Hybrid Forecasting Method for Wind Power Integrating Spatial Correlation and Corrected Numerical Weather Prediction

Shuai Hu[1], Yue Xiang[1,*], Hongcai Zhang[2], Shanyi Xie[3], Jianhua Li[4], Chenghong Gu[5], Wei Sun[6], Junyong Liu[1]

[1] College of Electrical Engineering, Sichuan University, Chengdu 610065, China
[2]The State Key Laboratory of Internet of Things for Smart City and Department of Electrical and Computer Engineering, University of Macau, Macau 999078, China
[3]Electric Power Research Institute of Guangdong Power Grid Corporation, Guangzhou 510080, China
[4]Southwest Electric Power Design Institute Co., Ltd. Of China Power Engineering Consulting Group, 610021, China
[5] Department of Electronic and Electrical Engineering, University of Bath, Bath BA2 7AY, UK
[6] School of Engineering, University of Edinburgh, Edinburgh EH9 3DW, UK

*** corresponding author: xiang@scu.edu.cn**

Abstract Wind power generation rapidly grows worldwide with declining costs and the pursuit of decarbonised energy systems. However, the utilization of wind energy remains challenging due to its strong stochastic nature. Accurate wind power forecasting is one of the effective ways to address this problem. Meteorological data are generally regarded as critical inputs for wind power forecasting. However, the direct use of numerical weather prediction in forecasting may not provide a high degree of accuracy due to unavoidable uncertainties, particularly for areas with complex topography. This study proposes a hybrid short-term wind power forecasting method, which integrates the corrected numerical weather prediction and spatial correlation into a Gaussian process. First, the Gaussian process model is built using the optimal combination of different kernel functions. Then, a correction model for the wind speed is designed by using an automatic relevance determination algorithm to correct the errors in the primary numerical weather prediction. Moreover, the spatial correlation of wind speed series between neighbouring wind farms is extracted to complement the input data. Finally, the modified numerical weather prediction and spatial correlation are incorporated into the hybrid model to enable reliable forecasting. The actual data in East China are used to demonstrate its performance. In comparison with the basic Gaussian process, in different seasons, the forecasting accuracy is improved by 7.02%–29.7% by using additional corrected numerical weather prediction, by 0.65–10.23% after integrating with the spatial correlation, and by 10.88–37.49% through using the proposed hybrid method.

Keywords wind power forecasting; hybrid model; Gaussian process; numerical weather prediction; spatial correlation; kernel function

Nomenclature

Abbreviations

NWP	Numerical weather prediction
SC	Spatial correlation
GP	Gaussian process
AR	Autoregressive model
ANN	Artificial neural network
SVM	Support vector machine
ARD	Automatic relevance determination
SE	Squared exponential kernel function
RQ	Rational quadratic kernel function
Mat	Matern kernel function
RMSE	Root mean square error
MAPE	Mean absolute percentage error
MAE	Mean absolute error

Symbols

X	Input set
Y	Output set
$m(x)$	Mean function
$k(x,x^{'})$	Covariance function
K	Kernel function
I	Identity matrix
σ_f	Signal variance
l	Length
f_*	Forecasting value of the testing sets
y	Output of the training sets
$\overline{f}_*$	Mean value of GP based forecasting model
$\sum_*$	Variance of GP based forecasting model
D	Dimension of the input vector
θ	Hyperparameter
L	Length-scale hyperparameters
$v_{\mathrm{w}}(t)$	Wind speed of target wind farm at time t
$v_{\mathrm{mi}}(t)$	Wind speed of the neighbouring wind farms at time t
τ_t	Time delay between neighbouring wind farms
$L_{\mathrm{w,mi}}$	Distance between wind farm w and m_i
τ	Kendall's τ rank correlation coefficient
N	Number of wind speed samples
$\overline{v_{\mathrm{w}}}$	The mean wind speed of the target wind farm
$\overline{v_{\mathrm{mi}}}$	The mean wind speed of the neighbouring wind farms
P_m	Forecasting wind power by method m
ω_m	Weighted value by method m
e_m	Forecasting errors by method m
e_c	Forecasting errors by hybrid model
$\sigma(e_c)$	The variance of hybrid model

$e(x)$	Gaussian white noise	λ	Lagrange multiplier
σ_n^2	Variance of Gaussian white noise	$f(\omega,\lambda)$	Lagrange function
$g(x)$	Sequence-added Gaussian white noise	T	Number of wind power samples

1 Introduction

Renewable energy has received significant attention due to the exhaustion of fossil energy and the deterioration of the environment [1]. Many national and international policies mandate the significant increase of renewable shares in the generation mix to achieve the 1.5 °C global warming target. Wind energy is one of the popular renewable energy resources that has rapidly developed worldwide and projected to be one of the dominant energy sources in the future [2]. Reports from the Global Wind Energy Council showed that the installed capacity of wind turbines worldwide had reached 591 GW by the end of 2018, increasing by 51.3 GW [3]. However, due to its variability and intermittency nature [4], wind energy can cause severe issues in maintaining a secure and stable electricity supply [5]. One important solution to manage wind variability is to provide accurate wind power forecasting so that dispatchers can schedule countermeasures and adjust the maintenance plans in time [6].

Wind power forecasting techniques can be broadly categorised into statistical and physical methods. Conventional statistical methods mainly include time series models, such as the autoregressive integrated moving average model [7], autoregressive model (AR) [8], and the autoregressive moving average model [9]. These conventional methods can perform well for systems with linearised simplification. However, they face difficulties in providing sufficient accurate forecasting for wind energy with strong inherent nonlinearity. In recent years, artificial intelligence methods, which are theoretically suitable for revealing complex nonlinear relationships in historical data, are extensively applied to wind power forecasting [10]. Many methods, including artificial neural networks (ANNs) [11], support vector machine (SVM) [12], have been successfully adopted for improving wind power forecasting. Catalão et al. [13] proposed an ANN model for a 3 h ahead wind power forecasting that used wavelet transform to decompose the original wind power series. Liu et al. [14] developed a novel SVM-based forecasting model, which used the genetic algorithm to ensure the generalisation of the SVM; the wavelet transform was used to decompose wind speed into two components. Ren et al. [15] discussed a forecasting method based on support vector regression and ANN with

improved empirical mode decomposition to decompose the original wind series into simple ones. However, the above statistical models have limitations as large size training sets are needed [16]. For instance, the insufficient historical data of newly-built wind farms can increase the difficulty of training the wind power forecasting model. In general, statistical methods are suitable for short-term forecasting.

Another kind of statistical method is Spatial correlation (SC); compared with a conventional statistical method (e.g., time series models), the SC could consider the interaction between adjacent wind farms based on temporal correlations [17], which is helpful to improve the forecasting accuracy [18]. SC methods use the data of neighbouring wind farms to establish the wind resource model in the target wind farm [19]. Recently, the adoption of the SC method in wind speed and power forecasting has been extensively studied. Li et al. [20] described a dynamic SC model between geographically distributed wind farms for forecasting short-term wind power with the backtracking framework built by the Kalman filter. Zhu et al. [21] investigated methods to forecast wind speed in multiple sites by adding the SC model. The spatial features were extracted by the convolutional neural network and long short-term memory. In the dynamic SC method, the geographical location and terrain of wind farms were considered, and the forecast limitations caused by a rapid variation of wind speed could be overcome [22]. Tastu et al. [23] considered the spatio-temporal dependencies of wind farms and proposed a forecasting model of wind power errors. Khodayar et al. [24] presented a wind speed forecasting model which used the deep learning method to learn the spatio-temporal features of neighbouring wind farms and the results show the effectiveness of the proposed model. Zhao et al. [25] considered the increasing number of wind farms and their interdependencies. A spatio-temporal wind power forecasting framework is introduced which demonmstartes the accuracy and efficiency of method than other benchmark methods. Some studies also used the spatial and temporal covariance functions for modelling the spatio-temporal data by considering its asymmetry [26], and built the corresponding matrix to reflect the spatial correlation of wind data [27].

Physical methods are mainly based on numerical weather prediction (NWP) and consider the manufacturer's power curves [28], which could mitigate the missing historical data. Considering the physical descriptions of surface roughness and obstacles, fluid dynamics and thermodynamics are used to obtain the key information of physical models at the hub height of wind turbines by observing initial conditions. Then, the refined wind speed data are fed into the corresponding manufacturer wind power curve to obtain the forecasted wind power. Nielsen et al. [29] used

meteorological forecasting data from three different global meteorological models to obtain improved performance of wind power forecasting. Bessac et al. [30] developed a forecasting model that combined multiple sources of physical model outputs and achieved improved accuracy over several months. In contrast with statistical methods, physical methods do not rely on historical data to train the forecasting model. Jung et al. [31] pointed out that physical methods perform far better than statistical methods in long-term wind power forecasting. Nevertheless, the fluid dynamics have a high correlation with climate phenomena, and the NWP is affected by initial conditions. Wind power forecasting could have a significant error because the NWP is slowly updated and lags behind actual changes. Its applicability will be limited for short-term wind power forecasting due to the high computational complexity in solving NWP models.

Recently, hybrid methods that incorporate the practical aspects of different forecasting methods have stimulated considerable research interest [32]. The combination of statistical and physical methods has received significant interest, which can achieve better forecasting performance than independently [33]. Chen et al. [34] presented a hybrid wind power forecasting model up to 24 h in advance based on the Gaussian process (GP) and NWP model. Li et al. [35] developed a novel hybrid model integrating support vector machine with an improved dragonfly algorithm, whose effectiveness was confirmed by actual datasets from a wind farm in France. Zhou et al. [36] used a combination of extreme-point symmetric mode decomposition, extreme learning machine and particle swarm optimisation to create a wind forecasting model. Azimi et al. [37] proposed a novel time-based K-means clustering method that combined discrete wavelet transform and harmonic analytical time series models to accelerate the forecasting. Dhiman et al. [38] developed a wind speed and power forecasting model based on the different variants of support vector regression and wavelet transform. These hybrid models generally showed improved performance compared with the individual conventional models because the shortcomings of each embedded model could be systematically tackled.

Table 1 summarises some recently developed forecasting methods for wind power.

Table 1. Summary of the recent wind power forecasting methods

Authors	**Year**	**Method type**	**Method**	**Type of input data**	**Data source**	**Forecasting horizon**
Karaku et al. [8]	2017	Statistical method	Polynomial AR	Wind speed	Cesme and Bandon	24 h
Wan et al. [39]	2017	Statistical method	Extreme learning machine that trains	Wind power	Denmark	3 h

			single hidden layer feed-forward neural network			
Chang et al. [11]	2017	Statistical method	Radial basis function neural network	Wind speed and power	Taiwan	72 h
Lu et al. [40]	2018	Statistical method	Improved radial basis function neural network	Temperature and wind speed and power	Taiwan	24 h
Viet et al. [41]	2018	Statistical method	Artificial neural network	Temperature and wind speed and power	Vietnam	24 h
Shahid et al. [42]	2020	Statistical method	Wavelets that utilise long short-term memory paradigm	Wind speed and direction and zonal and meridional components	European	250 h
Hu et al. [43]	2019	Statistical method	Convolution-based spatial–temporal wind power forecasting model	Wind power	Australian	30 min
Chen et al. [44]	2020	Statistical method	SC method using the related three neighbouring stations	Wind speed and direction, temperature, temperature gradient, pressure and relative humidity	South Africa	6 h
Jiao et al. [45]	2018	Hybrid method	Combined autoencoders and backpropagation algorithm	Wind power	EirGrid	24 h
Zhao et al. [25]	2018	Hybrid method	Sparsity-controlled vector AR and spatiotemporal models	Wind power	Denmark	1.5 h
Shen et al. [46]	2018	Hybrid method	Combined empirical mode decomposition and random forest	Wind power	America	100 h
Li et al. [47]	2018	Hybrid method	SVM based on the cuckoo search algorithm	Wind speed and direction	Northwest of China	72 h
Ju et al. [48]	2019	Hybrid method	Combined convolution neural network, lightGBM and SC	Temperature, fan status, generated power, wind speed and direction, motor speed, daily power generation and pitch angle	Northwest of China	6 h
Zhang et al. [12]	2019	Hybrid method	Combined least squares SVM and deep belief network	Wind power	Jiangsu	33 h
Zhang et al. [49]	2019	Hybrid method	Combined neural network and grey model	NWP and wind speed and power	China	60 h

In summary, the above literature review indicates that time series methods, artificial intelligence methods, SC methods and hybrid methods have been extensively used for wind power forecasting. Many studies have directly used historical or NWP data. However, the data derived from NWP are biased with the actual data [50], how to reduce its inaccuracy is an important issue [51]. Some studies proposed approaches considering the inaccuracy of NWP [52], but many pieces of research still disregarded [53]. The literature review also shows that ANN is one of the widely used statistical methods for wind power forecasting. However, the conventional ANN model has inherent drawbacks of over-fitting and slow convergence speed. The GP model holds several advantages in terms of its well-founded framework to identify the relationship between input variables and target variables compared with ANN [54]. GP, an effective nonlinear, nonparametric and probabilistic prediction method, contains fewer parameters than other statistical models to simplify forecasting [55]. Moreover, the GP-based model is self-adaptive to gain hyperparameters and is flexible to implement [56]. However, the conventional GP model with one type of kernel function may have limited forecasting capability due to the strong variability of wind speed. If the optimal kernel function scheme is chosen from a combination of several kernel function types, more accurate forecasting results could be produced.

