Research Article

Simplified-Boost Reinforced Model-Based Complex Wind Signal Forecasting

Qiushuang Lin and Chunxiang Li

Department of Civil Engineering, Shanghai University, Shanghai 200444, China

Correspondence should be addressed to Chunxiang Li; li-chunxiang@vip.sina.com

Received 9 September 2019; Revised 11 September 2020; Accepted 16 September 2020; Published 30 September 2020

Academic Editor: Chao Wu

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

Wind signal forecasting has become more and more crucial in the structural health monitoring system and wind engineering recently. It is a challenging subject owing to the complicated volatility of wind signals. The robustness and generalization of a predictor are significant as well as of high precision. In this paper, an adaptive residual convolutional neural network (CNN) is developed, aiming at achieving not only high precision but also high adaptivity for various wind signals with varying complexity. Afterwards, reinforced forecasting is adopted to enhance the robustness of the preliminary forecasting. The preliminary forecast results by adaptive residual CNN are integrated with historical observed signals as the new input to reconstruct a new forecasting mapping. Meanwhile, simplified-boost strategy is applied for more generalized results. The results of multistep forecasting for five kinds of nonstationary non-Gaussian wind signals prove the more excellent adaptivity and robustness of the developed two-stage model compared with single models.

1. Introduction

During the past decades, high-rise buildings, long-span roofs, and bridges were mushrooming continuously. As the main design load for high and flexibility buildings, wind load has attracted lots of attention from experts [1]. Wind load research is a hot topic currently, containing wind field measurement [2], theoretical research [3, 4], numerical simulation [5], and forecasting [6]. Concerning the field application and operation in wind engineering and structural health monitoring system, data loss is inevitable due to faulty sensors and the instability of power supply and data transmission in wireless sensing system [7]. Therefore, wind forecasting techniques have been developed for reconstructing wind signals in the event of data loss. Meanwhile, wind disasters forecasting including downburst and typhoon has become an essential topic currently [8]. Wind speed forecasting is also an important part in the application of wind energy resources [9]. Accurate wind forecasting is not only helpful in improving wind energy utilization but also helpful for distributed energy resources [10, 11].

Abundant approaches are developed for time series forecasting, which are divided into four categories: physical methods [12, 13], statistical methods [14, 15], artificial intelligent methods [16–28], and hybrid methods [29–39], which are summarized in Table 1.

Meanwhile, time and space often exist together in engineering applications. Hence, spatiotemporal pattern-based forecasting/detection methods have been a hot topic recently [40–42]. Besides, reinforcement learning is a popular technique in machine learning. In recent years, reinforcement learning is attracting extensive attention from experts and developing rapidly [43, 44].

Deep learning is another branch of machine learning broadly applied in various fields [45, 46]. As a popular deep learning network, convolutional neural network (CNN) is widely used in recognition and detection [26]. Input preprocessing is effective for CNN. Liu et al. [27] transformed the input of multivariate time series data into appropriate tensor representation for time series classification. Harbola and Coors [28] developed 1D multiple CNN combining several 1D single CNNs with different views of the same
input. In this paper, single CNN with multichannel input providing different views of the time series input is developed. The multichannel convert strategy can provide temporal patterns at additional timescales, improving the data utilization. In addition, compared to multiple CNNs with input of different time scales, the computation effort of single CNN with multichannel input is smaller.

Due to the complex and changeable nature of nonstationary wind, common forecasting methods are generally developed for a certain kind of wind signal or wind field without universality [19, 24]. Considering oversimplified models could hardly extract the inherent intermittency and turbulence effect of nonstationary signals, while too complex models may lead to overfitting, it is challenging to propose a generalized method. In this study, an adaptive residual CNN is developed as a base predictor inspired by the residual neural network [47]. The adaptive residual technique allows adaptively simplifying the model according to the various fluctuation complexity of wind signals.

Presently, many experts have focused on error correction methods for improving the forecasting accuracy. Most error correction methods built an additional model to forecast error components [35–38]. However, if the judgment on the characteristics of error components is inaccurate or the selection of error correction model was inappropriate, the forecasting deviation would be larger and bring interference to the preliminary forecast results. In this case, the accuracy might be reduced. In this paper, a reinforced strategy is adopted to optimize the preliminary forecasting. Specifically, the historical observed signals are integrated with the preliminary results as new input to reconstruct a new forecasting mapping, namely, reinforced forecasting. Different from the common error correction approach [35–38], the reinforced approach can enhance the forecasting ability by reconstructing mapping relationship upon actual historical components. The forecasting accuracy may not be improved in the least ideal case, but at least reinforced forecasting will not cause interference.

Wang et al. [39] introduced the measured wind speed at time $t$ to correct the forecast wind speed at $t + 1$. However, not only the observed component at last time step but also observed components at last few time steps may be useful. It is hard to judge how many historical observed components are effective for optimizing the preliminary forecasting. To solve the problem, simplified-boost technique is proposed to make full use of various historical observed components and improve the universality of reinforced strategy.

On the whole, an innovative hybrid model based on multichannel convert, adaptive residual CNN, reinforced forecasting strategy, and simplified-boost technique is proposed for multistep wind signal forecasting, aiming at promoting the robust and generalization of a predictor as well as precision.

The innovative contributions are expressed as follows: (1) Multichannel convert is first developed for preprocessing the input signals, leading to superior feature representation. (2) The developed adaptive residual CNN can be adjusted to different signals with varying complexity, greatly improving the adaptability of model. (3) For the pursuit of high precision and robustness, reinforced forecasting is developed to

