Research Article

Short-Term Wind Speed Forecasting Using Decomposition-Based Neural Networks Combining Abnormal Detection Method

Xuejun Chen,[1] Jing Zhao,[2] Wenchao Hu,[1] and Yufeng Yang[3]

[1] *Gansu Meteorological Service Center, Lanzhou, Gansu 730020, China*
[2] *School of Mathematics & Statistics, Lanzhou University, Lanzhou, Gansu 730000, China*
[3] *Gansu Meteorological Information & Technique Support & Equipment Center, Lanzhou, Gansu 730020, China*

Correspondence should be addressed to Jing Zhao; zhaoj_2010@lzu.edu.cn

As one of the most promising renewable resources in electricity generation, wind energy is acknowledged for its significant environmental contributions and economic competitiveness. Because wind fluctuates with strong variation, it is quite difficult to describe the characteristics of wind or to estimate the power output that will be injected into the grid. In particular, short-term wind speed forecasting, an essential support for the regulatory actions and short-term load dispatching planning during the operation of wind farms, is currently regarded as one of the most difficult problems to be solved. This paper contributes to short-term wind speed forecasting by developing two three-stage hybrid approaches; both are combinations of the five-three-Hanning (53H) weighted average smoothing method, ensemble empirical mode decomposition (EEMD) algorithm, and nonlinear autoregressive (NAR) neural networks. The chosen datasets are ten-minute wind speed observations, including twelve samples, and our simulation indicates that the proposed methods perform much better than the traditional ones when addressing short-term wind speed forecasting problems.

1. Introduction

Renewable energy is considered to be the most promising alternative energy resource and plays a significant role in securing a long-term sustainable energy supply while reducing global atmospheric emissions [1]. Renewable energies are highly expected to develop as clean, alternative energy resources. As the most active member of this group, wind power demonstrates considerable benefits and good potential. Because the development of wind power generation requires accurate information regarding wind resources, especially regional wind speeds, wind-related forecasting techniques have become a focal point for many studies. The literature shows that the error prediction costs can reach as high as 10% of a wind farm's annual total income from selling energy [2]. On the other hand, advanced estimates of wind speeds can provide useful information for the dispatching sector of a wind-power-connected system. Such a capacity may be able to influence the wind power connection and enhance grid stability. Due to these improvements, the penetration of wind power could be increased, and the existing energy structures would be greatly changed in the future.

In its actual generation, the main obstacle to the development of a wind-based industry is the variability of the turbines' output power, which seriously limits wind power penetration and threatens grid security. As one of the most important factors in estimating the output of wind turbines, wind speeds can easily be influenced by other meteorological factors, such as air temperature and air pressure, as well as obstacles and terrain. The uncertainty of wind speed forecasting results not only from the forecasting models but also from measurement errors in the meteorological factors. These measurement errors cannot be eliminated due to systematic errors of the instruments and methods of observation. As a result, wind speed forecasting is not easy to address; moreover, wind speed modeling now has become

one of the most difficult problems to tackle [3, 4]. Thus, this paper focuses on this meaningful area of research.

Researchers have put great effort into wind speed modeling and forecasting. Various wind speed forecasting methods have been proposed in the literature to predict wind speeds at different time horizons. Particularly for short-term wind speed forecasting, which is important for scheduling, controlling, and dispatching energy conversion systems [5], these methods can be classified into two categories: physical methods and statistical methods. Physical methods are often referred to as meteorological predictions of wind speed, involving the numerical approximation of models that describe the state of atmosphere [6] and including the WRF model. These models always use physical data, such as temperature, pressure, and topography information, to predict future wind speed [7, 8]. Unlike physical models, statistical methods make forecasts by uncovering the relationships within the observed wind speed time series. They use historical wind speed data and sometimes other variables (e.g., wind direction or temperature) to build the statistical structures from which the forecasts are derived. The data used are recorded at the observation site or at other nearby locations where data are available. In the literature, many statistical methods have been applied to this topic, such as the autoregressive integrated moving average (ARIMA) model, Kalman filters, the generalized autoregressive conditional heteroscedasticity (GARCH) model, and more. These statistical models can be used at any stage in the modeling and often combine various methods into one. Currently, grey models (GM) [9, 10] and some other new methods based on artificial intelligence (AI) techniques have been developed to address these problems. Examples include artificial neural networks (ANNs) of multilayer perceptrons (MLP) [11, 12], radial basis function (RBF) [13], recurrent neural networks (RNNs) [14, 15], and fuzzy logic [16, 17]. In fact, forecasting methods are conventionally not even classified as physical or statistical, as most modern methods include both.

Because each category of the abovementioned methods has its own strengths and weaknesses, complex modeling and forecasting problems cannot be well addressed by any single one. The combination of different methods has been shown, both by theoretical and empirical findings, to be an efficient and effective way of improving model performances [18–21]. Some early works provide many useful hybrid forecasting models that combine strengths from different models to enhance their forecasting performance. When considering the wind speed forecasting problem, it is difficult to obtain a high accuracy level with a single forecasting technique, mainly due to the strong and random fluctuations of wind speed series. In particular, if the focus falls on short-term wind speeds, data series always show irregular variations on a short time scale, increasing the complexity of modeling short-term wind speeds. Currently, combining different models has become a popular trend in wind speed estimations, and the following paper also adopts this concept to address short-term wind speed modeling and forecasting.

The main contribution of this paper is its development of two hybrid methods, consisting of abnormal value detection and modification, decomposition and reconstruction processes, and AI-based training and forecasting. The proposed approaches utilize a novel mixture of three technical tools, including the medians of five-three-Hanning (53H) weighted average smoothing method, the ensemble empirical mode decomposition (EEMD) method, and the nonlinear autoregressive (NAR) neural networks. The developed approaches are able to provide more accurate forecasting results than traditional techniques to address the tough but significant problem of short-term wind speed forecasting on wind farms.

It is well known that short-term wind speed series are difficult to model or predict, mainly due to their strong and random variation within a short time scale. For this reason, data processing is necessary to filter the abnormal values, which greatly impact the model's performance, and to extract valid information from the raw dataset for wind speed modeling and forecasting. In this paper, the 53H method is adopted as the first step of a hybrid wind speed modeling procedure and used to detect and remove abnormal values from the original raw series. Thus, a modified series can be obtained for use as the input data for the next modeling stage. After that, a decomposition and reconstruction process will be pursued according to the EEMD method to remove the noisy information contained in the data series because noise can impose a number of pseudo-variation requirements on models and may affect the correct understanding of data variations [22]. As an adaptive decomposition method, EEMD is based on the local characteristic time scale of the data series. This method constitutes a powerful enhancement of the original empirical model decomposition (EMD) by considering white noise along with the idea of an ensemble mean. This strategy is applicable to nonlinear and nonstationary processes, such as short-term wind speed series. The major advantage of EEMD may be that it automatically identifies the intrinsic time scales of the data without any assumptions regarding signal stationarity. After decomposition by the EEMD method, the local narrow band components, called the intrinsic model functions (IMFs), can be obtained. Each IMF has its own physical meaning and statistical characteristics. Thus, the denoising procedure can be easily accomplished by removing subcomponents with high frequencies and then by reconstructing the data series. In combination with the EEMD decomposition and reconstruction, a modeling and forecasting procedure based on neural networks is developed. As artificial intelligence becomes increasingly important in the technology industry, the interest in neural networks has grown to be substantial. These systems have a strong ability to mimic natural intelligence and are able to learn from examples due to their construction of an input-output mapping without any explicit derivation of the model equation. Their greatest strength is that no knowledge of the internal system parameters is required [23] to offer an acceptable solution. Neural networks fall into different categories, but the one selected for this paper is the NAR neural network, a type of RNNs with feedback arrangement based on a nonlinear autoregressive model.

This paper combines these three methodologies in two ways. Employing the 53H method in the first stage, a modified series removing abnormal values will be obtained and used

for the decomposition process. After addressing the decomposed IMFs by removing the subcomponents with high frequencies, the first goal of the hybrid method is to model and forecast each IMF using a NAR neural network. Then, the forecasted signal is reconstructed as the final result. The second method is to reconstruct the signal first by adding the noise-filtered IMFs, followed by the modeling and forecasting of the reconstructed signal with a NAR neural network to obtain the final forecasting result. Both approaches to model combination enhance the model's performance significantly with respect to short-term wind speed forecasting, although the second combination performs better in most cases. The major highlight of this paper is its novel mixture of abnormal detection, signal decomposition, and AI-based training and forecasting processes. As the first hybrid stage approaches, the 53H method contributes to the detection and modification of the abnormal values in the raw datasets; this strategy is quite effective when considering greatly fluctuating series, such as short-term wind speed observations. In addition, the use of two combined approaches to the decomposition and forecasting sectors enables the proposed hybrid methods to ably capture the complex characteristics of the original wind speed series, thus promoting their applicability to different datasets. The data sampled in this paper represent wind speed observations at ten-minute intervals that were collected from three stations in western China. For each station, observations in January, April, July, and October were chosen as the representations of winter, spring, summer, and autumn, respectively. The simulated results indicate that the developed hybrid models have a strong capacity for successful short-term wind speed modeling and forecasting, vastly outperforming the methods chosen for comparison.

The rest of this paper is organized as follows. Section 2 introduces the related works, and two proposed hybrid models. A case study and model discussion are provided in Section 3 in detail. Then, Section 4 reaches the conclusions.

2. Methods

This section reviews the methodologies and works related to the proposed hybrid models in the following text.

2.1. Medians of Five-Three-Hanning Weighted Average Smoothing.

The medians of five-three-Hanning weighted average smoothing, presented by Tukey [24, 25], are a statistical method of abnormal value removing. It is also termed the 53H method, where "5" is a method for a median of five smoothing, "3" is for a median of three smoothing, and "H" denotes Hanning smoothing. This method holds the principle that the median is a robust estimator of the mean to produce a smooth time sequence that can be extracted from the raw signal.

Let the raw time sequence be x_i, $i = 1, 2, \ldots, n$, where n is the total number of points in the signal. Then, the 53H method can be expressed as the following steps.

(i) Five-point moving average smoothing: construct a sequence $y_i^{(1)}$ from the median of five original data

points from x_{i-2} to x_{i+2}, $i = 3, 4, \ldots, n-3, n-2$. For a five-point smooth,

$$y_i^{(1)} = \frac{1}{9}\left(x_{i-2} + 2x_{i-1} + 3x_i + 2x_{i+1} + x_{i+2}\right). \tag{1}$$

(ii) Three-point moving average smoothing: construct a sequence $y_i^{(2)}$ from the median of five original data points from $y_{i-1}^{(1)}$ to $y_{i+1}^{(1)}$, $i = 4, 5, \ldots, n-4, n-3$. It can be defined as

$$y_i^{(2)} = \frac{1}{3}\left(y_{i-1}^{(1)} + y_i^{(1)} + y_{i+1}^{(1)}\right). \tag{2}$$

(iii) Hanning moving average smoothing: take Hanning filter to obtain the final smoothed signal, as

$$y_i^{(3)} = \frac{1}{4}\left(y_{i-1}^{(2)} + 2y_i^{(2)} + y_{i+1}^{(2)}\right),$$
$$i = 5, 6, \ldots, n-5, n-4. \tag{3}$$

(iv) Construct the sequence $z_i = |x_i - y_i^{(3)}|$, $i = 5, 6, \ldots, n-5, n-4$, and reject the point if $z_i > k\sigma$, where k is a predetermined threshold and σ is the standard deviation of x_i; replace the abnormal values in the raw time sequence by $y_i^{(3)}$.