This study, based on the GP model, proposes a novel hybrid wind power forecasting model by combining NWP with the SC of wind farms. The rolling mechanism is a technology to constantly update training data of the forecasting model. The wind speed data from NWP are corrected by selecting the key factors from high-dimensional data by using the automatic relevance determination (ARD) method. The optimal hyperparameters in the GP model are obtained by maximizing log-likelihood estimation. A combination of different kernel functions is used to establish the optimal scheme for the GP model. The forecasting models are developed for different seasons to effectively represent the seasonal variation of wind speed. The meteorologically and spatially detailed hybrid model shows improved forecasting performance with few errors when the strengths of these modelling techniques are appropriately combined. The main contributions of this study include the following:

- A novel ARD is developed to improve the accuracy of the NWP data via selecting key factors, which are used as the input to the corrected NWP wind speed model. This correction model effectively improves the accuracy of wind speed and the overall performance of the forecasting model.

- An SC method is developed by using the data of neighbouring wind in adjacent areas of the target wind farm. This method considers the geographical location and terrain of wind farms.
- A novel hybrid forecasting method with a rolling mechanism is created based on the GP model with combined kernel functions. The weighted values of sub-models are determined by the Lagrange multiplier method. The meteorological and spatial factors are also comprehensively considered in the hybrid model to obtain improved forecasting results.

The rest of this paper is organised as follows. Section 2 introduces the basic forecasting model based on the GP with different kernel functions, the corrected NWP wind speed model and the SC between the reference and the target wind farms. Section 3 implements the hybrid forecasting method based on the corrected NWP data and the SC and describes the detailed modelling. Section 4 presents the case study and compares the forecasting accuracy of the hybrid model with conventional models. Finally, Section 5 concludes the paper.

2 Basic model and methodology

The research methodology is presented in two parts. The first part presents the basic forecasting model based on GP. The second part describes the corrected NWP data and SC method in addition to the basic model. The relevant theory associated with the methods is also introduced in this section.

2.1 Basic forecasting model based on GP

2.1.1 Standard GP

The GP model is a supervised learning method that does not initially restrict the relationship between the target and input variables to a specific form [56]. The GP model can also provide forecasted distributions rather than merely point forecasting [57]. Given the complex patterns and relationships between wind power and meteorological data, the GP model could be a promising method for wind power forecasting.

Assuming that an input set $X = \left\{ x_i \in R^D \mid i = 1, \ldots, n \right\}$ and an output set $Y = \left\{ y_i \mid i = 1, 2, \ldots, n \right\}$ will be used as the training set, a GP $f(x)$ is completely specified by the mean and covariance functions such that

$$f(x) \sim N\left(m(x), k(x, x') \right), \tag{1}$$

The mean function and covariance function are defined as follows:

$$m(x) = E[f(x)], \tag{2}$$

$$k(x, x') = E[(f(x) - m(x))(f(x') - m(x'))]]. \tag{3}$$

Supposing that the training set d consists of n observations, $d = \{(x_i, y_i) \mid i = 1, ..., n\}$, the whole input matrix is a $n \times D$ matrix. Gaussian white noises are present, and the noise has zero mean and unit variance formulated as follows:

$$e(x) \sim N\left(0, \sigma_n{}^2\right), \tag{4}$$

where $\sigma_n{}^2$ is the variance of Gaussian white noise. The novel sequence-added Gaussian white noise is expressed as follows:

$$g(x) = f(x) + e(x). \tag{5}$$

Thus, $g(x)$ also obeys the Gaussian distribution, and it can be modelled as follows:

$$g(x) \sim N\left(m(x), K(X, X) + \sigma_n{}^2 I\right), \tag{6}$$

where K is the covariance matrix with elements $K_{ij} = k(x_i, x_j)$, $i, j = 1, 2, ..., n$. k is also called kernel function. To simplify the process of hyperparameters optimization, the mean function $m(x)$ is commonly set to zero, the covariance matrix considered the noise is represented by $K(X, X) + \sigma_n{}^2 I$, and I presents an identity matrix. The following four types of kernel functions [58] are considered in this study:

1) Squared exponential (SE) kernel function:

$$k_{\mathrm{SE}}(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{1}{2}\left\|x_i - x_j\right\|^2 \Big/ l^2\right) \tag{7}$$

2) Rational quadratic (RQ) kernel function:

$$k_{\mathrm{RQ}}(x_i, x_j) = \sigma_f^2 \left[1 + \frac{\left\|x_i - x_j\right\|^2}{2al^2}\right]^{-a} \tag{8}$$

3) Matern (Mat) kernel function ($v = 3/2$):

$$k_{\mathrm{Mat}_{v=3/2}}(x_i, x_j) = \sigma_f^2 \left[\left(1 + \frac{\sqrt{3}\left|x_i - x_j\right|}{l}\right) \exp\left(-\frac{\sqrt{3}\left|x_i - x_j\right|}{l}\right)\right] \tag{9}$$

4) Mat kernel function ($v = 5/2$):

$$k_{\mathrm{Mat}_{v=5/2}}(x_i, x_j) = \sigma_f^2 \left[\left(1 + \frac{\sqrt{5}|x_i - x_j|}{l} + \frac{5\|x_i - x_j\|^2}{3l^2} \right) \times \exp\left(-\frac{\sqrt{5}|x_i - x_j|}{l} \right) \right]. \quad (10)$$

where σ_f represents the signal variance, and l denotes the length-scale parameter. Conventional GP models use one type of kernel function, which may have certain issues in wind power forecasting, such as poor robustness and generalization performance. A combined kernel function is created in the GP model to integrate the advantages of different kernel functions.

The joint prior distribution of the forecasting value f_* of the testing sets and the output y of the training sets is formulated as follows:

$$\begin{bmatrix} y \\ f_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(X,X) + \sigma_n^2 I & k(X, x_*) \\ k(x_*, X) & k(x_*, x_*) \end{bmatrix} \right), \quad (11)$$

where x_* is a new input, and $k(X, x_*)$ is abbreviated as k_*, $k(X, x_*) = k(x_*, X)^T = [k(x_1, x_*), ..., k(x_n, x_*)]$. The mean value $\overline{f_*}$ and variance $\sum_*$ can be expressed as follows:

$$\overline{f_*} = k_*^T (K + \sigma_n^2 I)^{-1} y, \quad (12)$$

$$\sum\nolimits_* = k(x_*, x_*) - k_*^T (K + \sigma_n^2 I)^{-1} k_*. \quad (13)$$

2.1.2 ARD

The GP model is fully defined by its mean and covariance functions. In the conducted experiments, no fixed covariance function is present in the GP. By contrast, a parametric function is used, and the parameters are inferred by the observed values [56]. The process of inferring the parametric values, which is called learning hyperparameters, is completed by maximizing the log-likelihood function:

$$\ln p(f_* \mid \theta) = -\frac{1}{2} \ln |K| - \frac{1}{2} f_*^T K^{-1} f_* - \frac{n}{2} \ln(2\pi). \quad (14)$$

A standard nonlinear gradient optimisation method is used for maximizing this function. The separate lengthscale parameters can be combined with every input variable to extend this technique, and the relative importance of each input can be inferred from the observed values. The ARD can be used to accurately obtain the relative importance by its length-scale hyperparameters, which is formulated as follows:

$$k_{ARD}(x, x') = \theta_0 \left(1 + \sum_{i=1}^{D} l_i (x_i - x_i')^2 \right)^{-v} + b, \quad (15)$$

where D represents the dimension of the input vector, and b is the deviation. All the hyperparameters about ARD are contained in the vector $\theta=(\theta_0,L,b)^T$. The ARD method can be implemented through length-scale hyperparameters $L=\{l_1,\ldots,l_D\}$. The input variable becomes highly sensitive, and the corresponding importance is enhanced because the length l_i is short.

The wind power forecasting framework based on GP includes the following three steps:

Step 1: Select the inputs for the GP model by using the historical records and NWP data, such as wind speed, wind direction and temperature. The training and testing parts of the data are also determined.

Step 2: Establish an adequate covariance function for the given dataset. The optimal kernel function scheme can be obtained by combining different kernel functions.

Step 3: Obtain the hyperparameters of the GP model by maximizing the log-likelihood function of the GP model. Thereafter, the wind power in the testing sets is forecasted by the well-trained GP model, and the forecasted results are obtained.

2.2 Wind speed correction model

If wind power forecasting entirely relies on historical data of the target wind farm, then the time horizon of the forecasting model is generally less than 12 h. The forecasting horizon can be increased by using NWP data [59]. The NWP model simulates the weather condition by solving the mathematical models of the atmosphere with powerful computers. The NWP data (e.g. wind speed and direction, temperature, humidity and air pressure) with the high spatial resolution is needed for high forecasting accuracy due to the high spatial variability of wind resources. However, the uncertainties from the model initialisation of NWP is unavoidable, and its accuracy is also affected by the measuring technique. Many wind farms are located in remote regions with rich wind resources. However, these farms have some discrepancies with the location of stations that provide NWP data. Therefore, NWP data are not always consistent with actual values, and the existing error of NWP data cannot be neglected.

Considering that the little error of NWP data could cause a huge error of forecasted wind power, an NWP wind speed correction model is established by using the NWP data and actual values to improve the forecasting performance. Accordingly, the performance of wind power forecasting is improved. The NWP wind speed correction model consists of the following main steps:

Step 1: Determining statistically significant inputs. A set of ARD results is performed to examine the correlation between the wind speed correction model and the input factors. The factors that are relevant to the wind speed correction model are selected by the ARD model. These factors are used as inputs for the correction model.

Step 2: The NWP wind speed and the selected inputs from the first step are used as the inputs of the wind speed correction model, and the corrected wind speed is obtained based on the GP.

2.3 SC model

In geographically distributed wind farms, the wind speed in the target wind farm and neighbouring wind farms could be strongly correlated. Given that some wind farms are located in regions with complex topography, the mutation of wind speed is difficult to capture when forecasting is only based on the data of the target wind farm [25]. When using SC method, the data of the observation sites in different directions are required. However, the difficulty of data collection in different sites limits the application of SC. Assuming that the region is located in a flat terrain, a spatial translation for these sites could be adopted to address this problem. As the example shown in Fig. 1, the selected observation site in the target wind farm w and those in neighbouring wind farms m_1, m_2 and m_3 are not in a direct line. Accordingly, a line is drawn through the target wind farm w, which locates between neighbouring wind farms m_1, m_2 and m_3. Vertical lines are drawn through wind farms m_1, m_2 and m_3, and intersection points m_1', m_2' and m_3' are obtained. When the distance between wind farms and the corresponding intersection point is short, the data of the intersection points can be used as the data of the neighbouring wind farms. Fig. 2 illustrates the wind speed series in neighbouring wind farms m_1, m_2 m_3 and in target wind farm w.

