Hybrid forecasting system based on data area division and deep learning neural network for short-term wind speed forecasting

Zhuoyi Liu, Ryoichi Hara, Hiroyuki Kita

Graduate School of Information Science and Technology, Hokkaido University, Sapporo 0600814, Japan

Correspondence information:

Zhuoyi Liu

Power and Energy Systems Laboratory, Graduate School of Information Science and Technology, Hokkaido University, Sapporo 0600814, Japan,

Email address: liuzhuoyi1210@gmail.com

Telephone number: +8107047937490

Hybrid forecasting system based on data area division and deep learning neural network for short-term wind speed forecasting

Zhuoyi Liu[ab]*, Ryoichi Hara[ab], Hiroyuki Kita[ab]

a *Graduate School of Information Science and Technology, Hokkaido University, Sapporo 0600814, Japan*
b *Power and Energy Systems Laboratory, Hokkaido University, Sapporo 0600814, Japan*

* Corresponding author:
E-mail address: liuzhuoyi1210@gmail.com (Z. Liu)

ABSTRACT

Wind speed forecasting is essential for the dispatch, controllability, and stability of power grids, and its accuracy is vital to the effective use of wind resources. In this study, a novel hybrid wind speed forecasting system is developed based on the data area division (DAD) method and a deep learning neural network model. The system consists of three modules: extraction module, data preprocessing module, and forecasting module. In the data extraction module, a large amount of valid historical data is extracted, filtered, and classified from the forecast location and the surrounding locations. In the data preprocessing module, complementary ensemble empirical mode decomposition is used to decompose the wind speed data. In the forecasting module, a long short-term memory network optimized by using a genetic algorithm is used to forecast the decomposed wind speed data and integrate them into the final forecast results. Numerical simulation results show that (a) the forecast system maintains RMSE in the range of 0.2 m/s to 0.6 m/s and MAPE in the range of 3.0% to 7.0% for short-term wind speed forecasts at different locations for different time periods, showing good stability. (b) For wind speed forecasting at different time intervals, the accuracy of wind speed forecasting at 10-minute and 30-minute intervals is better, while the accuracy of forecasting at a 60-minute interval needs to be improved, but overall, the forecasting system shows good generalizability. (c) The forecast system improves the forecast accuracy of short-term wind speed forecasting more effectively than other conventional methods, and the improvement of RMSE and

MAPE remains in the range of 14% to 39% and 13% to 27% even compared with the hybrid forecast model that has better forecast accuracy. (d) For area-wide short-term wind power forecasting, the forecast deviation value of this forecasting system remains below 6% throughout the year, showing good practicality.

Keywords: Wind speed forecasting, Data area division, Complementary ensemble empirical mode decomposition, Long short-term memory, Genetic algorithm

1. Introduction

The use of renewable energy, such as wind power and solar power, has attracted widespread attention in the power industry in recent years because of the gradual increase in fossil fuel consumption and the intensification of the greenhouse effect. As a promising renewable energy source, wind energy has attracted the attention of many countries. A Global Wind Energy Council report indicates that the cumulative installed capacity of global wind energy has exceeded 700 GW as of 2020, and the global newly installed wind power capacity will exceed 300 GW in the next five years [1].

As the proportion of wind power generation in the power system gradually increases, it will have a steadily growing impact on the safe and stable operation of the power system [2]. The efficiency of wind power generation depends on wind energy, unlike that of traditional power generation methods; however, the randomness and intermittentness of wind intensify the uncontrollability of wind power generation [3], thereby resulting in unstable wind power generation. Accurate wind speed forecasting is crucial to ensure the safe and stable operation of power systems and to promote the further development of the wind power industry [4]. Research on wind speed forecasting has been attracting increasing attention from scholars and managers, and various wind speed forecasting models have been developed. These models can be roughly divided into three categories: physics, statistics, and artificial intelligence [5].

1.1 Traditional forecasting models

Numerical weather forecast (NWP) is a physical forecasting model that uses forecast results, which include wind speed, wind direction, temperature, and other weather data, of a weather forecast system [6]. The use of physical laws and boundary conditions to

simulate wind speed changes usually results in good long-term wind speed forecasting performance [7]. However, this approach relies on various types of historical data and consumes a great amount of computing resources [8].

Traditional statistical and artificial intelligence models are more widely used than physical models. Statistical forecasting models use historical or time–space samples to forecast wind energy or speed, thereby producing remarkable estimation results in short-term wind speed forecasting [9]. Autoregressive (AR), AR moving average (ARMA), and AR integrated moving average (ARIMA) models are typical examples of statistical models [10]. Radziukynas et al. used ARIMA [11] to forecast short-term wind speed after 1 h, while Rajesh et al. used *fractional*-ARIMA (*f*-ARIMA) to forecast wind speed two days in advance [12], obtaining remarkable forecasting performance results. However, the statistical forecasting model also has evident shortcomings. For instance, the time series data must be stable when the statistical model is used for forecasting. That is, capturing the law is difficult if the data are unstable. Thus, the model is less effective when used for longer-term wind speed forecasting or wind speed forecasting during periods of major wind speed changes [13].

To improve the accuracy of statistical models in wind speed forecasting, many scholars have proposed a forecasting method that combines artificial intelligence and statistical models [14]. In recent years, artificial intelligence has played an important role in various fields. Many scholars have applied artificial intelligence methods to the field of forecasting and achieved remarkable results. For example, general regression neural network (GRNN) [15], wavelet neural network (WNN) [16], generalized feed-forward neural network (GFNN) [17], and backpropagation neural network (BPNN) [18] are typical artificial neural networks (ANN) used in wind speed forecasting [19]. When most of the time series have nonlinear characteristics, ANN has good forecasting ability and can forecast results accurately [20]. However, ANN has some shortcomings that cannot be ignored, such as the tendency to fall into local optima and overfitting [21].

1.2 Hybrid forecasting model

The single forecasting method has advantages and limitations. First, achieving stable, highly accurate forecasting under different geographical conditions and periods of wind speed forecasting is difficult. To avoid such problems and improve the forecasting accuracy, a growing number of researchers have proposed two kinds of hybrid forecasting

models. The first kind of model involves optimizing the algorithm. Wang et al. proposed a hybrid forecasting model on the basis of WNN optimized by using a genetic algorithm (GA) [22]. Shukur et al. introduced a forecasting model [23] that combined Kalman and BPNN. They also optimized BPNN by using the Kalman principle and by considering the influence of coefficient fully, thus achieving good forecasting accuracy. Wang et al. proposed a neural network-based linear ensemble framework that combines multiple neural networks and used the advantages of different neural networks to forecast various types of wind speed. This method attained remarkable forecasting performance [24].

The other kind of hybrid forecasting model mainly focuses on data preprocessing. The intermittent original wind speed time series is decomposed into several fixed sublayers to improve the forecasting performance of the model effectively [25]. Widely recognized decomposition algorithms include wavelet decomposition (WD), wavelet packet decomposition (WPD) [26], empirical mode decomposition (EMD), integrated EMD (EEMD), and fast integrated EMD (FEEMD) [27]. Generally, WD and WPD have excellent time–frequency analysis ability, whereas EMD, EEMD, and FEEMD have good adaptive ability, thereby allowing random fluctuations to be eliminated. On this basis, Liu et al. proposed a hybrid forecasting model based on WD and support vector machine [28]. Wang et al. provided a short-term wind speed forecasting model using EMD and BPNNs [29]. Liu et al. proposed a hybrid forecasting model using EEMD and ARIMA [30].

1.3. Proposed forecasting system

The above analysis shows that most previous studies focused mainly on optimization algorithms and data processing to improve the accuracy of wind speed forecasting while ignoring the selection and extraction of wind speed data [31]. In addition, the established wind speed forecasting models concentrate on a single location only. Local historical data are used as training data for short-term forecasting of future wind speed. In this case, historical data that are too old are not informative. Moreover, providing accurate wind speed forecasting is difficult in areas where historical data are lacking or where the wind speed change cycle is short and the change range is large [32].

In the previous forecasting process, the forecasting accuracy of different areas varies greatly, and forecasting is even impossible in some areas, often because of drastic changes in wind speed and the lack of historical data of wind speed [43]. Therefore, on the basis of improving the precision, this research also focuses on improving the stability and

versatility of the forecasting model. Wind speed is another important indicator that affects wind power generation. The accuracy of wind speed forecasting plays an important role in the advance control of wind power generation and the deployment of power system supply and demand. To this end, this paper develops a new type of hybrid wind speed forecasting system. Numerical simulations prove that the proposed forecasting system can ensure high forecasting accuracy, stability, and versatility. The analysis of the degree of influence of wind power generation proves that the forecasting system is highly applicable to wind power generation forecasting. Moreover, the deviation of forecasting has a minimal effect on the wind power generation of the entire area.

The forecasting system proposed in this work includes three modules (**Fig.1**) and one evaluation indicator:

Data extraction module: This module proposes a data area division (DAD) method to extract historical wind speed data.

Data preprocessing module: This module adopts complementary ensemble empirical mode (CEEMD) to wind speed decomposition and denoising.

Forecasting module: GA-optimized long short-term memory (LSTM) neural network is used to forecast the decomposed submodules and to synthesize the final forecasting results.

Influence analysis: Unlike in other forecasting methods that evaluate the forecasting accuracy of the proposed forecasting system, the concept of wind power impact indicators is proposed to analyze the impact of wind power forecasting.

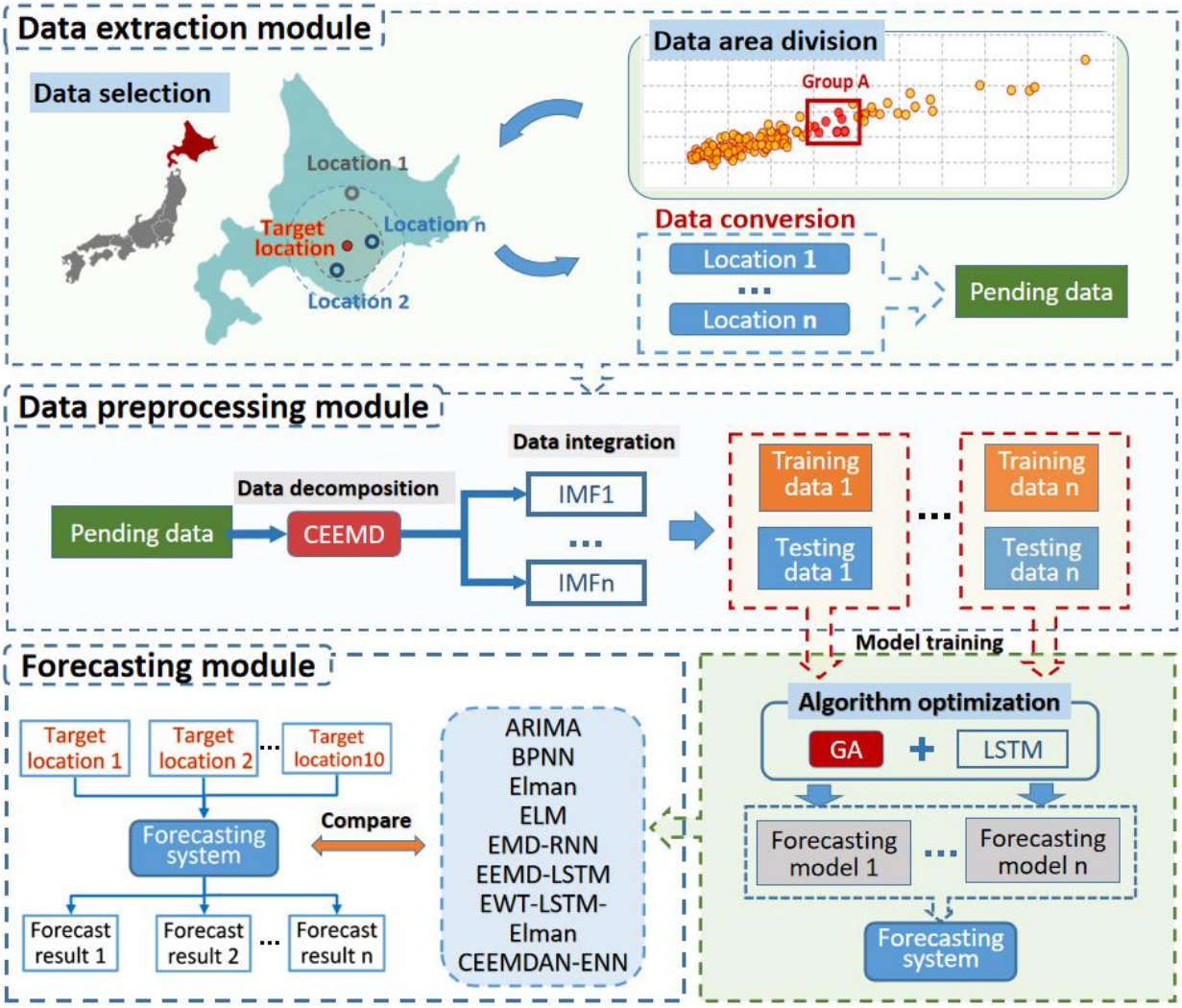


Fig. 1. Structure of the proposed hybrid forecasting system.