| Time series forecasting model | Advantage | Disadvantage | Reference papers |
|------------------------------|-----------|--------------|-----------------|
| Physical methods | Numerical weather prediction | Superior performance in long-term forecasting | High computing-consuming | [12, 13] |
| Statistical methods | Time series model | Obtain good performance in linear data; require less time to build | Difficult to handle the nonlinear nonstationary data owing to the linear nature | [14, 15] |
| Artificial intelligent methods | Artificial neural network | Handle nonlinear data | Fall into local optima; be influenced by the initial parameters | [16, 17] |
| | Support vector machine (SVM) | Strong generalization; suitable for small and medium datasets | The performance may be influenced by the kernel function and parameters | [18] |
| | Extreme learning machine | Fast learning speed and good generalization performance | | |
| | Decision tree (DT) | Simple; less data preprocessing | Instability | [19] |
| | Random forest (RF) | Stability; strong generalization | Instability; be sensitive to the dataset | [20] |
| | Fuzzy logic models | Robustness; strong fault tolerance | Overfitting due to data noise | [21] |
| | Recurrent neural network (RNN) | Suitable for time series | Low accuracy | [22] |
| | CNN | Stronger ability and flexibility | Computational complexity; time-consuming; overfitting | [23–25] |
| Hybrid methods | Decomposition-based methods | High accuracy | Overfitting | [26–28] |
| | Parameter optimization-based methods | High accuracy; stability | Time-consuming; computational complexity | [30] |
| | Weight-based forecasting methods | Robustness | Multi-collinearity | [31–34] |
| | Error correction-based methods | High accuracy | Be influenced by the selection of error correction model | [35–39] |
enhance the forecasting ability; meanwhile, simplified-boost technique is adopted for achieving more generalized results. (4) Multistep forecasting results of five kinds of complex wind signals show the advantages of the developed model. Both long-term trend and short-term precision can be guaranteed by the proposed two-stage frame, confirming that the hybrid model is the most responsive for wind signal forecasting in comparison with single models. (5) Scientific and comprehensive evaluation is conducted to testify the effectiveness of the developed model.

The remainder of this article is organized as follows. Section 2 presents the proposed novel framework and briefly introduces the related techniques. In Section 3, multistep forecasting for nonstationary non-Gaussian wind signals is applied to verify the effectiveness of the proposed model. In Section 4, insightful exploration and analysis are conducted. Section 5 makes a conclusion for this study.

2. Methodology

Given a set of observed historical wind signals with time span $X = \{x_1, x_2, \ldots, x_T\}$, the forecasting task is to obtain $x_{T+h}$ through automatically learning the potential critical characteristics of $X$ by neural network with multi-input multoutput (MIMO) strategy. Analogously, $x_{T+h+1}$ can be obtained based on $\{x_2, x_3, \ldots, x_{T+1}\}$. $h$ is the forecasting horizon. $T$ is the embedding dimension. The embedding dimension directly determines the dimension of input in training and testing sets. Namely, the previous $T$ data points are used to forecast next data point. The determination details of $T$ are shown in Section 2.2.

2.1. Frame of the Proposed Model. The whole framework of the proposed model is shown in Figure 1. The procedure is presented as follows.

Stage 1 (preprocessing): Divide original data into the training and testing parts before normalization. Then phase space reconstruction and multichannel convert for the network input are implemented.

Stage 2: Adaptive residual CNN is adopted for preliminary forecasting by building a mapping relationship between historical observed signals and signals to be forecasted.

Stage 3: Reinforced forecasting is implemented by reconstructing the mapping relationship. Specifically, the historical observed signals are integrated with fresh information (i.e., preliminary forecasting results) as new input.

Stage 4: Notice that the embedding dimension is $T$; namely, $T$ historical observed components exist for each signal to be forecasted. Thus, reinforced forecasting processes should be implemented for $T + 1$ times in total. To balance the effect of diverse historical observed signals, simplified-boost algorithm is adopted to obtain the final forecast values.

Stage 5: Evaluate the forecasting errors via the forecasting datasets.

2.2. Phase Space Reconstruction and Multichannel Convert.

Time series are usually transformed into phase space to reflect the system’s dynamics information preferably. In this paper, the embedding dimension is determined by minimizing the root mean square error (RMSE) using the forecast results of training set. Take dataset A in Section 3.2 as an example; the RMSE with embedding dimension $T$ is presented in Figure 2. The higher forecasting accuracy within training set is obtained when $T$ is between 8 and 12. The optimal embedding dimension should be between 8 and 10. Therefore, embedding dimension is selected as 10 in the following numerical examples.

Subsequently, multichannel convert is developed for the purpose of providing additional information by presenting temporal patterns at different time scales.

The original input is a one-dimensional vector with size of $T \times 1$. In the preprocessing stage, the first channel of input is the same as the original input. The second channel takes $T$ values but at an increment of two, starting from the second value. Similarly, the third channel takes $T$ values but at an increment of two, starting from the first value. Thusly, the second and third channels have half of the $T$ values, respectively. As presented in Figure 1, gray rectangles indicate excluded values. To get the same number of values as the first channel, the following and previous value are repeated as the current value in gray rectangles in the second and third channel, respectively. Thus, the size of input processed by multichannel convert is $T \times 1 \times 3$. The input can be regarded as a 3-channel image.

The convert strategy provides additional views of the input in terms of temporal resolution, so the network can automatically capture temporal patterns at different time scales. Furthermore, different time series may require feature representations at different time scales. The multichannel input can adaptively adjust the feature representation in suitable time scale.

2.3. Adaptive Residual CNN. Compared with fully connected neural networks, the relative progress of CNN is the introduction of convolution layer and pooling layer. The convolution layer extracts features by filter translation on the original time series. Pooling layer including maximum and average pooling reduces the learning parameters and network complexity by parameters sparse representation after features are aggregated. Sparse connection, parameter sharing and equivariant representation are three important characteristics of convolution operation.

CNN mainly performs as an implicit feature extraction with the convolutional filters detecting complex patterns as feature detectors. The deeper hidden layers have the ability to dig for more complex and abstract information. The number of hidden layers in CNN depends on the complexity of the target problem. On the one hand, more convolutional layers can dig out more potential features for wind signals with large variation and complicated fluctuation; on the other hand, excessive convolutional layers may fail in the forecasting target for slightly fluctuated signal due to overfitting. Therefore, an adaptive residual
CNN is proposed for adaptive wind signal forecasting inspired by the residual neural network [47]. The proposed adaptive residual technique aims at not only achieving high-accuracy, but also meeting the requirement to have high adaptivity for various wind signals with different complexity levels.