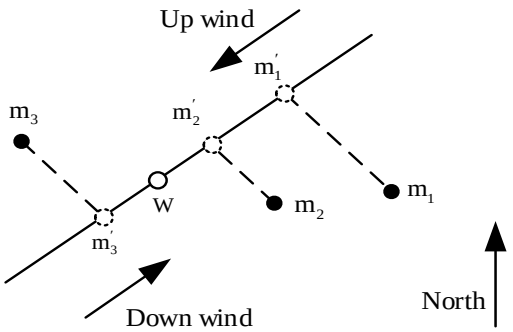


Fig. 1. Spatial translation of wind farms

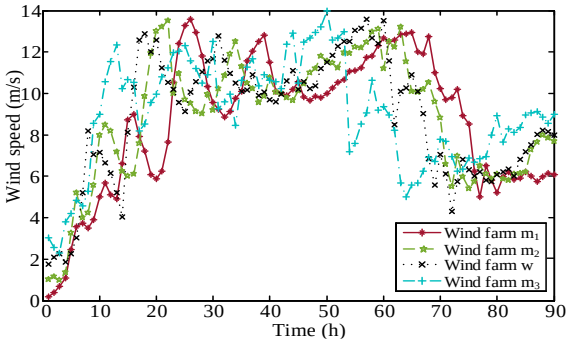


Fig. 2. Wind speed series of m_1, m_2, w and m_3

The similarity of the wind speed series suggests a high correlation with a certain time delay between the four wind farms (i.e. an SC exists between neighbouring wind farms). Accordingly, the wind speed or power of the target wind farm can be forecasted by using the data of the neighbouring wind farms. Wind power can be forecasted many hours ahead, and the adverse effect of the inaccuracy of NWP data can be mitigated. The relationship between wind farm m_i and target wind farm w is formulated as follows:

$$v_{\mathrm{w}}(t) = v_{\mathrm{m}i}(t-\tau), \tag{16}$$

where $v_{\mathrm{w}}(t)$ is the wind speed of target wind farm w at time t, $v_{\mathrm{m}i}(t)$ is the wind speed of wind farm m_i at time t, and τ_t represents the time delay of wind farm m_i relative to the target wind farm w. The relationship between the time delay and the wind speed could be expressed as follows:

$$\tau_t = L_{\mathrm{w,\,m}i} / v_{\mathrm{m}i}(t), \tag{17}$$

where $L_{\mathrm{w,m}i}$ is the distance between wind farm m_i and target wind farm w.

To analyse the SC of wind speed, the historical data of the target and neighbouring wind farms are used to identify the time delay when the highest correlation occurs between their wind speed series. The Kendall's tau (τ) rank correlation coefficient is an essential index to represent the strength of the relationship between two variables. For variables v_{w} and v_{m}, the correlation coefficient τ is formulated as follows:

$$\tau = \frac{N_c - N_d}{n(n-1)/2}, \tag{18}$$

where N_c is the number of concordant pairs, N_d is the number of discordant pairs, n is the number of samples, $\tau \in [-1,1]$. The correlation coefficient τ can be also expressed as the difference between the probability of concordant and discordant pairs:

$$\tau = P\{(v_{w,1} - v_{w,2})(v_{m,1} - v_{m,2}) > 0\} - P\{(v_{w,1} - v_{w,2})(v_{m,1} - v_{m,2}) < 0\} \quad (19)$$

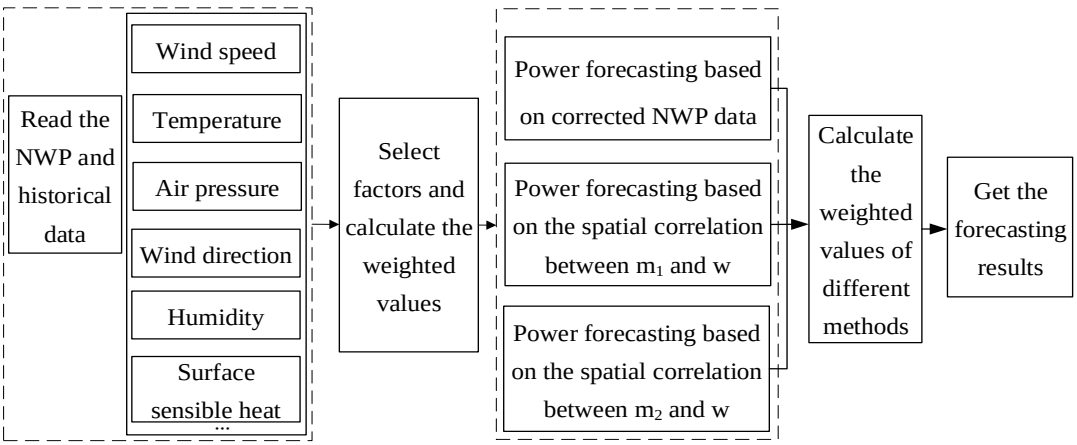
where $v_{w,1}$ and $v_{w,2}$ denote the first and second sample of the wind speed value at the target wind farm, respectively; and $v_{m,1}$ and $v_{m,2}$ are the first and second sample of the wind speed valueat neighbouring wind farms, respectively. The correlation is high when the coefficient is large. The Kendall's τ rank correlation coefficients of the wind speed series between the neighbouring and the target wind farms would be obtained with different time delays. With regard to a time period of wind speed data, the large Kendall's τ rank correlation coefficients correspond to the specific time delay reflecting the similarity of the wind speed series between two wind farms. Thus, the time delay with a large Kendall's τ rank correlation coefficient would be selected for forecasting the wind speed of the target wind farm by using the data of neighbouring wind farms.

3 Hybrid forecasting method

This section presents the study of short-term wind power forecasting by using a hybrid method involving the corrected NWP and SC. This section is divided into two: the process of creating a hybrid weighted model and the standard forecasting accuracy evaluation.

3.1 Process of hybrid weighted model

In the hybrid weighted model, the embedded forecasting methods have varying degrees of importance. To achieve a satisfactory forecasting performance, different weighted values are calculated and assigned to each forecasting method according to their relative importance. Fig. 3 shows the block diagram of the hybrid model.



Hybrid weighted forecasting model that combines the NWP data and the spatial correlation

Fig. 3. Hybrid weighted wind power forecasting model

Assuming that m types of forecasting methods are present, P_1, P_2, P_3, … P_m are the forecasting wind powers by each forecasting method, and ω_1, ω_2, ω_3, … ω_m represent the corresponding weighted values. The hybrid model can be formulated as follows:

$$\begin{cases} P = \omega_1 P_1 + \omega_2 P_2 + \omega_3 P_3 + ... + \omega_m P_m \\ \sum_{i=1}^{m} \omega_i = 1 \end{cases}, \tag{20}$$

where e_1, e_2, … e_m are the forecasting errors of each method, and σ_1, σ_2, … σ_i are the corresponding variance, i =1, 2, … m. The variance of the hybrid model is formulated as follows:

$$\sigma(e_c) = \sum_{i=1}^{m} \omega_i^2 \sigma(e_i) + \sum_{i=1}^{m} \sum_{j=1, j\neq i}^{m} Cov(\omega_i e_i, \omega_j e_j). \tag{21}$$

Given that the various methods are independent of each other, the covariance between them is zero. Therefore, the following equation is obtained:

$$\begin{cases} \sigma(e_c) = \sum_{i=1}^{m} \omega_i^2 \sigma(e_i) \\ \sum_{i=1}^{m} \omega_i = 1 \end{cases}. \tag{22}$$

The minimum forecasting error can be obtained by using the Lagrange multiplier method to minimise the variance. Thereafter, the weighted values of each method in the hybrid model are calculated. The Lagrange function is formulated as follows:

$$f(\omega, \lambda) = \sum_{i=1}^{m} \omega_i^2 \sigma(e_i) + \lambda \left(1 - \sum_{i=1}^{m} \omega_i\right). \tag{23}$$

The partial derivatives of ω_i and λ are obtained. The weighted values can be calculated when the partial derivatives are zero.

Fig. 4 shows the flowchart for implementing the developed hybrid wind power forecasting method, which involves the following key steps:

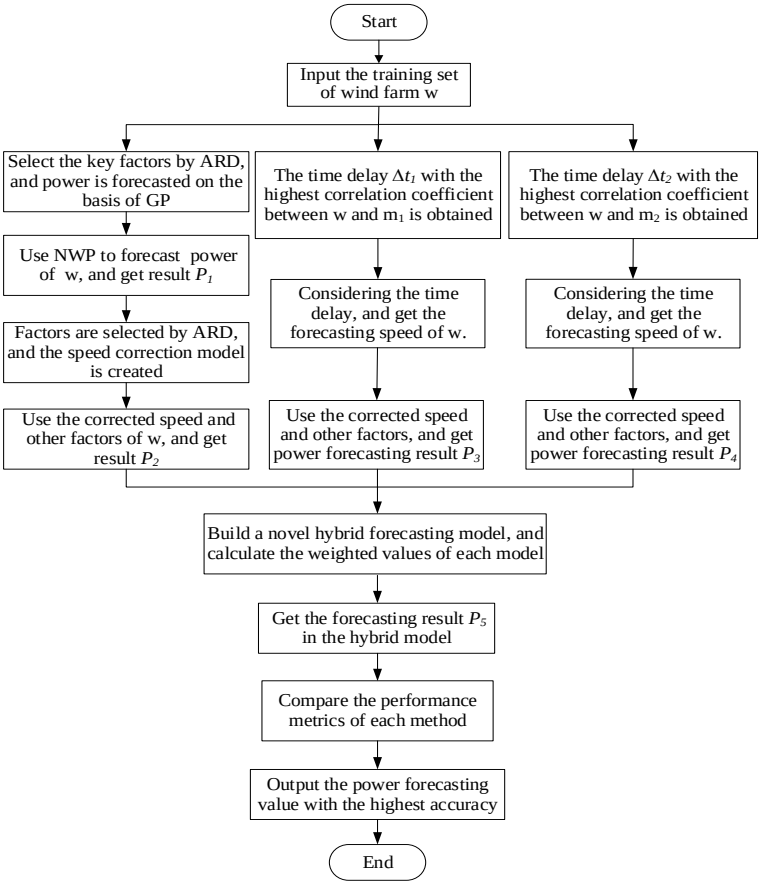


Fig. 4. Steps of wind power forecasting by the hybrid method

Step 1: Read the historical and NWP data, such as wind speed and direction, temperature, humidity and air pressure.

Step 2: Although wind power mainly depends on wind speed, the other meteorological factors may have effects. The ARD method is used to determine the appropriate input for improving the generalisation performance. The variables, which have a remarkable correlation with wind power, are used as inputs. The wind power forecasting model is established by GP, and the forecasted result P_1 is obtained.

Step 3: A wind speed correction model is created through the deviation between the forecasting and the actual values due to the low quality of NWP data. The corrected wind speed is used as input for the wind power forecasting model, and the forecasted result P_2 is obtained.

Step 4: The Kendall's τ rank correlation coefficients between the target wind farm w and its neighbouring wind farms are calculated, and the time delay with the highest correlation is then obtained. The wind speed of w at time t can be forecasted by the historical wind speed data of neighbouring wind farms m_1 and m_2. The

forecasted wind speed and other meteorological factors of w, such as wind direction and temperature, are used as inputs for the hybrid wind power forecasting model. The corresponding forecasted wind power P_3 and P_4 of w are then obtained.

Step 5: The weighted value of each method is calculated by the Lagrange multiplier method to develop the hybrid model. Thereafter, the forecasting wind power P_2, P_3 and P_4 are combined, and the novel hybrid forecasted result is obtained.

Step 6: The performance of each method is analysed, and the wind power forecasting value with high accuracy is obtained.

3.2 Rolling mechanism

To ensure the accuracy of the forecasting model, in the GP, only the recent data are used as the training sample. The rolling mechanism, which is under the condition that the length of the input data is fixed, can constantly update the training data. At the time t, the input data can be expressed as $s^{(i)}(t)=[x^{(i)}(t-L),x^{(i)}(t-L+1),\cdots,x^{(i)}(t-1)]$, where L is the length of the training sample, $x^{(i)}(t-1)$ is the data of the i th variable at a time $t-1$. When new input data are obtained, they are added into the training sample, and the oldest data with the same length are removed. At the time $t+1$, the training sample is expressed as $s^{(i)}(t+1)=[x^{(i)}(t-L+1),x^{(i)}(t-L+2),\cdots,x^{(i)}(t)]$.

3.3 Forecasting accuracy evaluation

Three performance metrics are used to evaluate the forecasting accuracy of different methods: root mean square error (RMSE), mean absolute percentage error (MAPE) and mean absolute error (MAE), which are defined as follows:

$$\mathrm{RMSE}=\sqrt{\frac{1}{T}\sum_{i=1}^{T}(P_i-\hat{P}_i)^2}\,, \tag{24}$$

$$\mathrm{MAPE}=\frac{1}{T}\sum_{i=1}^{T}\left|\frac{P_i-\hat{P}_i}{P_i}\times 100\right|, \tag{25}$$

$$\mathrm{MAE}=\frac{1}{T}\sum_{i=1}^{T}|(P_i-\hat{P}_i)|\,, \tag{26}$$

where P_i is the actual wind power, $\hat{P}_i$ is the forecasted wind power, and T represents the number of testing samples.

4 Case study

The proposed hybrid forecasting method is applied to wind farms in East China. Section 4.1 provides the data description, and Section 4.2 presents the simulation results and extended analysis.