The innovation and contribution of this work can be summarized as follows:

(1) The DAD method is proposed to divide wind speed data into areas. A large amount of effective wind speed historical data is extracted, thereby providing a reliable data basis for wind speed forecasting in different areas.

(2) Advanced data noise reduction and decomposition strategies are adopted to process wind speed series.

(3) The forecast interval of the wind speed time series will vary according to the different needs of managers. This work selects three wind speed time series (e.g., 10, 30, and 60 minutes) with different time intervals for forecasting to meet various requirements in practical applications.

(4) The concept of wind power impact indicator is proposed, and the power generation of the entire area of Hokkaido, Japan, is taken as the research target.

The remainder of this work is organized as follows: Section 2 introduces the main

methods and forecasting process of the proposed system. Section 3 introduces the relevant preparations before the experiment, including data selection, model parameter setting, and evaluation methods. Section 4 analyzes the experimental results under different data and discusses the forecasting results comprehensively. Section 5 provides the conclusion. **Fig. 1** shows the overall process framework of the proposed hybrid forecasting system.

2. Methodology

In this section, data extraction, data preprocessing, basic forecasting methods, and the entire forecasting process will be introduced.

2.1. Data extraction module

For wind speed forecasting, the extraction and use of historical wind speed data have always been important factors that affect the forecast accuracy. **Table 1** presents the data usage and brief descriptions of some of the methods mentioned in Sections 1.1 and 1.2. Previous studies focused on the selection of forecasting methods and the optimization of forecasting models, paying little attention to the selection of available historical data. Therefore, the previous forecasting models are mostly suitable for site-specific wind speed forecasting with sufficient historical data and are difficult to apply to a wider range of forecasting tasks.

Table 1
Data selection and description of each forecasting model.

Models	Datasets	Description
Physical forecasting methods		
NWP [7]	All hourly meteorological data for the four-year period 2012–2015 for New England, USA (single region)	For long-term wind speed forecasting, the forecasting accuracy based on physical methods is high, but it needs a large amount of historical meteorological data as the basis and is not suitable for short-term wind speed forecasting.
etc.		
Statistical forecasting methods		
ARIMA [11]	Wind speed data at 10-minute intervals from June to September 2007 at a wind farm in Lithuania (single site)	Compared with physical models, statistical models are easier to implement and more accurate for short-term wind speed forecasting, but they require stable wind speed historical data and are not suitable for wind speed forecasting during periods of drastic wind speed changes.
f-ARIMA [12]	Hourly average wind speed at four wind monitoring sites in North Dakota, USA, over a one-month period (forecast at each of the four locations)	
etc.		
Artificial intelligence methods		

ANN [18]	Wind speed data at 10-minute and 30-minute intervals for March–April 2011 in Shandong Province, China (single site)	Artificial neural networks can simulate complex nonlinear relationships and have high accuracy in short-term wind speed forecasting, but problems exist, such as the tendency to fall into local optima, overfitting, and the dependence of forecasting accuracy on the quality of historical data.
WNN [16]	Hourly average wind speed data for a wind farm in Xinjiang, China, from March–May and September–November 2014 (single site)	
etc.		
Hybrid forecasting models		
EMD-BPNNs [29]	Wind speed data at 10-minute and 1-hour time intervals for January and April 2011 at a wind farm in Inner Mongolia, China (single site)	Hybrid forecasting models that combine data processing models and forecasting models have higher accuracy than other models. However, each hybrid model tends to be designed for a specific situation, and the same method used for a different location or time period may yield poor forecasting results.
EEMD-ARIMA [30]	Wind speed data for a one-minute time interval at two wind monitoring sites along the Qinghai–Tibet Railway in China (two sites)	
etc.		

In the process of wind speed forecasting, the author's team found that wind speed data from other areas that are highly correlated with the wind speed time series of the location to be forecasted can be used as training data after processing. Considering that the wind speed time series at these locations have similar changes, the wind speeds are also very close. The extraction of wind speed data from multiple locations within a short period of time as training data can ensure a higher correlation between the selected historical data and the wind speed to be forecasted while avoiding the problem of insufficient historical data to train the forecasting model. Thus, a DAD method is proposed based on the above point of view. The flowchart is shown in **Fig. 2**.

Step 1. Divide according to the wind speed： On the basis of the average and standard deviation of the historical wind speed time series at each location, the area to be forecasted was divided into several groups using the k-means method. With 141 locations in Hokkaido taken as an example, the specific steps are:

(1) k locations are selected randomly as initial clustering centers, $\alpha = \alpha_1, \alpha_2, \cdots, \alpha_k$.

(2) For each location xi in the dataset, its distance to the k cluster centers is calculated and assigned it to the group corresponding to the cluster center with the smallest distance.

(3) For each group, recalculate its cluster center $\alpha_j = \frac{1}{|c_i|}\sum_{x \in c_i} x$ is recalculated. (the center of mass of all samples belonging to the group)

(4) Steps 2 and 3 are repeated until the clustering centers no longer change significantly.

The only parameter that needs to be set for the k-means method is the number of clusters K. Here, K is set to 10 (the trial-and-error method suggests that this scale is reasonable). As shown in **Fig. 3**, all locations in Hokkaido can be divided into 10 groups.

Groups A, B, and C are taken as examples; group A includes 10 locations, group B includes 6 locations, and group C includes only 1 location. The wind speed data used for the grouping in this example are the historical wind speed data for various locations in January 2019; in fact, the wind speed data used for the grouping when forecasting using the proposed forecasting system is the historical wind speed data for 30 days prior to the forecast time point.

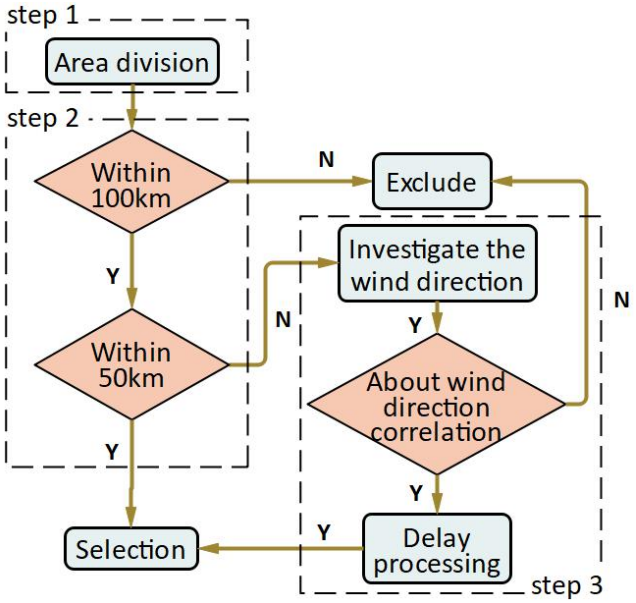


Fig. 2. Flowchart of the DAD method.

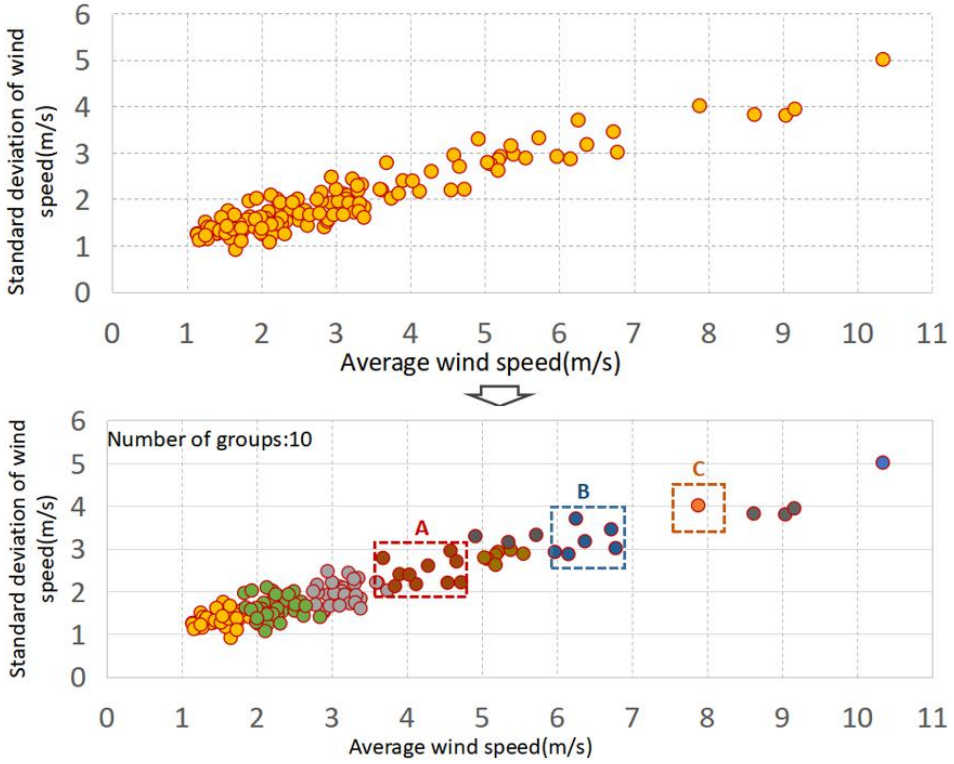


Fig. 3. Grouping results for the Hokkaido area.

Step 2. Divide by geographic distance: On the basis of Step 1, the available wind speed history data are selected according to the geographic distance. For example, Muroran is taken as the center, and a search for areas within radii of 0, 50, 100, and 150 km is performed. The results are shown in **Table 1**. The Pearson correlation coefficient between each location and Muroran's wind speed time series is calculated by using Equation (1). The results are shown in **Table 2**. The wind speed time series of locations within 50 km have an obvious correlation with Muroran, and the wind speed time series of locations within 100 km and Muroran exhibit a certain relevance. Therefore, on the basis of the division result of Step 1, the wind speed data of the locations within 50 km of the center to be forecasted directly are used as the training data, and the wind speed data of locations within 50–100 km need to be investigated further.

$$r(V_1, V_2) = \frac{Cov(V_1, V_2)}{\sqrt{Var[V_1]Var[V_2]}} \tag{1}$$

where V_i is the wind speed time series at location i, Cov(V_1 ,V_2) is the covariance of V_1 and V_2, and $Var[V_i]$ is the variance of V_i.

Table 2
Locations within each radius around Muroran.

Observation radii (km)	Locations	Number of locations
0	Muroran	1
50	Date, Noboribetsu	3
100	Tomakomai, Rankoshi, Hokuto, Suttsu	7
150	Okushiri, Iwamizawa	9

Table 3
Correlation coefficient of wind speed time series between locations.

Correlation coefficient	Muroran
Date	0.6021
Noboribetsu	0.5766
Tomakomai	0.4123
Rankoshi	0.4526
Hokuto	0.3314
Suttsu	0.4021
Okushiri	0.2421
Iwamizawa	0.1719

Step 3. Divide by wind direction: Our investigation shows that within a certain period of time, the wind speed series of adjacent locations with roughly the same wind direction have a strong correlation. Muroran is taken as an example. **Fig. 4** shows the wind direction distribution of Muroran in January, April, July, and October 2019. The change in wind direction shows evident seasonality. With January taken as an example, the wind direction in Muroran is mainly northwest wind. At this time, Suttsu, which is in the 50–100 km

range around Muroran, is compared with Muroran. The comparison of the wind speed time series in January is shown in **Fig. 5 (a)**. The wind speed time series at the two locations show a certain correlation. After several hours of delay processing, Suttsu's wind speed time series has a stronger correlation with that of Muroran. The correlation coefficients between the wind speed time series of Suttsu after each hourly time delay and the wind speed time series of Muroran, and the autocorrelation coefficients between the wind speed time series of Suttsu after each hourly time delay and the original wind speed time series are shown in **Table 4**, where the autocorrelation coefficients are calculated as shown in Equation (2).

$$r_a(h) = \frac{\sum_{i=1}^{n-h}(y_i - \hat{u})(y_{i+h} - \hat{u})}{\sum_{i=1}^{n}(y_i - \hat{u})} \quad (2)$$

where h is the number of hours of delay, n is the total length of the wind speed time series, and $\hat{u}$ is the average wind speed of the wind speed series.

