve the forecasting of wind speed in the future by excavating the hidden change rules in historical wind speed data. Early statistical models include the autoregressive model (AR) [8], the autoregressive moving average model (ARMA) [9], the grey prediction model (GM) [10], etc. With the development of artificial intelligence techniques, support vector machines (SVM) and a large number of artificial neural network (ANN) algorithms have been introduced into wind speed forecasting. For example, Mohandes et al. [11] introduced SVM to wind speed prediction and compared its performance with the multilayer perceptron (MLP) neural networks. Cadenas et al. [12] used the ANN to forecast the wind speed in the region of La Venta, Oaxaca, Mexico, and the results showed very good accuracy for the short-term wind speed forecasting. Flores et al. [13] used back propagation neural network (BPNN) to forecast the wind speed. Li et al. [14] compared three different ANNs (adaptive linear element NN, BPNN, and radial basis function NN) in one-hour ahead wind speed forecasting, and the results showed that even for the same wind dataset, the choice of best-performing model might not be the same with different evaluation metrics.

Hybrid forecasting methods are formed by hybridizing physical methods and statistical methods together. In the hybrid methods, NWP data is also used as the input data of the statistical models. Salcedo-Sanz et al. [15] used the output of the fifth generation mesoscale model (MM5 model) as the input variables of the ANN, and the experiment results showed that the hybrid MM5-neural network approach was able to obtain good short-term predictions of wind speed at specific points. Ortiz-García et al. [16] then made improvements to the system proposed in Ref. [15], and used regression SVM instead of ANN to obtain the final prediction. Giorgi et al. [17] trained the ANNs based on the measured wind speed series and on some NWP parameters, and an improvement of the performance had been reached, especially with the longer time horizons.

However, the wind speed series has multiscale characteristics. The multiple frequency components existed in the wind speed series are always the challenging parts in forecasting. Currently, the concept of "decomposition and ensemble" (or "divide and conquer") [18,19] has been used to solve this problem, and a number of various decomposition-ensemble based hybrid models have been developed and widely used. For instance, Liu et al. [20] selected wavelet decomposition (WD) and wavelet packet decomposition (WPD) to decompose an original wind speed series respectively, and then used the ANN models to do the multistep forecasting in each subseries. Guo et al. [21] combined empirical mode decomposition (EMD) and standard feed-forward neural network (FNN), and Liu et al. [22] combined EMD and ANN. The new hybrid model all outperformed the conventional FNN or ANN model. But the WD is sensitive to the wavelet base function and decomposition level, and for the EMD, one major challenge is the frequent appearance of mode mixing [23]. Fortunately, there exists an improved method named ensemble empirical mode decomposition (EEMD), which makes up

for the deficiency of EMD. The EEMD is an empirical, intuitive, direct and self-adaptive data processing method created especially for nonlinear and nonstationary signal sequences [3]. For example, Wang et al. [24] proposed a wind speed forecasting method based on EEMD and optimized BPNN (GA-BPNN) for short-term wind speed forecasting, and computational results had shown the good performance of EEMD.

In this study, we mainly focus on the wind speed statistical forecasting methods for short-term forecast. The first step is to determine the specific forecasting time horizon. When using statistical methods for direct multistep forecasting of wind speed series, in general, the longer the forecasting length, the lower the forecasting accuracy. Therefore, it is necessary to determine the predictable length of direct multistep forecasting. In this paper, the autocorrelation analysis method based on the maximal information coefficient (MIC) is introduced to measure the predictability of wind speed series. Then, the predictable time of statistical forecasting methods based on historical wind speed data is analyzed. The next step is to choose the wind speed statistical forecasting method. When the ordinary neural networks mentioned above process the wind speed series, the reading and processing of the input wind speed data are independent at each moment. These neural networks cannot fully consider the correlation of the wind speed series itself. When forecasting the wind speed at the next moment, they are unable to share the features learned from the previous input wind speed data. Therefore, the forecasting accuracy of these neural network models is limited. In order to make full use of the correlation between the data of wind speed series at each moment, this paper introduces a new long short-term memory neural network (LSTMNN) model for wind speed forecasting. The LSTMNN has a unique memory and forgetting mode. It can handle the long-term dependence of wind speed series very well, and effectively use the historical input information of the wind speed series. The LSTMNN has been widely used in many fields such as traffic forecasting [25], solar energy forecasting [26], stock price volatility prediction [27] and water table depth prediction [28]. These predictions all get good results [29].

In order to improve the forecasting accuracy of the LSTMNN model, this paper proposes a novel wind speed statistical forecasting method based on the EEMD-FuzzyEn-LSTMNN model. The EEMD-FuzzyEn-LSTMNN model is developed through combining EEMD, Fuzzy Entropy (FuzzyEn) and LSTMNN. In the proposed model, the original wind speed series is firstly decomposed into a series of components by using EEMD. Then, the FuzzyEn is introduced to calculate the complexity of each component, and the components are classified according to the calculated FuzzyEn values. The components with similar FuzzyEn values are classified into one group, and the components of each group are superimposed to obtain a new subsequence. This process can avoid cumbersome calculations caused by forecasting each component separately. Finally, the LSTMNN model is used to forecast each subsequence after classification. The forecasting result of the original wind speed series is obtained by aggregating the forecasting result of each subsequence.

The remainder of this paper is organized as follows: Section 2 briefly describes the fundamental methods including EEMD, FuzzyEn and LSTMNN. Section 3 introduces the proposed EEMD-FuzzyEn-LSTMNN model. Section 4 uses MIC to analyze the predictable time of statistical forecasting methods for direct multistep forecasting. Section 5 firstly describes the wind speed data of three cases and selects error evaluation indexes, and then provides the forecasting results of the BPNN model, the SVM model, the LSTMNN model and the proposed EEMD-FuzzyEn-LSTMNN model. Finally, Section 6 gives the conclusions and discusses the future work.

2. Methodology

The research methodology used in this study includes EMD, EEMD, FuzzyEn, RNN and LSTMNN. The brief description of those methods is stated as follows.