4.1 Data description

The actual meteorological data from the wind farms in East China, acquired from the national meteorological information centre [60], are used in the forecasting model. One hundred twenty sample points at 1 h time interval are chosen to analyse and verify the proposed model. Approximately 80% of data are used for training, and 20% of data are utilised for testing. Different forecasting modules are used for four seasons because the wind speed in four seasons significantly varies. Fig. 5 shows the mean monthly wind speed from 2013 to 2018 in the target wind farm. Fig. 6 shows the typical annual distribution of wind speed. These figures demonstrate that the average wind speed from March to May is higher than those of the rest of the year. The distributions of wind speeds also vary in high and low wind speed seasons. Thus, the analysed data are divided into four periods according to seasons—spring (from March to May), summer (from June to August), autumn (from September to November) and winter (from December to next February).

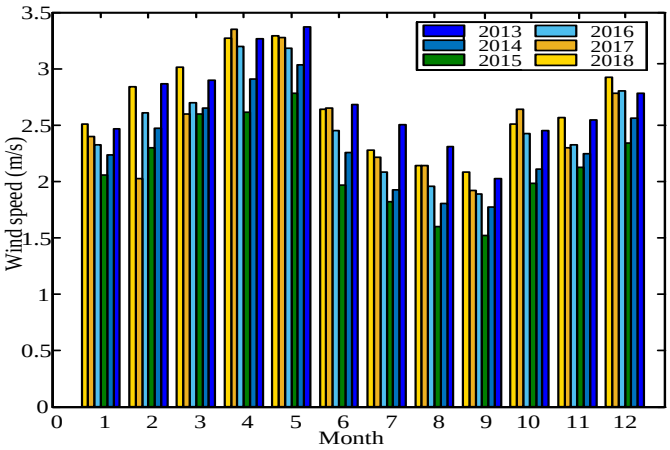


Fig. 5. Mean monthly wind speed

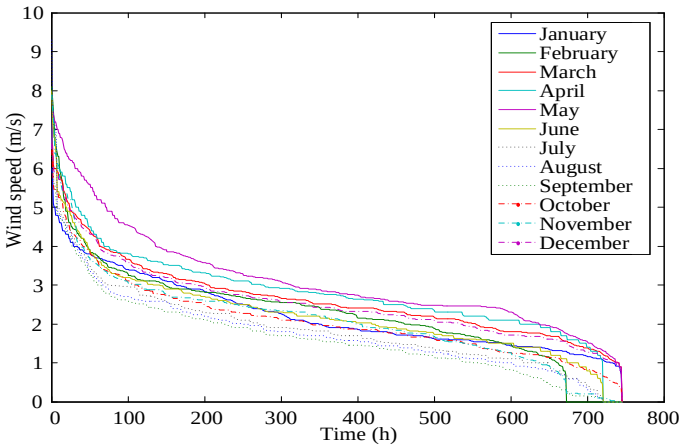


Fig. 6. Typical annual distribution of wind speed

The choice of the optimal kernel function amongst the several candidates as per Section 2.1 should be motivated by the intended usage of the forecasts. The forecasting results using various combinations of kernel functions are also compared. The combined form of kernel functions is given as follows:

$$k(x,x') = k_1(x,x') + k_2(x,x')\,. \tag{27}$$

Ten types of wind power forecasting results are generated and shown in Fig. 7, which uses either a single type of kernel function or their combinations based on SE, RQ, Mat ($v = 5/2$) and Mat ($v = 3/2$). The corresponding performance metrics are represented in Table 2.

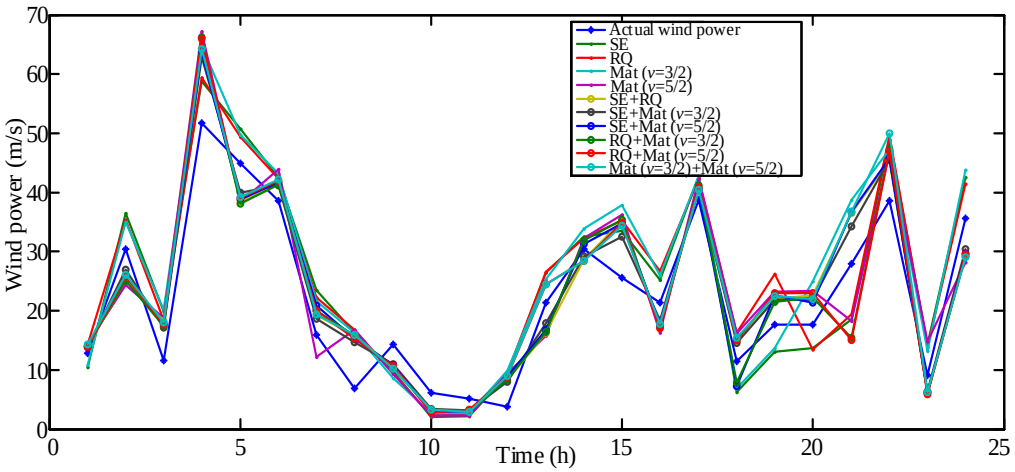


Fig. 7. Wind power forecasting based on the different kernel functions

Table 2. Performance metrics of the forecasting model based on the GP with different kernel functions

Kernel function	RMSE (MW)	MAPE	MAE (MW)
SE	6.069902	0.395229	5.653775

RQ	6.156618	0.39465	5.647139
Mat ($\nu = 3/2$)	6.77634	0.405166	6.091823
Mat ($\nu = 5/2$)	6.812558	0.405619	6.11798
SE + RQ	5.730346	0.327676	5.01388
SE + Mat ($\nu = 3/2$)	4.24602	0.25317	4.807119
SE + Mat ($\nu = 5/2$)	5.586123	0.330217	4.914534
RQ + Mat ($\nu = 3/2$)	6.291651	0.337668	5.354872
RQ + Mat ($\nu = 5/2$)	6.196443	0.341961	5.305759
Mat ($\nu = 3/2$) + Mat ($\nu = 5/2$)	5.899431	0.337354	5.108981

Table 2 illustrates that the error of the forecasting model with combined kernel functions is generally less than that with a single function type because the combination can thoroughly obtain data characteristics. The best forecasting performance is found when the kernel function SE and Mat ($\nu = 3/2$) are combined. In the following study, this optimal combination is used for forecasting by the GP.

4.2 Analysis of the simulation results

4.2.1 Impact of the meteorological factors on wind power

(1) Analysis of ARD

The ARD method is applied to determine the relevance between the input and the output data. This study takes the wind power forecasting model as an example. The most influential meteorological factor to the forecasted wind power can be identified by ARD. In the given dataset, the relevance values of wind speed, wind direction, temperature, air pressure and humility are found as 1.0681, 5.8007, 8.6758, 8.3907 and 7.3568, respectively. This finding indicates that wind speed is the most significant input for the forecast, followed by wind direction.

(2) Analysis of heat map

A heat map is introduced to further investigate the different meteorological factors' correlation to wind power. This map displays the relationship between all measured factors. The columns and rows in the heat map are organised according to the meteorological factors that affect wind power. The deep colour depicts an increased correlation, whereas the lighter one denotes a decreased correlation. The correlation between different factors can be effectively displayed. The visual clustering is also reduced [61], and the heat map analysis of various factors is shown in Fig. 8. In this figure, the correlation coefficient between wind speed and wind power reaches a maximum of 0.9, followed by that between wind direction and wind power (0.31). However, the correlation coefficients between wind power and the other three factors are low. This finding is consistent with the analysis of ARD. All meteorological factors

have some correlations with each other. The correlation coefficient between wind speed and wind direction is up to 0.34.

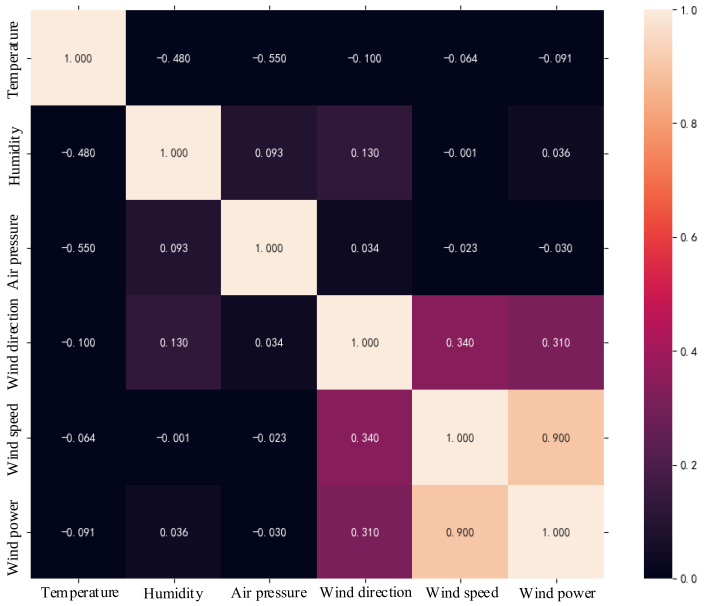


Fig. 8. Heat map of various factors

(3) Analysis with multi-factors

An input factor is eliminated each time, and wind power is forecasted to verify the conclusion of the ARD results and heat map. The results of the different cases are shown in Fig. 9.

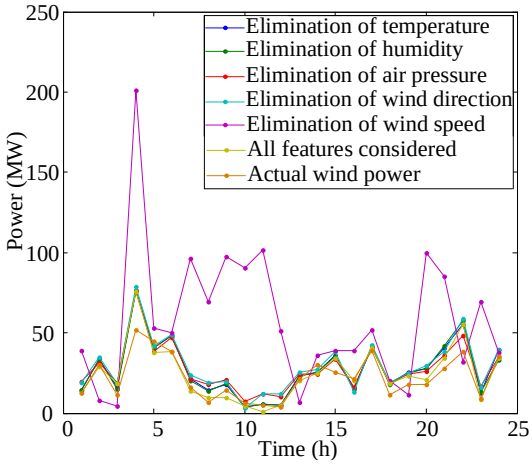


Fig. 9. Wind power forecasting value with eliminating input factor

The difference between the forecasting wind power and the actual value is the highest when the wind speed is eliminated. This notion indicates that wind speed has the greatest influence on forecasting wind power. The second-largest deviation is shown when the wind direction is excluded, which is considerably lower than the wind speed case and closer to the other cases. The values of RMSE, MAPE and MAE and their variances are listed in Table 3.

The variances of performance metrics are just marginal when temperature, humidity and air pressure are eliminated. The greatest variance of RMSE, MAPE and MAE (48.5144 MW, 3.46059 and 35.48947 MW, respectively) is seen when the wind speed is eliminated. From these results, the dominating role of wind speed for wind power forecasting is further confirmed.

Table 3. Comparison of the performance metrics with different methods

Methods	RMSE (MW)	Variance of RMSE	MAPE	Variance of MAPE	MAE (MW)	Variance of MAE
Elimination of temperature	8.406627	1.16060	0.32550	0.07232	6.440736	1.63361
Elimination of humidity	8.543051	1.29702	0.31779	0.06461	6.417485	1.61036
Elimination of air pressure	8.282949	1.43692	0.48369	0.23051	6.881641	2.07452
Elimination of wind direction	10.12328	2.87725	0.57038	0.31720	8.473269	3.66614
Elimination of wind speed	55.76044	48.5144	3.71377	3.46059	40.29659	35.48947
Considering all factors	4.24602	\	0.25317	\	4.807119	\

Note: Variance of RMSE/MAPE/MAE represents the difference of RMSE/MAPE/MAE between the condition with a certain factor elimination and that includes all factors.

4.2.2 Wind power forecasting results based on the corrected NWP

The quality of input data is crucial for accurate and reliable forecasting. However, these complexities are often found in the dataset, which may significantly affect the forecast. In this study, a correction model is created between the NWP wind speed data and their actual value to mitigate their adverse effect. The meteorological factors, which play an essential role in the NWP wind speed correction model, are chosen as the input for the correction model. The wind speed data from the NWP are also used as the input. The accuracy of wind speed can be improved by the correction model proposed in Section 3. A comparison of the original NWP wind speed and the corrected value is shown in Fig. 10.