Table 4 shows that the autocorrelation coefficients of the wind speed time series of Suttsu and the original wind speed time series gradually become smaller as the number of delay hours increases. From this finding, the wind speed time series after the delay treatment can be judged to no longer have an autocorrelation with the original wind speed time series after the delay hours reach a certain range. Meanwhile, the correlation coefficients of the wind speed time series of Suttsu and Muroran after the delay treatment show that the wind speed time series of the two places have the strongest correlation after the five-hour delay treatment. The autocorrelation coefficient of the wind speed time series of Suttsu is 0.5176 at this time; thus, the wind speed time series at this time still maintains the characteristics of the original wind speed time series to some extent. The comparison chart of the wind speed time series of the two places in this case is shown in **Fig. 5 (b)**. On the basis of the above analysis, Suttsu's wind speed time series after a five-hour delay is selected when forecasting Muroran's wind speed.

Through the screening of the three aforementioned steps, when Muroran is the forecasted target location and January is the forecasted time period, the wind speed data of Date, Noboribetsu, Suttsu, and Rankoshi can be used as training data.

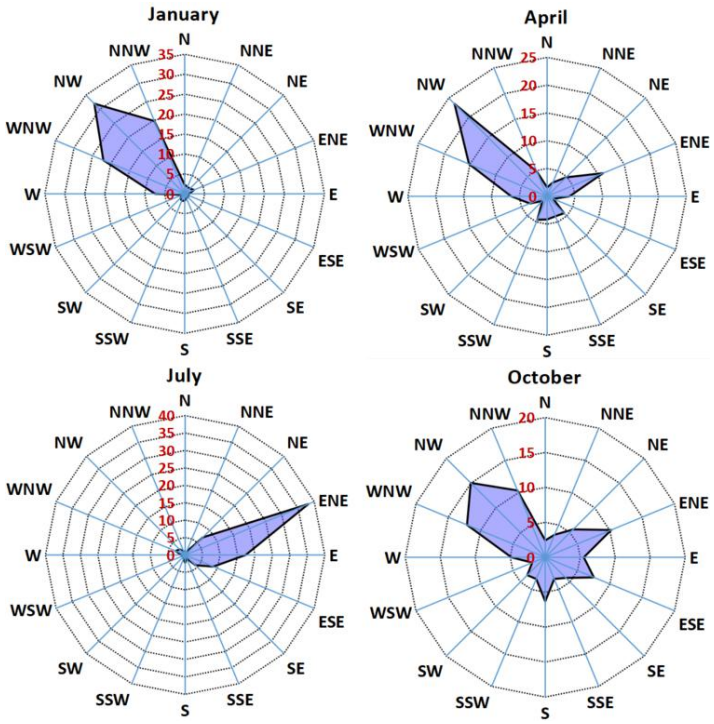
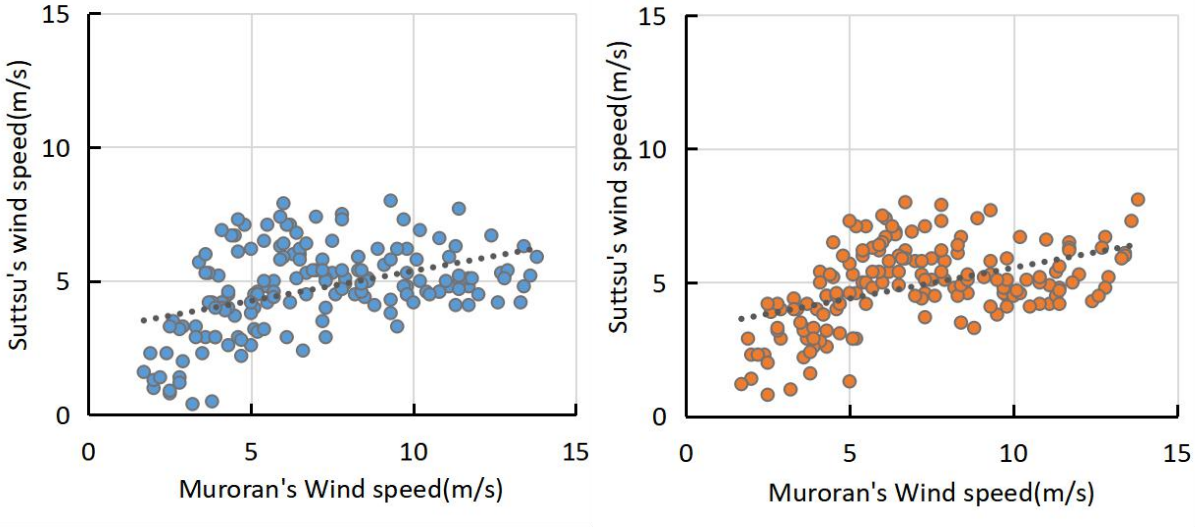


Fig. 4. Wind direction distribution of Muroran in January, April, July, and October 2019.



(a) (b)

Fig. 5. Comparison of Muroran's and Suttsu's wind speed time series.

Table 4
Changes in correlation and autocorrelation coefficients with time-delayed processing

Time delay	Correlation coefficient (Muroran and Suttsu)	Autocorrelation coefficient (Suttsu)
1h	0.4429	0.6986
2h	0.4365	0.6463
3h	0.4863	0.5774
4h	0.5118	0.5203
5h	0.5711	0.5176

6h	0.5321	0.4592

2.2. Data preprocessing module

Although the aforementioned data collection method provides sufficient training data for the forecasting model, thereby improving the versatility of the forecasting model greatly, the original data that are directly used will lead to poor training effects because of the noise and random fluctuations in the original wind speed series. To reduce the random volatility of wind speed time series, many researchers have proposed data preprocessing strategies using EMD [33] or EEMD [34].

EMD [33] refers to the nonlinear and non-stationary signal analysis method developed by Huang et al. in 1998. This method extracts a series of intrinsic mode functions (IMF) from the original signal through stepwise filtering and processes them separately. However, EMD has an evident shortcoming called mode mixing [35], in which a single IMF composed of signals with significantly different scales is easily generated or signals of the same scale appear in different IMF components. To solve this problem, Wu and Wang proposed EEMD in 2008 [34], which avoids mode mixing effectively by adding white noise to the original signal. However, it also has low decomposition efficiency and residual noise.

To decompose the original wind speed data more efficiently and accurately, an improved version of EEMD called CEEMD is used to process wind speed data [36].

The description of the CEEMD algorithm is shown in **Fig. 6**, with *NE* and *Nstd* being discussed in detail in Section 3.2.

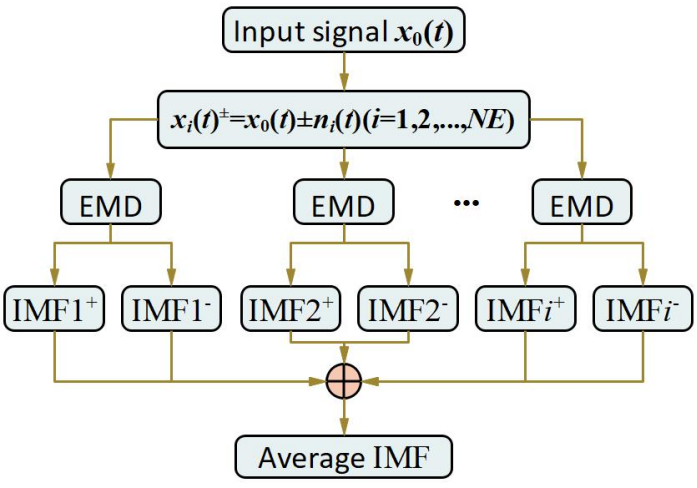


Fig. 6. CEEMD algorithm flowchart.

1) In the original wind speed signal $x_0(t)$, two groups of white noise sequence $\pm n$(t) with opposite signs are added to obtain $x_i(t)^{\pm}$ (i is the sequence number of the added white noise).
2) The signal $x_i(t)^{\pm}$ with the added white noise is decomposed into IMFi by using EMD.
3) Steps 1 and 2 are repeated using different white noises, and the corresponding IMF$i^{\pm}$ components are obtained. The number of repeated procedures is called the ensemble number (*NE*).
4) The average value of IMF$i^{\pm}$ components is taken as the final result or IMF.

In EMD, the combination of all IMFs corresponds to the original data. However, in CEEMD, the combined result is no longer the original data because white noise has been added.

With the addition of white noise, the original wind speed signal becomes more stable, and continuous signals can be decomposed easily, thereby avoiding the problem of mode mixing. The reason is that each time a pair of white noise $\pm n(t)$ with opposite sign is added, the effect of some white noise on the original signal can be eliminated when the IMF$i^{\pm}$ is superposed, thus restoring the characteristics of the original wind speed signal.

2.3 Forecasting module

In the forecasting module, the LSTM neural network is used, which has an improved effect on time series forecasting in the deep learning neural network. To improve forecasting efficiency, GA is used to optimize the internal parameters of LSTM.

LSTM refers to an improved version of the recurrent neural network (RNN) proposed by Schmidhuber and Hochreiter in 1997 [37]. LSTM is suitable for dealing with time series-related problems because of its unique design structure. At present, it is widely used in speech recognition, text forecasting, and sentiment analysis.

The function of RNN to realize long-term memory calculation is related to the previous calculation of n times. RNN has a very weak memory ability because of limited function. The increase in the amount of calculation leads to increased model training time, the problem of gradient disappearance, and gradient explosion. Moreover, as the complexity of the neural network increases, mismatch may occur easily between the order of magnitude

of the training data and the complexity of the network, or too much noise interference may occur in the samples, causing the neural network to remember the features of the noise and ignore the relationship between the real inputs and outputs. All these conditions can easily cause overfitting problems. On the basis of the improvement of the RNN, LSTM has a complex dynamic structure, which can learn long-term calculation and memory information from the input data by introducing the LSTM block in the chain structure [38]. **Fig. 7** shows the input gate i_t, forget gate f_t, output gate o_t, unit state c_t, and hidden state h_t in the calculation unit of LSTM. Activation function and point-by-point multiplication are used to form the threshold so that the LSTM unit has a computationally long ability to distance historical information.

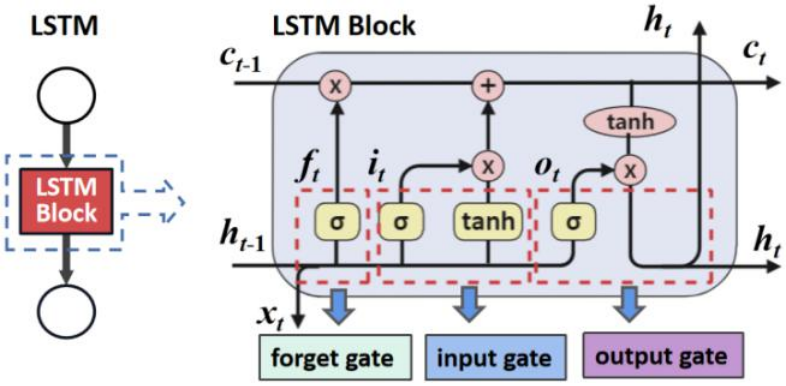


Fig. 7. LSTM structure.

The input gate i_t determines the unit that needs to be updated

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \tag{3}$$

where W and b are the weight matrix and the offset, respectively. The forget gate f_t determines the historical information that needs to be retained

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right) \tag{4}$$

Subsequently, the state of the model is updated as

$$c_t = f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{5}$$

The output gate o_t outputs the calculation result in the LSTM unit as

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \tag{6}$$

where σ is the activation function. The hidden state information is obtained as

$$h_t = o_t \tanh(c_t) \tag{7}$$

where tanh is the activation function.

The complexity of the network structure is kept in a reasonable range in each step of the training process because of the inclusion of dropout layer in the LSTM network

structure; thus, the overfitting problem can be alleviated to some extent [21]. The conceptual diagram of the dropout method is shown in **Fig. 8**. During the training process, for a certain layer of the neural network in one iteration, some neurons are first randomly hidden temporarily before this training and optimization. In the next iteration, some neurons continue to be hidden randomly and so on until the end of training. As a result of the random hiding, each iteration is training a different network with a relatively simple structure, and the forecasting result is the average of the output of each neural network.

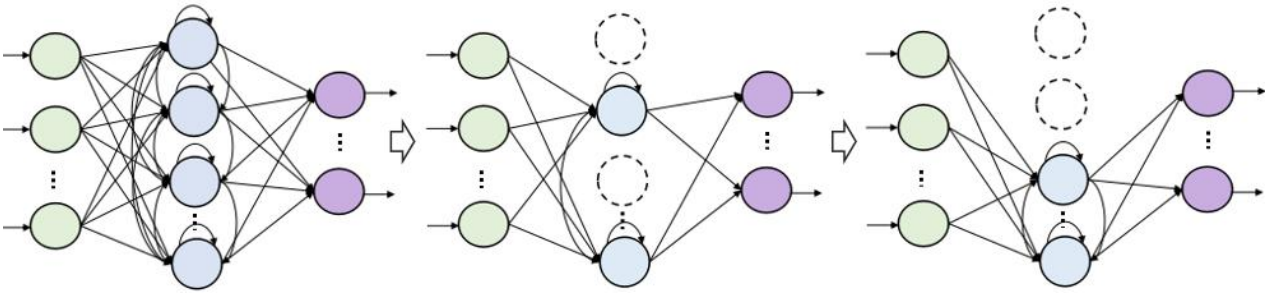


Fig. 8. Conceptual diagram of dropout algorithm.