2.2. Empirical Mode Decomposition (EMD).

In the EMD process, the raw signal can be decomposed into locally narrow band components called intrinsic model functions (IMFs), which can be regarded as hidden oscillation models embedded in the original data series [26].

Given a raw data series $x(t)$, the procedure of EMD method can be expressed as follows.

(i) Identify all the local extrema and determine the upper and lower envelops. Their mean is denoted as $m_1(t)$, and the first component is defined as

$$h_1(t) = x(t) - m_1(t). \tag{4}$$

(ii) Ideally, $h_1(t)$ should satisfy the definition of an IMF. However, that is usually not the case. Therefore, it is necessary to repeat the above sifting [27]. Treat $h_1(t)$ as a proto-IMF and define

$$h_{11}(t) = h_1(t) - m_{11}(t). \tag{5}$$

(iii) The above sifting process will be repeated k times, as

$$h_{1k}(t) = h_{1(k-1)}(t) - m_{1k}(t). \tag{6}$$

Until $h_{1k}(t)$ satisfies the stoppage criterion, let $c_1(t) = h_{1k}(t)$. The stoppage criterion is determined by Cauchy type of convergence test, defined as

$$SD_k = \sum_{t=0}^{T} \frac{\left|h_{1(k-1)}(t) - h_{1k}(t)\right|^2}{h_{1(k-1)}^2(t)}. \tag{7}$$

Separate $c_1(t)$ from the rest of the data by

$$r_1(t) = x(t) - c_1(t). \tag{8}$$
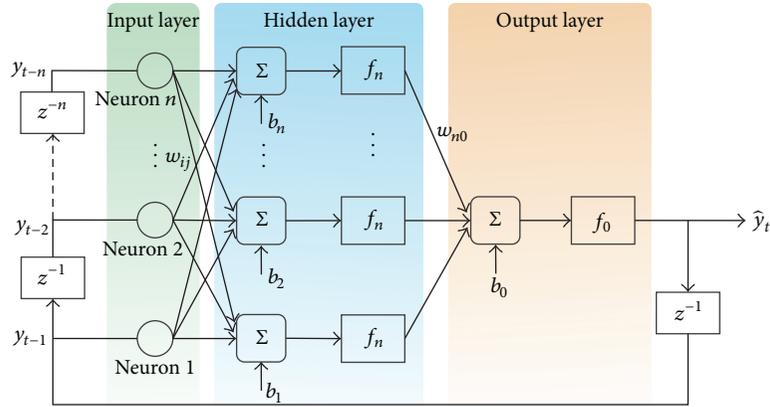
FIGURE 1: The structure of NAR neural networks.

(iv) Repeat the above process with all the subsequences, as

$$r_i(t) = r_{i-1}(t) - c_i(t), \quad i = 2, 3, \dots, n. \tag{9}$$

The process should be stopped when the residue $r_n(t)$ becomes a monotonic function from which no more IMF can be extracted.

(v) Reconstruct the original signal as

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t). \tag{10}$$

When using an EMD method, it does not require any a priori known basis [26, 28]; this indicates that EMD method is completely adaptive to the signal itself.

2.3. Ensemble EMD (EEMD). The EEMD method is a noise-assisted enhancement of the EMD method proposed by Huang et al. [26, 29], since the EMD method has several known difficulties. The first major weakness of the original EMD is the frequent occurrence of model mixing, which is defined as a single IMF [26, 30]. IMFs can also consist of widely disparate scales or can consist of a similar signal residing in different IMF components. Model mixing is often a consequence of an intermittent signal, which can not only cause serious aliasing in the frequency distribution but can also make the physical meaning of individual IMFs seriously unclear. To overcome the scale separation issue without introducing a subjective intermittence test, the ensemble EMD method was presented [30–33] as a powerful modification of the original EMD method with the idea of ensemble mean. The EEMD method adds white noise to the original dataset, as

$$X_i(t) = x(t) + w_i(t). \tag{11}$$

The noise is treated as possible random noise that would be encountered in the measurement process. Then, decompose the dataset with added white noise by using the EMD procedure, and obtain the ensemble means of the corresponding IMFs of the decompositions as the final result.

The main effect of decomposition using EEMD is that the added white noise series cancel each other in the final mean of the corresponding IMFs. This means that the IMFs stay within the natural dyadic filter windows and thus significantly reduce the chance of mode mixing and preserve the dyadic property.

2.4. Nonlinear Autoregressive (NAR) Neural Network. Artificial neural network (ANN), a widely used category of neural networks, is considered an intelligent system that has strong ability to recognize time series patterns and nonlinear characteristics [34, 35]. ANN combines artificial neurons to process information; simple neurons are connected by weight links and this sets up a network. Each input is multiplied by those weights computed by a mathematical function which defines the activation of the neuron. There also is another activation function, which computes the output of the artificial neural; it depends on a certain threshold.

The output of a neuron can be written as

$$y = f\left(b + \sum_i w_i x_i\right), \tag{12}$$

where b is the bias for the neuron. The bias input to a neuron can be regarded as an offset value. It helps the signal to exceed the threshold of activation function, where the activation function is denoted as f. y represents the output; x_i and w_i are the inputs and weights, respectively.

Neural networks can be classified into dynamic and static categories. Static networks have no feedback elements and contain no delays; the output is calculated directly from the input through feedforward connections. Relatively in dynamic networks, the output depends not only on the current input to the network, but also on the current or previous inputs, outputs, or states of the network. Among them, recurrent neural network (RNN) is an essential class where connections between units form a directed cycle (Figure 1).

It allows exhibiting dynamic temporal behavior by creating an internal state of the network. The NAR neural network is a type of recurrent network with feedback arrangement based on a nonlinear autoregressive model. Thus, in a NAR network, there is a feedback of the true output instead of the estimated one in the input. This allows using only static backpropagation when training the network and also makes the feedforward architecture more accurate.

Currently, it is commonly used in multistep ahead time series forecasting. The forecasting value is determined by the following equation:

$$\widehat{y}(t) = f\left(y(t-1) + y(t-2) + \cdots + y(t-d)\right), \quad (13)$$

where f is a nonlinear function. Specifically, the function values depend only on regressed d previous values of the output signal.

2.5. Proposed Models. This section develops two hybrid methods for short-term wind speed modeling and forecasting. The proposed methods consist of three stages: abnormal value removing and modification, signal decomposition and denoising, and signal reconstruction and NAR-based forecasting. Figure 2 introduces the workflow of the proposed approaches, briefly showing the three stages as follows.

Stage 1: Abnormal Value Removal and Modification. The accurate forecasting of the short-term wind speeds heavily depends on the reliability of the observed data series. Thus, in the first stage, the 53H method is chosen as a detector to find and remove abnormal values from the raw observation series. This stage is composed of a 53H-modified series.

Stage 2: Signal Decomposition and Denoising. During the actual operations of wind farms and power systems, several uncertain factors may influence the data acquisition process, including measurement, recording, conversion, and transmission. Most of these are beyond control, and any of these factors can introduce noises and uncertainties into the wind speed series, leading to poor generalization and undesirable forecasting performances. Thus, wind speed modeling and forecasting is a difficult proposition. To address this problem, the EEMD-based signal filtering method is applied as a denoising process. This stage decomposes the 53H-modified signal into IMFs using the EEMD method and removes the subcomponents with high frequencies.

Stage 3: Signal Reconstruction and NAR-Based Prediction. After addressing the decomposed IMFs by removing the subcomponents with high frequencies, there are two modeled concepts. The first involves the modeling and forecasting of each IMF by the NAR neural network, after which the forecast signal is reconstructed as the final result. This model is named HEN1 (short for the hybrid model 1 combined 53H detection, EEMD filter, and NAR networks). The second involves an initial reconstruction of the signal by adding the noise-filtered IMFs, followed by the modeling and forecasting of the reconstructed signal with a NAR neural network, obtaining the final forecast result. Similarly, this model is denoted as

HEN2 (the hybrid model 2 combined 53H detection, EEMD filter, and NAR networks).

3. Case Study

3.1. Collection of Data. The wind speed data used for the model construction and performance testing in this paper includes observations at ten-minute time intervals that were collected from three sites in the Hexi Corridor of western China. For each station, observations in January, April, July, and October were selected to represent winter, spring, summer, and autumn, respectively, in order to evaluate the applicability of the proposed models. The wind speeds from the different seasons have their own ranges of values and variation; generally, wind speeds in April have larger mean values. The datasets used are displayed in Figure 3; the first 755 points are used for model construction and then the following 144 points are chosen as a testing set. For each station, the figure consists of four parts, mainly included in the information on the selected data series and site descriptions. Basic statistics show that wind speed observations collected from the different sites have their own ranges of variation and statistical characteristics. Generally, the wind speeds from April, the representation of spring, have larger values, while winter always contains lower wind speeds.

3.2. Criteria for Evaluating Forecasting Performance. The experiments in this paper adopt three error criteria, which are commonly used to evaluate the forecasting performances of different models. They are the root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). Let x_i, $i = 1, 2, \ldots, n$, be the observation series with the corresponding forecast $\widehat{x}_i$. These criteria are defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \widehat{x}_i)^2},$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \widehat{x}_i\right|, \quad (14)$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i - \widehat{x}_i}{x_i}\right| \times 100\%.$$

3.3. Simulation Process of Hybrid 53H-EEMD-NAR Method. As described in Section 3, this paper develops two hybrid models for short-term wind speed modeling and forecasting. The guiding concept of hybrid approaches is to construct a mixed procedure, combining an abnormal detection, a modification method, a signal filter method, and NAR-based neural network forecasting. Both of the hybrid models are three-stage procedures, as introduced in Figure 2. This section aims to show the simulation process for each stage and then to provide the forecasting results.

3.3.1. Abnormal Value Removing and Modification. In the first stage of the hybrid approaches, the 53H method is employed
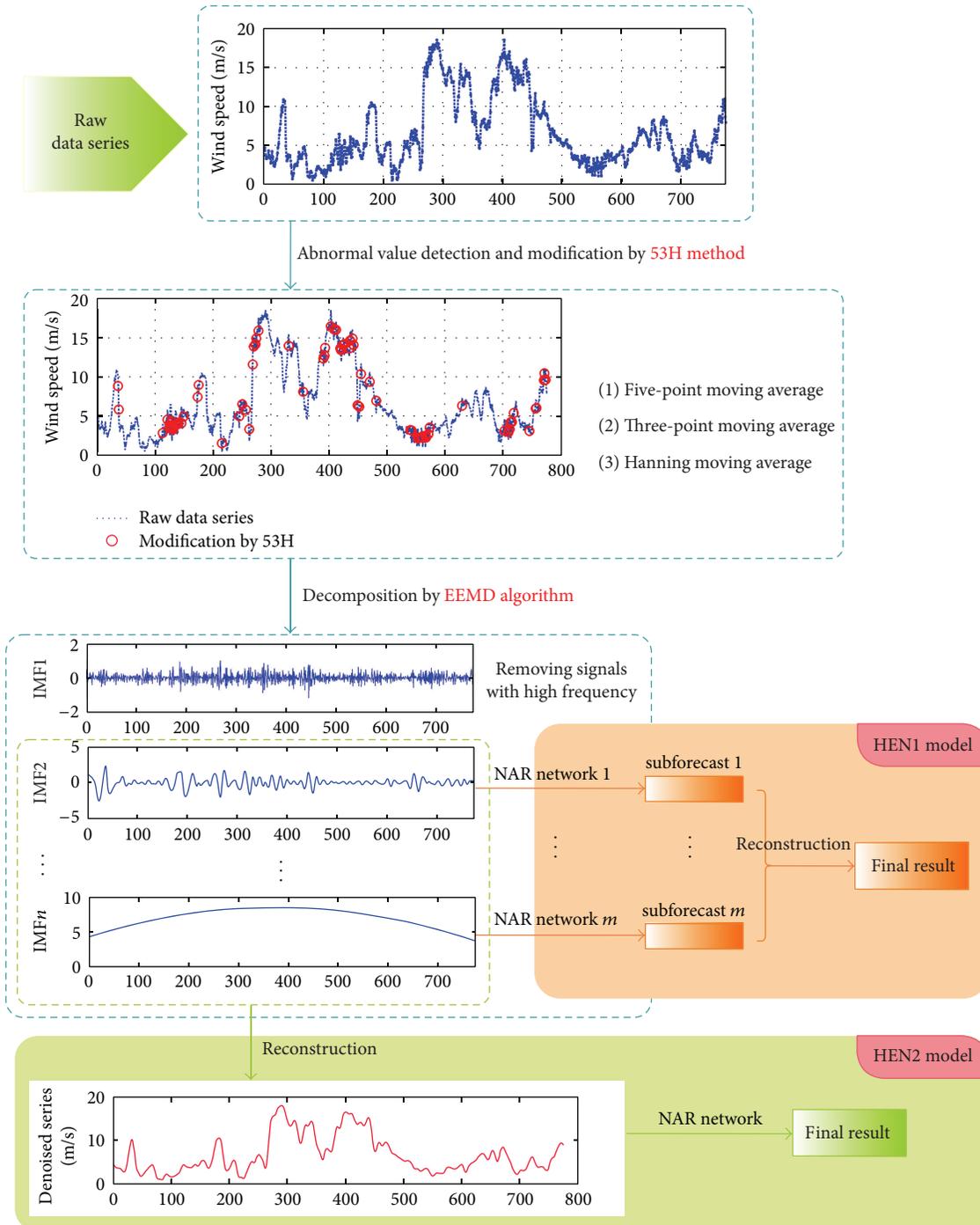
Figure 2: The flowchart of the proposed models.