As the original input of CNN, a sequence of wind signals are transmitted forward layer by layer. Data features are
extracted and enhanced by each convolution layer progressively. In the following numerical examples in this study, three convolutional layers are adopted in the adaptive residual CNN for guaranteeing the expressive ability of output features. Meanwhile, in order to map weak nonlinear input-output relation, it is customary to have an additional connection from the first convolutional layer to the output layer, with an additional dense layer as shown in Figure 1. Assuming that the original stacked convolutional layers fit a residual mapping, it is easier to push the residual to zero than to fit the weak nonlinear mapping by the original stacked layers, if one convolutional layer is optimal enough. The simpler path is more perfect. The finding submits to the parsimony rule that the simplest solution tends to be the right one [47]. The adaptive residual technique allows adaptively simplifying the model and being adjusted to different wind signals with various fluctuation characteristics, which greatly improves the generalization ability of model.

From another point of view, traditional CNNs only utilize the last convolutional layer for regression task without considering the information included in previous convolutional layers. Actually, different convolutional layers contain feature information in different scales. The proposed adaptive residual technique allows exploiting the information among different convolutional layers and making full use of hierarchical features in previous convolutional layers. Note that the additional connection is not merely limited to begin from the first convolutional layer.

2.4. Reinforced Forecasting. In this section, a reinforced forecasting algorithm is utilized to optimize the adaptive residual CNN. The preliminary forecast values considered as additional fresh information are integrated with historical observed components as input variables so as to reconstruct a completely new forecasting mapping. The mapping formulation can be expressed as follows:

\[
x_{T+h}^{\text{final}, k+1} = f(x_{T+h}^{\text{pre}}, \ldots, x_{T+1}^{\text{pre}}, x_{T}^{\text{observe}}, \ldots, x_{T-(k-1)}^{\text{observe}}),
\]

\[0 \leq k \leq T.
\]

\[
x_{T+h}^{\text{pre}}, \ldots, x_{T+1}^{\text{pre}}
\]

are preliminary forecast values based on adaptive residual CNN, where MIMO strategy is adopted. \(x_{T}^{\text{observe}}, \ldots, x_{T-(k-1)}^{\text{observe}}\) are historical observed signals. \(k\) denotes the number of historical observed signals inserted as input and \(k = 0\) means no historical observed signals are inserted as input. \(x_{T+h}^{\text{final}, k+1}\) indicates reinforced forecasting results by different \(k\).

The reinforced strategy makes full use of intrinsic connections between the current preliminary forecast values, the historical information, and the current actual values to be forecasted, using the superior nonlinear fitting capability of the reinforced model. Reinforced forecasting is a key segment to promote not only the accuracy but also the robustness. The predictor used for reinforced forecasting can be different from preliminary forecasting. Many predictors can be options, e.g., RNN, long short-term memory neural network (LSTM), and SVM. In the following numerical examples, adaptive residual CNN is still selected as the reinforced forecasting model, which can receive the best results by trial and error.

The two-stage model includes preliminary forecasting and reinforced forecasting, which work as two temporal dependency captures and could be regarded as encoder and decoder, respectively. It is worth mentioning that the two-stage framework has complementary properties on forecasting ability so as to enhance the robustness of the algorithm.

2.5. Simplified-Boost Reinforced Forecasting. However, it is difficult to decide exactly how many historical observed components (i.e., the value of \(k\)) are suitable to the reinforced forecasting for a specific forecasting task. In fact, if the value of \(k\) is too small—i.e., the input features are insufficient—the mapping relationship would be stochastic excessively. Quite the reverse, redundant information would be introduced.

\(k\) is a key point for reinforced forecasting. Different \(k\) may bring different effects for the same task. Take 1-step forecasting as an example; the detailed description of data is presented in Section 3.2. As shown in Figure 3, it is clear that the performance is the best when \(k = 4\) for dataset A, where “pr” denotes the preliminary forecast results. However, the optimal \(k\) is different in terms of different forecasting tasks. For instance, the best results are obtained when \(k = 4\) for dataset A, while \(k = 10\) for dataset B.

In order to solve the problem, a suitable solution is to combine advantages of models with different \(k\). Thus, a simplified-boost algorithm is applied for balancing the effects of \(k\). Boosting algorithm builds a committee of predictors that may be superior to an arbitrary single predictor [48]. The predictor weights and sample weights are updated according to the historical performance in each iteration. Simpler than conventional boosting algorithm, \(T + 1\) base predictors are constructed during training process of the simplified-boost reinforced forecasting. The \(T + 1\) base predictors are built based on the input variables \(x_{T+h}^{\text{pre}}, x_{T+1}^{\text{pre}}, \ldots, x_{T-(k-1)}^{\text{observe}}\) \(k = 0, \ldots, T\), respectively. Then the predictors are weighted according to their forecasting performance. Predictors with more confidence about their predictions are weighted more heavily, which guarantees well-behaved predictors to be used. Finally, predictors are combined by weighted sum. Supposing that the sample number in training set is \(n\), the procedure of the simplified-boost reinforced forecasting algorithm is shown in Algorithm 1.

The simplified-boost algorithm can balance the effect of input with diverse historical observed signals. More importantly, the generalization ability has been greatly improved so as to forecast various kinds of wind signals adaptively.

3. Experiments

In this section, multistep forecasting for five kinds of nonstationary non-Gaussian wind signals is implemented to
testify the high accuracy and robustness of the proposed model. Six models are adopted for comparison, including DT, RF, back propagation neural network (BPNN), RNN [23], gated recurrent unit (GRU) [24], and LSTM [25]. All of the simulations are operated in python 3.5 NVIDIA GeForce GTX 1060 3GB, 3.40 GHz CPU E3-1230 v5, 64 bits, 8 GB RAM, Windows 7.