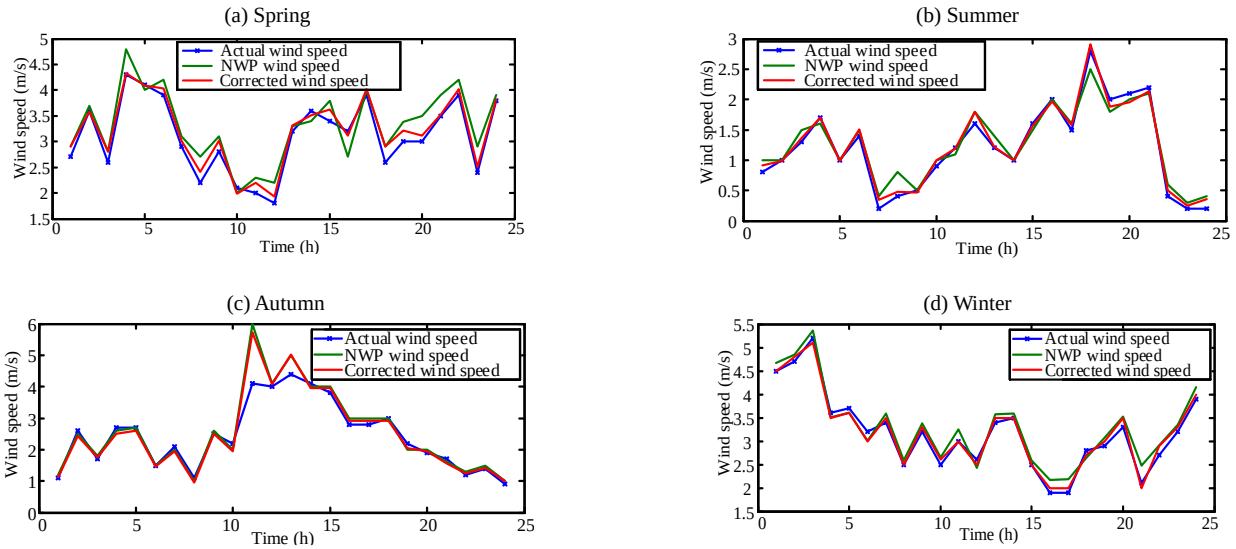


Fig. 10. Wind speed series before and after correction in four seasons

The wind power forecasting results before and after wind speed correction in different seasons are shown in Figs. 11–14. In these figures, the forecasting value after the correction is generally more close to the actual value, thereby indicating the performance improvement by using the wind speed correction model. The results also show that the accuracy of wind power forecasting reduces during high wind speed periods. In terms of seasonal variation, the forecasting error in summer is higher than that in winter because the former has the highest average wind speed in the year. During the day, the deviation between the forecasting and the actual power is relatively larger at the 4th and 22nd h in spring, at the 18th and 20th h in summer, from 12th h to 15th h in autumn and at the 3rd h in winter because these hours coincide with higher wind periods.

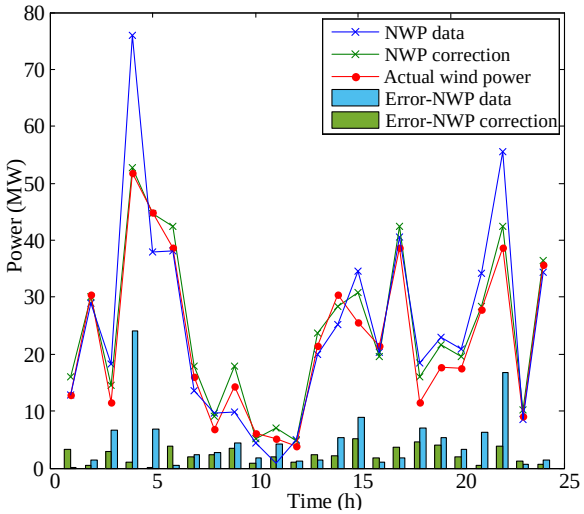


Fig. 11. Wind power forecasting value based on the corrected NWP data in spring

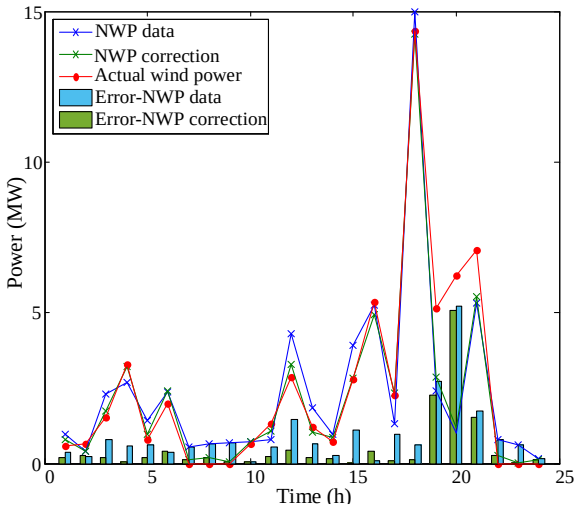


Fig. 12. Wind power forecasting value based on the corrected NWP data in summer

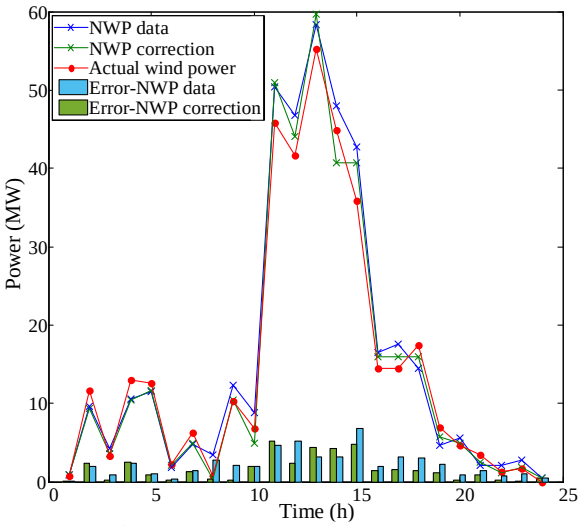


Fig. 13. Wind power forecasting value based on the corrected NWP data in autumn

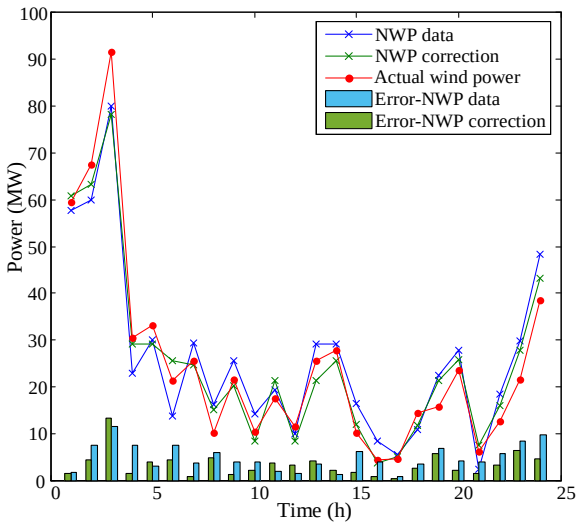


Fig. 14. Wind power forecasting value based on corrected NWP data in winter

4.2.3 Wind power forecasting results with consideration of the SC

The wind speed series of wind farms within a certain geographic range have correlations with each other. The SC between the target wind farm and farms nearby can be calculated using Kendall's τ rank correlation coefficients with different time delays. Fig. 15 shows the comparison of wind speed series between wind farm m_1 and target wind farm w in each season and the correlation coefficients over different time delays. The amount of time delay over which the correlation coefficient reaches its peak value exhibits seasonal variations. From spring to winter, such times delays are −4, −6, −2 and −1 h. Nevertheless, these negative values indicate that the wind speed series in wind

farm m_1 is always ahead of that in wind farm w.

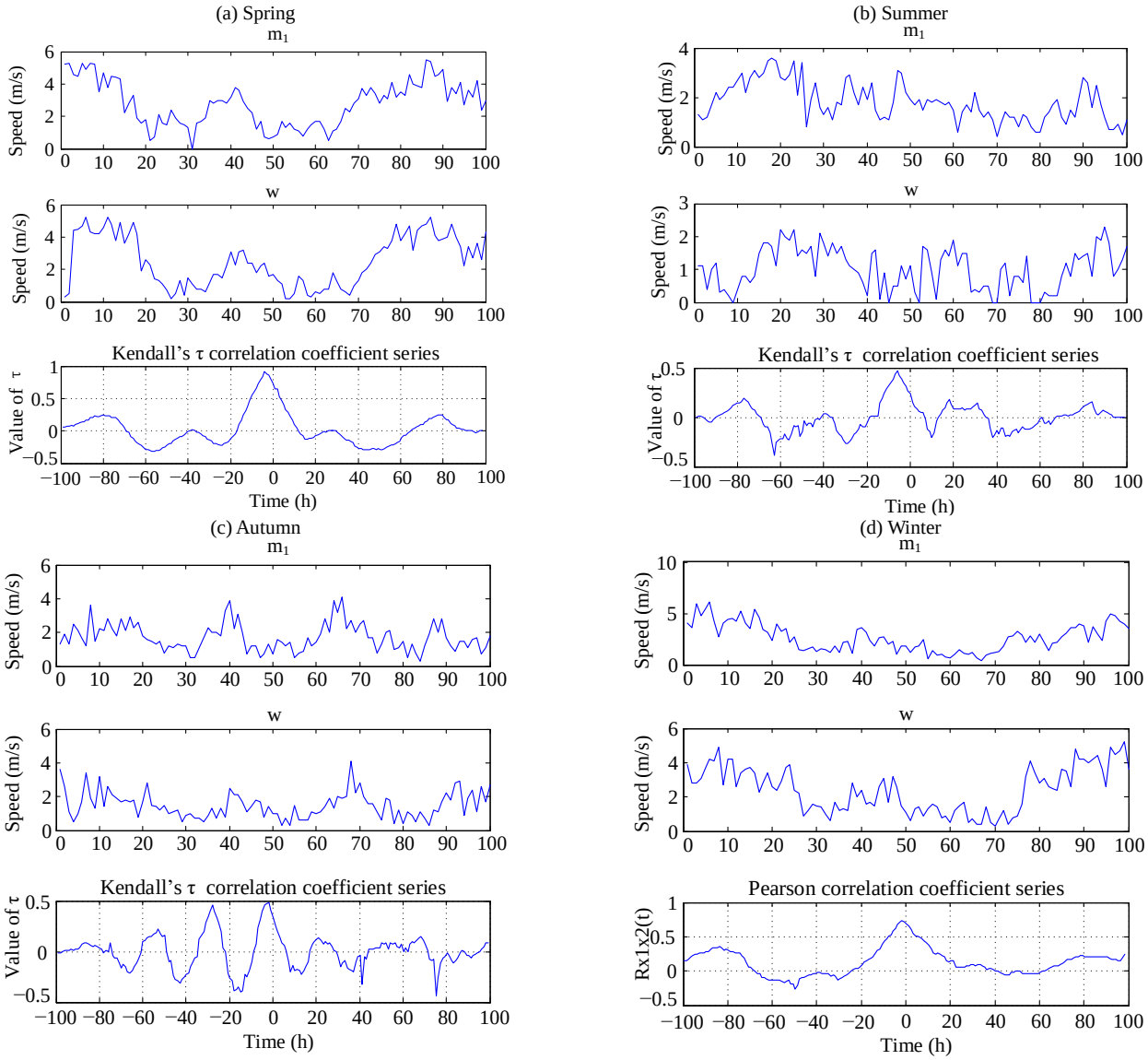
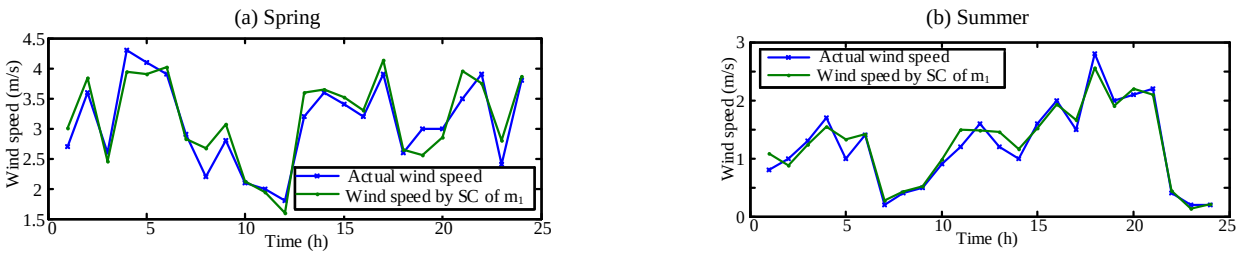


Fig. 15. Kendall's τ correlation coefficient series of wind speed between m_1 and w in different seasons

Fig. 16 presents the forecasted wind speed of the target farm w by using the SC method with a wind farm m_1 in different seasons. The adopted SC method is effective because the forecasted value is close to the actual value in the figure.