to remove and modify the abnormal values in the raw data series. This step is necessarily the first stage of the following model construction and estimation because short-term wind speeds always show strong and random variations within a short time interval. Great fluctuations heavily impact the model performance, resulting in poor forecasting accuracy and even model invalidation. As a result, data detection and smoothing modification methods are adopted here to

help the hybrid approaches overcome the interference from abnormal and strong variations.

Figure 4 provides the results obtained from the 53H method, using the wind speed series at Station 1 collected in July 2011 as an example. These data encompass not only the results of the whole dataset but also three enlargements, which are the detection and modification of the strong-variation data segments chosen from the whole set. It is

(A) Selected data series

(B) Frequency distribution
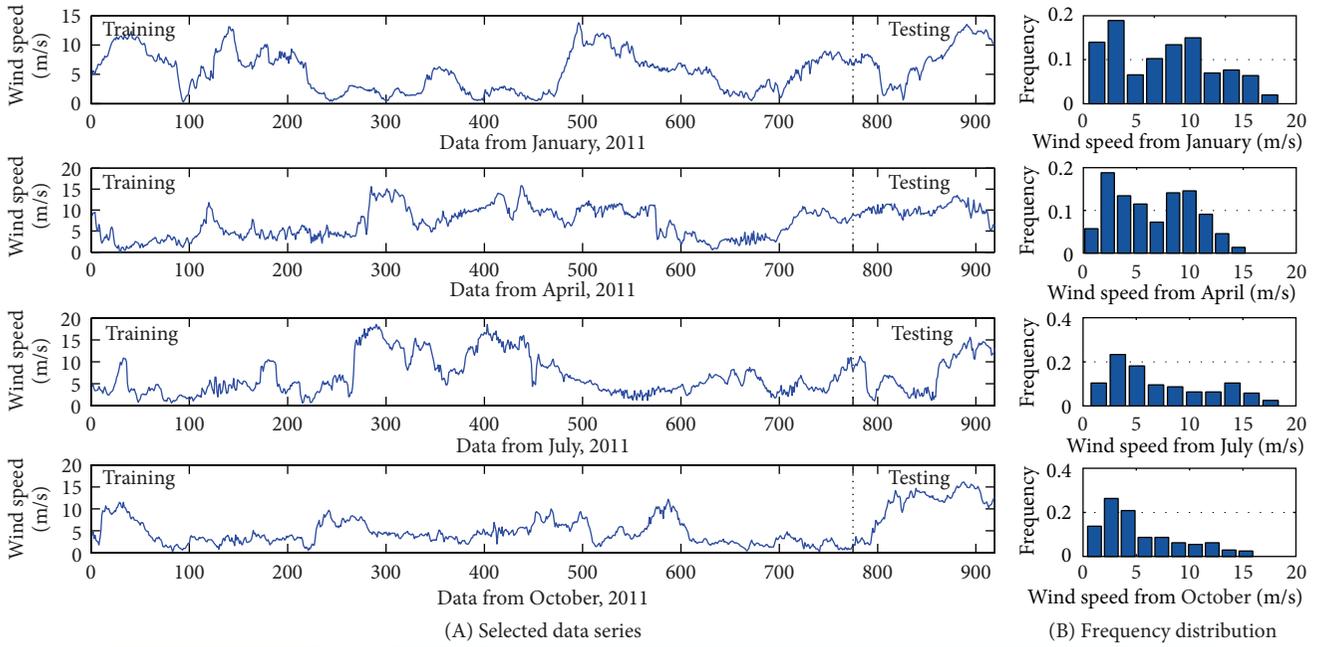
Station 1: Shengdiwan
Longitude: 98.37′E
Latitude: 40.26′N
Altitude: 1230 m
Observation height: 70 m

(C) Site information

	Mean (m/s)	Maximum (m/s)	Minimum (m/s)	Standard deviation	Skewness	Kurtosis
January	5.72	13.80	0.30	3.49	0.271	−0.983
April	6.32	15.30	0.10	3.63	0.211	−1.125
July	6.82	18.50	0.60	4.39	0.823	−0.434
October	5.39	16.10	0.30	3.72	1.032	0.139

(D) Basic statistics of wind speed observation

(a)

(A) Selected data series

(B) Frequency distribution
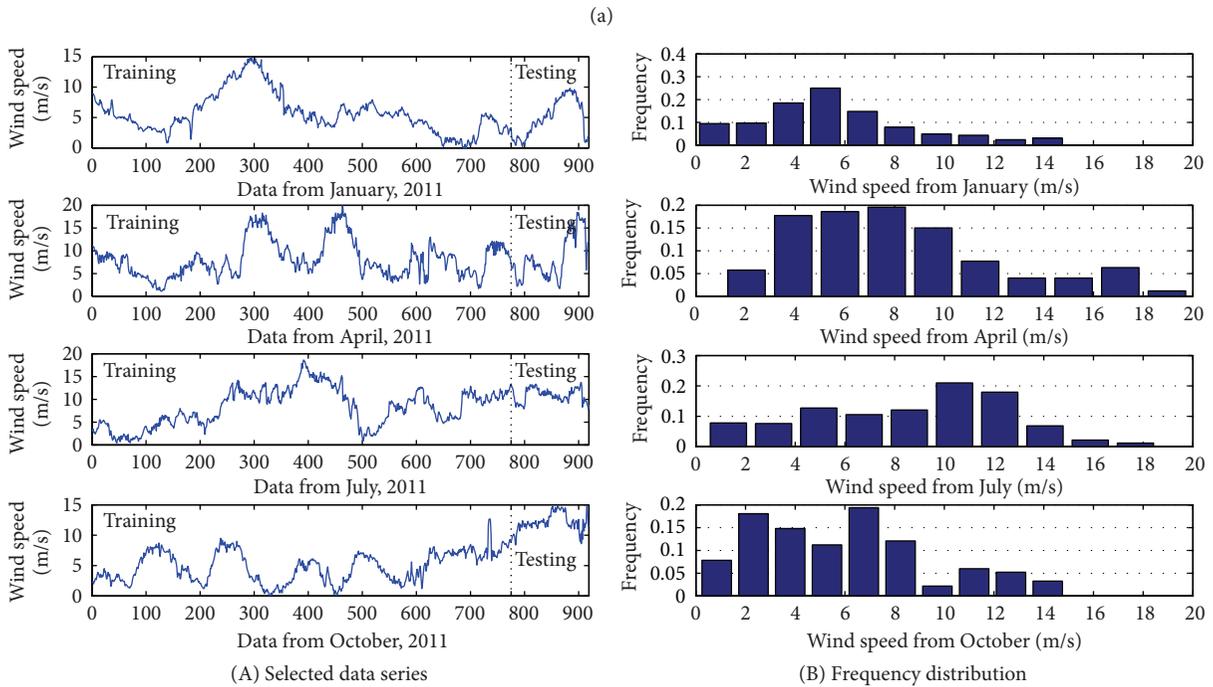
Station 2: Qiaowan
Longitude: 96.67′E
Latitude: 40.70′N
Altitude: 1447 m
Observation height: 100 m

(C) Site information

	Mean (m/s)	Maximum (m/s)	Minimum (m/s)	Standard deviation	Skewness	Kurtosis
January	5.61	14.90	0.10	3.21	0.750	0.345
April	8.17	19.90	1.10	4.09	0.777	−0.075
July	6.84	18.60	0.40	3.95	−0.187	−0.789
October	5.95	14.90	0.10	3.55	0.559	−0.429

(D) Basic statistics of wind speed observation

(b)

Figure 3: Continued.

(A) Selected data series

(B) Frequency distribution

Station 3: Shengdiwan
Longitude: 98.38′E
Latitude: 40.26′N
Altitude: 1230 m
Observation height: 70 m

(C) Site information

	Mean (m/s)	Maximum (m/s)	Minimum (m/s)	Standard deviation	Skewness	Kurtosis
January	5.43	12.50	0.10	2.84	0.287	−0.617
April	7.94	26.90	0.50	5.16	1.094	0.634
July	8.87	16.10	1.20	2.90	−0.182	−0.431
October	7.21	23.10	0.30	3.86	1.377	2.261

(D) Basic statistics of wind speed observation

(c)

FIGURE 3: (a) Selected datasets from Station 1: Tianshuijing. (b) Selected datasets from Station 2: Qiaowan. (c) Selected datasets from Station 3: Shengdiwan.

apparent that the points modified by the 53H method are centralized into several segments and marked by the dashed rectangular boxes. A detailed version of these special segments is provided by three enlargements, which demonstrate the clear view that the raw observations in these segments always have strong variations. Furthermore, the modified data series effectively avoid the disordered fluctuations, instead containing major information on the trends and changes. Figure 4 also illustrates the differences in the basic statistics between the raw data series and the 53H-modified series. The mean values of the 53H-modified series are found to be nearly equal to the mean values of the raw data, all of them located within the narrow interval of ±0.01 m/s. When considering the difference in the standard deviations of the two groups, the 53H method helps to cut down the standard deviations from the raw series to the eleven chosen samples, with their varying degrees of reduction. All of these contribute to the enhanced performance of the hybrid models and to the development of hybrid forecasting accuracy. Consequently, it is of great importance that the 53H-modified series will be used as the input of the EEMD filter in the following stages for the hybrid approaches.