GA optimizes the internal parameters of the LSTM (**Fig. 9**).

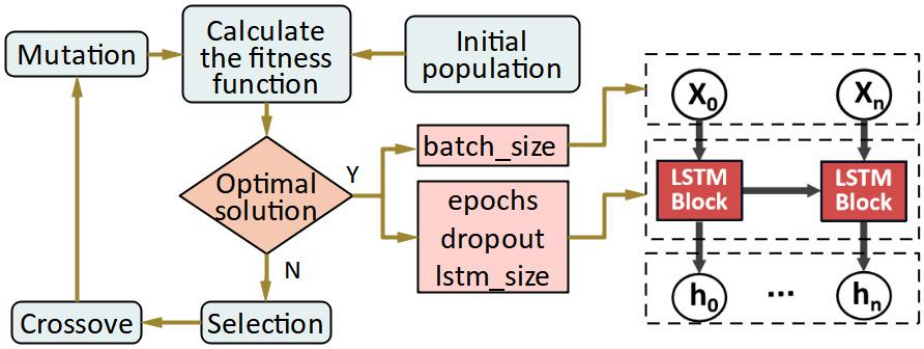


Fig. 9. Flowchart of GA optimization of LSTM.

GA is used to optimize the data time window step size (batch_size), the number of hidden units in the LSTM Network (lstm_size), the number of training times (epochs), and the dropout ratio (dropout). Four parameters are optimized in the search space to determine the best parameter combination.

The main process is presented as follows:

1) The population is initialized and decoded.
2) The mean square error of the LSTM neural network is taken as the fitness function.
3) The selected cross mutation operation is performed on the solved individual.
4) If the target value of the fitness function reaches the optimal value, then the next step is performed; otherwise, Step 3 is repeated.

5) The process is followed step by step to obtain the fitness target value and optimal parameters.
6) The forecasting mean square error is calculated based on the best parameters.
7) The termination conditions are judged as follows: If the number of population iterations is satisfied, then the calculation is stopped. At this time, the LSTM network determines the global optimal parameter combination [batch_size, lstm_size, epochs, dropout]; otherwise, Step 6 is repeated.

3. Forecast preparation

In this section, the specific implementation conditions of this research are described. These conditions mainly include the selection of forecasting locations and forecasting time periods, the introduction of using historical wind speed data, the preparation of forecasting models and the setting of related parameters, and the introduction of forecasting accuracy evaluation methods.

3.1. Data description

Hokkaido is located in the northernmost part of Japan. It is rich in solar, wind and biomass resources, and has the greatest potential for new energy utilization in Japan. The introduction rate of wind power generation in Hokkaido ranks first in Japan (approximately 55% and 29% for onshore and offshore wind power respectively), and in 2018, the total power generation of wind power plants was 441,185 kW, with 54 power plants and 329 wind turbines. Most of the power plants are concentrated in coastal areas with abundant wind resources, and their main distribution is shown in **Fig. 10**.

On the basis of the above analysis, 10 locations in Hokkaido (i.e., Wakkanai, Embetsu, Ishikari, Otaru, Esashi, Matsumae, Muroran, Hiroo, Atsutoko, and Nemuro) (**Fig. 10**) are selected as the target locations for forecasting. The wind speed data from a total of 141 locations in the Hokkaido area, including the target location and locations near the target location where wind speed data can be collected, published on the official website of the Japan Meteorological Agency are used as training and validation data to train the proposed forecasting system. The forecast target time period selected for this study is from January to December 2019. Thus, the actual wind speed data used in this study, including the training data, are wind speed data from 141 locations in Hokkaido from November 2018 to

December 2019, and the forecast system is evaluated in terms of forecast accuracy at three time intervals of 10, 30, and 60 minutes. The description of the forecast accuracy evaluation criteria will be given in Section 3.3.

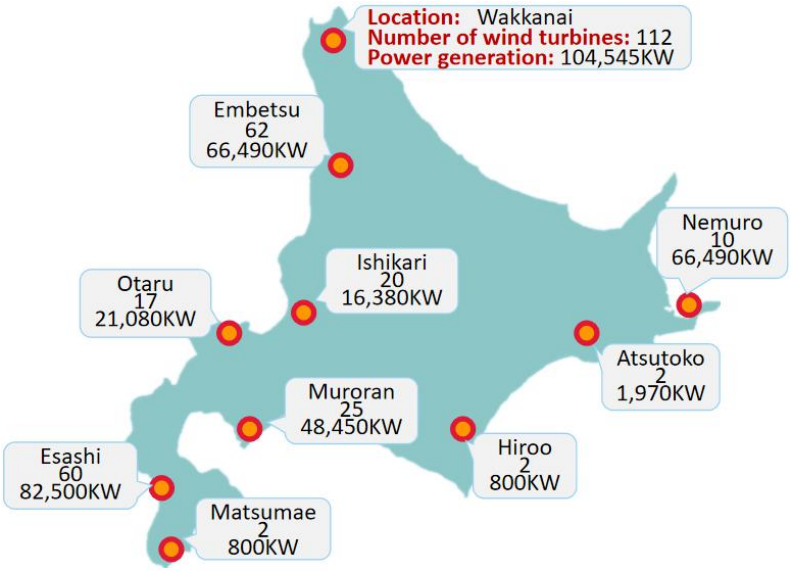


Fig. 10. Distribution of wind power plants in Hokkaido.

The example of Muroran is illustrated for the specific data selection method. As mentioned in Section 2.1, when Muroran is used as the target location for forecasting, the historical wind speed data for five locations—Muroran, Date, Noboribetsu, Suttsu, and Rankoshi—can be used as training data. The data selection principles for the forecasting of wind speeds from January 1 to 31, 2019 are shown in **Table 5**.

Table 5
Principles of data selection

Time interval	Training data				Validation data			
	Number of locations	Time Period	Number of days	Number of data	Number of locations	Time Period	Number of days	Number of data
10-min	5	2018.12.2~ 2018.12.31	30	21600	1	2019.1.1~ 2019.1.31	31	4464
30-min	5	2018.12.2~ 2018.12.31	30	7200	1	2019.1.1~ 2019.1.31	31	1488
60-min	5	2018.11.2~ 2018.12.31	60	7200	1	2019.1.1~ 2019.1.31	31	744

For forecasting the wind speed at each time interval, the training data and validation data are selected for the same time interval, and the number of days was set to 30 and 60 days because, as mentioned in our published paper [43], the amount of data should be as sufficient as possible considering the timeliness and training efficiency of the data; this number was finally determined by trial and error.

In accordance with the above-mentioned data selection principle, this paper selects the training and verification data of 10 locations to be forecasted, and the wind speed in 2019 is forecasted.

3.2. Parameter setting of the model

CEEMD: When CEEMD is used, two parameters need to be set: the ensemble number (*NE*) and the amplitude of the white noise (*Nstd*). In this paper, *NE* is initially set to 300 and *Nstd* to 0.3 based on experience. The value was adjusted after passing the forecast result. With Muroran taken as an example (**Table 6**), the mean absolute percentage error (MAPE) of wind speed forecasts at 10-minute intervals in January 2019 varies with the changes in the values of *NE* and *Nstd*.

Table 6
Forecasting accuracy when setting different *NE* and *Nstd*.

MAPE (%)	*Nstd*				
NE	0.1	0.2	0.3	0.4	0.5
100	8.06	7.87	7.27	7.15	7.71
200	8.21	7.62	7.02	7.23	7.34
300	7.92	7.37	7.01	7.02	7.31
400	7.12	6.91	6.11	6.18	6.75
500	6.94	6.92	6.07	6.14	6.68
600	7.23	7.26	6.68	6.58	7.53

NE is in the range of 400–500, and the best forecasting effect can be achieved when *Nstd* is 0.3 or 0.4. Therefore, the actual decomposition effect is combined, NE is set to 500, and *Nstd* is set to 0.3.

LSTM: The conceptual diagram of the LSTM structure is shown in **Fig. 8**, where input_size is the number of input layer units, lstm_size is the number of units in each hidden layer, and output_size is the number of output layer units. The input_size and output_size must first be determined. The principle of setting input_size was explained in detail in a paper previously published by our team [43]. On the basis of considering the continuity characteristics of the time series, the selection of data with higher correlation performance with the time series of wind speed to be forecasted was prioritized. On the basis of previous experiences, the wind speed that has a high correlation with the wind speed to be forecasted is usually the wind speed at 3 to 5 time points before the forecast time point [43]. With Muroran's 10-minute interval wind speed data taken as an example, let $Y(t)$ be the wind speed time series to be forecasted and $Y(t\text{-}i)$ be the wind speed time series before i time at the time point to be forecasted. The correlation between $Y(t)$ and

each time series is shown in **Table 7**. For the wind speed data at 10-minute intervals, the wind speed of the four time points before the forecast time point and the wind speed at the forecast time point have a relatively high correlation. Therefore, when the wind speed is forecast at 10-minute intervals, the input_size is set to 4. Similarly, the input_size for the 30-minute interval wind speed forecasting is set to 4, and the input_size for the 60-minute interval wind speed forecasting is set to 3. However, the proposed forecasting system does not forecast the original wind speed data directly and forecasts the decomposed components instead. Yet, the decomposed components still inherit the correlation between the original wind speed time series in each time interval. Thus, the previous method is still used to set it. Output_size is set to 1 because the purpose of this research is to forecast the wind speed at the next moment.

Table 7
Correlation coefficient between $Y(t)$ and wind speed time series at each time point.

Correlation coefficient	$Y(t\text{-}1)$	$Y(t\text{-}2)$	$Y(t\text{-}3)$	$Y(t\text{-}4)$	$Y(t\text{-}5)$	$Y(t\text{-}6)$
$Y(t)$	0.7154	0.6963	0.6253	0.5232	0.3123	0.2104

In the process of numerical simulation using LSTM, batch_size (dividing the data into several groups for training with the amount of data in each group), epochs (rounds of training the model with all data), dropout (inactivation probability of dropout layer), and lstm_size (the total number of units in the hidden layer) also need to be set. The setting of lstm_size will affect the accuracy of the forecasting model directly. The batch_size and epoch settings will affect the training efficiency and convergence of the model greatly because of the large number of training data samples. The function of dropout is to inactivate the input and recursive connections of LSTM units with a certain probability in the process of forward transfer and weight update. Setting the dropout probability value correctly can avoid overfitting and improve model performance effectively. Thus, GA is used to optimize the batch_size, epochs, dropout, and lstm_size of the forecasting model. The optimization process is described in the summary of Section 2.3, and the initial values are set as shown in **Table 8**.

Table 8
Initial parameter setting of LSTM.

Parameter	Initial Value
batch_size	100
epochs	50
dropout	0.2
lstm_size	50

3.3. Evaluation criteria of forecast performance

1) Wind speed forecasting accuracy evaluation indicators

To evaluate the forecasting results comprehensively from the perspective of accuracy and stability, several evaluation indicators are selected to compare the performance of competitive models. MAE, MAPE, and RMSE are used to evaluate the forecasting accuracy. MAE and RMSE can display the average deviation between the forecasted value and the actual value, and the use of MAE and RMSE can solve the problem of offsetting positive and negative deviations. MAPE is a common indicator and can be used to test the overall forecasting of this combined model effect. Equations (8) – (10) indicates the specific definition of the evaluation criteria. y_t and $\hat{y}_t$ represent the actual and forecasted value at time t, respectively. N represents the total number of samples.