2.1. EMD and EEMD

EMD was proposed by Huang in 1998 [30]. It is a self-adaptive and efficient method for analyzing nonlinear and nonstationary signals. Since the wind speed series is nonlinear and nonstationary, EMD is efficient to analyze the wind speed signal. The basic idea of EMD is to smooth the fluctuating signals adaptively and to decompose them into fluctuations or trends of different scales. After decomposition, a finite and small number of intrinsic mode functions (IMFs) and a residual component are obtained. An IMF is a function that satisfies the following two conditions: (a) in the whole data set, the number of extrema and the number of zero crossings must either be equal or differ at most by one; and (b) at any point, the average of the envelopes defined by the local maxima and the local minima must be zero [31]. The specific steps for decomposing time series by using the EMD algorithm are detailed in Ref. [30].

Mode mixing is the most significant drawback of EMD [23], which means that a single IMF consists of signals with dramatically disparate scales or a signal of the same scale appears in different IMF components. The mode mixing compromises the stationarity of IMFs and therefore limits the effectiveness of the EMD algorithm. To solve the mode mixing problem in EMD, a new noise-assisted data analysis method EEMD is proposed. In EEMD, the white Gaussian noise is added into the original time series, and the EMD is then applied on noise added time series to obtain IMFs that are free from mode mixing [32,33]. However, the resulting IMFs include white noise, which is then removed by obtaining the mean of multiple trials. Therefore, the true IMF components are defined as the mean of an ensemble of trails and each trail consists of the decomposition results of the signal plus a white Gaussian noise of finite amplitude [3].

For the wind speed series $\{v(t), t = 1, 2, \dots, N\}$, the main steps of the EEMD algorithm are described as follows:

Step 1:　Add the white Gaussian noise series $\varepsilon_j(t)$ to the original wind speed series $v(t)$ and obtain a new series $V_j(t)$.

Step 2:　Decompose the new series $V_j(t)$ into several IMFs and a residue by using the EMD algorithm.

Step 3:　For $j = 1, 2, \dots, N_E$, repeat Step 1 and Step 2, and add different white Gaussian noise series each time. N_E is the number of repeated procedures.

Step 4:　Take the mean of all IMF components and the mean of residual components as the final results.

After the decomposition by using EEMD algorithm, the wind speed series $\{v(t)\}$ can be expressed as:

$$v(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{1}$$

where $\{c_i(t), I = 1, 2, \dots, n\}$ represents the different IMF component, and $r_n(t)$ is the residue after n IMFs are derived. These IMFs contain components of different time characteristic scales of the wind speed series. Their scales range from small to large, and the frequencies range from high to low.

2.2. Fuzzy Entropy

Using the EEMD algorithm to decompose the wind speed series, a series of components are obtained. If a forecasting model is built separately for each component, the process is slightly redundant. Moreover, when superimposing the forecasting results of all components, the forecasting errors of each component are also superimposed. In order to reduce the calculation scale and the accumulation of forecasting errors, the components obtained by EEMD can be classified according to certain criteria. Then, the components under each category are superimposed to form a new subsequence, and each subsequence is forecasted separately. The forecasting result of the original wind speed series can be obtained by superimposing the forecasting result of each subsequence.

FuzzyEn is a metric of the complexity of time series [34–36]. It measures the complexity of the series by the probability that the time series produces a new pattern as the embedding dimension

changes. The larger the FuzzyEn value, the greater the probability that the sequence will produce a new pattern, which means the sequence is more complex. Each IMF and the residue obtained by EEMD contain components of the wind speed series at different time characteristic scales. Their frequencies are different, so the complexity of each component is also considered to be different. In this paper, FuzzyEn is introduced to calculate the complexity of all components, and then the components are classified according to the obtained FuzzyEn value. For the components with similar FuzzyEn values, they can be considered to have similar complexity and wind speed characteristics, so they can be classified into one group.

The specific steps for calculating the FuzzyEn of time series are detailed in Ref. [34].

2.3. RNN and LSTMNN

LSTMNN is developed on the basis of Recurrent Neuron Network (RNN). Unlike ordinary neural networks, RNN adds a self-connected hidden layer spanning time steps. Therefore, RNN can memorize the input information in front of the time series and apply it to the calculation of the current output.

A simple RNN consists of three layers: an input layer, a hidden layer and an output layer. Given m as the number of forecasting steps of direct multistep forecasting, the structure diagram of the RNN forecasting model for wind speed series is shown in Figure 1. In this figure, $\{v_i, i = 1, 2, \ldots, t\}$ is the measured input wind speed; $\{V_{i+m}, i = 1, 2, \ldots, t\}$ is the forecasted output wind speed; a_i is the activation value from time-step i; W_{av}, W_{va} and W_{aa} are the connection weights between neurons, which are the same at each time step.
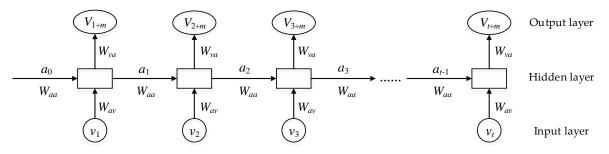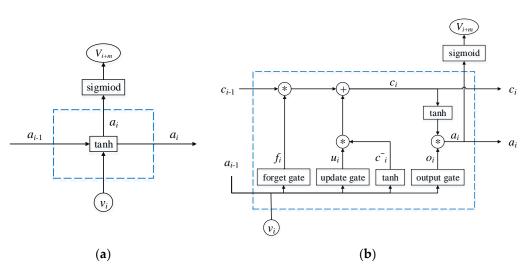

Figure 1. The structure of the RNN forecasting model for wind speed series.

In the RNN model, the hidden layers of the front and back time steps are connected. It can make full use of the information of the wind speed series before the forecasted time, thereby improving the forecasting accuracy. However, the RNN model may have the vanishing gradient problem during the training process [37]. The LSTMNN proposed by Hochreiter and Schmidhuber [38] can solve this problem of the RNN model through its special structural design. Moreover, the LSTMNN can better handle the long-term dependence of time series and effectively utilize the historical input information of time series.