### 3.1. Evaluation Criteria

Six evaluation indicators, consisting of RMSE, mean absolute error (MAE), mean absolute percent error (MAPE), Pearson correlation coefficient (R), symmetric mean absolute percentage error (SMAPE), and mean absolute scaled error (MASE), are chosen to evaluate the forecasting performance. The value of R needs to be close to 1, which means the correlation difference or the delay time is small. MASE is less than one if it arises from a better forecast than the average one-step naïve forecast computed in-sample. Conversely, it is greater than one if the forecast is worse than the average one-step naïve forecast computed in-sample [49]:

\[
\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x}_i)^2},
\]

\[
\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |x_i - \bar{x}_i|,
\]

\[
\text{MAPE} = \frac{1}{N} \sum_{i=1}^{N} \frac{|x_i - \bar{x}_i|}{x_i} \times 100\%,
\]

\[
R = \frac{\text{cov}(x_i, \bar{x}_i)}{\sqrt{\text{cov}(x_i, x_i) \times \text{cov}(\bar{x}_i, \bar{x}_i)}} \tag{2}
\]

\[
\text{SMAPE} = \frac{1}{N} \sum_{i=1}^{N} \frac{|x_i - \bar{x}_i|}{(x_i + \bar{x}_i)/2} \times 100\%,
\]

\[
\text{MASE} = \frac{1}{N} \sum_{i=1}^{N} \frac{x_i - \bar{x}_i}{(1/(T-1))\sum_{j=1}^{T-1} |x_{i-j} - x_{i-(j+1)}|},
\]

---

**Figure 2:** Effect of embedding dimension on RMSE.

**Figure 3:** RMSE with different number of historical observed signals.
where \(x_i\) is the observed value, \(\bar{x}_i\) is the forecast value, \(N\) is the number of forecast samples, and \(\text{cov}\) refers to calculating the covariance matrix.

### 3.2. Dataset Description

Without loss of generality, five kinds of datasets are adopted for experiments. Namely, wind speed is on a super high-rise building roof near the coast of the East China Sea in Xiamen at 06:37-07:37 on 8th August 2015 before typhoon Soudelor landing. The sampling frequency is 2 Hz, denoted by A. Wind pressure is on the surface of a 28-storey building, located in the west coast of Qingdao, Shandong province. The sampling frequency is 1 Hz, denoted by B. Wind pressure is that of Leqings sports center, a long-span membrane structure whose maximum cantilever span is 57 m [50]. The sampling frequency is 5 Hz, denoted by C. Downburst wind speed was recorded during the 2002 thunderstorm outflow experiment, which was conducted by the department of atmospheric science of Texas Tech University from 20th May 2002 to 15th July 2002 [51]. The sampling frequency is 1 Hz, denoted by D. The downburst often occurs with thunderstorm weather, which may lead to extreme wind loads. Wind speed is from 00:00 on 24th August to 06:00 on 25th August in 2015 before and after Super Typhoon Goni passing through 10 m height of the measured base located at 31°11′46.36″ N and 121°47′8.29″ E. The sampling frequency is 1 Hz, denoted by E.

In total, 500 samples are selected for each dataset as shown in Figure 4. The 1st-400th samples of original datasets are considered as training set, which are used to construct the forecasting models. The 401th-500th samples are utilized as testing set to verify the effectiveness of models marked in red. Statistics indicators are presented in Table 2. Kurtosis and skewness are adopted to quantify the non-Gaussian characteristics. The Runest method is utilized to measure the nonstationary characteristics. It can be inferred that dataset C is provided as nonstationary, non-Gaussian series. Datasets A, B, D, and E are analyzed as nonstationary series.

### 3.3. Comparison Results

The forecast results and errors criteria of datasets A and B for multistep forecasting are presented in Figures 5 and 6 and Tables 3 and 4, respectively. The forecast errors criteria of datasets C, D, and E are shown in Tables 5–7. The detailed discussion according to the charts is expressed in the following.

Based on the forecast results in Figures 5 and 6, the developed two-stage forecasting frame obtains more satisfactory forecast values compared with single models. It is demonstrated that the developed model can track the volatility trend and range more sensitively. The forecast values of the proposed model are much closer to the observed values, which can also be reflected clearly in the curves of forecast errors. The forecast errors \(x_i - \bar{x}_i\) of all the compared models are shown in Figures 5 and 6. It is clear that the forecast error magnitudes of the proposed model are the smallest among the compared models, especially for 3-step and 5-step forecasting. All in all, the forecasting performance of the hybrid model exhibits a considerable improvement over the single models.

The probability distributions with respect to forecast errors in Figures 5 and 6 describe the different forecasting abilities. The fluctuations range around zero is smaller and more concentrated regarding the proposed model, which deduces the advantageous performance compared to other models.

According to the fitting curves of the observed and forecast values, the forecast values of the proposed model fit the observed values preferably. They are obviously closer to the observed values than the compared models, which testifies the effectiveness of adaptive residual strategy and reinforced forecasting strategy once again.

It is clear in accordance with the forecasting evaluation indicators in Tables 3–7 that RMSE, MAE, MAPE, SMAPE, and MASE of multistep forecasting by the proposed model are mostly the minimum. Meanwhile, the correlation coefficient \(R\) is mostly the maximum, which indicates the strongest positive correlation between the forecast values and observed values. Regarding MASE, most of the forecast results are less than 1 except for the 1-step forecasting of datasets C, D, and E. Anyway, the evaluation criterion MASE of the proposed model is always the smallest among the compared models.

As clearly inferred from Tables 3–7, the proposed two-stage model achieves significant promotion in the forecasting accuracy and stability compared with other models. For instance, 1-step forecasting for dataset D by the proposed model achieves 15.31%, 3.89%, 10.21%, 49.47%, 18.64%, and 21.67% reductions in the MAPE, respectively, in
comparison with DT, RF, BPNN, RNN, GRU, and LSTM. 3-step forecasting for dataset $E$ by the proposed model achieves 41.96%, 35.73%, 31.56%, 30.56%, 50.90%, and 48.81% reductions in the MAPE, respectively, in comparison with DT, RF, BPNN, RNN, GRU, and LSTM. 5-step forecasting for dataset C by the proposed model achieves 28.50%, 27.92%, 20.19%, 19.50%, 43.56%, and 26.97% reductions in the RMSE, respectively, in comparison with DT, RF, BPNN, RNN, GRU, and LSTM.

Remark 1. The proposed model outperforms the other compared models by a large margin. It can be speculated that no common single forecasting model is always suitable for any case. RNNs (including RNN, LSTM, and GRU) do not act well compared to the other models. The main reason for the fails of the advanced models may be that they are equipped with many adaptable parameters, causing complex training. Especially in the case of small sample learning, the training datasets are insufficient for training the complex construction, which would lead to overfitting easily.