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2} \tag{8}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|y_t - \hat{y}_t| \tag{9}$$

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100\% \tag{10}$$

To facilitate the accuracy comparison among different models, this paper uses three improvement percentage indicators, namely, P_{RMSE}, P_{MAE}, and P_{MAPE}. The calculation method is shown in Equations (11) – (13).

$$P_{RMSE} = \left|\frac{RMSE_1 - RMSE_2}{RMSE_2}\right| \times 100\% \tag{11}$$

$$P_{MAE} = \left|\frac{MAE_1 - MAE_2}{MAE_2}\right| \times 100\% \tag{12}$$

$$P_{MAPE} = \left|\frac{MAPE_1 - MAPE_2}{MAPE_2}\right| \times 100\% \tag{13}$$

2) Indicators of impact on wind power generation in the entire area

Wind speed is the most important factor that affects wind power generation. This paper mainly studies wind speed forecasting technology with the expectation that it can be applied to wind power generation forecasting in the future. Therefore, analyzing the impact of wind speed forecasting accuracy on the wind power of a single location or even the entire area is important. Such analysis determines whether the wind speed forecasted by this forecasting system has practical application value for the control of wind power generation. In this paper, the concept of wind power impact indicator for the entire area is presented on the basis of the above analysis. **Fig. 11** shows the power curve of the more

common wind turbines in Hokkaido. The cut-in wind speed is 2 m/s, the rated wind speed is 9 m/s, and the cut-out wind speed is 25 m/s. The relationship between wind speed and wind power generation when the wind speed is in the range of 2–9 m/s is shown in Equation (14), where C_p (=0.593) is the Betz coefficient, S is the contact area between the windmill blade and the air (m^2), ρ is the air density (kg/m^3), and V is the wind speed (m/s). Although different wind turbines have dissimilar contact areas S between wind turbine blades and air, the S of the wind turbines at the locations to be forecasted in this paper is assumed to be exactly the same to facilitate the evaluation. On this basis, the area wind power impact indicator K is proposed, as shown in Equation (15), and the calculation formula of the weight coefficient ω_i is shown in Equation (16), where P_i is the wind power generation at location i during the test period. M is the number of locations, k_i is the influence coefficient of location i that is calculated from Equation (17), where $\hat{y}_t$ is the forecasted wind speed, y_t is the actual wind speed at time t, and y_{rated} is the rated output speed (in **Fig. 11**, y_{rated}=9 m/s in this study), and the definition of $f(x)$ is shown in Equation (18).

The impact of wind speed forecasting accuracy on the forecast of wind power generation is analyzed by calculating the value of K for the entire area at different time periods to re-evaluate the forecasting system in terms of practicality.

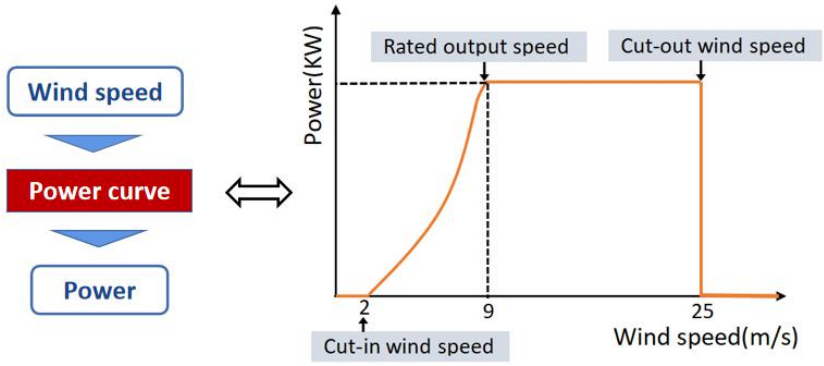


Fig. 11. Power curve of the more common wind turbines in Hokkaido.

$$P = \frac{1}{2} C_p (\rho S) v^3 \tag{14}$$

$$K = \sum_{i=1}^{M} \omega_i \cdot k_i \tag{15}$$

$$\omega_i = P_i / \sum_{i=1}^{M} P_i \tag{16}$$

$$k_i = |f(y_i) - f(\hat{y}_i)|/f(y_{rated}) \tag{17}$$

$$f(x) = \begin{cases} x^3 & 2 \le x < 9 \\ y_{rared}^3 & 9 \le x < 25 \\ 0 & x < 2, x > 25 \end{cases} \tag{18}$$

4. Experiments and analysis

For the analysis of the versatility and practicality of the proposed hybrid forecasting system, experiments are conducted to investigate the performance of the proposed forecasting system in different situations. For the comparison of the performance of the proposed forecasting system and other forecasting models, different types of forecasting models are selected and the same historical wind speed data are used as verification data. The experimental results and related analysis will be introduced in this section.

4.1 Experiment I: Forecast of annual wind speed in different locations

As mentioned in the summary of 3.1, the 10 locations in Hokkaido, namely, Wakkanai, Embetsu, Ishikari, Otaru, Esashi, Matsumae, Muroran, Hiroo, Atsutoko, and Nemuro, are used as the forecast target locations. The wind speed at 10-minute intervals is used as the forecast target to forecast the wind speed for the whole year of 2019. MAE, RMSE, and MAPE are calculated on the basis of the forecast results. The monthly average is computed as shown in **Table 9**. The comparisons of MAE, RMSE, and MAPE are shown in **Figs. 12–14**, respectively. **Figs. 12** and **13** show that the RMSE is basically maintained in the range of 0.2 to 0.6 in the annual wind speed forecast for the whole site. The maximum value of RMSE is approximately 0.6, which appears in the location forecasting of Locations 1 and 5 in January. From November to February, the RMSE value is relatively high, mainly concentrated in the range of 0.40 to 0.55, while the range of MAE is between 0.35 and 0.50, and the overall deviation is not obvious. The minimum RMSE value appeared in the forecasting for Location 6 in July. In other locations, except for Location 3, the RMSEs in July and August are also small, mainly concentrated in the range of 0.15 to 0.25. The range of MAE is between 0.10 and 0.30. Thus, the preliminary conclusion is that the stability of the forecasting model will be affected by the season and the geographic location of the forecasted location, but the overall impact is insignificant. The forecasting results obtained from the 10 locations are relatively similar. MAE shows a certain periodic

change, but not much difference is observed because of the change in location.

Fig. 14 shows that MAPE has obvious seasonal changes. Although certain differences are observed in the monthly averages of various locations, the overall performance suggests that the forecast results from July to September have the highest accuracy, and MAPE is in the range of 3.0% to 4.5%, whereas the forecast accuracy from April to June is unstable. MAPE is more common in the range of 3.0%–5.0%. The forecast results from November to March are relatively poor. MAPE is generally in the range of 4.0%–7.0%, with the maximum value reaching 7.33%.

Table 9
Forecast accuracy of different locations and different periods.

Locations		1	2	3	4	5	6	7	8	9	10
Jan	MAE	0.4585	0.4280	0.4554	0.4379	0.5195	0.4044	0.4831	0.5066	0.4343	0.4530
	RMSE	0.6233	0.5215	0.5493	0.5685	0.6145	0.5127	0.5434	0.5955	0.5137	0.5443
	MAPE	6.1160	5.5197	5.9051	5.7018	6.7265	5.6281	6.3686	6.5672	7.3334	5.7401
Feb	MAE	0.4195	0.4322	0.4428	0.4154	0.4160	0.4437	0.4425	0.3874	0.4201	0.3809
	RMSE	0.5697	0.5147	0.5732	0.5072	0.4850	0.5262	0.5213	0.4905	0.4487	0.4525
	MAPE	5.7734	5.6510	5.7722	5.3628	5.4349	5.7044	5.9386	5.1983	5.5680	5.8191
Mar	MAE	0.3397	0.3524	0.3563	0.3721	0.3776	0.4091	0.3666	0.3602	0.3299	0.3547
	RMSE	0.4224	0.4322	0.3933	0.4756	0.4461	0.4744	0.4525	0.4274	0.4167	0.4660
	MAPE	5.0368	5.9311	5.9994	6.3376	6.4091	4.6228	5.3917	6.1228	5.5742	5.6271
Apr	MAE	0.1771	0.2171	0.2424	0.2849	0.2387	0.1952	0.2885	0.2753	0.1932	0.1880
	RMSE	0.2227	0.2891	0.3238	0.3572	0.3137	0.2425	0.3024	0.3398	0.2723	0.2018
	MAPE	3.4580	3.7372	4.1021	4.8524	4.1009	4.0399	5.5705	5.3642	3.8030	4.5353
May	MAE	0.2497	0.2544	0.2421	0.2751	0.2782	0.2378	0.2523	0.2517	0.2552	0.2379
	RMSE	0.3109	0.3381	0.3340	0.3131	0.3179	0.2908	0.3372	0.3328	0.3321	0.3217
	MAPE	4.6770	4.4950	4.3270	4.7720	4.9402	4.5515	4.8709	4.8005	4.9569	4.5753
Jun	MAE	0.2178	0.2321	0.2462	0.2224	0.2263	0.2840	0.2577	0.2356	0.2424	0.2519
	RMSE	0.3281	0.3186	0.2981	0.3136	0.3116	0.3377	0.3196	0.3238	0.3228	0.3413
	MAPE	3.4168	4.2014	4.3916	4.0103	3.9997	3.5614	4.5829	4.2572	4.3195	4.4589
Jul	MAE	0.1606	0.1922	0.2551	0.1561	0.1847	0.1238	0.1380	0.1822	0.1601	0.1597
	RMSE	0.2081	0.2474	0.2942	0.1988	0.2311	0.1621	0.1739	0.2232	0.2259	0.2071
	MAPE	3.3331	4.0364	4.4927	3.2846	3.9551	3.3957	2.8916	3.7694	3.3308	4.0216
Aug	MAE	0.1701	0.1820	0.2943	0.1460	0.2243	0.2168	0.1992	0.1718	0.1534	0.1981
	RMSE	0.2187	0.2247	0.3156	0.1902	0.2550	0.2495	0.2273	0.2404	0.1957	0.2139
	MAPE	3.6372	4.0689	5.3377	3.2608	4.9829	3.7635	4.3800	3.7793	3.3321	3.9537
Sep	MAE	0.2393	0.2543	0.2363	0.2808	0.2453	0.2727	0.2719	0.2404	0.3028	0.2956
	RMSE	0.2571	0.3262	0.3568	0.3304	0.3207	0.3206	0.3619	0.2958	0.3542	0.3457
	MAPE	4.5054	4.3351	3.9605	4.6963	4.1705	4.5815	4.5817	4.1256	4.3090	4.5701
Oct	MAE	0.3070	0.2843	0.3018	0.3014	0.3108	0.2736	0.2934	0.2822	0.3433	0.2752
	RMSE	0.3988	0.3836	0.3812	0.3802	0.3253	0.3280	0.3486	0.3645	0.4168	0.3162
	MAPE	4.9011	4.8518	5.2358	5.1665	5.2966	4.7693	5.0125	4.8538	5.9785	4.7609
Nov	MAE	0.4949	0.4280	0.3976	0.4164	0.4995	0.4456	0.3592	0.3978	0.4698	0.3750
	RMSE	0.5314	0.5334	0.5605	0.5196	0.5703	0.5183	0.4122	0.4673	0.5521	0.4645
	MAPE	6.1775	6.4359	6.0433	6.2986	6.5360	6.1467	5.3917	6.9815	6.3727	5.6925
Dec	MAE	0.4141	0.4262	0.4456	0.4212	0.4216	0.5037	0.4151	0.4112	0.4192	0.3943

RMSE	0.4965	0.5351	0.5274	0.4932	0.5130	0.5469	0.4968	0.5151	0.5116	0.4801
MAPE	6.0393	6.4716	6.8401	6.5408	6.4071	6.0770	5.6745	6.2500	6.5137	6.2792

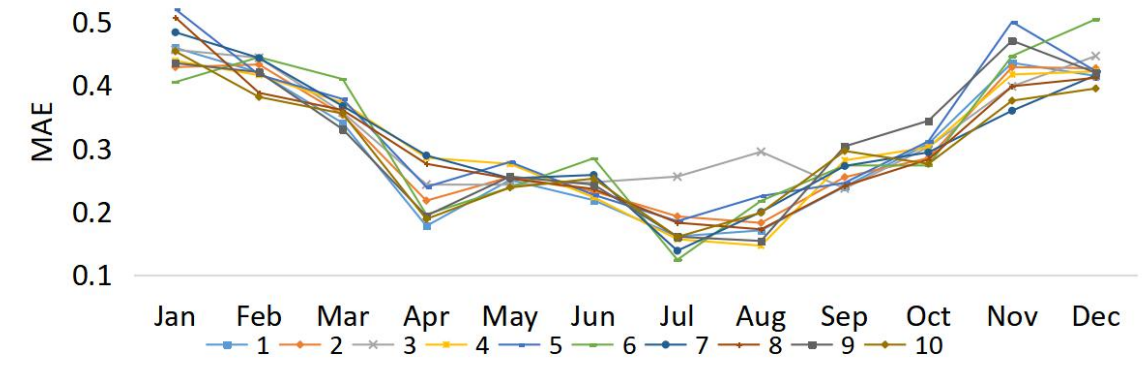


Fig. 12. Distribution of MAE in 10 locations in different periods.