The basic LSTMNN also consists of three layers: an input layer, a hidden layer and an output layer. But compared with the RNN, the hidden layer of the LSTMNN adds some threshold units for controlling information transfer, which makes the neural network have a unique memory mode [38,39]. The structures of the RNN forecasting model and the LSTMNN forecasting model at the time-step i are shown in Figure 2.

In the forward propagation process of the LSTMNN forecasting model, besides the activation value a_i, a memory cell c_i is transferred from the previous hidden layer to the latter. In addition, three gate structures are added to the hidden layer in LSTMNN, namely forget gate, update gate, and output gate. The three gate structures can control the preservation, reading and modification of the memory cell in the LSTMNN forecasting model.

Figure 2. The structure of forecasting model for wind speed series at the time-step *i*: (**a**) the RNN forecasting model; (**b**) the LSTMNN forecasting model.

In the forward propagation process of the LSTMNN forecasting model, the forecasted output wind speed of the time-step *i* is calculated as follows:

$$\widetilde{c}_i = \tanh(W_{ca} \cdot a_{i-1} + W_{cv} \cdot v_i + b_c) \tag{2}$$

$$f_i = \sigma\left(W_{fa} \cdot a_{i-1} + W_{fv} \cdot v_i + b_f\right) \tag{3}$$

$$u_i = \sigma(W_{ua} \cdot a_{i-1} + W_{uv} \cdot v_i + b_u) \tag{4}$$

$$o_i = \sigma(W_{oa} \cdot a_{i-1} + W_{ov} \cdot v_i + b_o) \tag{5}$$

$$c_i = f_i c_{i-1} + u_i \widetilde{c}_i \tag{6}$$

$$a_i = o_i \tanh c_i \tag{7}$$

$$V_{i+m} = \sigma(W_{va} \cdot a_i + b_v) \tag{8}$$

where the subscript *i* represents the time-step *i*; $\widetilde{c}_i$ is the candidate for replacing the memory cell; c_i is the memory cell; f_i, u_i and o_i are the values of forget gate, update gate and output gate, respectively; σ is the sigmoid function; W_{ca}, W_{cv}, W_{fa}, W_{fv}, W_{ua}, W_{uv}, W_{oa}, W_{ov}, W_{va} are weight matrices; b_c, b_f, b_u, b_o, b_v are bias vectors.

After the forward propagation process of the LSTMNN forecasting model is completed, the back propagation process is followed. The model is firstly expanded into a deep network in chronological order. Then, the BPTT algorithm [40] and chain rule are used to iteratively update the connection weights and thresholds in the model until the optimal solution is obtained.

3. The EEMD-FuzzyEn-LSTMNN Model

In this section, the EEMD-FuzzyEn-LSTMNN model is established for wind speed forecasting based on the concept of "decomposition and ensemble". The flowchart of the proposed model is shown in Figure 3. The main structure of the EEMD-FuzzyEn-LSTMNN model includes the following four steps:

Step 1:　Use EEMD to decompose the original wind speed series into a number of IFM components and a residual component. These components are respectively denoted by IMF$_1$, IMF$_2$, . . . , IMF$_n$ and R$_n$.

Step 2:　Calculate the FuzzyEn of each component and classify all components according to the calculated FuzzyEn values. The components with similar FuzzyEn values are classified into

　　　　　one group, and the components in one group are superimposed to obtain a new subsequence. All subsequences are respectively denoted by $S_1, S_2, \ldots, S_N$.

Step 3:　　Use the LSTMNN model to forecast each subsequence separately.

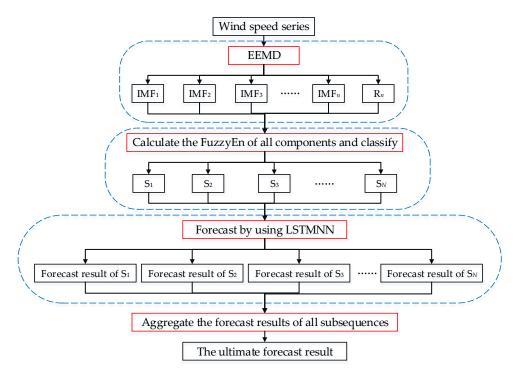Step 4:　　Aggregate the forecasting result of each subsequence to obtain the ultimate forecasting series of wind speed.

Figure 3. The flowchart of the EEMD-FuzzyEn-LSTMNN model.

4. The Predictable Time of Wind Speed Series

　　　　The MIC is a statistic that measures the correlation between two variables [41]. Based on mutual information, the MIC can measure both the linear correlation and the nonlinear correlation between two variables. It has generality and equitability. The calculation method of MIC is detailed in [41]. Since the wind speed series is a well-known nonlinear time series, the MIC with many excellent characteristics can be used to measure its autocorrelation. In this section, we use MIC to measure the predictable time of statistical forecasting methods based on historical wind speed data.

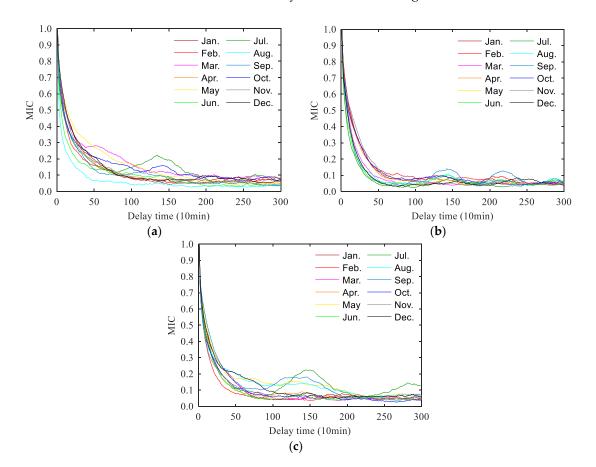### 4.1. Selection of Wind Speed Series

　　　　In this paper, we selected wind speed series from three anemometer towers in China. The three anemometer towers are in three wind farms which are located in Hunan province, Henan province and Zhejiang province, respectively. The location, local terrain condition, selected time and measurement height of the three anemometer towers are shown in Table 1. These wind speed series were used to analyze the predictability of direct multistep forecasting. On the three anemometer towers, the calibrated NRG measuring instruments were used to automatically record the wind speed data, and the recorded time interval was 10 min. After processing the missing and invalid data of the anemometer towers, the integrity rate of the wind speed data can reach 100%. And these wind speed data have passed the reasonableness test.