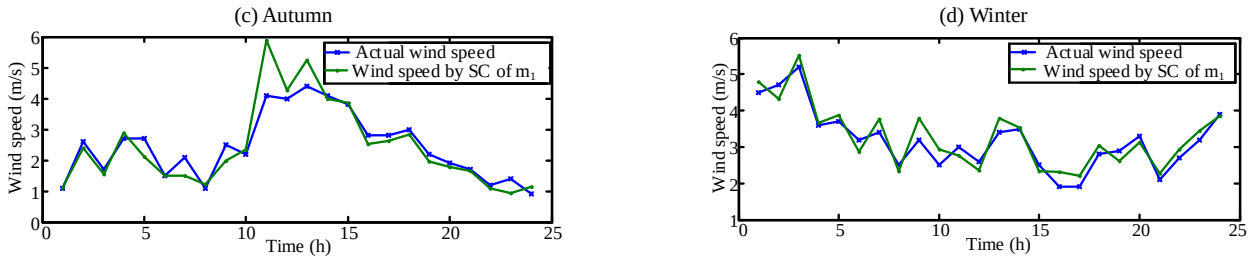
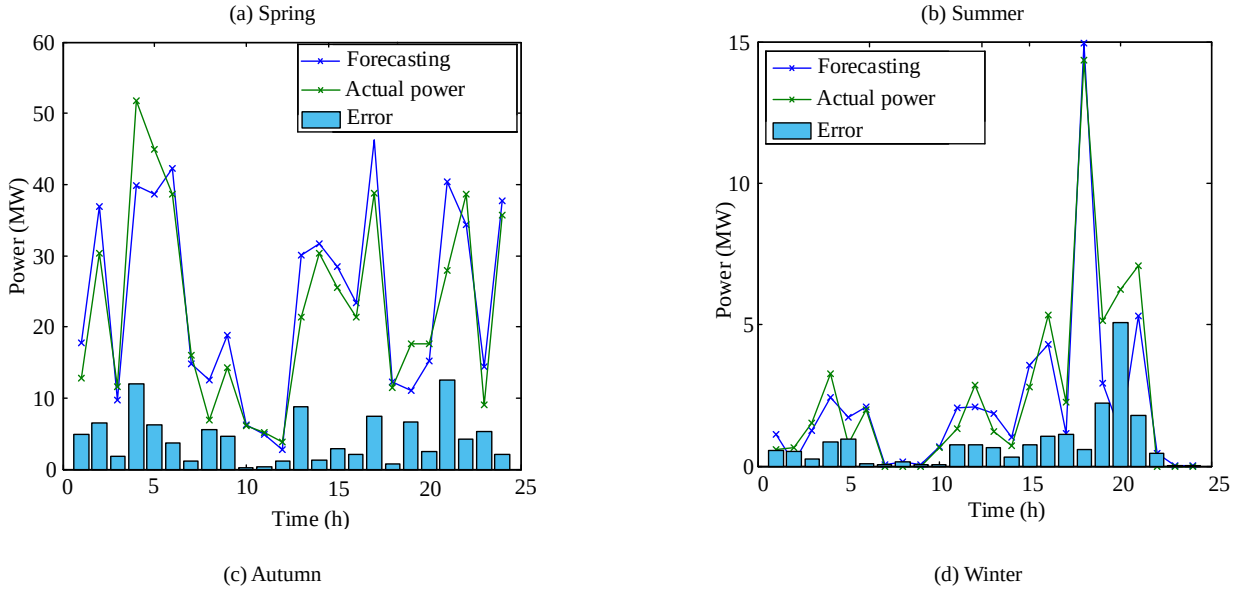


Fig. 16. Forecasted wind speed by SC of m_1 and w in different seasons

The wind speed of the wind farm m_1 from the $(k-\Delta t)$th h to the kth h is used as an input for the wind speed forecasting model. Accordingly, the wind speed of the target wind farm w from the kth h to the $(k+\Delta t)$th h is defined as the output. Other factors, such as temperature and humidity, are also considered to forecast wind power. Fig. 17 shows the wind power forecasting results of the target wind farm w in different seasons based on the GP model. The blue curve is the forecasted wind power of the farm w by using the SC method, and the green curve is the actual value. The histogram represents forecasting errors. Fig. 17 demonstrates that the forecasted wind power is generally close to the actual value even though the errors in spring and winter are much larger than those in summer and autumn because of the high wind speed and high connection between the wind power and the wind speed.



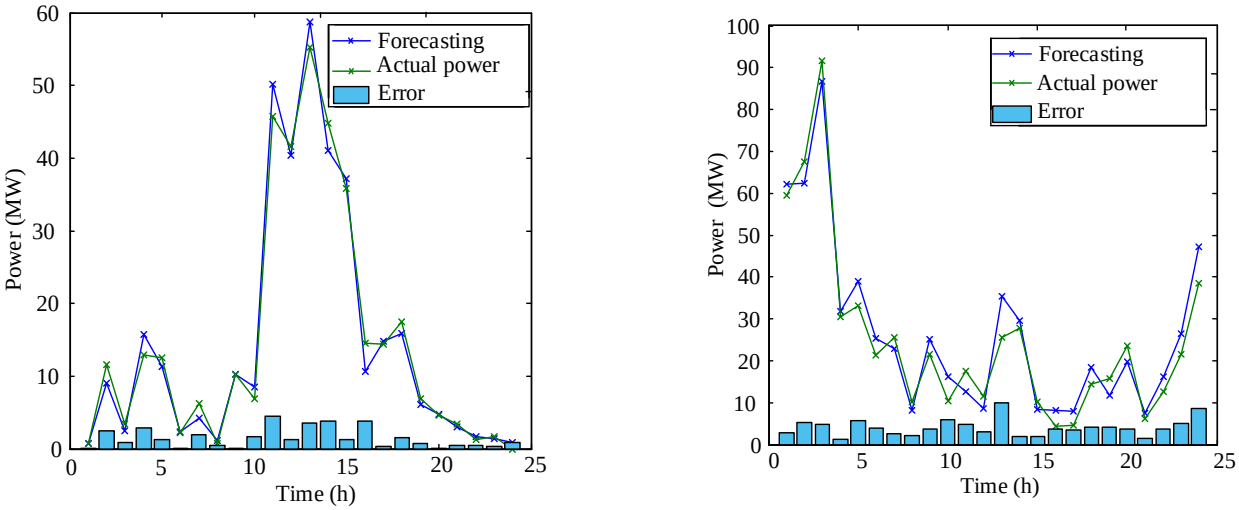
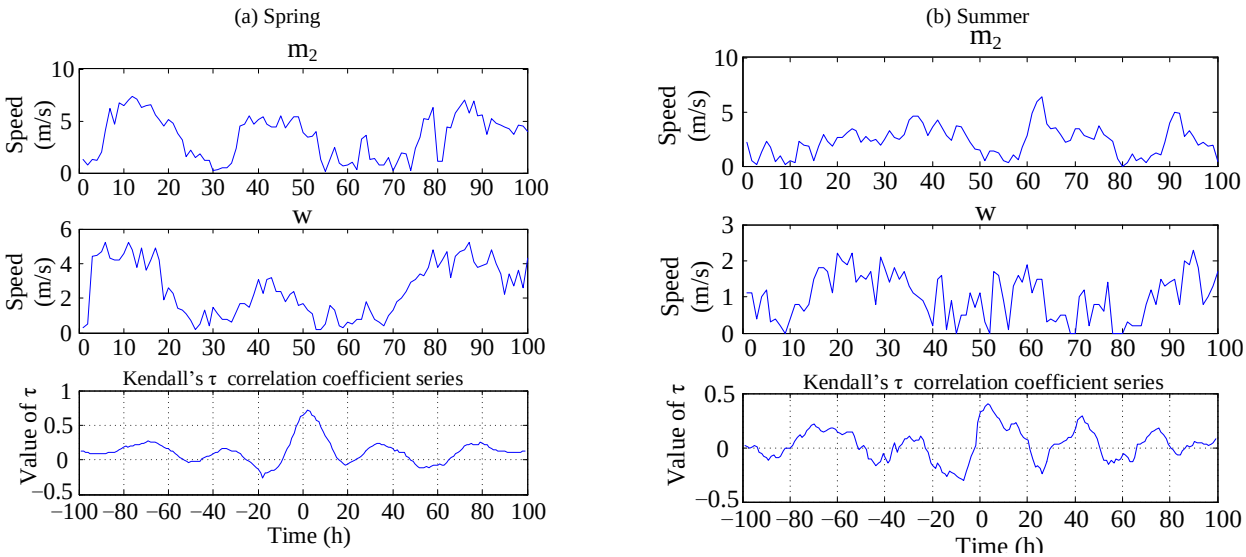


Fig. 17. Wind power forecasting value based on the SC between m_1 and w in different seasons

Fig. 18 shows Kendall's τ correlation coefficient series between wind farm m_2 and target wind farm w in different seasons. The figure presents that the time delays for the highest correlation coefficient are 3, 3, 1 and 1 h from spring to winter. These positive values indicate that the wind speed series in wind farm m_2 lag behind target wind farm w.



(c) Autumn

(d) Winter

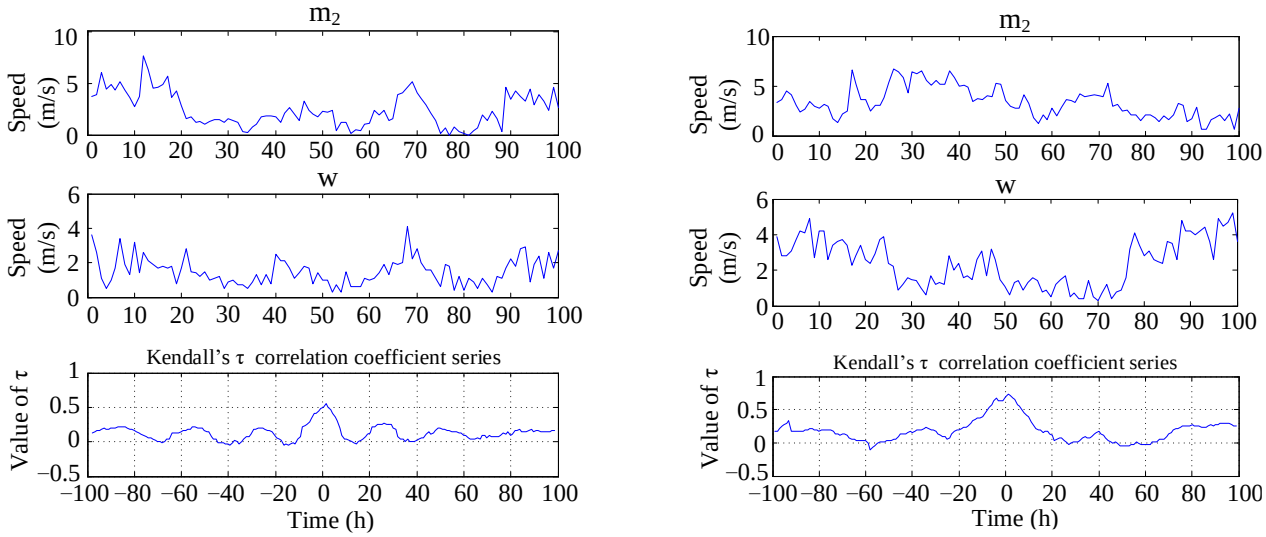


Fig. 18. Kendall's τ correlation coefficient series of wind speed between m_2 and W in different seasons

Fig. 19 depicts the corresponding forecasted value of wind speed based on SC between the wind farm m_2 and target wind farm W. The result shows that the forecasted wind speed is close to the actual data in four seasons, thereby proving the effectiveness of the SC method between wind farms within a certain geographic scale.

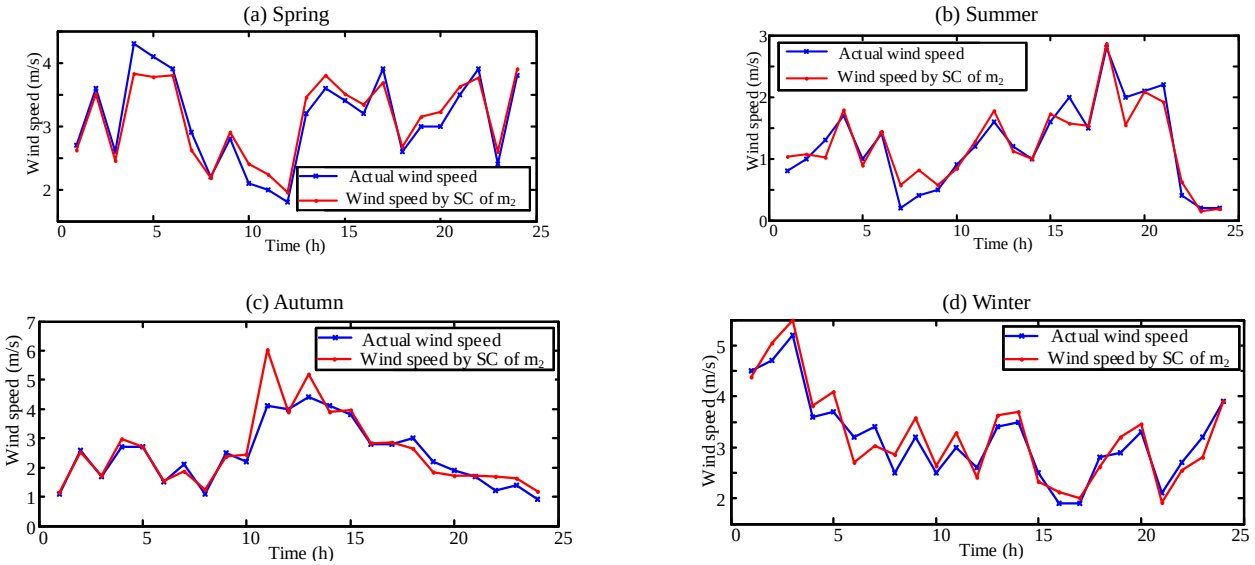


Fig. 19. Forecasted wind speed by SC of m_2 and W in different seasons

The wind speed from the kth h to the $(k+\Delta t)$th h of wind farm m_1 is used as the input for the wind speed forecasting model, with the wind speed of the target wind farm W from the $(k-\Delta t)$th h to the kth h as the output. The other meteorological factors, such as temperature and humidity, are also considered for the wind power forecasting model. Fig. 20 shows the forecasting results of the target wind farm W. The forecasting errors are much high in spring and winter, which coincide with large wind power.