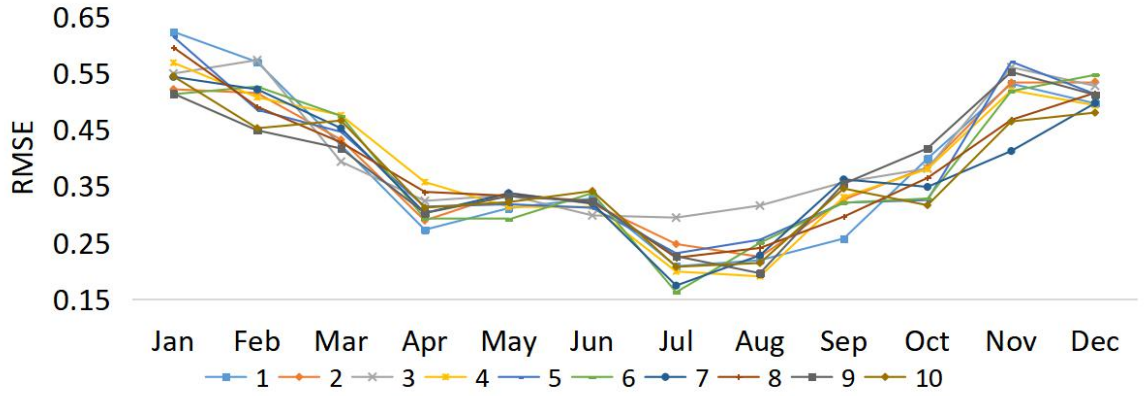


Fig. 13. Distribution of RMSE in 10 locations in different periods.

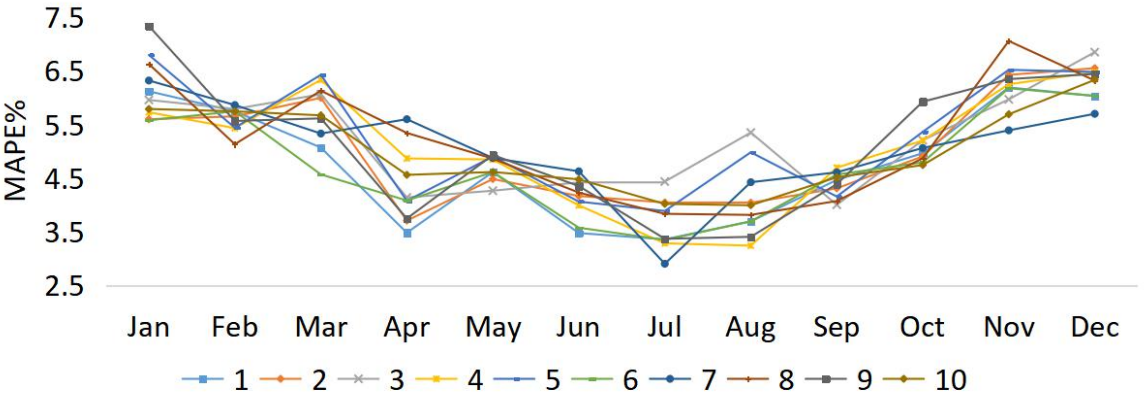


Fig. 14. Distribution of MAPE in 10 locations in different periods.

The monthly wind speed distribution is studied to investigate the cause of this phenomenon. With Location 1 (Wakkanai) taken as an example, the monthly instantaneous wind speed distribution in mid-2019 is organized into a box plot (**Fig. 15**). In Hokkaido,

the average wind speed from November to March is relatively high, and the range of wind speed variation is relatively wide. The wind speed is mostly concentrated in the range of 6–10 m/s, whereas the average wind speed from July to September is relatively low, and the wind speed changes smoothly. The proposed wind speed forecasting system has high stability and can deal with wind speed forecasting under different conditions better when concentrated near the average wind speed of 4–6 m/s. However, the specific forecasting conditions will be affected by the wind speed situation at the target location. The deviation range of MAPE is 3%–7%.

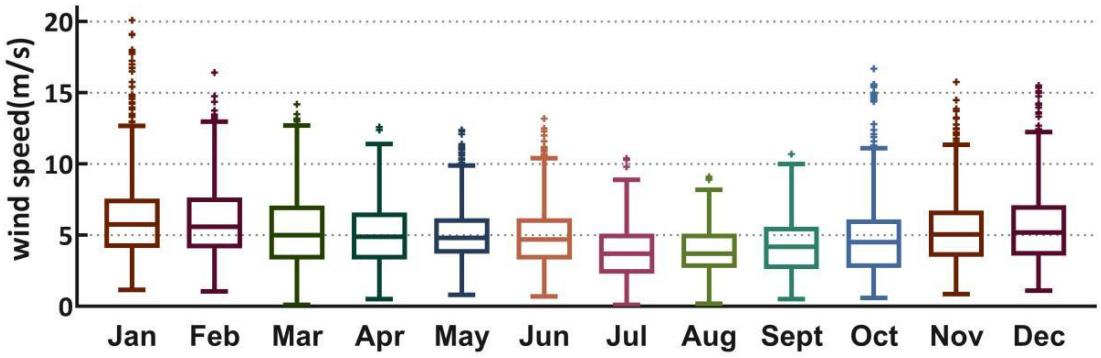


Fig. 15. Monthly instantaneous wind speed distribution of Wakkanai.

4.2 Experiment II: Accuracy comparison of different time intervals

In this section, the wind speed at the three time intervals of 10, 30, and 60 minutes is forecasted by using the data selection method in Section 3.1 to verify the accuracy of the proposed forecasting system in wind speed forecasting at different time intervals. Experiment 1 shows that the forecasting accuracy for 10-minute wind speed forecasting is greatly affected by wind conditions but is slightly affected by geographic location. Thus, only Wakkanai is selected as the forecast target location, and the selected forecast time periods are January, April, July, and October. The MAE and RMSE of the forecasted results are shown in **Table 10**. The comparisons of MAE and RMSE are shown in **Figs. 16** and **17**. **Fig. 18** shows a box plot chart of MAPE. As the time interval increases, the forecasting accuracy gradually decreases. Overall, the forecasting accuracy of the 30-minute interval can still be maintained at a good level. The RMSE is relatively stable and maintained in the range of 0.3 to 0.7. The MAPE can be maintained in the range of 4% to 9%. Similar to the forecasting results at a 10-minute interval, in July, when the wind speed is relatively stable, a good forecasting accuracy is achieved, and the minimum MAPE can reach 4.6%. In January, when the wind speed changes sharply, the forecasting

accuracy is relatively poor but still acceptable. The MAPE is maintained in the range of 6% to 9%. The forecasting accuracy of the 60-minute wind speed is significantly reduced, and the RMSE is in the range of 0.5 to 0.8, with relatively large fluctuations. The MAPE is in the range of 6%–12%, and the highest is 11.6%. However, most cases are concentrated around 8%. The forecasting accuracy needs to be improved, but it is still within an acceptable range.

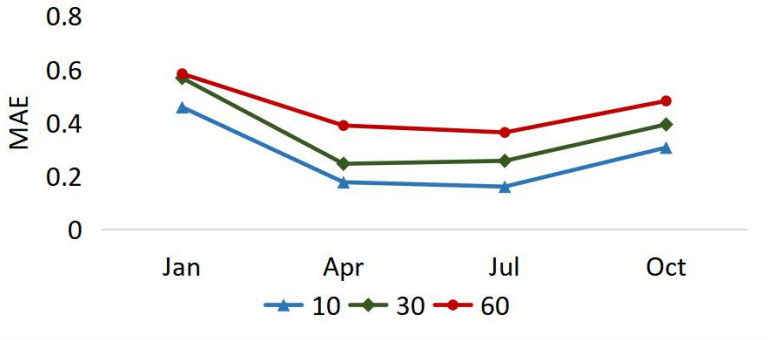


Fig. 16. Distribution of MAE of different time intervals.

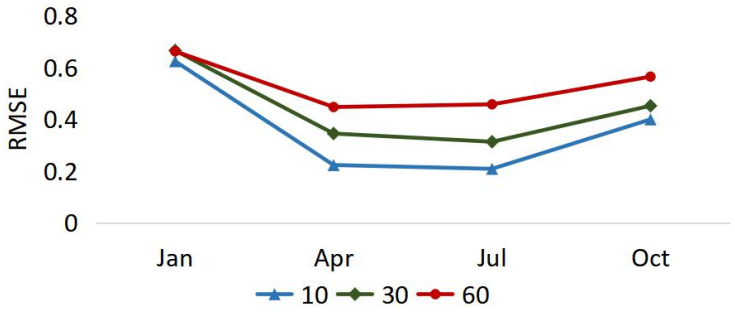


Fig. 17. Distribution of RMSE of different time intervals.

Table 10
Forecast accuracy at different time intervals.

	Jan			Apr			Jul			Oct		
	10	30	60	10	30	60	10	30	60	10	30	60
MAE	0.4585	0.5675	0.5836	0.1771	0.2462	0.3897	0.1606	0.2572	0.3637	0.3070	0.3936	0.4818
RMSE	0.6233	0.6647	0.6617	0.2227	0.3443	0.4466	0.2081	0.3125	0.4569	0.3988	0.4514	0.5638

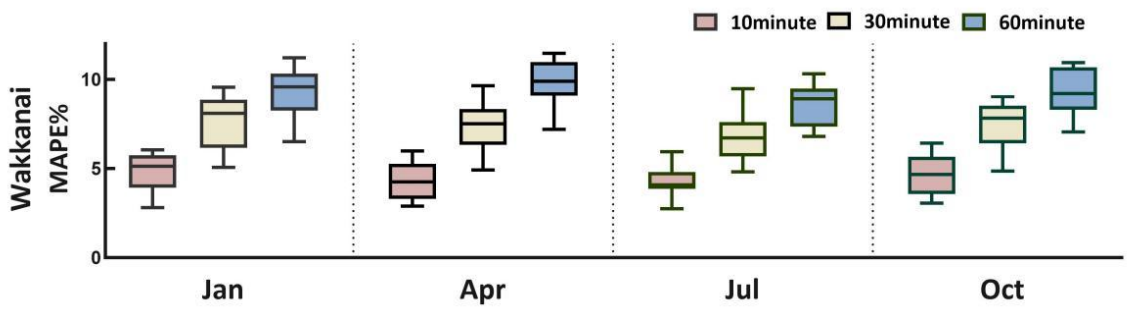


Fig. 18. Comparison of MAPE at different time intervals.

4.3 Experiment III: Accuracy comparison of different forecasting models

In this section, several typical single forecasting models, such as ARIMA, BPNN [20], Elman [39], ELM [40], and the widely used hybrid model EMD-RNN EEMD-LSTM, EWT-LSTM-Elman [41], and CEEMDAN-ENN [42], are selected.

The same historical wind speed data are used to train these forecasting models. For the proposed forecasting system, the DAD method introduced in Section 2.1 is used to extract data and train the forecasting model.

With Wakkanai taken as the forecasted target location and the wind speed at 10-minute intervals used as the forecast target, the wind speeds in January, April, July, and October 2019 are forecasted. **Table 11** and **Fig. 19** show the comparison of the typical single forecasting model results and the proposed model, and **Table 12** and **Fig. 20** show the comparison of hybrid forecasting model results. The results show the following:

Table 11
Comparison of forecasting accuracy with typical single model.

	Jan			Apr			Jul			Oct		
	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE
ARIMA	0.6622	9.2890	0.9321	0.4318	9.4578	0.5933	0.3903	9.6159	0.4491	0.5353	9.6440	0.8087
BPNN	0.5769	8.1650	0.6616	0.3707	7.2106	0.5926	0.3121	7.8228	0.5509	0.3838	8.8544	0.5716
Elman	0.7249	9.1513	0.8186	0.4293	8.7521	0.6176	0.3047	8.2564	0.4405	0.4908	8.8776	0.6110
ELM	0.6625	8.1591	0.7840	0.4521	8.2349	0.5292	0.3668	8.2442	0.4381	0.4795	8.7584	0.6929
Proposed model	**0.4585**	**6.1188**	**0.6233**	**0.1771**	**3.4817**	**0.2227**	**0.1606**	**3.5519**	**0.2081**	**0.3070**	**4.8564**	**0.3988**

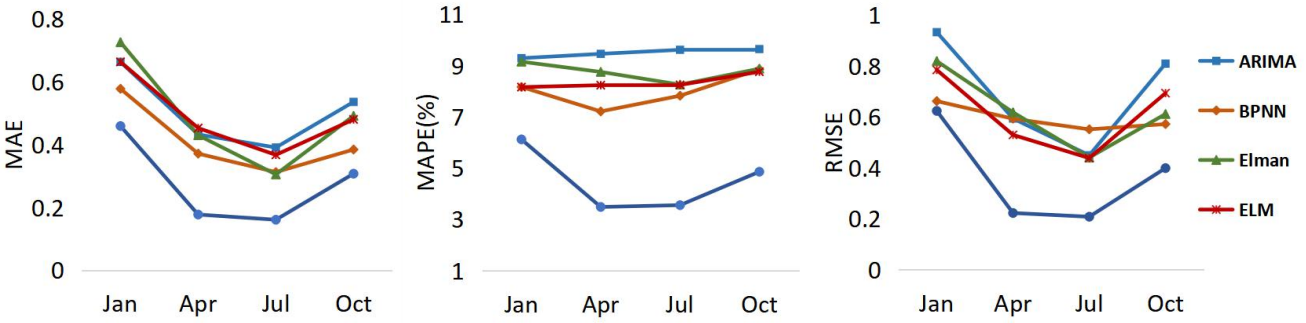


Fig. 19. Comparison of forecasting accuracy with typical single model.