We can conclude that the proposed two-stage frame method is efficient and achieves a superior performance among the compared models.

4. Results Analysis

In this section, comprehensive analysis from a variety of aspects is conducted for further verification of the performance of the developed model.

4.1. DM Test. As one type of hypothesis testing approaches, the DM test [52, 53] is applied to demonstrate the forecasting validity of the developed hybrid model based on statistical thinking.

Under a given significance level $\alpha$, the null hypothesis $H_0$ represents that there is no significant difference between the proposed model and the compared model in terms of their forecasting performance. In contrast, $H_1$ means that there exists significant difference between the two models. The related formulas can be written as

$$\begin{align*}
H_0: & \quad E[I(e^i_1)] = E[I(e^i_2)], \\
H_1: & \quad E[I(e^i_1)] \neq E[I(e^i_2)],
\end{align*}$$

wherein $I$ is the loss function of the forecasting error. $e^i_1$, $e^i_2$ indicate the forecasting errors of the compared model and the proposed model, respectively.

The DM test statistics can be described as

$$DM = \frac{\sum_{i=1}^{N} (I(e^i_1) - I(e^i_2))/N}{\sqrt{s^2/N}}.$$  

**Table 2: Statistics indicators of wind signal datasets A-E.**

| Dataset | Mean (m/s) | Std (m/s) | Kurtosis | Skewness | Nonstationary |
|---------|------------|----------|----------|----------|--------------|
| A       | 12.2154    | 5.4263   | 2.8661   | 0.1211   | Yes          |
| B       | 232.3584   | 37.3457  | 2.8617   | 0.3615   | Yes          |
| C       | 86.7584    | 11.0850  | 2.5459   | 0.8090   | Yes          |
| D       | 9.2556     | 2.2242   | 3.4244   | 0.3695   | Yes          |
| E       | 2.9962     | 1.0915   | 2.9061   | 0.4937   | Yes          |

**Figure 4: Original wind signals.**
Figure 5: Results of multistep forecasting of dataset A.
Figure 6: Results of multistep forecasting of dataset B.
### Table 3: Multistep forecasting performance of dataset A.

| Dataset A | RMSE (m/s) | MAE (m/s) | MAPE (%) |
|-----------|------------|-----------|----------|
| **Model** | **1 step** | **3 steps** | **5 steps** | **1 step** | **3 steps** | **5 steps** | **1 step** | **3 steps** | **5 steps** |
| DT        | 1.103      | 1.370     | 1.623     | 0.829      | 1.128      | 1.384      | 8.435      | 12.318     | 15.935      |
| RF        | 0.904      | 1.307     | 1.604     | 0.684      | 1.079      | 1.282      | 7.210      | 11.770     | 14.423      |
| BPNN      | 1.158      | 1.370     | 1.719     | 0.930      | 1.103      | 1.398      | 8.728      | 11.659     | 15.114      |
| RNN       | 1.081      | 1.561     | 1.695     | 0.866      | 1.187      | 1.342      | 8.479      | 12.938     | 14.621      |
| GRU       | 1.050      | 1.749     | 1.906     | 0.807      | 1.366      | 1.531      | 8.341      | 15.544     | 18.497      |
| LSTM      | 0.955      | 1.597     | 1.793     | 0.755      | 1.272      | 1.444      | 7.931      | 14.556     | 16.377      |
| Proposed  | 0.766      | 1.055     | 1.092     | 0.585      | 0.776      | 0.817      | 6.089      | 8.112      | 8.564       |

### Table 4: Multistep forecasting performance of dataset B.

| Dataset B | RMSE (Pa) | MAE (Pa) | MAPE (%) |
|-----------|-----------|----------|----------|
| **Model** | **1 step** | **3 steps** | **5 steps** | **1 step** | **3 steps** | **5 steps** | **1 step** | **3 steps** | **5 steps** |
| DT        | 16.489    | 23.055   | 29.053   | 12.237    | 17.162    | 20.852    | 5.754      | 7.991      | 9.729       |
| RF        | 13.073    | 21.800   | 29.502   | 9.072     | 15.656    | 22.119    | 4.295      | 7.368      | 10.248      |
| BPNN      | 10.278    | 20.472   | 28.494   | 7.351     | 15.297    | 21.762    | 3.455      | 7.121      | 10.068      |
| RNN       | 11.642    | 19.547   | 27.292   | 7.652     | 14.335    | 20.234    | 3.505      | 6.590      | 9.357       |
| GRU       | 12.772    | 18.888   | 25.046   | 9.460     | 13.243    | 18.920    | 4.379      | 5.932      | 8.572       |
| LSTM      | 10.783    | 20.887   | 26.942   | 7.849     | 15.276    | 19.791    | 3.666      | 7.162      | 9.020       |
| Proposed  | 8.874     | 10.216   | 12.766   | 6.143     | 7.464     | 9.292     | 2.889      | 3.475      | 4.258       |

### Table 5: Multistep forecasting performance of dataset C.

| Dataset C | RMSE (Pa) | MAE (Pa) | MAPE (%) |
|-----------|-----------|----------|----------|
| **Model** | **1 step** | **3 steps** | **5 steps** | **1 step** | **3 steps** | **5 steps** | **1 step** | **3 steps** | **5 steps** |
| DT        | 0.806     | 0.789    | 0.939    | 0.660     | 0.614      | 0.700      | 0.843      | 0.786      | 0.895       |
| RF        | 0.728     | 0.759    | 0.944    | 0.606     | 0.588      | 0.717      | 0.775      | 0.753      | 0.916       |
| BPNN      | 0.644     | 0.722    | 0.878    | 0.516     | 0.557      | 0.672      | 0.661      | 0.713      | 0.859       |
| RNN       | 1.209     | 0.872    | 0.924    | 0.998     | 0.673      | 0.738      | 1.279      | 0.860      | 0.944       |
| GRU       | 0.759     | 0.943    | 1.078    | 0.621     | 0.736      | 0.847      | 0.795      | 0.944      | 1.080       |
| LSTM      | 0.854     | 0.896    | 0.912    | 0.645     | 0.697      | 0.719      | 0.828      | 0.892      | 0.917       |
| Proposed  | 0.636     | 0.696    | 0.734    | 0.522     | 0.543      | 0.550      | 0.668      | 0.697      | 0.704       |
\[ s^2 \text{ is an estimation for the variance of } d_i = \hat{y}_i - \bar{y}_i \] and \( N \) is the number of forecast samples. If \(|\text{DM}| > Z_{\alpha/2}\), the null hypothesis \( H_0 \) will be rejected. This case means that there is a difference in the forecasting performance between the proposed model and the compared model.