Table 1. Information of wind speed series from three anemometer towers.

Number	Location	Local Terrain Condition	Selected Time of Wind Speed Series	Height of Wind Speed Measurement (m)
1	Hunan province	Mountainous area	From April 2014 to March 2015	80
2	Henan province	Plain area	From June 2016 to May 2017	120
3	Zhejiang province	Coastal area	From August 2011 to July 2012	100

4.2. Autocorrelation Analysis of Wind Speed Series Based on MIC

The MIC was used to measure the autocorrelation of the selected wind speed series. There are 36 months of wind speed series. Taking a month's wind speed series $\{v(t), t = 1, 2, \ldots, N\}$ as an example, the MIC of $\{v(t), t = 1, 2, \ldots, N\}$ and $\{v(t), t = 1, 2, \ldots, N\}$ was firstly calculated. Then, the MIC of $\{v(t), t = 1, 2, \ldots, N - 1\}$ and $\{v(t), t = 2, 3, \ldots, N\}$ was calculated. Continuing this step until the MIC of $\{v(t), t = 1, 2, \ldots, N - \tau\}$ and $\{v(t), t = \tau + 1, \tau + 2, \ldots, N\}$ was calculated. In the above example, N represents the number of wind speed data; τ represents the number of delay time (unit is 10 min) and this paper took τ as 300. After calculating the auto-correlated MIC of all wind speed series, the variation law of the MIC with the delay time is shown in Figure 4.

Figure 4. The variation law of the MIC with the delay time: (**a**) wind speed series of anemometer tower 1; (**b**) wind speed series of anemometer tower 2; (**c**) wind speed series of anemometer tower 3.

It can be seen from Figure 4 that as the delay time increases, the auto-correlated MIC of the wind speed series gradually decreases and finally reaches a steady trend. With the decrease of MIC, the correlation between future and historical wind speed series also decreases. At this time, the forecasting error will gradually increase when forecasting the future wind speed based on the historical wind speed. Therefore, in order to ensure the forecasting accuracy of the direct multistep forecasting, a fixed MIC value can be determined, and the delay time corresponding to the MIC value

can be used as the predictable time of the wind speed series. According to the calculation results of the three anemometer towers in Figure 4, the variation law of MIC with delay time tends to be gentle when the MIC drops to 0.2. Therefore, we can take the corresponding delay time (or correlation length) as the predictable time of the wind speed series when the MIC is equal to 0.2. The calculation results are shown in Table 2.

Table 2. Correlation length of wind speed series when MIC equals 0.2.

Month	Correlation Length (h)		
	Anemometer Tower 1	**Anemometer Tower 2**	**Anemometer Tower 3**
January	8.17	6.17	6.67
February	5.83	5.83	3.67
March	14.83	4.50	6.83
April	7.50	6.00	4.67
May	12.50	4.17	6.17
June	4.50	3.67	4.83
July	7.00	3.33	5.33
August	3.00	4.00	8.67
September	6.33	3.83	6.67
October	9.67	6.00	5.17
November	7.67	6.50	5.00
December	9.00	4.17	8.50

4.3. Analysis of Predictable Time

Because the selected wind speed series in different regions are different, the correlation length of each month's wind speed series obtained by taking MIC equal to 0.2 is also different. Therefore, this paper compared the influence of different correlation lengths on wind speed forecasting error to obtain a more general wind speed predictable length.

In this paper, the basic time series forecasting method, Persistence Model [42], was used to forecast the corresponding correlation length of each month's wind speed series. Assuming that the wind speed series of a certain month is $\{v(t), t = 1, 2, \ldots, N\}$, and the correlation length obtained when the MIC equals 0.2 is denoted by l (unit is h). Since the recorded time interval of the three anemometer towers is 10 min, the number of wind speed data with a correlation length of l is $6l$. The wind speed series was segmented according to $6l$, and then the wind speed of the previous period was used to forecast the wind speed of the next period. That is to say, $\{v(t), t = 1, 2, \ldots, 6l\}$ was used to forecast $\{v(t), t = 6l + 1, 6l + 2, \ldots, 2 \times 6l\}$; $\{v(t), t = 6l + 1, 6l + 2, \ldots, 2 \times 6l\}$ was used to forecast $\{v(t), t = 2 \times 6l + 1, 2 \times 6l + 2, \ldots, 3 \times 6l\}$ and so on until the end of the series.

After using the persistence model to forecast the 36 wind speed series in different regions, the errors between the measured and forecasted wind speed of each month were counted. Mean Absolute Error (*MAE*) and Root Mean Squared Error (*RMSE*) were used as the error evaluation index. The formulas for calculating *MAE* and *RMSE* are as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|v_i - v_{pi}\right| \tag{9}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(v_i - v_{pi}\right)^2} \tag{10}$$

where N is the number of forecasted wind speed data; v_i is the measured wind speed, m/s; v_{pi} is the forecasted wind speed, m/s.

The variation law of *MAE* and *RMSE* with the correlation length is shown in Figure 5. As the correlation length of the wind speed series increases, the forecasting error tends to increase overall.

However, when the correlation length is less than four hours, the forecasting error is small and stable, and the fluctuation range is small. Therefore, four hours can be used as the predictable time of the wind speed series for direct multistep forecasting based on historical wind speed data. The forecasting time of the wind speed series in the following section was taken as four hours.
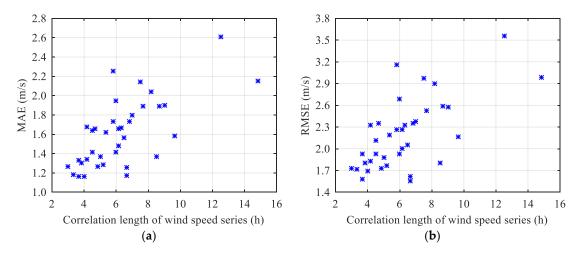

Figure 5. The variety law of *MAE* and *RMSE* with correlation length: (**a**) *MAE*; (**b**) *RMSE*.