Table 12
Comparison of forecasting accuracy with typical hybrid forecasting model.

	Jan			Apr			Jul			Oct		
	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE
EMD-RNN	0.6149	7.9848	0.7782	0.2829	6.6105	0.4011	0.2306	5.0210	0.4626	0.3709	6.6625	0.4425
EEMD-LSTM	0.5448	7.7728	0.7577	0.2970	6.9082	0.4155	0.2670	5.9450	0.3555	0.3598	7.1698	0.5262
EWT-LSTM-Elman	0.6307	7.6084	0.7764	0.2829	6.9394	0.4887	0.2418	5.2369	0.2904	0.4047	7.7360	0.5135
CEEMDAN-ENN	0.5711	7.1008	0.7072	0.2916	6.9850	0.5350	0.2397	4.7566	0.4264	0.3999	7.4260	0.5758
Proposed model	**0.4585**	**6.1188**	**0.6233**	**0.1771**	**3.4817**	**0.2227**	**0.1606**	**3.5519**	**0.2081**	**0.3070**	**4.8564**	**0.3988**

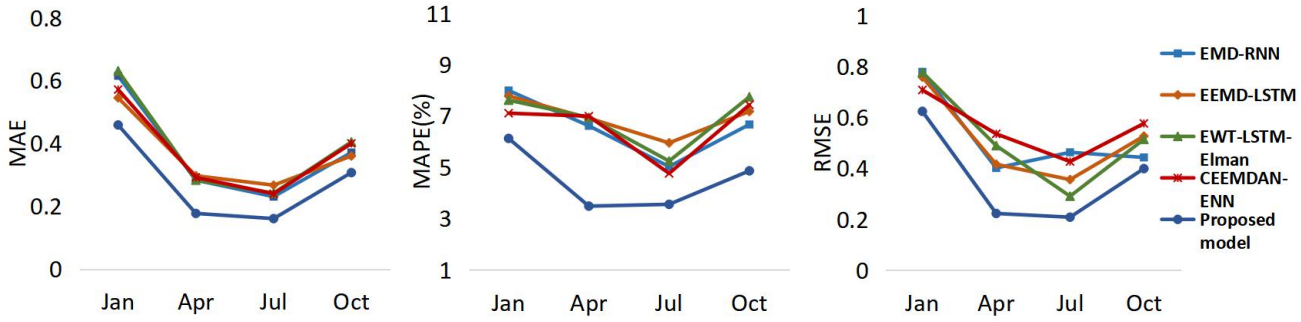


Fig. 20. Comparison of forecasting accuracy with typical hybrid forecasting model

(1) The classic single forecasting model has relatively poor versatility, and the forecasting accuracy varies greatly in different locations. For Wakkanai, which is the forecasted target location selected in this section, the fluctuation range of MAPE is not large and can basically be maintained in the range of 7%–9%. However, the overall accuracy is low. For forecasts in different seasons, January has the lowest accuracy, whereas July has the highest accuracy. From the perspective of RMSE, the stability of the classic single forecasting model is generally low, and RMSE fluctuates in the range of 0.5 to 1.0 with a large range of change. Unlike the classic single forecasting model, the proposed forecasting system reaches the optimal value of MAE, MAPE, and RMSE in different periods.

(2) The overall forecasting accuracy of the hybrid forecasting model is higher than that of the single forecasting model, and the versatility of the model is relatively good. The results in **Table 12** show that the RMSE remains in the range of 0.4 to 0.8, and the stability is higher than that of the single model. However, due to the poor wind speed decomposition performance of the EMD-RNN model, the EWT-LSTM-Elman model cannot deal with high-frequency components well when the Elman neural network model is used. Thus, the forecasting accuracy is low, especially in January when the wind speed changes frequently and the average wind speed is high, and the MAPE of both models is in the range of 7% to 8%. By contrast, the proposed forecasting system has the highest stability and is least affected by external conditions. Although the forecasting accuracy of the CEEMDAN-ENN model can reach a high level under certain circumstances, the minimum MAPE is 4.7%, but compared with the proposed model, the RMSE is relatively divergent, and the stability is not as good as that of the proposed forecasting system. In the location forecast for Wakkanai, due to insufficient wind speed data published by the Meteorological Agency, the wind speed time series appears discontinuous at some time points, and the average wind speed of Wakkanai is higher than those of other locations, which resulted in lower forecasting accuracy than

others. However, the forecasting accuracy is not significantly reduced when the proposed forecasting system is used, thus proving that the proposed DAD method is effective and has an obvious effect on improving the stability of the forecasting model.

To facilitate the observation of how much the forecasting accuracy of the proposed forecasting system has improved compared with that of other models, the forecasting result of Wakkanai is taken as an example and P_{MAE}, P_{MAPE}, and P_{RMSE} are calculated, as shown in **Fig. 21**.

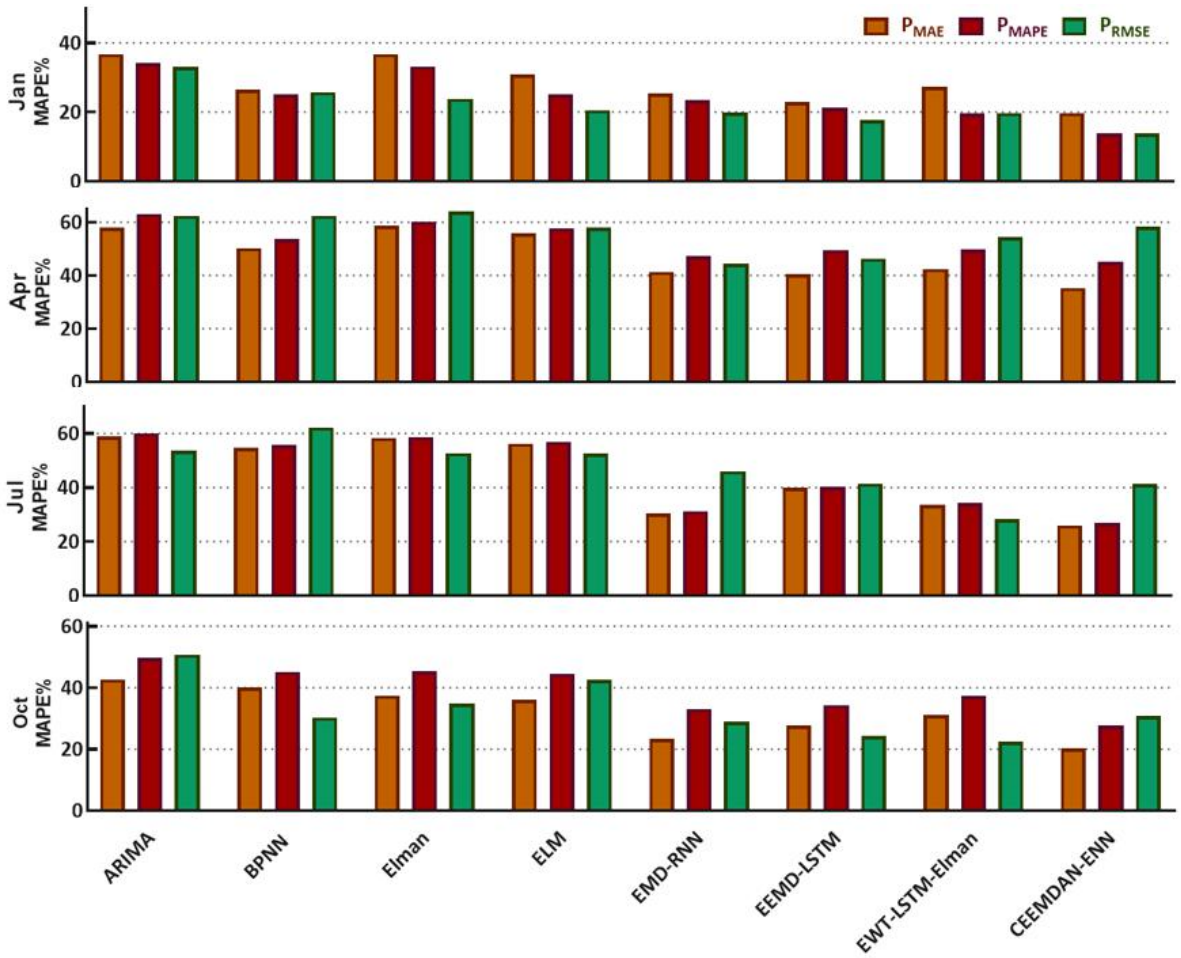


Fig. 21. Improved accuracy compared with other forecasting models.

(1) Compared with a single model, the proposed forecasting system has a greater improvement in P_{MAE}, P_{MAPE}, and P_{RMSE}. For example, in the wind speed forecast in January, for the ARIMA model, P_{MAE}, P_{MAPE}, and P_{RMSE} reached 36.76%, 34.12%, and 33.12%, respectively. For the BPNN model, P_{MAE}, P_{MAPE}, and P_{RMSE} reached 26.52%, 25.06%, and 25.78%, respectively.

(2) For other hybrid models, CEEMDAN-ENN has the highest forecasting accuracy, but the proposed forecasting system still improved compared with CEEMDAN-ENN. For example, in the forecast of wind speed in July, the values of P_{MAE}, P_{MAPE}, and P_{RMSE}

are 25.99%, 26.82%, and 38.19%, respectively. The values of P_{MAE}, P_{MAPE}, and P_{RMSE} in the forecast of wind speed in January reached 19.71%, 13.82%, and 13.86%, respectively, indicating that the proposed method has better forecasting ability and versatility than the other hybrid forecasting models.

(3) Compared with other models, the proposed combined model improves the forecasting accuracy to a large extent, because the percentage of improvement in each season is more obvious than that of all comparison models.

4.4 Impact of forecasting accuracy on wind power generation in the entire area

In Section 3.3, the concept of wind power impact indicator *K* is proposed. For the entire Hokkaido area, the impact indicator *K* at 10-minute intervals is calculated by using Formula (9), and the monthly average is calculated using the forecasting result of Experiment I in 4.1. The result is shown in **Fig. 22**. imilar to the change in the accuracy of wind speed forecasting, the impact on power generation also has obvious seasonality. From November to March, the wind is strong and changes frequently. At this time, the error is relatively large when the proposed forecasting system is used to forecast wind speed. *K* is in the range of (2%, 6%), and the difference between the maximum value and the minimum value is large, but most of them are concentrated in the range of (4%, 4.5%). When the proposed forecast model is used to forecast the wind speed from November to March, the forecast result will have an average deviation of 4% to 4.5% in the estimation of wind power generation.

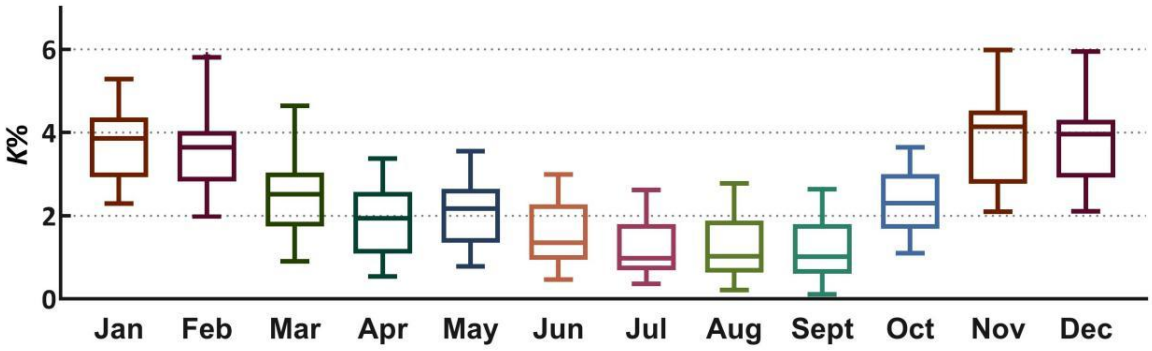


Fig. 22. Monthly distribution of wind power impact indicator *K*.

Similarly, from April to June, *K* is in the range of (1.5%, 3.5%), and it is mostly concentrated near the intermediate value. Therefore, the proposed forecast model is used to forecast the wind speed from April to June. The forecasting result will have a deviation of

approximately 2.5% for the evaluation of wind power generation.

For July to October, K is very small and will not diverge because it is mostly concentrated in the range of (0.5%, 1.5%). Therefore, the forecasting result will have a deviation of approximately 1% in the estimation of wind power generation when the proposed forecasting system is used to forecast the wind speed from July to October.