Take datasets A and B as examples. Table 8 displays the DM statistical values. It can be seen that (1) for 1-step forecasting, the DM statistical values are well above the critical value of 10% significance level, \( Z_{0.10/2} = 1.64 \). The null hypothesis could be rejected. (2) For 3-step forecasting, the DM statistical values are larger than the critical value of 5% significance level, \( Z_{0.05/2} = 1.96 \), which shows the competitive forecasting accuracy of the proposed model compared with the single models. (3) For 5-step forecasting, all of the DM statistical values are larger than \( Z_{0.01/2} = 2.58 \), showing that our proposed hybrid model is superior to the compared methods for multistep forecasting. In short, the proposed two-stage model significantly outperforms the comparative forecasting models.

4.2. Stability Test. Variance of the forecast errors \( \text{Var}(\hat{y}_i - \bar{y}_i) \) is employed to assess the stability of the proposed model. It is well known that the stable ability of a forecasting model depends on the performance variance, which represents the model’s variability. Smaller variance indicates more stable forecast results. The comparison results are shown in Table 9. It is clear that the proposed model obtains the smallest variance of forecast errors, testifying that the proposed model is more stable than other models.

4.3. Effect of Adaptive Residual CNN. The comparison results between adaptive residual CNN and traditional CNN for 1-step forecasting of datasets A, B, and E are shown in Figure 7. It is obvious that the forecasting performances of adaptive...
residual CNN are superior to those of traditional CNN for all the three cases.

According to Figure 7, we can draw a conclusion that not the more complex the model is, the better. In fact, the model structure should be appropriate to the complexity of the forecast target. 3-layer CNN performs worse than 1-layer model regarding datasets A and B, but it performs better regarding dataset E. Besides, it is clear in Figure 7 that in terms of dataset E, though the traditional 3-layer CNN performs better than adaptive residual CNN sometimes, the latter shows greater stability than the former. It is further verified that adaptive residual CNN can provide more accurate and stability forecasting.

Model capacity refers to its ability to fit various functions. If the model capacity is insufficient, it would lead to underfitting. Quite the reverse, if the model capacity is too high, overfitting phenomenon would occur easily. The proposed adaptive residual CNN is equipped with adjustable model capacity. It can be adjusted adaptively according to the complexity of the task. Various inherent fluctuation characteristics of wind signals can be captured by the adaptive residual strategy, improving the great generalization ability of model.

4.4. Effects of Reinforced Forecasting Strategy and Simplified-Boost Strategy. The reinforced forecasting performances for 1-step forecasting of datasets A and B are indicated in Figure 7. The experiment is implemented for 50 times for each value of k. “pr” denotes preliminary forecast results, and “final” indicates the simplified-boost reinforced forecast results.

It is declared that (1) not all reinforced models take effect. For instance, the reinforced model of dataset A acts awfully
when $k = 9$. (2) In terms of the same wind signal, the reinforced forecasting performances differ greatly with different input variables, which means the change of $k$ brings about different effects. (3) Regarding different wind signals, the optimal value of $k$ is different. (4) The simplified-boost reinforced forecasting acts best among all the models.

To see more clearly, the performances of simplified-boost reinforced forecasting and preliminary forecasting for 20 times experiments are shown in Figure 9. It is demonstrated that the forecasting ability is enhanced by the simplified-boost reinforced strategy.

5. Conclusion

The accuracy and robustness of models are essential for wind signal forecasting. However, the intermittency, uncertainty, and diversity of complex wind signals lead to enormous challenge for forecasting by a generalized model.

The multichannel convert provides additional views of the same input about temporal resolution, presenting additional information at different time scales. The proposed adaptive residual strategy allows forecasting for multiple kinds of wind signals with varying complexity by the same neutral network, which not only reduces the parameter redundancy and computational complexity, but also promotes the generalization capability of the model. Subsequently, the preliminary forecasting is enhanced by the simplified-boost reinforced forecasting, improving the forecasting accuracy and robustness to a great extent. The results of multistep forecasting for five kinds of nonstationary non-Gaussian wind signals verify the efficiency of the proposed model.

Though the preliminary forecasting and reinforced forecasting are two seemingly independent stages of the ensemble model, they complement each other actually. The deficiency of reinforced forecasting is that if the preliminary results are awful, the fault information or noise may be introduced into the reinforced model, resulting in unsatisfactory behavior of reinforced forecasting. Ideal preliminary results will bring quite effective information. Thus, satisfactory results could be obtained by the secondary forecasting, which creates a beneficent cycle, and vice versa. The phenomenon is particularly evident in multistep forecasting. Regarding multistep forecasting, more fresh
information from preliminary forecasting will be introduced into the reinforced model, which has a more significant impact on the enhancing performance. In conclusion, the reinforced forecasting is the icing on the cake for preliminary forecasting. The complementary nature makes them interdependent. Therefore, the forecasting accuracy of preliminary forecasting is still important.

The developed hybrid model performs strong forecasting ability. It can be considered for application in long-term forecasting and other fields, such as traffic forecasting, economic forecasting, and house price tendency forecasting. In terms of long-term forecasting, the correlation between long-term data is relatively weak due to the long interval of data compared with short-term data. Thus, the proposed model can be combined with data decomposition technology as a forecasting module for the long-term data with high volatility. Further study will focus on seeking for more efficient reinforced forecasting model different from the preliminary model, so as to better compensate the inherent defects of the preliminary model. What is more, more effective boosting strategy will be studied in the future.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant no. 51778354).