5. Case Study

The three forecasting cases of wind speed series selected in this paper are from three anemometer towers in mountainous area, plain area and coastal area, respectively. Then the EEMD-FuzzyEn-LSTMNN model was established to forecast the wind speed series for the future four hours. Furthermore, the applicability and effectiveness of the EEMD-FuzzyEn-LSTMNN model were verified by comparing with the BPNN model, SVM model and LSTMNN model.

5.1. Wind Speed Data Description of the Cases

The wind speed data selected in Case A, Case B and Case C is the wind speed series from October to December of the anemometer tower 1, 2 and 3 respectively. The detailed information of the three anemometer towers is shown in Table 1.

The direct multistep forecasting time of the wind speed series was taken as four hours according to the conclusion of Section 4. In each case, 864 wind speed data from 6 days in December were randomly selected as the test set of the forecasting model, and the wind speed data from the previous two months were used as the training set of the forecasting model. Taking the wind speed series $\{v(t), t = 1, 2, \ldots, N\}$ of a certain case as an example, the training and test set of the forecasting model are illustrated as follows. The training set is $\{v(t), t = 1, 2, \ldots, N - 864 - 24\}$, in which $\{v(t), t = 1, 2, \ldots, N - 864 - 24 \times 2\}$ is the training input sample and $\{v(t), t = 24 + 1, 24 + 2, \ldots, N - 864 - 24\}$ is the training output sample. Similarly, the test set is $\{v(t), t = N - 864 - 24 + 1, N - 864 - 24 + 2, \ldots, N\}$, in which $\{v(t), t = N - 864 - 24 + 1, N - 864 - 24 + 2, \ldots, N - 24\}$ is the test input sample and $\{v(t), t = N - 864 + 1, N - 864 + 2, \ldots, N\}$ is the test output sample.

5.2. Error Evaluation Index

The selection of evaluation indexes has an important impact on the assessment of wind speed forecasting methods. The evaluation indexes selected in this paper are *MAE*, *RMSE* and Mean Absolute Percentage Error (*MAPE*). The calculation formulas of *MAE* and *RMSE* are given by Formulas (9) and (10), and the formula for *MAPE* is as follows:

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \frac{|v_i - v_{pi}|}{v_i} \tag{11}$$

5.3. Parameter Settings of the BPNN Model, the SVM Model and the LSTMNN Model

When the BPNN model, the SVM model and the LSTMNN model were used to forecast the wind speed series in the cases, a series of parameter settings for the models are shown in Table 3. In addition to the parameters specified in the table, the default values given by the corresponding software toolbox were used for other parameters. The wind speed data should be normalized before being input into the forecasting model. The normalization function used in this paper is the mapminmax function, which normalizes the wind speed series to $[-1, 1]$.

Table 3. Parameter settings of wind speed forecasting model.

Forecasting Models	Parameters	Number or Type
BPNN model	Number of neurons in the hidden layer	20
	Learning rate of training	0.001
	Training target	0.00001
SVM model	Type of svm	epsilon-SVR
	Type of kernel function	linear kernel function
LSTMNN model	Number of neurons in the LSTM layer	20
	Type of activation function of the output layer	tanh
	Optimizer	adam
	Learning rate	0.0001

5.4. Decomposition Results by EEMD

In order to improve the overall forecasting performance, this paper firstly adopted EEMD to decompose the original wind speed series of the three cases separately, and the decomposition results are listed in Figure 6. It is obvious that each wind speed series is decomposed into 13 components which are respectively denoted by IMF_1, IMF_2, ... , IMF_{12} and R from top to bottom.

(a)

Figure 6. *Cont.*

Figure 6. The decomposition results of wind speed series: (**a**) Case A; (**b**) Case B; (**c**) Case C.

5.5. Classification of Components Based on FuzzyEn

The original wind speed series of the three cases are all decomposed into 13 components by EEMD. Then, the FuzzyEn values of each component were calculated respectively, and the results are shown in Figure 7.
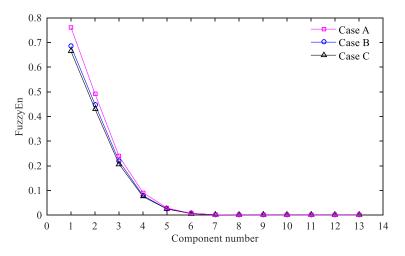

Figure 7. The FuzzyEn of each component obtained by EEMD.

In the three cases, the FuzzyEn values of the 1st component to the 13th component gradually decrease firstly, and then become smooth. The FuzzyEn of the high-frequency component is large, which means it is more complicated. On the contrary, the FuzzyEn of the low-frequency component is small, indicating that its complexity is small. The components with similar FuzzyEn values can be classified into one group. In Figure 7, the FuzzyEn values of the components at the front descending curve segment are quite different, so each component can be regarded as a subsequence. The FuzzyEn values of the 7th component to 13th component at the final stationary curve segment are almost the same, so they can be grouped together and superimposed to form a new subsequence. After the classification was completed, 7 subsequences were finally obtained.