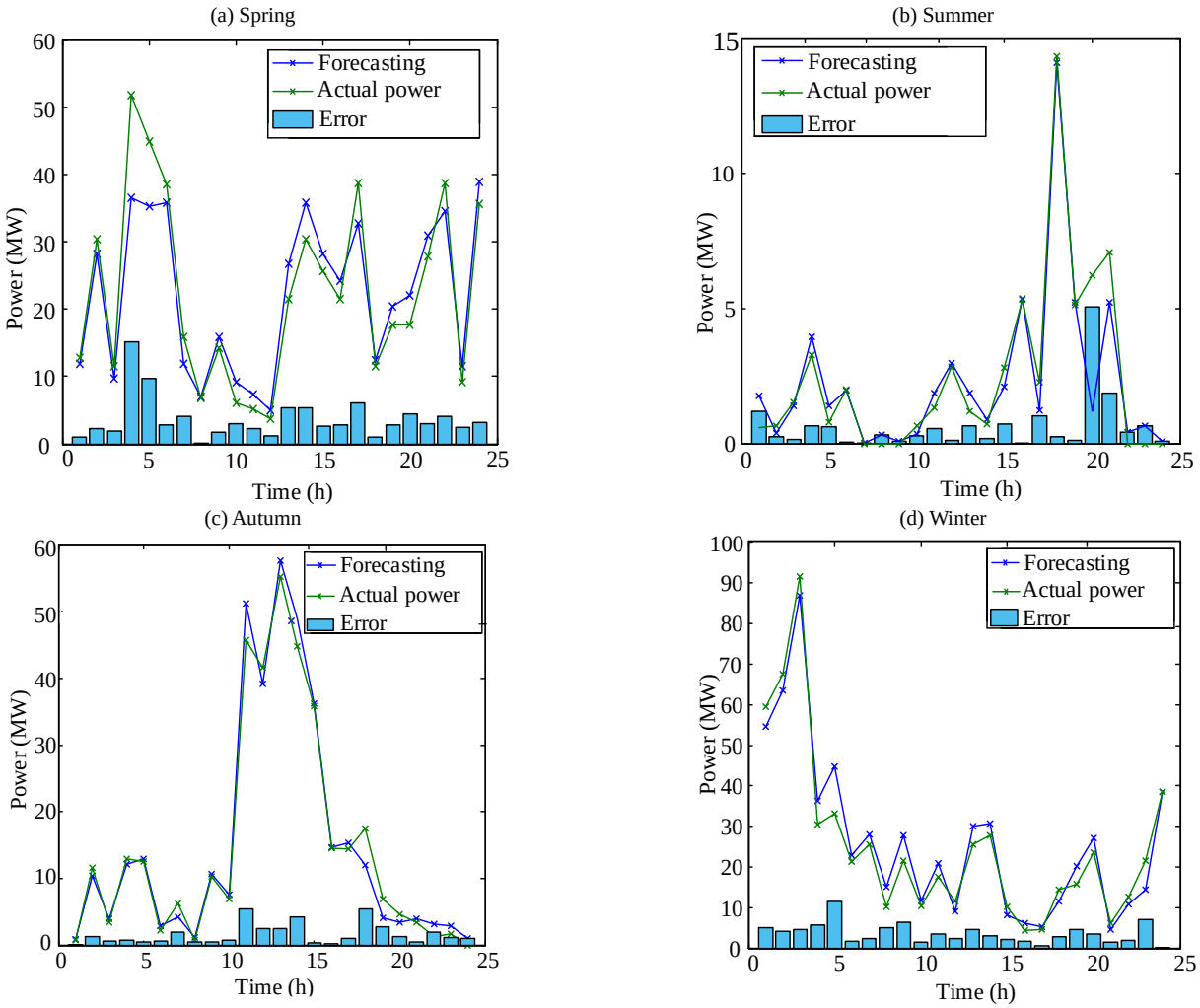


Fig. 20. Wind power forecasting value based on the SC between m_2 and w in different seasons

Table 4 lists the indexes that reflect the correlations between the target wind farm w and neighbouring wind farms m_1 and m_2. The average wind speed in spring is the highest, followed by those in winter and summer. The correlation coefficient between the target and the neighbouring wind farms reaches the peak value in spring, followed by those in winter and autumn, and the smallest value is in summer.

Table 4. The indexes about SC in the neighbouring wind farms

Indexes	**Spring**		**Summer**		**Autumn**		**Winter**	
	m_1	**m_2**	**m_1**	**m_2**	**m_1**	**m_2**	**m_1**	**m_2**
Average wind speed	2.7854	3.5947	1.4552	2.1417	1.7364	2.3969	2.5781	3.0218
Time delay	4 h	3 h	6 h	3 h	2 h	1 h	1 h	1 h
Correlation coefficient	0.9253	0.7696	0.4692	0.4573	0.4988	0.5271	0.8063	0.7467

4.2.4 Results from the hybrid weighted forecasting model

The novel hybrid wind power forecasting model is created by combining the corrected NWP and SC. The

combination of these sub-models is achieved by assigning each one with a specific weighted value (Table 5). These weighted values vary amongst models and also change in different seasons. Thus, the annual forecasting can be divided into four scenarios.

Table 5. Weighted values of the hybrid model in four seasons

Index	**Spring**			**Summer**			**Autumn**			**Winter**		
	ω_1	ω_2	ω_3	ω_1	ω_2	ω_3	ω_1	ω_2	ω_3	ω_1	ω_2	ω_3
Weighted value	0.489	0.239	0.272	0.344	0.314	0.342	0.353	0.316	0.331	0.343	0.322	0.335

Figs. 21–24 show the forecasted wind power by seasons and the corresponding error. The blue curve shows the forecasted wind power based on the corrected wind speed. The green curve shows the forecasted wind power based on the SC between wind farm m_1 and target wind farm w. The red curve shows the forecasted wind power based on the SC between wind farm m_2 and target wind farm w. The cyan curve shows the forecasted wind power based on the novel hybrid model. The purple curve shows the actual wind power. The histograms from right to left show the forecasting errors of the four methods. In these figures, the hybrid model outperforms all the other mentioned methods by much higher forecasting accuracy, thereby indicating the considerable advantage that can be gained from the hybrid forecasting method.

When all inputs (e.g the corrected wind speed, the obtained wind speed by SC method, wind direction, temperature, humidity and air pressure) are incorporated into a single forecasting model, there would be another forecasting result. To justify the advantages of the hybrid forecasting method furtherly, the forecasting results when incorporating all inputs into the GP model are compared with the hybrid model.

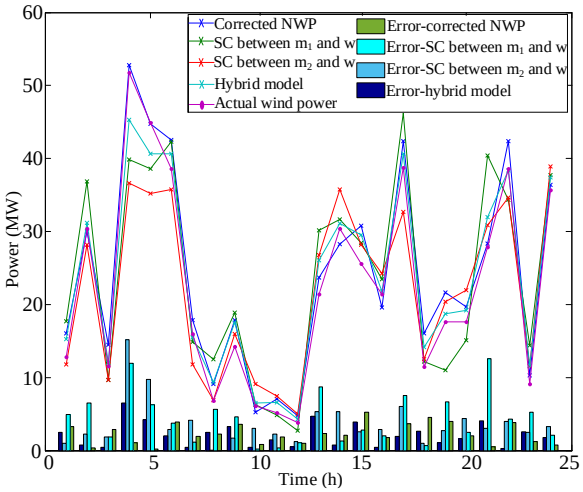


Fig. 21. Forecasted wind power and corresponding error in spring

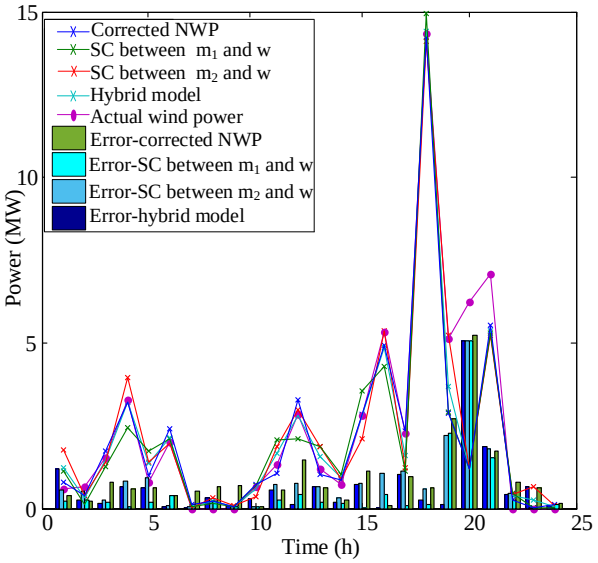


Fig. 22. Forecasted wind power and corresponding error in summer

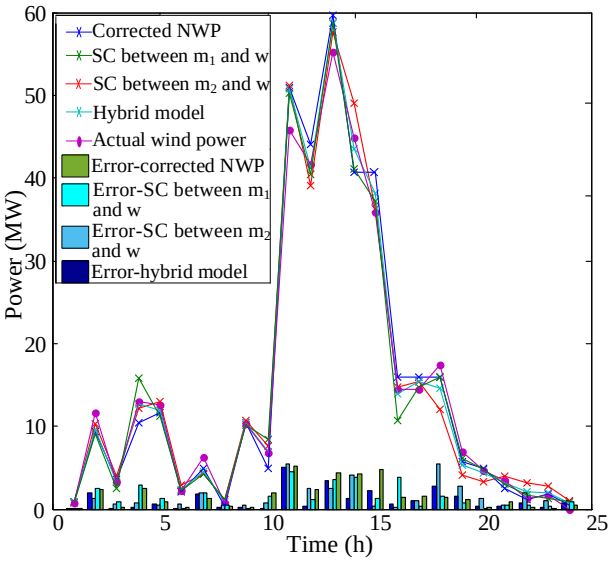


Fig. 23. Forecasted wind power and corresponding error in autumn

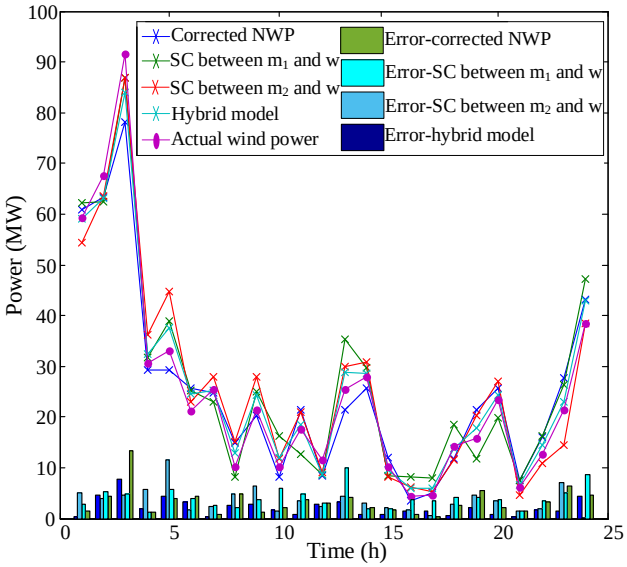


Fig. 24. Forecasted wind power and corresponding error in winter

Table 6 provides the performance metrics of different forecasting methods. The novel hybrid model significantly performs better than conventional models. The RMSE values are reduced by 0.33078–1.5919, 0.01102–0.15111, 0.31536–0.76779 and .46431–1.53667 MW in spring, summer, autumn and winter, respectively. The MAPE values are reduced by 0.0287–0.12918, 0.0494–0.21913, 0.03481–0.2738 and 0.07372–0.16959 in spring, summer, autumn and winter, respectively. The MAE values are reduced by 0.17644–2.69465, 0.01448–0.28154, 0.35463–1.14045 and 1.15365–2.72491 MW in spring, summer, autumn and winter, respectively. The assessment of the improvement in the forecasting accuracy is formulated as follows:

$$k = \frac{e_{NWP} - e_{mix}}{e_{NWP}} \times 100\%, \tag{28}$$

where e_{NWP} is the RMSE value by the Gaussian forecasting method based on the corrected NWP data, and e_{mix} is the RMSE value by the hybrid model. The accuracy of using the hybrid method is 37.49% higher than that by utilising the original NWP data in spring. The accuracy values in summer, autumn and winter increase by 10.88%, 31.88% and 35.67%, respectively. The average increase in accuracy is 28.98%.