Although actual wind power forecasting is much more complex, more potential factors can affect power forecasting compared with wind speed forecasting. However, a deviation of 0.5% to 4.5% compared with the rated power will be observed only when the proposed wind speed forecasting system is used to forecast the wind speed and calculate the wind power generation in the power curve. The deviation value will change with the season, but an overall deviation value less than 5% indicates a small deviation, which also confirms that the proposed wind speed forecasting system is practical for forecasting wind power generation.

5. Conclusion

In this work, a novel hybrid forecasting system is proposed for short-term wind forecasting based only on the wind speed history data from the forecast location and the surrounding locations. The innovations of this forecasting system are reflected in (1) the proposed DAD method, which extracts a large amount of valid historical data from the surrounding locations to provide a reliable database for the training of the forecasting model, (2) the use of the decomposition and integration strategy in the data preprocessing stage to effectively reduce noise pollution in the original data, (3) the use of GA to optimize the important parameters in the LSTM network, thus improving the forecasting efficiency and forecasting accuracy of the model, and (4) the proposed concept of wind power impact indicators to analyze the deviation values of wind power generation forecasts in the Hokkaido region of Japan.

The numerical simulation results show that the accuracy of this system for short-term wind speed forecasting at different locations remains in a stable range for all periods of the year, where MAE and RMSE are basically maintained in the range of 0.1 m/s to 0.6 m/s, and MAPE is maintained in the range of 3% to 7%. For wind speed forecasting at different time intervals, this forecast system performs well in wind speed forecasting at 10-minute and 30-minute intervals with MAPE ranging from 3% to 9%, while it has an average wind speed forecasting performance at 60-minute intervals with MAPE ranging from 6% to 12%.

Compared with other forecasting models, the forecasting accuracy of this forecasting system has improved. In particular, the improvement of MAE, RMSE, and MAPE compared with that of the traditional single forecasting model is in the range of 26% to 40%, 25% to 35%, and 25% to 40%, respectively. Even compared with the hybrid forecasting model, the improvement of MAE, RMSE, and MAPE is in the range of 18%–28%, 14%–39%, and 13%–27%, respectively. The analysis of wind power impact indicators shows that the forecast deviation of the system is below 6% throughout the year for short-term wind power forecasting in the Hokkaido region, which is of high practical value.

The proposed forecasting system can provide short-term wind speed forecasting results with high accuracy and can therefore be considered for real-time control of wind turbines and load allocation planning. With regard to the generalizability of the forecasting system, this paper uses 10 locations in Hokkaido as forecasting targets, which proves to some extent that the forecasting system has good generalizability. However, the usability of the forecasting system has not been verified for countries or regions outside Hokkaido, Japan; such verification will be conducted in future studies.

Acknowledgments

The authors thank the power system laboratory of Hokkaido University for its support to our research. We thank Prof. Ryoichi Hara for his valuable opinions on research methods and data processing. We also thank Prof. Hiroyuki Kita for his constructive comments during the review process.

References

[1] Global Wind Energy Counci lhttp://gwec.net/global-figures/graphs/.
[2] Zhao J, Wang J, Su Z. Power generation and renewable potential in China. Renew Sustain Energy Rev 2014; 40:727–40. https://doi.org/10.1016/j.rser.2014.07.211.
[3] Wang H, Wang G, Li G, Peng J, Liu Y. Deep belief network based deterministic and probabilistic wind speed forecasting approach. Apply Energy 2016; 182:80–93. https://doi.org/10.1016/ j. apenergy.2016.08.108.
[4] Imal M, Sekkeli M, Yildiz C, Kececioglu F. Wind energy potential estimation and evaluation of electricity generation in Kahramanmaras, Turkey. Energy Educ Sc Technol Part A-Energy Sci Res 2012;30(1):661–72.
[5] Tian C, Hao Y, Hu J. A novel wind speed forecasting system based on hybrid data preprocessing and multi-objective optimization. Appl Energy 2018. https://doi.org/10.1016/j.apenergy. 2018.09.012.
[6] Allen DJ, Tomlin AS, Bale C, Skea A, Vosper S, Gallani ML, et al. A boundary layer scaling

technique for estimating near-surface wind energy using numerical weather prediction and wind map data. Appl Energy 2017. https://doi.org/10.1016/j.apenergy.2017.09.029.
[7] James EP, Benjamin SG, Marquis M. Offshore wind speed estimates from a high resolution rapidly updating numerical weather prediction model forecast dataset. Wind Energy 2018; 21:264–84. https://doi.org/10.1002/we.2161.
[8] Tom H, Peter C. Correction and downscaling of NWP wind speed forecasts. Meteorol Appl 2010; 14:105–16. https://doi.org/10.1002/met.12.
[9] Ouyang T, Zha X, Qin L. A combined multivariate model for wind power prediction. Energy Convers Manage 2017; 144:361–73. https://doi.org/10.1016/j.enconman.2017.04.077.
[10] Wang J, Hu J. A robust combination approach for short-term wind speed forecasting and analysis - Combination of the ARIMA (Autoregressive Integrated Moving Average), ELM (Extreme Learning Machine), SVM (Support Vector Machine) and LSSVM (Least Square SVM) forecasts using a GPR (Gaussian Process Regression) model. Energy 2015; 93:41–56. https://doi.org/10.1016/j.energy.2015.08.045.
[11] Radziukynas V, Klementavicius ABT-ISC on P and EE of RTU. Short-term wind speed forecasting with ARIMA model 2014:145–9. https://doi.org/10.1109/RTUCON.2014.6998223
[12] Kavasseri RG, Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. Renew Energy 2009; 34:1388–93. https://doi.org/10.1016/j.renene.2008.09.006.
[13] Yang Z, Wang J. A hybrid forecasting approach applied in wind speed forecasting based on a data processing strategy and an optimized artificial intelligence algorithm. Energy 2018; 160 :87-100.https://doi.org/10.1016/j.energy.2018.07.005.
[14] Cadenas E, Rivera W. Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model. Renew Energy 2010;35:2732 – 8. https://doi.org/10.1016/j.renene. 2010.04.022.

[15] Song J, Wang J, Lu H. A novel combined model based on advanced optimization algorithm for short-term wind speed forecasting. Appl Energy 2018; 215:643–58. https://doi.org/10.1016/j. apenergy.2018.02.070.
[16] Moreno SR, dos Santos Coelho L. Wind speed forecasting approach based on singular spectrum analysis and adaptive neuro fuzzy inference system. Renew Energy 2018; 126:736–54. https://doi.org/10.1016/j.renene.2017.11.089.
[17] Ramesh Babu N, Jagan Mohan B. Fault classifcation in power systems using EMD and SVM. Ain Shams Eng J 2015:1–9. https://doi.org/10.1016/j.asej.2015 .08.005. https://doi.org/10.1016/ j. asej.2015.08.005.
[18] Yang Z, Wang J. A combination forecasting approach applied in multistep wind speed forecasting based on a data processing strategy and an optimized artifcial intelligence algorithm. Appl Energy 2018; 230:1108–25. https://doi.org/10.1016/j.apenergy.2018.09.037.
[19] Marugán AP, Márquez FPG, Perez JMP, Ruiz-Hernández D. A survey of artifcial neural network in wind energy systems. Appl Energy 2018; 228:1822–36. https://doi.org/10.1016/ j. apenergy.2018.07.084.
[20] Chen L, Lai XBT-A-PP and EEC. Comparison between ARIMA and ANN models used in short-term wind speed forecasting 2011:1–4. https://doi.org/10.1109/APPEEC.2011.5748446.
[21] Aly HH Hamed. A proposed intelligent short-term load forecasting hybrid models of ANN, WNN and KF based on clustering techniques for smart grid. Int J Electric Power Syst Res 2020. https://doi.org/10.1016/j.epsr.2019.106191.
[22] Wang D, Luo H, Grunder Olivier, Lin Yanbing. Multi-step ahead wind speed forecasting using an improved wavelet neural network combining variational mode decomposition and phase space reconstruction. Renew Energy 2017; 113:1345–58. https:// doi.org/ 10.1016/ j. renene.2017.06.095.
[23] Shukur OB, Lee MH. Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA. Renew Energy 2015; 76:637–47. https://doi.org/10.1016/j.renene.2014.11.084.
[24] Wang L, Wang Z, Qu H, Liu S. Optimal forecast combination based on neural networks for time series forecasting. Appl Soft Comput 2018; 66:1–17. https://doi.org/ 10.1016/j.asoc. 2018.02.004.
[25] Zhang C, Wei H, Zhao J, et al. Short-term wind speed forecasting using empirical mode decomposition and feature selection. Renew Energy 2016; 96:727–37. https://doi.org/10.1016/ j. renene.2016.05.023
[26] Zhang K, Qu Z, Wang J, Zhang W, Yang F. A novel hybrid approach based on cuckoo search optimization algorithm for short-term wind speed forecasting. Environ Prog Sustain Energy

2017;36. https://doi.org/10.1002/ep.12533.
[27] Liu H, Tian H, Liang X, Li Y. Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks. Appl Energy 2015; 157:183–94. https://doi.org/10.1016/ j. apenergy.2015.08.014.
[28] Liu D, Niu D, Wang H, Fan L. Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. Renew Energy 2014;62:592‑7.https://doi.org/10.1016/j.renene.2013.08.011.
[29] Wang S, Zhang N, Wu L, Wang Y. Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method. https://doi.org/10.1016/ j. renene.2016.03.103.
[30] Liu H, Tian H, Li Y. An EMD-recursive ARIMA method to predict wind speed for railway strong wind warning system. J Wind Eng Ind Aerodyn 2015; 141:27–38. https://doi.org/10.1016/ j. jweia.2015.02.004.
[31] Zhang W, Qu Z, Zhang K, Mao W, Ma Y, Fan X. A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. Energy Convers Manag 2017;136:439‑51.https://doi.org/10.1016/j.enconman.2017.01.022.
[32] Wang H, Li G, Wang G, Peng J, Jiang H, Liu Y. Deep learning based ensemble approach for probabilistic wind power forecasting. Appl Energy 2017;188:56‑70. https://doi.org/10.1016/ j. apenergy.2016.11.111
[33] Liu H, Chen C, Tian H, et al. A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. Renew Energy 2012;48(6):545‑56. https://doi.org/ 10.1016/ j. renene.2012.06.012.
[34] Sun N, Zhou J, Chen L, Jia B, Tayyab M, Ping T. An adaptive dynamic short-term wind speed forecasting model using secondary decomposition and an improved regularized extreme learning machine. Energy 2018;165:939‑57. https://doi.org/10.1016/j.energy.2018.09.180
[35] Niazy R, Beckmann C, Brady J, Smith S. Performance evaluation of ensemble empirical mode decomposition. Adv Adapt Data Anal 2009; 01:231-242. https://doi.org/10.1142/ S179353690 9000102
[36] Zhao X, Wang C, Su J, Wang J. Research and application based on the swarm intelligence algorithm and artifcial intelligence for wind farm decision system. Renew Energy 2019;134:681‑97. https://doi.org/10.1016/j.renene.2018.11.061
[37] Yu R, Gao J, Yu M, Lu W, Xu T, Zhao M, et al. LSTM-EFG for wind power forecasting based on sequential correlation features. Futur Gener Comput Syst 2019;93:33‑42. https://doi.org/ 10.1016/ j. future.2018.09.054
[38] Zhang J, Yan J, Infeld D, Liu Y, Fue-sang L. Short-term forecasting and uncertainty analysis of wind turbine power based on long shortterm memory network and Gaussian mixture model. J Appl Energy 2019;241:229‑44. https://doi.org/10.1016/j.apenergy.2019.03.044.
[39] Liu H, Tian H, Liang X feng, Li Y fei. Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks. Appl Energy 2015;157:183‑94.Citehttps://doi.org/10.1016/ j. apenergy.2015.08.014.
[40] Salcedo-Sanz S, Pastor-Sánchez A, Prieto L, Blanco-Aguilera A, García-Herrera R. Feature selection in wind speed prediction systems based on a hybrid coral reefs optimization - Extreme learning machine approach. Energy Convers Manage 2014; 87:10–8. https://doi.org/10.1016/ j. enconman.2014.06.041
[41] Liu H, Mi X, Li Y. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. Energy Convers Manage 2018; 156:498-514. ttps://doi.org/10.1016/j.enconman. 2017.11.053
[42] Zhang W, Qu Z, Zhang K, Mao W, Ma Y, Fan X. A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. Energy Convers Manage 2017; 136:439–51. https://doi.org/10.1016/j.enconman.2017.01.022
[43] Liu Z, Hara R, Kita H. Wind Speed Forecast System using Recurrent Neural Network, Joint Technical Meeting on Power Engineering & Power Systems Engineering, IEE of Japan, PE-19–189, PSE-19 -201, pp.198–203,2019(in Japanese).