5.6. Forecasting Results and Analysis

The BPNN model, the SVM model, the LSTMNN model and the EEMD-FuzzyEn-LSTM model were used to forecast the 864 wind speed data randomly selected from the three cases. The forecasting results are shown in Figure 8 and the forecasting error is shown in Figure 9.
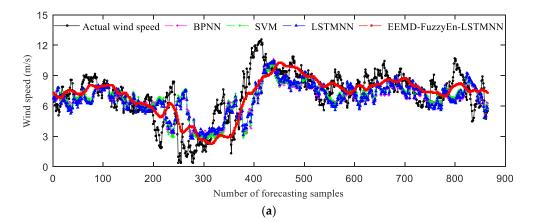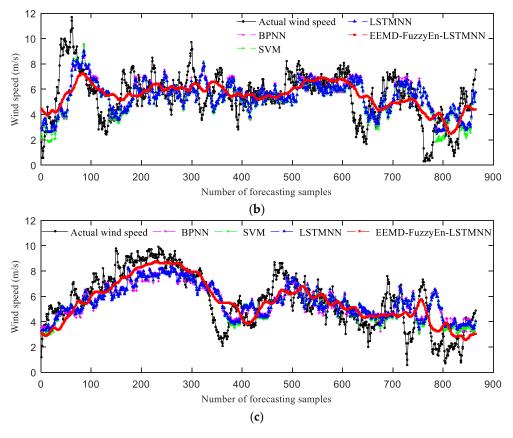

(a)

Figure 8. *Cont.*

(**b**)

(**c**)

Figure 8. Forecasting results of wind speed series: (**a**) Case A; (**b**) Case B; (**c**) Case C.
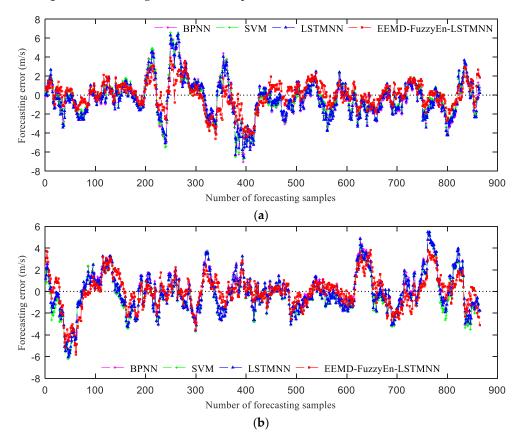

(**a**)

(**b**)

Figure 9. *Cont*.
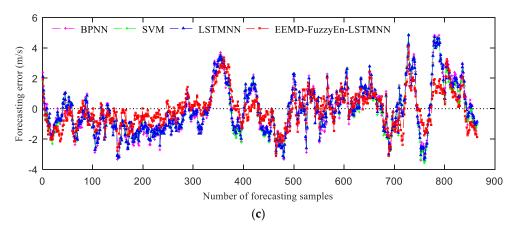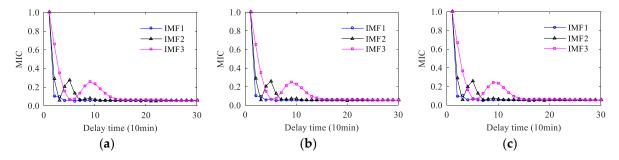
(c)

Figure 9. Forecasting error of wind speed series: (**a**) Case A; (**b**) Case B; (**c**) Case C.

As is shown in Figure 8; Figure 9, the forecasting effect of the BPNN model is similar to that of the SVM model, and the forecasting effect of the LSTMNN model is slightly better than that of the BPNN model and the SVM model for wind speed series in different regions. However, the forecasting effect of the EEMD-FuzzyEn-LSTMNN model is greatly improved compared with the LSTMNN model, and the forecasting error is obviously reduced.

The subsequences are forecasted separately in the EEMD-FuzzyEn-LSTMNN model, and the fluctuation of wind speed forecasting results obtained by the superposition is smoother. By analyzing each subsequence, it was found that the first few high-frequency subsequences have strong randomness and their regularity is very low. MIC was used to analyze the autocorrelation of the first three high-frequency subsequences, and the variation law of MIC with delay time is shown in Figure 10. As the delay time increases, the MIC of autocorrelation decreases rapidly. In the forecasting of wind speed series for the future four hours (24 × 10 min), the autocorrelation of these high-frequency subsequences is poor, which leads to large forecasting errors. When the forecasting results of all subsequences are superimposed, the high-frequency part of the wind speed series is weakened due to the low forecasting accuracy, which results in more gentle fluctuation of the forecasting results.

(a)　　　　　　　　　　　**(b)**　　　　　　　　　　　**(c)**

Figure 10. The variation law of MIC with the delay time for high-frequency components: (**a**) Case A; (**b**) Case B; (**c**) Case C.

The error evaluation indexes of the four forecasting models are calculated and presented in Table 4. In the table, the evaluation indexes of the EEMD-FuzzyEn-LSTMNN model are the best and its forecasting accuracy is the highest.

Comparing the LSTMNN model with the BPNN model, the *MAE*, *MAPE* and *RMSE* of the three cases are reduced by 0.01–0.04 m/s (1–3%), 0.004–0.01 (1–3%), 0.02–0.03 m/s (1–2%), respectively. In the case study of this paper, the default values given by the toolbox were used for most parameters of the LSTMNN model, which may lead to a small increase in the evaluation indexes. Since the training process of the LSTMNN involves the adjustment of many parameters, the forecasting accuracy of this model will be further improved after systematic parameter optimization. Comparing the

EEMD-FuzzyEn-LSTMNN model with the LSTMNN model, the error evaluation indexes are greatly improved. The *MAE*, *MAPE* and *RMSE* of the Case A are reduced by 0.4413 m/s (29.18%), 0.11 (29.7%), 0.6014 m/s (30.04%), respectively. The *MAE*, *MAPE* and *RMSE* of the Case B are reduced by 0.2773 m/s (19.74%), 0.1051 (21.74%), 0.3076m/s (16.81%), respectively. The *MAE*, *MAPE* and *RMSE* of the Case C are reduced by 0.3397m/s (26.99%), 0.0882 (26.27%), 0.4009m/s (25.66%), respectively. The values in the parentheses above are all relative values.

Table 4. Calculation results of error evaluation indexes for the four forecasting models.