References

[1] S. Gao and S. L Wang, "Progressive collapse analysis of latticed telecommunication towers under wind loads," Advances in Civil Engineering, vol. 2018, Article ID 3293506, 15 pages, 2018.
[2] Z. Shu, L. Tang, S. Jiang, Z. Li, R. Zhao, and J. Zheng, "Improving the efficiency of Doppler lidar receiver for upper atmospheric wind field measurement," Optik, vol. 149, pp. 169–173, 2017.
[3] Y. Zhou, Y. Q. Li, Y. Y. Zhang, and A. Yoshida, "Characteristics of wind load on spatial structures with typical shapes due to aerodynamic geometrical parameters and terrain type," Advances in Civil Engineering, vol. 2018, Article ID 9738038, 17 pages, 2018.
[4] X. L. Li, D. K. Yu, and Z. L. Li, "Parameter analysis on wind-induced vibration of UHV cross-rod suspension tower-line," Advances in Civil Engineering, vol. 2017, Article ID 8756019, 9 pages, 2017.
[5] F. Ubertini and F. Giuliano, "Computer simulation of stochastic wind velocity fields for structural response analysis: comparisons and applications," Advances in Civil Engineering, vol. 2010, Article ID 749578, 2010.
[6] L. Casella, "Wind speed reconstruction using a novel Multivariate Probabilistic method and Multiple Linear Regression: advantages compared to the single correlation approach," Journal of Wind Engineering and Industrial Aerodynamics, vol. 191, pp. 252–265, 2019.
[7] S. K. Perepu and A. K. Tangirala, "Reconstruction of missing data using compressed sensing techniques with adaptive dictionary," Journal of Process Control, vol. 47, pp. 175–190, 2016.
[8] X. F. Zhao, L. G. Niu, Q. H. Chen, Z. Zhang, and Y. X. Luan, "Restoration and rebuilding of the wind disturbed ecosystems at the wind disaster region in the national nature reserve of changbai mountain," Journal of Northeast Forestry University, vol. 4, 2004.
[9] Z. Zhang, L. Ye, H. Qin et al., "Wind speed prediction method using shared weight long short-term memory network and Gaussian process regression," Applied Energy, vol. 247, pp. 270–284, 2019.
[10] M. Singh, "Protection coordination in distribution systems with and without distributed energy resources - a review," Protection and Control of Modern Power Systems, vol. 2, no. 3, pp. 294–310, 2017.
[11] S. Fan, G. He, X. Zhou, and M. Cui, "Online optimization for networked distributed energy resources with time-coupling constraints," IEEE Transactions on Smart Grid, vol. 99, p. 1, 2020.
[12] D. J. Allen, A. S. Tomlin, C. S. E. Bale, A. Skea, S. Vosper, and M. L. Gallani, "A boundary layer scaling technique for estimating near-surface wind energy using numerical weather prediction and wind map data," Applied Energy, vol. 208, pp. 1246–1257, 2017.
[13] X. Zhao, J. Liu, D. Yu, and J. Chang, "One-day-ahead probabilistic wind speed forecast based on optimized numerical weather prediction data," Energy Conversion and Management, vol. 164, pp. 560–569, 2018.
[14] M. Lydia, S. Suresh Kumar, A. Immanuel Selvakumar, and G. Edwin Prem Kumar, "Linear and non-linear autoregressive models for short-term wind speed forecasting," Energy Conversion and Management, vol. 112, pp. 115–124, 2016.
[15] S. N. S. Aasim, S. N. Singh, and A. Mishaputra, "Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting," Renewable Energy, vol. 136, pp. 758–768, 2019.
[16] A. Solgi, V. Nourani, and A. Pourbaghi, "Forecasting daily precipitation using hybrid model of wavelet-artificial neural network and comparison with adaptive neurofuzzy inference system (case study: verayneh station, nahavand)," Advances in Civil Engineering, vol. 2014, Article ID 279368, 15 pages, 2014.
[17] S. Golnaraghi, Z. Zangenehmador, O. Moselhi, and S. Alkass, "Application of artificial neural network(s) in predicting formwork labour productivity," Advances in Civil Engineering, vol. 2019, Article ID 5972620, 11 pages, 2018.
[18] H. S. Dhiman, D. Deb, and J. M. Guerrero, "Hybrid machine intelligent SVR variants for wind forecasting and ramp events," Renewable and Sustainable Energy Reviews, vol. 108, pp. 369–379, 2019.
[19] Z. Li, L. Ye, Y. Zhao et al., "Short-term wind power prediction based on extreme learning machine with error correction," Protection and Control of Modern Power Systems, vol. 1, no. 1, pp. 9–16, 2016.
[20] J. Man, P. Li, R. Ran, Y. Che, and Y. Zhou, "A short-term photovoltaic power prediction model based on the gradient boost decision tree," Applied Sciences, vol. 8, no. 5, p. 689, 2018.
[21] R. Feng, H.-j. Zheng, H. Gao et al., "Recurrent Neural Network and random forest for analysis and accurate forecast of...."
atmospheric pollutants: a case study in Hangzhou, China,”
Journal of Cleaner Production, vol. 231, pp. 1005–1015, 2019.
[22] S. Tivari, R. Babbar, and G. Kaur, “Performance evaluation of two
ANFIS models for predicting water quality index of river satluj
(India),” Advances in Civil Engineering, vol. 2018, Article ID
8971079, 10 pages, 2018.
[23] G. Barbounis and T. Thanasis, “Long-term wind speed and
power forecasting using local recurrent neural network
models,” IEEE Transactions on Energy Conversion, vol. 21, no. 1,
pp. 273–284, 2006.
[24] M. Ding, H. Zhou, H. Xie et al., “A gated recurrent unit neural
networks based wind speed error correction model for short-
term wind power forecasting,” Neurocomputing, vol. 365,
pp. 54–61, 2019.
[25] F. A. Gers, J. Schmidhuber, and F. Cummins, “Learning to
forget: continual prediction with LSTM,” Neural Computation,
vol. 12, no. 10, pp. 2451–2471, 2000.
[26] S. Y. Li and X. F. Zhao, “Image-based concrete crack detection
using convolutional neural network and exhaustive search
technique,” Advances in Civil Engineering, vol. 2019, Article
ID 6520620, 12 pages, 2019.