3.3.2. Decomposition and Noise Filter. Moving on to the second stage of the proposed hybrid approaches, the EEMD

method is applied as a signal filter for data denoising. Figure 5 displays this noise reduction process for the data collected at Station 1 in July 2011. Decomposition for the selected sample generates nine IMFs and one residual series. The extracted IMFs represent a range of frequencies, from high to low. The IMFs with higher frequencies represent the pattern of shorter periods, whereas the IMFs with lower frequencies represent the pattern of longer periods. That is, as IMF1 has the highest frequency, it most likely also has the noise information; IMF9 can thus be regarded as the trend term. Generally, if an IMF represents a very short-term pattern, it should be discarded to achieve denoising for the original data series. From Figure 5, it is clear that IMF1 and IMF2 are subsignals with high frequencies, and the patterns corresponding to these two IMFs are of a very short period. Considering the short-term wind speed forecasting problem studied in this paper, the first two IMFs can be extracted to represent very short-term patterns. Figure 5 also displays the removed signal information and the denoised data series, indicating that EEMD-based decomposition and signal filtering are quite effective in removing the interference from noisy signals.

For the other examples used in this paper, Figure 6 provides a simple view of the EEMD-based decomposition and signal filter results. For each sample, the input of EEMD-based decomposition is the 53H-modified data series,

(a) Detection and modification result by 53H method

- • - Actual observation
- ○ Modification by 53H method

(b) Segment enlargement 1

(c) Segment enlargement 2

(d) Segment enlargement 3

		Number of modified points	Mean (m/s)		Standard deviation	
			Value	Variation	Value	Variation
Station 1	Jan.	7	5.73	↗0.01	3.49	0.00
	Apr.	61	7.19	↗0.01	3.56	↘1.97%
	Jul.	75	6.81	↘0.10	4.36	↘0.69%
	Oct.	42	5.38	0.00	3.70	↘0.54%
Station 2	Jan.	18	5.60	↘0.01	3.19	↘0.93%
	Apr.	117	8.18	↗0.01	4.05	↘0.98%
	Jul.	41	8.63	↘0.01	3.93	↘0.76%
	Oct.	19	5.96	↗0.01	3.54	↘0.28%
Station 3	Jan.	23	5.42	↘0.01	2.84	↘0.35%
	Apr.	103	7.94	0.00	5.12	↘0.78%
	Jul.	70	8.89	↗0.02	2.88	↘0.69%
	Oct.	68	7.21	0.00	3.83	↘0.78%

Notice: "Variation" means the range and trend of variations compared with the raw datasets.

(e) Basic statistics between the raw series and 53H-modified series

Figure 4: Data detection and modification by the 53H method, selected from July 2011 at Station 1.

as introduced in Section 3.3.1. "Noise signals" in Figure 6 represent the sum of all of the removed subsignals, while "denoised data series" stand for the filtered series that have discarded the noise information. Obviously, the denoised series is much smoother and is able to contain the major information in the original data series. This step is significant within the proposed hybrid approaches because the EEMD method helps in overcoming the interference from a noisy signal. In the next stage, the denoised data series will be treated as the inputs of the NAR-based procedures.

3.3.3. Reconstruction and Prediction. As introduced above, two methods of combination were designed at this stage to mix the EEMD-filtered signal series and NAR-based training and forecasting procedures.

The first method (denoted as HEN1) was intended to model and forecast each IMF using its respective NAR neural network; the forecast signal would then be reconstructed as the final result. The second method (denoted as HEN2) was intended to reconstruct the signal first by summing the noise-filtered IMFs, followed by the model from the NAR neural network, in order to obtain the final forecasting result. These constructed approaches are employed for one-day-ahead predictions; for each adopted NAR neural network, the number of hidden neurons is determined by an optimal selection process using the training dataset. The

adjustment of data training was performed according to the Levenberg-Marquardt algorithm, which is commonly used to solve nonlinear least squares problems [36]. Additionally, the performance was evaluated by error criteria chosen as MSE for this paper. Figure 7 shows the forecasting results of both proposed models, and data displayed was sampled from Station 1 in July 2011.

Obviously, the forecasting results from the two proposed models perform similarly and approach the original data series. This can be observed in the correlation plots of the predicted wind speeds versus the actual wind speed, which are also displayed in Figure 7. In the correlation plots, the dashed straight line indicates that the predicted wind speed is equal to the actual observation; meanwhile, the farther the points get away from the straight line, the larger the forecasting error is introduced into the hybrid approaches. It is clear that most points are located in a narrow range near to the straight line. As indicated, the proposed hybrid approaches perform well when faced with short-term wind speed forecasting. The frequency distributions also suggest that the forecasting errors concentrate around the zero point; the distributions from the two hybrid approaches seem similar to each other.

The results obtained from the two proposed models were similar; however, each model also demonstrated distinct characteristics and benefits for hybrid forecasting procedures.

Notice: IMF1 and IMF2 are determined as the noisy signal in this sample.
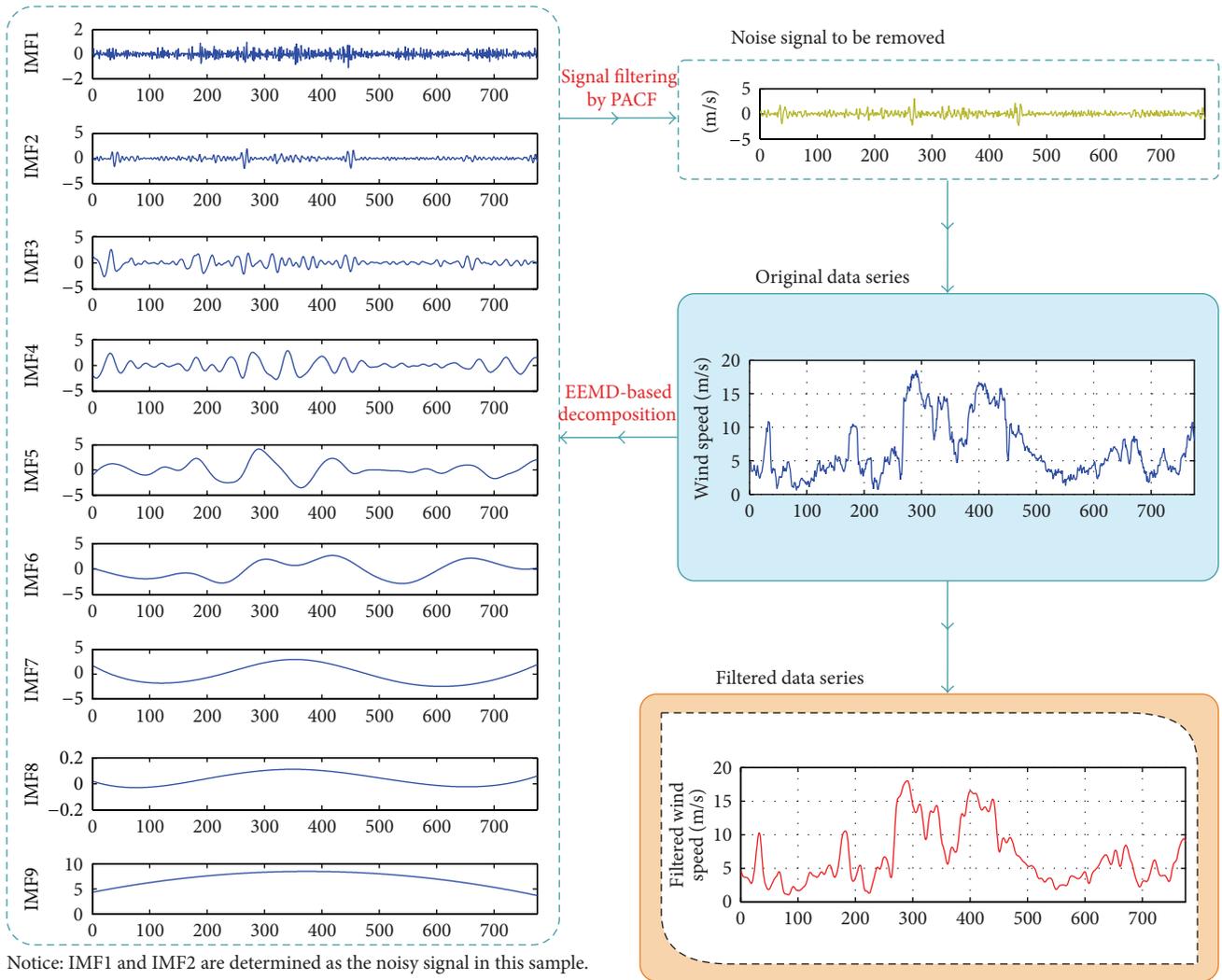
FIGURE 5: The decomposed signals by EEMD method.

The major benefit from the EEMD-based data decomposition model was its capacity to break down the original data series into several subsignals, called IMFs, to satisfy the two requirements introduced in [26]. Therefore, while an IMF does represent a simple oscillatory mode as a counterpart to simple harmonic function, it is also much more general. Instead of the constant amplitude and frequency of a simple harmonic component, an IMF can have variable amplitude and frequency along the time axis. As a result, its modeling and forecasting for each IMF provides more satisfactory results and better accuracy. In another respect, the number of decomposed IMFs cannot be controlled; when the number is comparatively large, the HEN1 model may also generate a large forecasting error because the modeling and forecasting for each IMF introduces model error into the entire hybrid procedure. Just as in the problem discussed in this paper, wind speed series with ten-minute time intervals show a complex mixture of nonlinearity, volatility, and strong randomness. As shown in Figure 5, nine IMFs were extracted

from the original data series using EEMD-based decomposition. Applying the HEN1 model, NAR-based forecasting introduces a large model error in stage 3, especially in several partitions. This is shown in Figure 7; as at several specific points, the HEN1 model has comparatively large forecasting errors. In this setting, the HEN2 model is always able to show its strength by summing the filtered signal series together as the input of a NAR neural network. By removing the noisy signals, the power of the HEN2 model is mainly derived from the smoother nature of the denoisy series, which retains the major information from the original data series. Thus, the following NAR-based forecasting procedure provides a final result.