Case	Forecasting models	*MAE* (m/s)	*MAPE*	*RMSE* (m/s)
Case A	BPNN model	1.5308	0.3800	2.0371
	SVM model	1.5264	0.3742	2.0211
	LSTMNN model	1.5122	0.3704	2.0022
	EEMD-FuzzyEn-LSTMNN model	1.0709	0.2604	1.4008
Case B	BPNN model	1.4406	0.4875	1.8626
	SVM model	1.4523	0.4861	1.8694
	LSTMNN model	1.4045	0.4835	1.8295
	EEMD-FuzzyEn-LSTMNN model	1.1272	0.3784	1.5219
Case C	BPNN model	1.2697	0.3397	1.5798
	SVM model	1.2733	0.3374	1.5767
	LSTMNN model	1.2585	0.3358	1.5624
	EEMD-FuzzyEn-LSTMNN model	0.9188	0.2476	1.1615

In addition, compared with the anemometer towers in other areas (Case B and Case C), the forecasting accuracy of the mountainous anemometer tower (Case A) is improved more greatly. Therefore, the EEMD-FuzzyEn-LSTMNN model is more advantageous for forecasting the wind speed series which has large forecasting errors by using ordinary neural networks.

6. Conclusions

This paper first introduces the LSTM neural network for the forecasting of wind speed series. Then, in order to further improve the forecasting accuracy, a wind speed statistical forecasting method based on the EEMD-FuzzyEn-LSTMNN model is proposed. And the autocorrelation analysis based on the MIC is used to obtain the predictable time of wind speed series for direct multistep forecasting. The main conclusions of this paper are as follows:

- MIC is used to analyze the autocorrelation of wind speed series from different terrain conditions, and some suitable correlation lengths are obtained. As the correlation length of the wind speed series increases, the forecasting error tends to increase overall. The forecasting error analysis shows that four hours can be taken as the predictable time of the wind speed series for direct multistep forecasting based on historical wind speed data.
- The wind speed series from different terrain conditions is forecasted for the future four hours, and the forecasting accuracy of the LSTMNN model is slightly higher than that of the BPNN model and the SVM model. It shows that the LSTM neural network can make better use of the historical input information of the wind speed series, and it is more suitable for the wind speed statistical forecasting method.
- Under different terrain conditions, the forecasting accuracy of the EEMD-FuzzyEn-LSTMNN model is much higher than that of the LSTMNN model. Comparing the EEMD-FuzzyEn-LSTMNN model with the LSTMNN model, the *MAE*, *MAPE* and *RMSE* of the three cases are reduced by 0.2773–0.4413 m/s (19.74–29.18%), 0.0882–0.11 (21.74–29.7%), 0.3076–0.6014 m/s (16.81–30.04%), respectively. Moreover, the EEMD-FuzzyEn-LSTMNN model has more advantages for forecasting the wind speed series which has large forecasting errors by using ordinary neural networks.

The current forecasting method in this paper is to directly forecast the same steps for each subsequence. However, due to the strong randomness of the high-frequency subsequence, it is not appropriate for the high-frequency subsequence to make a direct multistep forecasting of the same steps as the low-frequency subsequence. Therefore, the autocorrelation of each subsequence can be analyzed to determine its own appropriate steps for the direct multistep forecasting. Then, each subsequence is forecasted separately according to its corresponding forecasting steps, and all the forecasting results of subsequences are superimposed to get the final forecasting results. This method will help to further improve the forecasting accuracy of the EEMD-FuzzyEn-LSTMNN model, and it will become our next research goal.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, H.; Tian, H.Q.; Chen, C.; Li, Y.F. An experimental investigation of two Wavelet-MLP hybrid frameworks for wind speed prediction using GA and PSO optimization. *Electr. Power Energy Syst.* **2013**, *52*, 161–173. [CrossRef]
2. Jiang, P.; Wang, Y.; Wang, J.Z. Short-term wind speed forecasting using a hybrid model. *Energy* **2017**, *119*, 561–577. [CrossRef]
3. Hu, J.M.; Wang, J.Z.; Zeng, G.W. A hybrid forecasting approach applied to wind speed time series. *Renew. Energy* **2013**, *60*, 185–194. [CrossRef]
4. Chang, G.W.; Lu, H.J.; Chang, Y.R.; Lee, Y.D. An improved neural network-based approach for short-term wind speed and power forecast. *Renew. Energy* **2017**, *105*, 301–311. [CrossRef]
5. Landberg, L. Short-term prediction of local wind conditions. *J. Wind Eng. Ind. Aerodyn.* **2001**, *89*, 235–245. [CrossRef]
6. Martí, I.; Cabezón, D.; Villanueva, J.; Sanisidro, M.J.; Loureiro, Y.; Cantero, E.; Sanz, J. LocalPred and RegioPred. Advanced tools for wind energy prediction in complex terrain. In Proceedings of the 2003 European Wind Energy Conference and Exhibition, (EWEC'03), Madrid, Spain, 16–19 June 2003.
7. Focken, U.; Lange, M.; Waldl, H.P. Previento-A wind power prediction system with an innovative upscaling algorithm. In Proceedings of the 2001 European Wind Energy Conference and Exhibition, (EWEC'01), Copenhagen, Denmark, 2–6 July 2001.
8. Brown, B.G.; Katz, R.W.; Murphy, A.H. Time series models to simulate and forecast wind speed and wind power. *J. Clim. Appl. Meteorol.* **1984**, *23*, 1184–1195. [CrossRef]
9. Torres, J.L.; García, A.; De Blas, M.; De Francisco, A. Forecast of hourly average wind speed with ARMA models in Navarre (Spain). *Sol. Energy* **2005**, *79*, 65–77. [CrossRef]
10. El-Fouly, T.H.M.; El-Saadany, E.F.; Salama, M.M.A. Grey predictor for wind energy conversion systems output power prediction. *IEEE Trans. Power Syst.* **2006**, *21*, 1450–1452. [CrossRef]
11. Mohandes, M.A.; Halawani, T.O.; Rehman, S.; Hussain, A.A. Support vector machines for wind speed prediction. *Renew. Energy* **2004**, *29*, 939–947. [CrossRef]
12. Cadenas, E.; Rivera, W. Short term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks. *Renew. Energy* **2009**, *34*, 274–278. [CrossRef]
13. Flores, P.; Tapia, A.; Tapia, G. Application of a control algorithm for wind speed prediction and active power generation. *Renew. Energy* **2005**, *30*, 523–536. [CrossRef]
14. Li, G.; Shi, J. On comparing three artificial neural networks for wind speed forecasting. *Appl. Energy* **2010**, *87*, 2313–2320. [CrossRef]
15. Salcedo-Sanz, S.; Pérez-Bellido, Á.M.; Ortiz-García, E.G.; Portilla-Figueras, A.; Prieto, L.; Paredes, D. Hybridizing the fifth generation mesoscale model with artificial neural networks for short-term wind speed prediction. *Renew. Energy* **2009**, *34*, 1451–1457. [CrossRef]