[27] C. Liu, W. H. Hsiao, and Y. C. Tu, “Time series classification
with multivariate convolutional neural network,” IEEE
Transactions on Industrial Electronics, vol. 66, no. 6,
pp. 4788–4797, 2018.
[28] S. Harbola and V. Coors, “One dimensional convolutional
neural network architectures for wind Prediction,” Energy
Conversion and Management, vol. 195, pp. 70–75, 2019.
[29] X. Niu and J. Wang, “A combined model based on data
preprocessing strategy and multi-objective optimization al-
gorithm for short-term wind speed forecasting,” Applied
Energy, vol. 241, no. 1, pp. 519–539, 2019.
[30] P. Jiang, R. Li, and H. Li, “Multi-objective algorithm for the
design of prediction intervals for wind power forecasting
model,” Applied Mathematical Modelling, vol. 67, pp. 101–122,
2019.
[31] P. Jiang and Z. Liu, “Variable weights combined model based
on multi-objective optimization for short-term wind speed
forecasting,” Applied Soft Computing, vol. 82, Article ID
105587, 2019.
[32] D. S. d. O. Santos Júnio, J. F. L. de Oliveira, and P. S. G. de Mattos Neto, “An intelligent hybridization of ARIMA with machine learning models for time series fore-
casting,” Knowledge-Based Systems, vol. 175, pp. 72–86, 2019.
[33] T. Ouyang, H. Huang, Y. He, and Z. Tang, “Chaotic wind
power time series prediction via switching data-driven
modes,” Renewable Energy, vol. 145, pp. 270–281, 2020.
[34] N. Korprasertsak and T. Leephakpreeda, “Robust short-term
prediction of wind power generation under uncertainty via
statistical interpretation of multiple forecasting models,”
Energy, vol. 180, pp. 387–397, 2019.
[35] Y. Jiang and G. Huang, “Short-term wind speed prediction:
hybrid of ensemble empirical mode decomposition, feature
selection and error correction,” Energy Conversion and
Management, vol. 144, pp. 340–350, 2017.
[36] Y. Jiang, G. Huang, X. Peng, Y. Li, and Q. Yang, “A novel wind
speed prediction method: hybrid of correlation-aided DWT,
LSSVM and GARCH,” Journal of Wind Engineering and
Industrial Aerodynamics, vol. 174, pp. 28–38, 2018.
[37] Y. Li, H. Wu, and H. Liu, “Multi-step wind speed forecasting
using EWT decomposition, LSTM principal computing,
RELM subordinate computing and IEWT reconstruction,”
Energy Conversion and Management, vol. 167, no. 1,
pp. 203–219, 2018.
[38] Y. Hao and C. Tian, “A novel two-stage forecasting model
based on error factor and ensemble method for multi-step
wind power forecasting,” Applied Energy, vol. 238, no. 15,
pp. 368–383, 2019.
[39] H. Wang, S. Han, Y. Liu, J. Yan, and L. Li, “Sequence transfer
correction algorithm for numerical weather prediction wind
speed and its application in a wind power forecasting system,”
Applied Energy, vol. 237, no. 1, pp. 1–10, 2019.
[40] D. Rahul, S. R. Samantaray, and B. K. Panigrahi, “An spa-
tiotemporal information system based wide-area protection
fault identification scheme,” International Journal of Electrical Power & Energy Systems, vol. 89, pp. 136–145, 2017.
[41] M. Cui, J. Wang, and B. Chen, “Flexible machine learning-
based cyberattack detection using spatiotemporal patterns for
distribution systems,” IEEE Transactions on Smart Grid,
vol. 11, no. 2, pp. 1805–1808, 2020.
[42] S. Sun, Q. Yang, and W. Yan, “Optimal temporal-spatial PEV
charging scheduling in active power distribution networks,”
Protection and Control of Modern Power Systems, vol. 2, no. 4,
pp. 379–388, 2017.
[43] E. Oh and H. Wang, “Reinforcement-learning-Based energy
storage system operation strategies to manage wind power
forecast uncertainty,” IEEE Access, vol. 8, pp. 20965–20976,
2020.
[44] C. Chen, M. Cui, F. Li, S. Yin, and X. Wang, “Model-free
emergency frequency control based on reinforcement
learning,” IEEE Transactions on Industrial Informatics,
pp. 1–11, 2020.
[45] G. Yao, F. J. Wei, Y. Yang, and J. Y. Sun, “Deep-learning-based
buhgole detection for concrete surface image,” Advances in
Civil Engineering, vol. 2019, Article ID 8582963, 12 pages,
2019.
[46] P. S. Crawford, M. A. Al-Zarrad, A. J. Graettinger et al., “Rapid
disaster data dissemination and vulnerability assessment
through synthesis of a web-based extreme event viewer and
deep learning,” Advances in Civil Engineering, vol. 2018,
Article ID 7258156, 15 pages, 2018.
[47] K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, “Deep residual
learning for image recognition,” in Proceedings of the 2016
IEEE Conference On Computer Vision And Pattern
Recognition, Las Vegas, NV, USA, June 2016.
[48] Y. Freund and R. E. Schapire, “A decision-theoretic gener-
alization of on-line learning and an application to boosting,”
Journal of Computer and System Sciences, vol. 55, no. 1,
p. 119–139, 1997.
[49] R. J. Hyndman and A. B. Koehler, “Another look at measures of
forecast accuracy,” International Journal of Forecasting,
vol. 22, no. 4, pp. 679–688, 2006.
[50] Z. H. Zhang, Z. H. Liu, and S. L. Dong, “Field measurement of
wind pressure and wind-induced vibration of large-span
spatial cable-truss system under strong wind or typhoon,”
Journal of Shanghai Normal University (Natural Sciences),
vol. 5, no. 42, pp. 546–550, 2013.
[51] K. D. Gast, “A comparison of extreme wind events as sampled
in the 2002 thunderstorm outflow experiment,” Master’s
Thesis, Texas Tech University, Lubbock, TX, USA, 2003.
[52] F. X. Diebold and R. S. Mariano, “Comparing predictive
accuracy,” Journal of Business & Economic Statistics,
vol. 13, no. 3, pp. 253–263, 1995.
[53] Y.-L. Hu and L. Chen, “A nonlinear hybrid wind speed
forecasting model using LSTM network, hysteric ELM and
Differential Evolution algorithm,” Energy Conversion and
Management, vol. 173, pp. 123–142, 2018.