Figure 8 then provides a brief overview of the predicted results from the two hybrid approaches for all twelve samples used in this paper. Correlation plots indicate that the distances between the hybrid predictions and the actual wind speed are quite close for both of the proposed approaches. Meanwhile, wind speed series with large variations can also
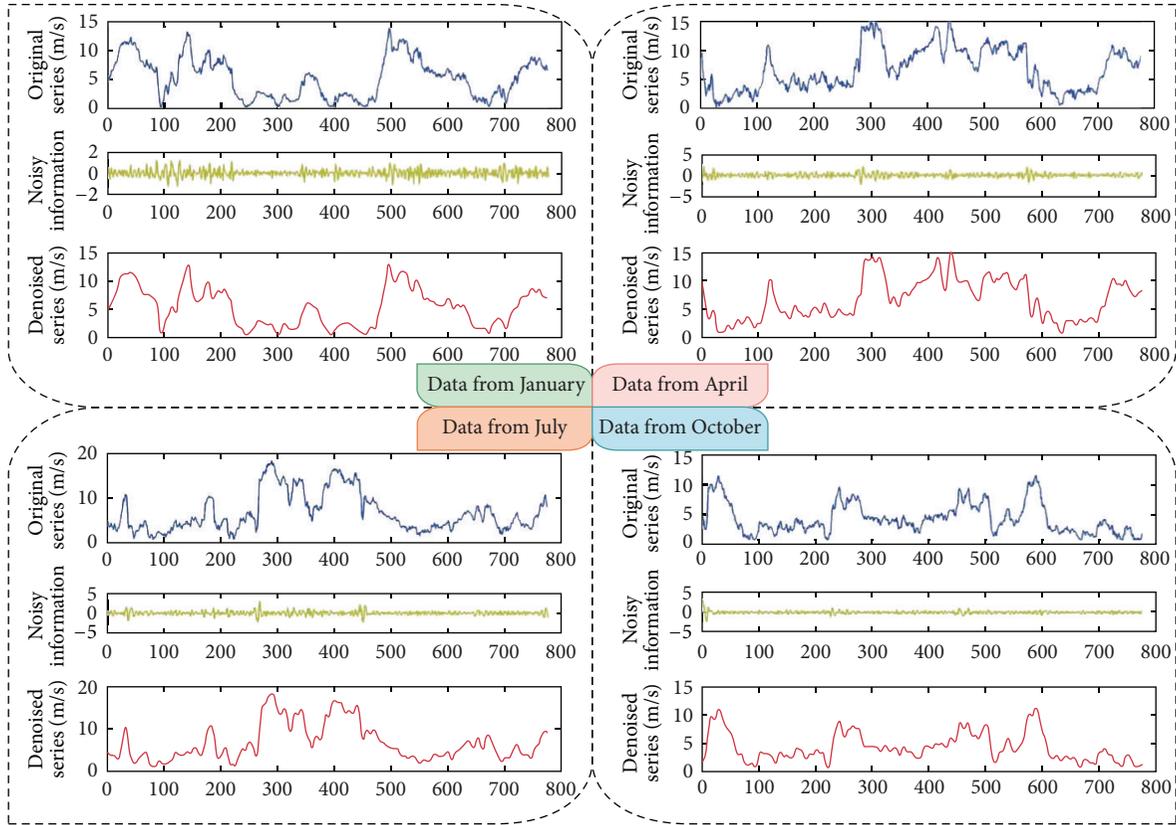
(a)
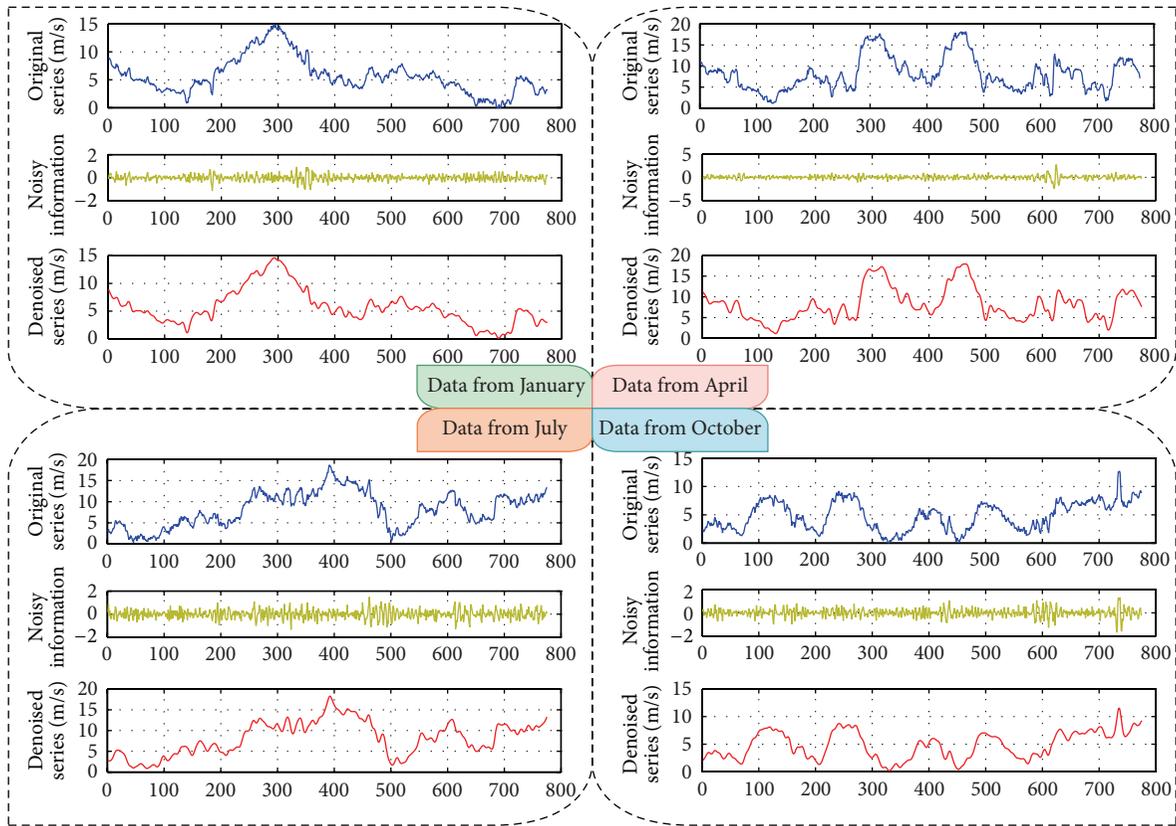
(b)

Figure 6: Continued.

(c)
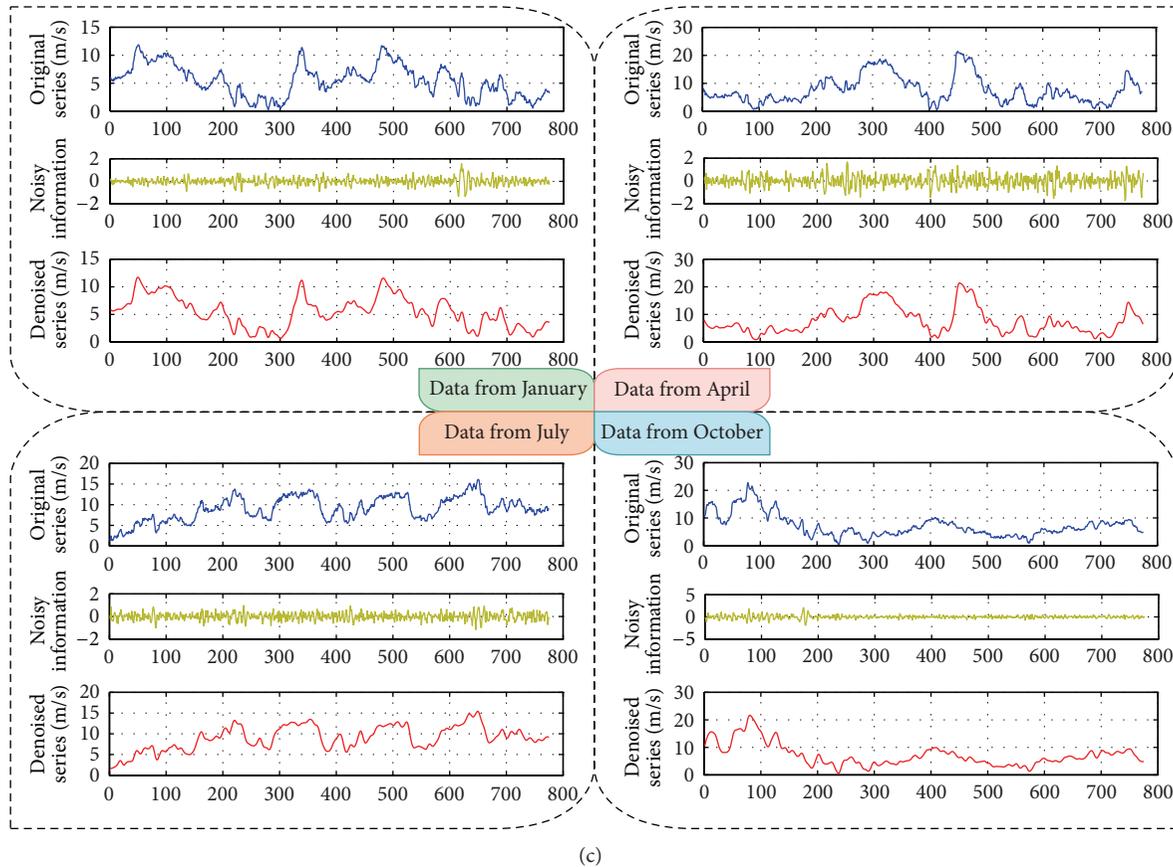
FIGURE 6: (a) Results of EEMD-based decomposition and denoising at Station 1. (b) Results of EEMD-based decomposition and denoising at Station 2. (c) Results of EEMD-based decomposition and denoising at Station 3.

be found, such as the data collected in April at Station 1; in that month, both of the hybrid methods had comparatively large errors. Despite of this, the hybrid approaches maintained the trend of variation, retaining the major information of actual wind speed in this sample. In the corresponding correlation plot, all of the points were concentrated around the given dashed straight line, which indicates that the methods are effective and performed stably for this strongly fluctuating data series. In the following section, the model comparison and performance analysis will be introduced. It will be shown that the proposed hybrid methods perform much better than the other selected models, mainly due to their combination of the strengths from different approaches.

3.4. Analysis and Performance Comparison. To evaluate the forecasting performance of the proposed hybrid approaches, several methods are chosen for model comparisons, including single forecasting methods, hybrid approaches, and commonly used benchmarks. According to the proposed methods in this paper, the single NAR-based neural network, hybrid 53H-NAR method, and hybrid EEMD-NAR method were compared. The generally widely used time series model, called the ARIMA process, was also chosen, as well as the persistent model that is commonly considered a benchmark for short-term forecasting problems.

Figure 9 displays the forecasting results of the different chosen methods, including the two proposed approaches and five selected methods; each method is shown as applied to the data collected at Station 1 from July 2011. The proposed approaches perform much better than the others, particularly in the segments with complex variation and strong fluctuations. Here, three segments are selected for a more detailed explanation, as marked by the dashed rectangular boxes in Figure 9. Segment 1 is representative of data with complex variation; the wind speed in this period has large variance. The single NAR-based neural network reveals an apparently weaker ability with an abnormal final result. This is mainly due to the inferences from both the abnormal values and the noise-signals within the raw data series. This result also indicates that it is difficult to obtain reliable and stable forecasting results using a single method (despite the AI-based forecasting tool chosen here), especially when considering any complex processes similar to short-term wind speed series. The marked Segment 2 follows a downward trend of the wind speed series immediately after a sharp increase; this situation is common in actual observations and is important in an accurate forecasting of short-term wind speeds. The two proposed approaches hold well and follow the changing trends of the actual series. While the single NAR-based neural network and 53H-NAR provide fluctuating results,
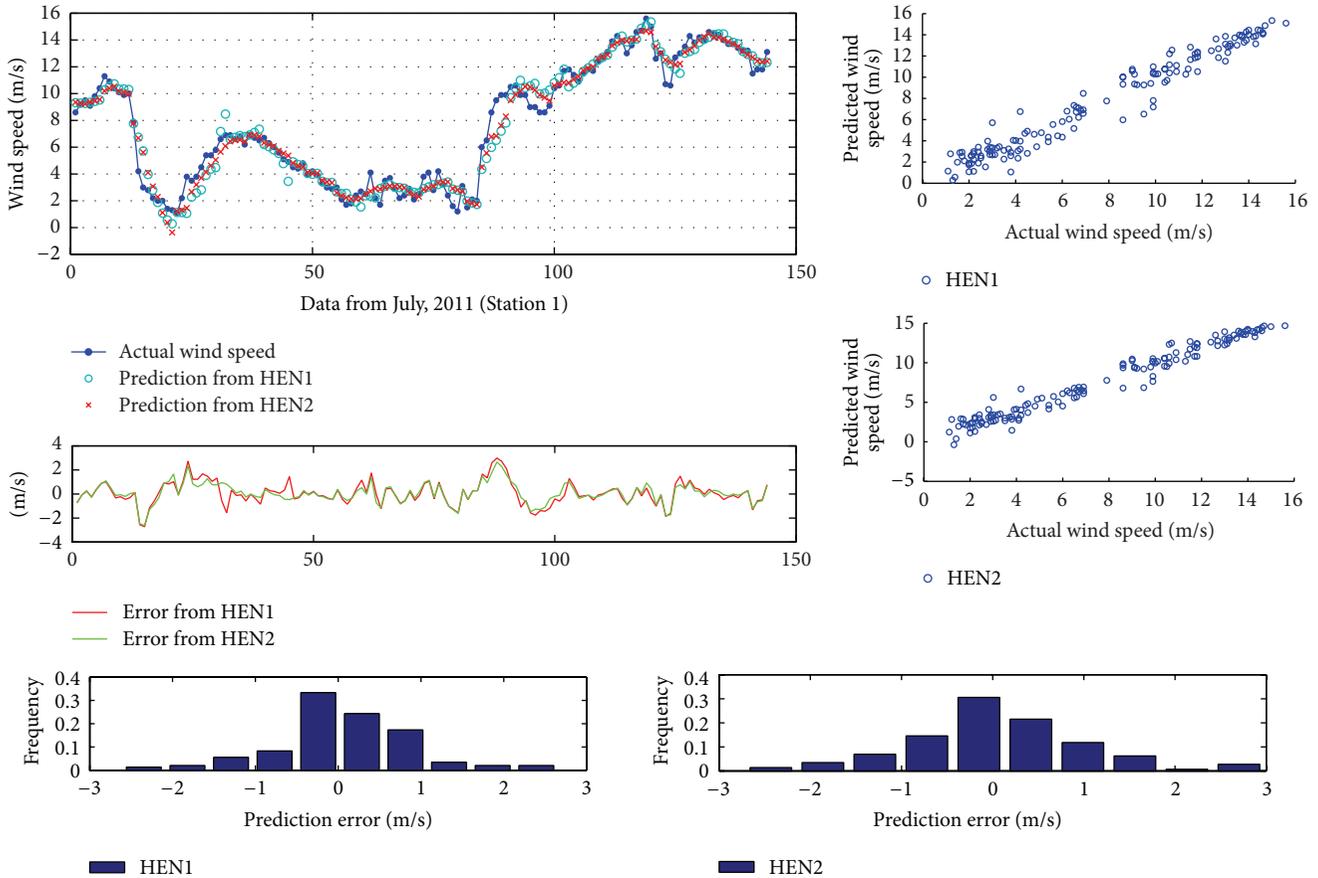
FIGURE 7: The forecasting results from both proposed hybrid models, exemplified by data from July 2011 at Station 1.