16.　Ortiz-García, E.G.; Salcedo-Sanz, S.; Pérez-Bellido, Á.M.; Gascón-Moreno, J.; Portilla-Figueras, J.A.; Prieto, L. Short-term wind speed prediction in wind farms based on banks of support vector machines. *Wind Energy* **2011**, *14*, 193–207. [CrossRef]

17.　Giorgi, M.G.D.; Ficarella, A.; Tarantino, M. Assessment of the benefits of numerical weather predictions in wind power forecasting based on statistical methods. *Energy* **2011**, *36*, 3968–3978. [CrossRef]

18.　Wang, D.Y.; Luo, H.Y.; Grunder, O.; Lin, Y.B. Multi-step ahead wind speed forecasting using an improved wavelet neural network combining variational mode decomposition and phase space reconstruction. *Renew. Energy* **2017**, *113*, 1345–1358. [CrossRef]

19.　Yang, H.F.; Jiang, Z.P.; Lu, H.Y. A hybrid wind speed forecasting system based on a 'decomposition and ensemble' strategy and fuzzy time series. *Energies* **2017**, *10*, 1422. [CrossRef]

20.　Liu, H.; Tian, H.Q.; Pan, D.F.; Li, Y.F. Forecasting models for wind speed using wavelet, wavelet packet, time series and Artificial Neural Networks. *Appl. Energy* **2013**, *107*, 191–208. [CrossRef]

21.　Guo, Z.H.; Zhao, W.G.; Lu, H.Y.; Wang, J.Z. Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. *Renew. Energy* **2012**, *37*, 241–249. [CrossRef]

22.　Liu, H.; Chen, C.; Tian, H.Q.; Li, Y.F. A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. *Renew. Energy* **2012**, *48*, 545–556. [CrossRef]

23.　Huang, N.E.; Shen, Z.; Long, S.R. A new view of nonlinear water waves: The Hilbert spectrum. *Annu. Rev. Fluid Mech.* **1999**, *31*, 417–457. [CrossRef]

24.　Wang, S.X.; Zhang, N.; Wu, L.; Wang, Y.M. Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method. *Renew. Energy* **2016**, *94*, 629–636. [CrossRef]

25.　Zhao, Z.; Chen, W.H.; Wu, X.M.; Chen, P.C.Y.; Liu, J.M. LSTM network: A deep learning approach for short-term traffic forecast. *IET Intell. Transp. Syst.* **2017**, *11*, 68–75. [CrossRef]

26.　Srivastava, S.; Lessmann, S. A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. *Sol. Energy* **2018**, *162*, 232–247. [CrossRef]

27.　Kim, H.Y.; Won, C.H. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Syst. Appl.* **2018**, *103*, 25–37. [CrossRef]

28.　Zhang, J.F.; Zhu, Y.; Zhang, X.P.; Ye, M.; Yang, J.Z. Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. *J. Hydrol.* **2018**, *561*, 918–929. [CrossRef]

29.　Yu, J.; Kim, S. Locally-weighted polynomial neural network for daily short-term peak load forecasting. *Int. J. Fuzzy Log. Intell. Syst.* **2016**, *16*, 163–172. [CrossRef]

30.　Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [CrossRef]

31.　Wang, J.J.; Zhang, W.Y.; Li, Y.N.; Wang, J.Z.; Dang, Z.L. Forecasting wind speed using empirical mode decomposition and Elman neural network. *Appl. Soft Comput.* **2014**, *23*, 452–459. [CrossRef]

32.　Wu, Z.H.; Huang, N.E. A study of the characteristics of white noise using the empirical mode decomposition method. *Proc. Math. Phys. Eng. Sci.* **2004**, *460*, 1597–1611. [CrossRef]

33.　Wu, Z.H.; Huang, N.E. Ensemble Empirical Mode Decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41. [CrossRef]

34.　Chen, W.T.; Wang, Z.Z.; Xie, H.B.; Yu, W.X. Characterization of surface EMG signal based on fuzzy entropy. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2007**, *15*, 266–272. [CrossRef] [PubMed]

35.　Grzegorzewski, P. On separability of fuzzy relations. *Int. J. Fuzzy Log. Intell. Syst.* **2017**, *17*, 137–144. [CrossRef]

36.　Novák, V. Detection of structural breaks in time series using fuzzy techniques. *Int. J. Fuzzy Log. Intell. Syst.* **2018**, *18*, 1–12. [CrossRef]

37.　Bengio, Y.; Simard, P.Y.; Frasconi, P. Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Netw.* **1994**, *5*, 157–166. [CrossRef] [PubMed]

38.　Hochreiter, S.; Schmidhuber, J. Long Short-term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]

39.　Gers, F.A.; Schmidhuber, J.; Cummins, F. Learning to forget: Continual prediction with LSTM. In Proceedings of the 9th International Conference on Artificial Neural Networks, (ICANN'99), Edinburgh, UK, 7–10 September 1999.

40. Werbos, P.J. Backpropagation through time: What it does and how to do it. *Proc. IEEE* **1990**, *78*, 1550–1560. [CrossRef]

41. Reshef, D.N.; Reshef, Y.A.; Finucane, H.K.; Grossman, S.R.; Mcvean, G.; Turnbaugh, P.J.; Lander, E.S.; Mitzenmacher, M.; Sabeti, P.C. Detecting novel associations in large data sets. *Science* **2011**, *334*, 1518–1524. [CrossRef]

42. Madsen, H.; Pinson, P.; Kariniotakis, G.; Nielsen, H.A.; Nielsen, T.S. Standardizing the performance evaluation of short-term wind power prediction models. *Wind Eng.* **2005**, *29*, 475–489. [CrossRef]