the persistent and ARIMA models show apparently delayed trends. Next, Segment 3 was selected because the actual series experienced a rapid decline here before sharply trending upward. The NAR and 53H-NAR methods enlarge the range of ups and downs within the actual series, mainly due to the abnormal values and noisy signals contained in the original series. The EEMD-NAR method performs well in the first half of the decreasing section, but then it moves far away from the trend of the actual observations. Comparing these data to the results of the proposed HEN1 model, which performed much better in this segment, the contribution of the 53H-based modification is obvious. Generally speaking, the two proposed hybrid approaches can provide more reliable and stable results regarding short-term wind speed forecasting problems.

Table 1 and Figure 10 show the performance comparisons among different models for all twelve samples addressed in this paper. Three error criteria are considered, including the RMSE, MAE, and MAPE, introduced in Section 3.2. After that, Table 2 provides the performance comparisons according to data collected from different seasons and from different stations. On the whole, the proposed two approaches provide the final results with minimal statistical error; the HEN2 model performed best in most cases. Analyzing the examinations at Station 1, the two proposed approaches

perform similarly; the values of all three error criteria grow quite close between the two hybrid models. While at Station 3, the HEN2 model shows a stronger ability to forecast short-term wind speeds, compared with the HEN1 model. This may result from difference in the series information between the observations of these two stations. As such, a model's performance may become distinct by inputting different datasets. The average MAPE of the HEN2 model is 8.71%, which is the lowest MAPE among the considered approaches as calculated for the twelve simulated samples. This outcome leads to reductions of 38.25%, 36.57%, and 13.73%, according to MAPE, respectively, when it is compared with the NAR, 53H-NAR, and EEMD-NAR results. This indicates the strong capacity of the proposed HEN2 method to forecast short-term wind speeds, suggesting the benefits of combining these three statistical tools. More specifically, although the HEN2 model maintains the lowest MAPE, its forecasting errors range from 3.63% to 17.20% for different samples. This implies that the model's performance is heavily based on the original input dataset, mainly regarding its internal information and the statistical characteristics of the historical observations. Moreover, this persistent method is always taken as a benchmark in short-term forecasting problems, as in this paper. In Table 1, the forecasting errors from the NAR and ARIMA models get close to the errors

Figure 8: Continued.

Actual wind speed
Prediction from HEN1
Prediction from HEN2

○ HEN1
× HEN2
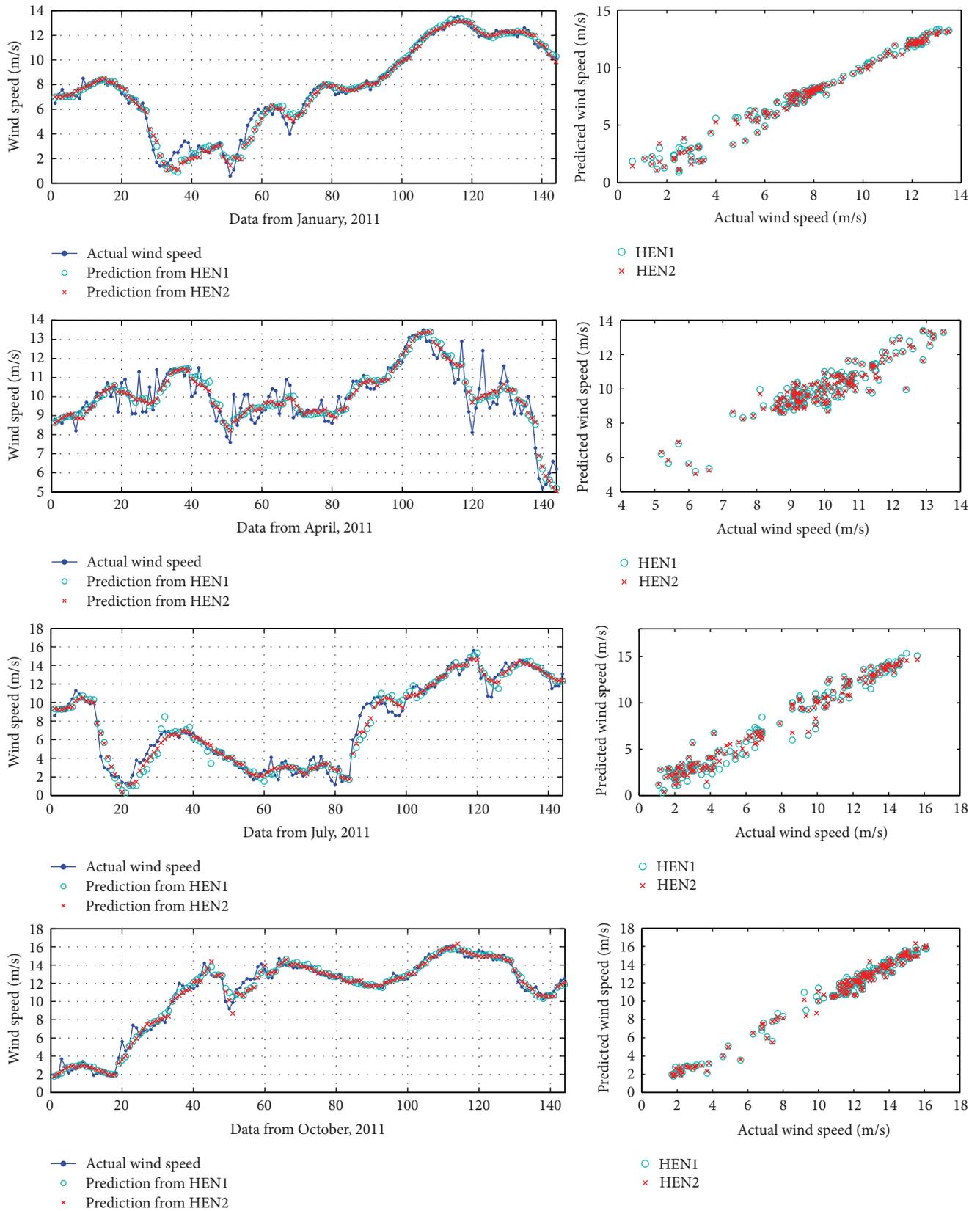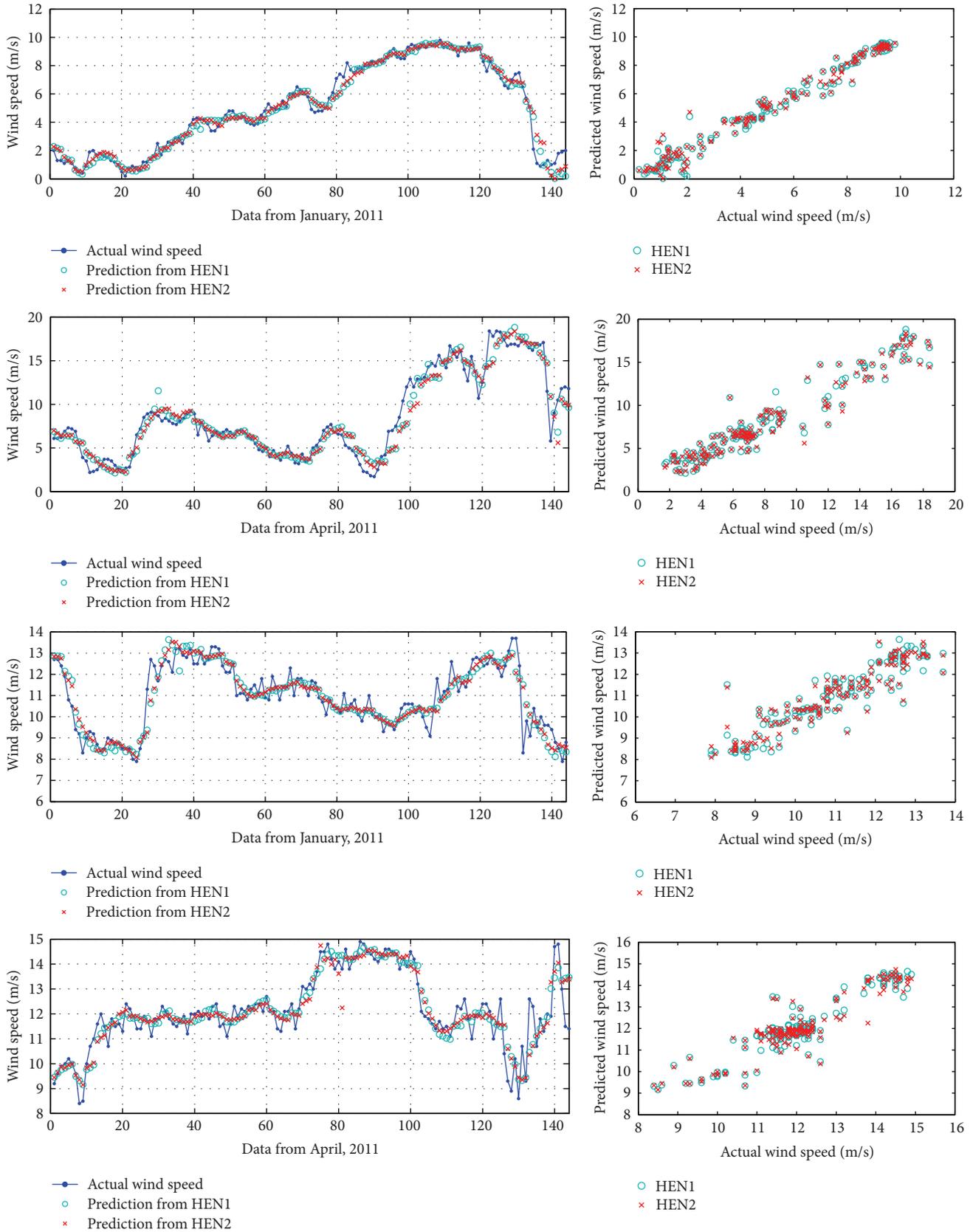
(b)

Figure 8: Continued.
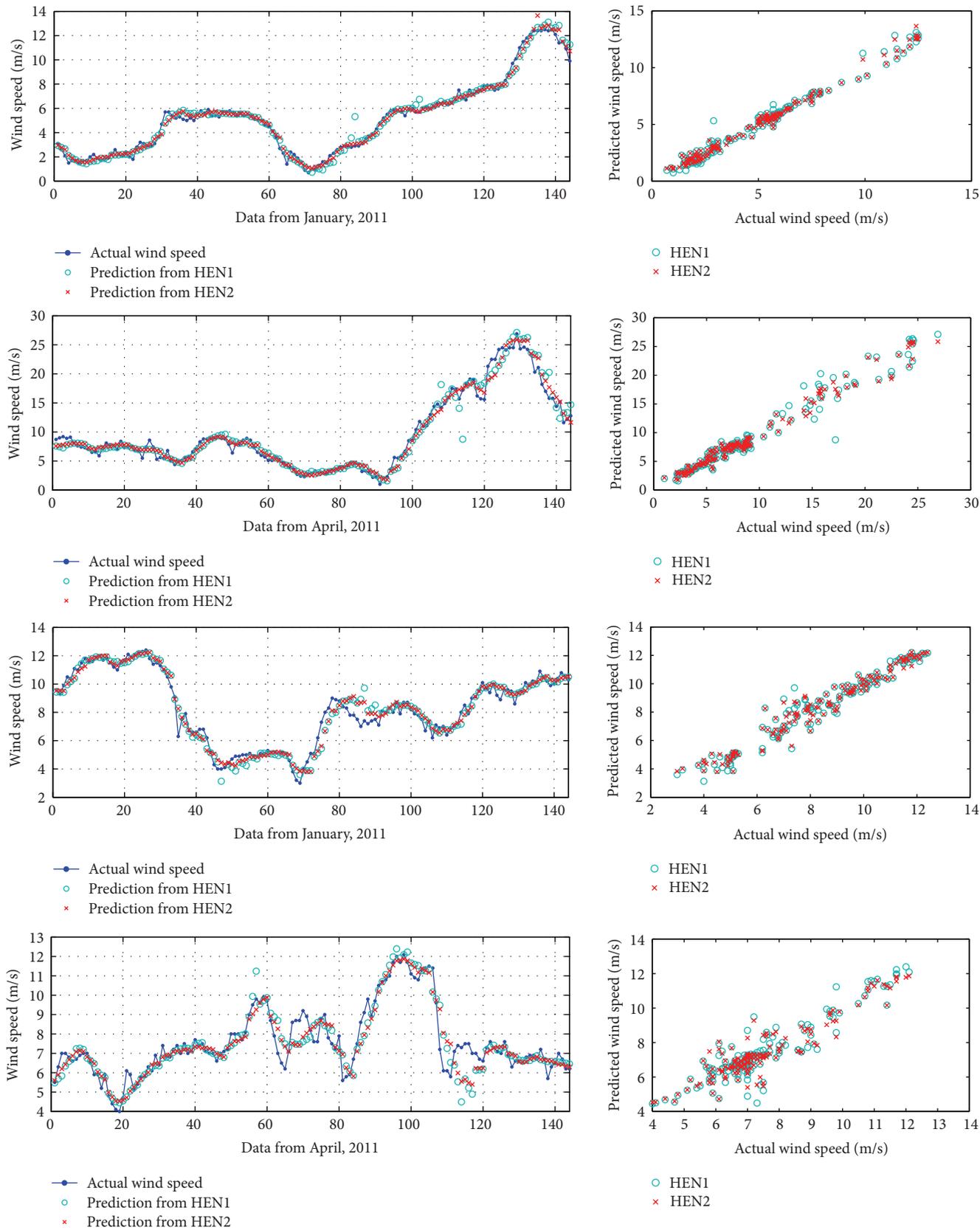
FIGURE 8: (a) Results of hybrid approaches for Station 1. (b) Results of hybrid approaches for Station 2. (c) Results of hybrid approaches for Station 3.
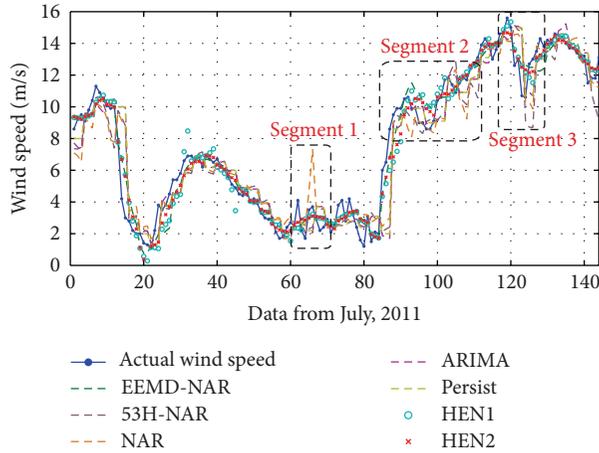
FIGURE 9: Forecasting results from different models, exampled by data at Station 1 from July 2011.
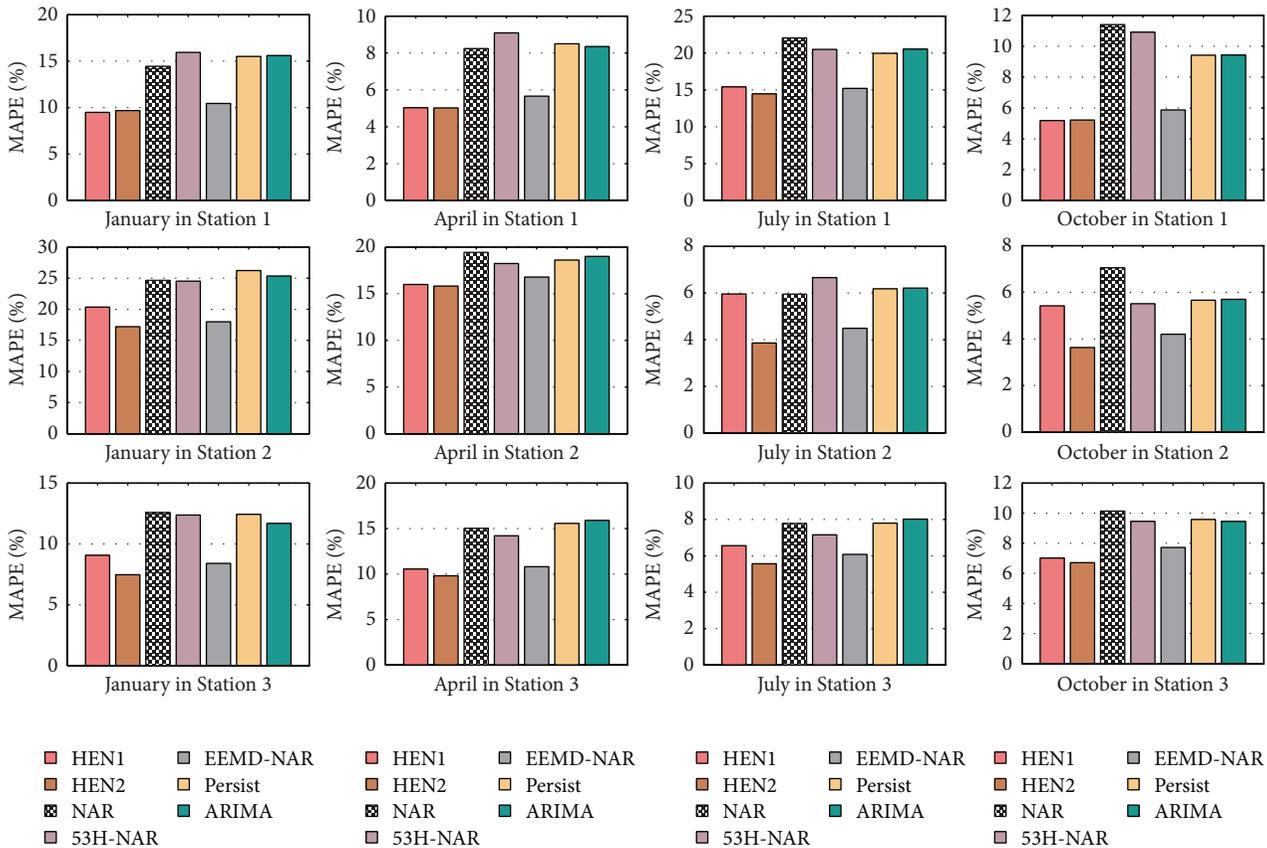
FIGURE 10: Error comparison among different models.

from the benchmark, demonstrating that a single method cannot meet the need for accurate short-term wind speed forecasting.

Table 2 provides a clear comparison among different seasons and stations. Clearly, each model performs similarly across different stations, with slightly higher errors at Station 2. However, there is a large gap in the forecasting errors when the results are considered according to different seasons, as shown in the table. The data from January and April includes significantly greater forecasting errors higher than that from July and October. This situation is identical in all three chosen stations, with the exception of Station 1.

Table 1: Forecasting errors of different models.

Sample	HEN1			HEN2			NAR			53H-NAR			EEMD-NAR			Persist			ARIMA		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MEPE
Site 1																					
Jan.	0.53	0.36	**9.49%**	0.54	0.36	9.68%	0.79	0.56	14.44%	0.79	0.56	15.94%	0.54	0.37	10.45%	0.79	0.56	15.51%	0.78	0.54	15.60%
Apr.	0.63	0.47	5.03%	0.62	0.47	**5.02%**	1.10	0.78	8.24%	1.30	0.87	9.09%	0.62	0.47	5.66%	1.17	0.80	8.50%	1.14	0.78	8.35%
Jul.	0.95	0.70	15.40%	0.83	0.61	**14.46%**	1.44	1.04	22.04%	1.52	1.05	20.49%	0.87	0.65	15.20%	1.45	0.96	19.96%	1.49	1.00	20.53%
Oct.	0.56	0.41	**5.18%**	0.57	0.42	5.22%	1.21	0.82	11.40%	1.19	0.81	10.91%	0.59	0.43	5.87%	1.00	0.70	9.42%	1.02	0.72	9.43%
Site 2																					
Jan.	0.57	0.46	20.35%	0.55	0.37	**17.20%**	0.71	0.52	24.65%	0.77	0.55	24.52%	0.55	0.38	17.99%	0.78	0.57	26.22%	0.77	0.57	25.36%
Apr.	1.50	1.11	16.00%	1.38	1.01	**15.82%**	1.68	1.20	19.43%	1.68	1.16	18.24%	1.60	1.11	16.79%	1.68	1.19	18.61%	1.70	1.22	19.00%
Jul.	0.73	0.55	5.96%	0.60	0.41	**3.86%**	0.88	0.62	5.95%	1.05	0.70	6.66%	0.61	0.43	4.49%	0.93	0.65	6.18%	0.93	0.65	6.21%
Oct.	0.74	0.56	5.42%	0.59	0.42	**3.63%**	2.14	0.86	7.05%	0.93	0.66	5.51%	0.58	0.43	4.20%	0.96	0.66	5.66%	0.94	0.67	5.70%
Site 3																					
Jan.	0.45	0.31	9.07%	0.35	0.26	**7.48%**	0.58	0.43	12.58%	0.57	0.41	12.38%	0.36	0.27	8.41%	0.59	0.43	12.44%	0.56	0.41	11.69%
Apr.	1.38	0.90	10.56%	0.99	0.75	**9.81%**	1.73	1.22	15.03%	1.68	1.14	14.21%	1.01	0.77	10.81%	1.84	1.29	15.57%	1.88	1.32	15.90%
Jul.	0.66	0.44	6.56%	0.54	0.38	**5.57%**	0.75	0.56	7.78%	0.71	0.51	7.16%	0.55	0.40	6.08%	0.78	0.54	7.80%	0.78	0.56	8.02%
Oct.	0.74	0.51	7.02%	0.68	0.48	**6.71%**	1.04	0.74	10.14%	0.89	0.67	9.46%	0.75	0.52	7.72%	0.92	0.69	9.58%	0.92	0.68	9.45%
Mean	0.78	0.56	9.67%	0.69	0.50	**8.71%**	1.17	0.78	13.23%	1.09	0.76	12.88%	0.72	0.52	9.47%	1.08	0.75	12.95%	1.08	0.76	12.94%

TABLE 2: Performance comparison according to data collected from different seasons and from different stations.

Category	HEN1			HEN2			NAR			53H-NAR			EEMD-NAR			Persist			ARIMA		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MEPE
Jan.	0.51	0.37	12.97%	0.48	0.33	11.45%	0.69	0.50	17.22%	0.71	0.51	17.61%	0.48	0.34	12.28%	0.72	0.52	18.06%	0.70	0.51	17.55%
Apr.	1.17	0.82	10.53%	0.99	0.74	10.22%	1.50	1.07	14.23%	1.55	1.06	13.85%	1.08	0.78	11.09%	1.56	1.09	14.23%	1.57	1.11	14.42%
Jul.	0.78	0.56	9.30%	0.65	0.47	7.96%	1.02	0.74	11.92%	1.09	0.75	11.44%	0.68	0.49	8.59%	1.05	0.72	11.31%	1.07	0.74	11.59%
Oct.	0.68	0.49	5.88%	0.61	0.44	5.19%	1.46	0.81	9.53%	1.00	0.71	8.63%	0.64	0.46	5.93%	0.97	0.68	8.22%	0.96	0.69	8.19%
Station 1	0.67	0.49	8.78%	0.64	0.47	8.60%	1.14	0.80	14.03%	1.20	0.82	14.11%	0.66	0.48	9.30%	1.10	0.76	13.35%	1.11	0.76	13.48%
Station 2	0.88	0.67	11.93%	0.78	0.55	10.13%	1.35	0.80	14.27%	1.11	0.77	13.73%	0.84	0.59	10.87%	1.09	0.77	14.17%	1.09	0.78	14.07%
Station 3	0.81	0.54	8.30%	0.64	0.47	7.39%	1.03	0.74	11.38%	0.96	0.68	10.80%	0.67	0.49	8.26%	1.04	0.74	11.35%	1.03	0.74	11.27%

Combining the data description (as Figure 3 shows) into analysis, if the input datasets perform with stronger variation and more complex changing trends, the forecasting results will always be accompanied by higher statistical errors.

4. Conclusions

Short-term wind speed forecasting, an essential support to the regulatory actions and short-term load dispatching planning during the operation of wind farms, is currently regarded as one of the most difficult problems to be solved. This paper contributes to the goal of short-term wind speed forecasting by developing two three-stage hybrid approaches (named HEN1 and HEN2, resp.). Both are three-stage models, with an abnormal value detection and modification process based on the 53H method, a signal decomposition and noise filtering sector dealt with by the EEMD algorithm, and a training and forecasting stage handled by NAR-based neural networks. The chosen datasets were ten-minute wind speed observations from three stations in western China, including twelve samples. Both the simulation and the comparison indicate a strong capacity for the proposed methods to address short-term wind speed forecasting problems.

(i) The HEN2 model performs best in most cases. The average MAPE of the HEN2 model is 8.71%, which is the lowest MAPE out of all the considered approaches, based on the twelve samples simulated. This approach results in reductions of 38.25%, 36.57%, and 13.73% according to MAPE compared with the NAR, 53H-NAR, and EEMD-NAR, respectively.

(ii) The HEN2 performs better than the HEN1 in most cases. This may results from both the different model structures of them and the particularity of data in different stations. When the number of IMFs obtained from EEMD is relatively large, the HEN2 performs better. The reason may be that each IMF can introduce extra uncertainty into the final forecast, during the HEN1 process.

(iii) According to comparisons of different seasons and stations, each model was shown to perform similarly across different stations, although a large gap of forecasting errors was found when considering the results according to different seasons. This indicates that model performance is heavily reliant on the inner information of an input data series, even though the proposed hybrid approaches maintain the lowest forecasting error in all twelve samples.

In summary, the developed methods, especially the HEN2 method, can provide a significant enhancement of the model's performance in short-term wind speed forecasting. This is of great significance in the actual regulation of active decision making and short-term load dispatching planning during the operation of wind farms.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] I. Fyrippis, P. J. Axaopoulos, and G. Panayiotou, "Wind energy potential assessment in Naxos Island, Greece," *Applied Energy*, vol. 87, no. 2, pp. 577–586, 2010.

[2] A. Fabbri, T. G. S. Román, J. R. Abbad, and V. H. M. Quezada, "Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1440–1446, 2005.

[3] I. J. Ramírez-Rosado and L. A. Fernández-Jiménez, "An advanced model for short-term forecasting of mean wind speed and wind electric power," *Control and Intelligent Systems*, vol. 32, no. 1, pp. 21–26, 2004.

[4] A. Sfetsos, "A novel approach for the forecasting of mean hourly wind speed time series," *Renewable Energy*, vol. 27, no. 2, pp. 163–174, 2002.

[5] A. Sfetsos, "A comparison of various forecasting techniques applied to mean hourly wind speed time series," *Renewable Energy*, vol. 21, no. 1, pp. 23–35, 2000.

[6] M. Lange and U. Focken, *Physical Approach to Short-Term Wind Power Prediction*, Springer, New York, NY, USA, 2009.

[7] L. Landberg, "Short-term prediction of the power production from wind farms," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 80, no. 1-2, pp. 207–220, 1999.

[8] M. Negnevitsky and C. W. Potter, "Innovative short-term wind generation prediction techniques," in *Proceedings of the IEEE PES Power Systems Conference and Exposition (PSCE '06)*, pp. 60–65, Atlanta, Ga, USA, October-November 2006.

[9] S. J. Wu and S. L. Lin, "Intelligent web-based fuzzy and grey models for hourly wind speed forecast," *International Journal of Computers*, vol. 4, pp. 235–242, 2010.

[10] J. F. Li, B. H. Zhang, G. L. Xie, Y. Li, and C. X. Mao, "Grey predictor models for wind speed-wind power prediction," *Power System Protection and Control*, vol. 38, no. 19, pp. 151–159, 2010.

[11] L. Lin, J. T. Eriksson, H. Vihriala, and L. Soderlund, "Predicting wind behavior with neural networks," in *Proceedings the European Wind Energy Conference*, pp. 655–658, 1996.

[12] M. C. Alexiadis, P. S. Dokopoulos, H. S. Sahsamanoglou, and I. M. Manousaridis, "Short-term forecasting of wind speed and related electrical power," *Solar Energy*, vol. 63, no. 1, pp. 61–68, 1998.

[13] H. G. Beyer, T. Degner, J. Haussmann, M. Hoffman, and P. Rujan, "Short term forecast of wind speed and power output of a wind turbine with neural networks," in *Proceedings of the 2nd European Congress on Intelligent Techniques and Soft Computing*, pp. 349–352, 1994.

[14] G. Kariniotakis, G. S. Stavrakakis, and E. F. Nogaret, "Wind power forecasting using advanced neural network models," *IEEE Transactions on Energy Conversion*, vol. 11, no. 4, pp. 762–767, 1996.

[15] A. More and M. C. Deo, "Forecasting wind with neural networks," *Marine Structures*, vol. 16, no. 1, pp. 35–49, 2003.

[16] G. Kariniotakis, G. S. Stavrakakis, and E. F. Nogaret, "A fuzzy logic and neural network based wind power model," in *Proceedings of the European Wind Energy Conference*, pp. 596–599, 1996.

[17] X. Wang, G. Sideratos, N. Hatziargyriou, and L. H. Tsoukalas, "Wind speed forecasting for power system operational planning," in *Proceedings of the International Conference on Probabilistic Methods Applied to Power Systems*, pp. 470–474, Ames, Iowa, USA, September 2004.

[18] S. Makridakis, "Why combining works?" *International Journal of Forecasting*, vol. 5, no. 4, pp. 601–603, 1989.

[19] F. C. Palm and A. Zellner, "To combine or not to combine? Issues of combining forecasts," *International Journal of Forecasting*, vol. 11, no. 8, pp. 687–701, 1992.

[20] R. L. Winkler, "Combining forecasts: a philosophical basis and some current issues," *International Journal of Forecasting*, vol. 5, no. 4, pp. 605–609, 1989.

[21] P. Newbold and C. W. J. Granger, "Experience with forecasting univariate time series and the combination of forecasts," *Journal of the Royal Statistical Society*, vol. 137, no. 2, pp. 131–165, 1974.

[22] N. An, W. Zhao, J. Wang, D. Shang, and E. Zhao, "Using multi-output feedforward neural network with empirical mode decomposition based signal filtering for electricity demand forecasting," *Energy*, vol. 49, pp. 279–288, 2013.

[23] Y. Huang, "Advances in artificial neural networks—methodological development and application," *Algorithms*, vol. 2, no. 3, pp. 973–1007, 2009.

[24] J. W. Tukey, *Exploratory Data Analysis*, Addison-Wesley, Reading, Mass, USA, 1977.

[25] P. A. Tukey and J. W. Tukey, "Graphic display of data sets in 3 or more dimensions," in *Interpreting Multivariate Data*, Wadsworth, Belmont, Calif, USA, 1981.

[26] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis," *Proceedings of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[27] Y.-C. Wei, C.-J. Lee, W.-Y. Hung, and H.-T. Chen, "Application of Hilbert-Huang transform to characterize soil liquefaction and quay wall seismic responses modeled in centrifuge shaking-table tests," *Soil Dynamics and Earthquake Engineering*, vol. 30, no. 7, pp. 614–629, 2010.

[28] S. Kizhner, K. Blank, T. Flatley, N. E. Huang, D. Petrick, and P. Hestnes, "On certain theoretical developments underlying the Hilbert-Huang transform," in *Proceedings of the IEEE Aerospace Conference*, Big Sky, Mont, USA, 2006.

[29] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.

[30] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.

[31] J.-R. Yeh, J.-S. Shieh, and N. E. Huang, "Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method," *Advances in Adaptive Data Analysis*, vol. 2, no. 2, pp. 135–156, 2010.

[32] P. H. Tsui, C. C. Chang, C. C. Chang, N. E. Huang, and M. C. Ho, "An adaptive threshold filter for ultrasound signal rejection," *Ultrasonics*, vol. 49, no. 4-5, pp. 413–418, 2009.

[33] P. H. Tsui, C. C. Chang, and N. E. Huang, "Noise-modulated empirical mode decomposition," *Advances in Adaptive Data Analysis*, vol. 2, no. 1, pp. 25–37, 2010.

[34] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, New York, NY, USA, 2nd edition, 1998.

[35] L. Ljung, *System Identification: Theory for the User*, Prentice Hall PTR, 2nd edition, 1998.

[36] K. Levenberg, "A method for the solution of certain problems in least squares," *Quarterly of Applied Mathematics*, vol. 2, pp. 164–168, 1944.