Table 6. Performance metrics by each wind power forecasting method

Method	Spring			Summer			Autumn			Winter		
	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE
NWP data	4.2460	0.2531	4.8071	1.3891	0.4029	0.8746	2.4085	0.3925	2.2018	4.3083	0.2873	4.8938
Corrected NWP data	2.9849	0.1527	2.2889	1.2699	0.3744	0.7147	2.2394	0.1535	1.6105	3.2360	0.1692	3.3226
SC between m_1 and w	4.0969	0.2268	4.3515	1.3801	0.5441	0.9817	1.9561	0.1571	1.4159	3.5224	0.2523	4.0443
SC between m_2 and w	3.8117	0.1836	3.6731	1.2490	0.4809	0.7578	2.1617	0.2324	1.5187	3.3428	0.1914	3.6061
Inputs	3.5728	0.1772	2.9637	1.2932	0.4261	0.7332	1.8362	0.1978	1.3466	3.1832	0.1736	2.9373

incorporated into GP												
Hybrid model	2.6541	0.1240	2.1124	1.2380	0.3250	0.7002	1.6407	0.1187	1.0613	2.7717	0.1177	2.1689

A comparison of the hybrid model with other existing models in Table 1 is discussed to evaluate the forecasting performance of the proposed model. The proposed hybrid forecasting method has an RMSE of 2.6541, a MAPE of 0.1240 and an MAE of 2.1124 in spring at 24 h ahead (Fig. 21). The forecasting accuracy by the hybrid method is higher than that by other methods in Table 6. The performance metric is lower than those of some methods (Table 1). For example, the wind power forecasting model of the long short-term memory embedded with wavelet kernels has a MAPE between 0.4212 and 0.4983 for forecasting with 250 h ahead [43]. A novel convolution-based spatial–temporal wind power forecasting model reports an RMSE of 5.9–10.1 at 30 min ahead [44]. The forecasting model combining autoencoders and the backpropagation algorithm has an average MAPE of 15.96% at 24 h ahead [46]. The combined forecasting model based on the neural network and grey model has a MAPE of 16.2% at 60 h ahead [50]. The constructed convolution neural network and lightGBM model reported a MAE of 2.28 at 6 h ahead [60]. Various datasets need different forecasting accuracy values; however, the forecasting effect cannot be judged by only using these performance metrics. Nevertheless, considering the range of errors, the proposed model still has improved performance.

5 Conclusion

A novel hybrid model built on the GP model is developed for short-term wind power forecasting in this study. Unlike the original GP model, the hybrid model also integrates the corrected NWP data and SC between geographically distributed wind farms. The basic GP model is established using the optimal combination of different kernel functions. The SC method, calculated using the Kendall's τ correlation coefficient is adopted to strengthen input data of the target wind farm for wind power forecasting. When integrating these methods into the hybrid forecasting model, the weighted values are adaptively assigned according to their importance using the Lagrange multiplier method, which could ensure the small error of the hybrid forecasting model. In the case study, the proposed hybrid forecasting method is applied to the actual data of wind farms in East China. In comparison with the conventional GP model, the forecasting performance in terms of RMSE is improved by 7.02–29.7% with the corrected NWP, 0.65–10.23% with the SC method and 10.88–37.49% with the hybrid model. The hybrid model has proven its significant value in improving the forecast accuracy and reducing the operational risk of real wind farms

given that it clearly outperformed the other methods.

The presented wind power forecasting model is generally suitable for short-term wind power forecasting because of the applicability of quarterly data training. The observations and training data are updated constantly to improve the model performance. For applications purpose, the presented forecasting model could utilize the information of spatial factors comprehensively, and is further capable for improving the accuracy of NWP data when considering the meteorological factors. The effectiveness and advantages can be inferred from the results of simulation. This has important implications for the real operation of wind farms.

In future work, advanced artificial intelligence and machine learning methods would be introduced to assist the automatic scene division on the complex input data. According to the concept of data dependence, a similar type of input data with similar spatial and temporal characteristics is collected in the same scenario. The scene division method will use central semantic to reflect the complex data characteristics and classify the data into different scenarios. In this way, the forecasting results of the next period can be obtained to effectively improve the applicability of wind power forecasting models.

Acknowledgment

This work was supported by the Fundamental Research Funds for the Central Universities (YJ201654).

References

[1] Rosales-Asensio E, Borge-Diez D, Blanes-Peiro J, Pérez-Hoyos A, Comenar-Santos A. Review of wind energy technology and associated market and economic conditions in Spain. Renewable and Sustainable Energy Reviews 2019;101:415-27.

[2] Hu T, Wu W, Guo Q, Sun H, Shi L, Shen X. Very short-term spatial and temporal wind power forecasting: A deep learning approach. CSEE Journal of Power and Energy Systems 2020;6(2):434-43.

[3] Wang J, Wang Y, Li Z, Li H; Yang H. A combined framework based on data preprocessing, neural networks and multi-tracker optimizer for wind speed prediction. Sustainable Energy Technologies and Assessments 2020;40.

[4] Croonenbroeck C, Stadtmann G. Renewable generation forecast studies – Review and good practice guidance. Renewable and Sustainable Energy Reviews 2019;108:312-322.

[5] Yan J, Zhang H, Liu Y, Han S, Li L, Lu Z. Forecasting the high penetration of wind power on multiple scales using multi-to-multi mapping. IEEE Transactions on Power Systems 2017;33(3):3276-84.

[6] Saroha S, Aggarwal S. Wind power forecasting using wavelet transforms and neural networks with tapped delay. CSEE Journal of Power and Energy Systems 2018;4(2):197-209.

[7] Aasim, Singh S, Mohapatra A. Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting. Renewable Energy 2019;136:758-68.

[8] Karaku O, Kuruolu E, Altnkaya M. One-day ahead wind speed/power prediction based on polynomial autoregressive model. IET Renewable Power Generation 2017;11(11):1430-9.

[9] Jiang Y, Chen X, Kun Y, LIAO Y. Short-term wind power forecasting using hybrid method based on enhanced boosting algorithm. Journal of Modern Power Systems and Clean Energy 2017;5(1):126-33.

[10] Dhiman H, Deb D, Guerrero J. Hybrid machine intelligent SVR variants for wind forecasting and ramp events. Renewable and Sustainable Energy Reviews 2019;108:369-79.

[11] Chang G, Lu H, Chang Y, Lee Y. An improved neural network-based approach for short-term wind speed and power forecast. Renewable Energy 2017;105:301-11.

[12] Zhang Y, Le J, Liao X, Zheng F, Li Y. A novel combination forecasting model for wind power integrating least square support vector machine, deep belief network, singular spectrum analysis and locality-sensitive hashing. Energy 2019;168:558-72.

[13] Catalão J, Pousinho H, Mendes V. Short-term wind power forecasting in Portugal by neural networks and wavelet transform. Renewable Energy 2010;36(4):1245-51.

[14] Liu D, Niu D, Wang H, Fan L. Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. Renewable Energy 2014;62:592-7.

[15] Ren Y, Suganthan P, Srikanth N. A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods[J]. IEEE Transactions on Sustainable Energy 2017;6(1):236-244.

[16] Hoolohan V, Tomlin A, Cockerill T. Improved near surface wind speed predictions using Gaussian process regression combined with numerical weather predictions and observed meteorological data. Renewable Energy 2018;126:1043-54.

[17] Ezzat A, Jun M, Yu D. Spatio-temporal short-term wind forecast: A calibrated regime-switching method. The annals of applied statistics 2019, 13(3): 1484.

[18] Browell J, Drew D, Philippopoulos K. Improved very short‐term spatio‐temporal wind forecasting using atmospheric regimes. Wind Energy 2018, 21(11): 968-979.

[19] Fang X, Hodge B, Du E, Zhang N, Li F. Modelling wind power spatial-temporal correlation in multi-interval optimal power flow: A sparse correlation matrix approach. Applied Energy 2018;230:531-9.

[20] Li P, Guan X, Wu J, Zhou X. Modeling dynamic spatial correlations of geographically distributed wind farms and constructing ellipsoidal uncertainty sets for optimization-based generation scheduling. IEEE Transactions on Sustainable Energy 2015;6(4):1594-605.

[21] Zhu Q, Chen J, Shi D, Zhu L, Bai X, Duan X, Liu Y. Learning temporal and spatial correlations jointly: a unified framework for wind speed prediction. IEEE Transactions on Sustainable Energy 2019;11(1):509-23.

[22] Li S, Wang P, Goel L. Wind power forecasting using neural network ensembles with feature selection. IEEE Transactions on Sustainable Energy 2015;6(4):1447-56.

[23] Tastu J, Pinson P, Kotwa E, Madsen H, Nielsen H. Spatio‐temporal analysis and modeling of short‐term wind power forecast errors. Wind Energy 2011, 14(1): 43-60.

[24] Khodayar M, Wang J. Spatio-Temporal Graph Deep Neural Network for Short-Term Wind Speed Forecasting, IEEE Transactions on Sustainable Energy, 2019, 10(2): 670-681.

[25] Zhao Y, Ye L, Pinson P, Tang Y, Lu P. Correlation-constrained and sparsity-controlled vector autoregressive model for spatio-temporal wind power forecasting. IEEE Transactions on Power Systems 2018, 33(5): 5029-5040.

[26] Gneiting T, Genton M, Guttorp P. Geostatistical space-time models, stationarity, separability, and full symmetry. Monographs On Statistics and Applied Probability 2006, 107: 151.

[27] Ezzat A, Jun M, Ding Y. Spatio-temporal asymmetry of local wind fields and its impact on short-term wind forecasting. IEEE transactions on sustainable energy 2018, 9(3): 1437-1447.

[28] Wang H, Han S, Liu Y, Yan J, Li L. Sequence transfer correction algorithm for numerical weather prediction wind speed and its application in a wind power forecasting system. Applied Energy 2019;237:1-10.

[29] Nielsen H, Nielsen T, Madsen H, San M, Marti I. Optimal combination of wind power forecasts. Wind energy 2007;10(5):471-82.

[30] Bessac J, Constantinescu E, Anitescu M. Stochastic simulation of predictive space–time scenarios of wind speed using observations and physical model outputs. The Annals of Applied Statistics 2018;12(1):432-58.

[31] Jung J, Broadwater RP. Current status and future advances for wind speed and power forecasting. Renewable and Sustainable Energy Reviews 2014;31:762-77.

[32] Chen N, Xue Y, Ding J, Chen Z, Wang W, Wang N. Ultra-short term wind speed prediction using spatial correlation. Automation of Electric Power Systems 2017;41(12):124-30.

[33] Wang D, Luo H, Grunder O, Lin Y. Multi-step ahead wind speed forecasting using an improved wavelet neural network combining variational mode decomposition and phase space reconstruction. Renewable Energy 2017;113:1345-58.

[34] Chen N, Qian Z, Nabney IT, Meng X. Wind power forecasts using Gaussian processes and numerical weather prediction. IEEE Transactions on Power Systems 2014;29(2):656-65.

[35] Li L, Zhao X, Tseng M, Tan R. Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. Journal of Cleaner Production 2020;242.

[36] Zhou J, Yu X, Jin B. Short-term wind power forecasting: A new hybrid model combined extreme-point symmetric mode decomposition, extreme learning machine and particle swarm optimization. Sustainability 2018;10(9):1-18.

[37] Azimi R, Ghofrani M, Ghayekhloo M. A hybrid wind power forecasting model based on data mining and wavelets analysis. Energy Conversion and Management 2016;127:208-25.

[38] Dhiman H, Deb D, Guerrero J. Hybrid machine intelligent SVR variants for wind forecasting and ramp events. Renewable and Sustainable Energy Reviews 2019;108:369-79.

[39] Wan C, Lin J, Wang J, Song Y, Dong Z. Direct quantile regression for nonparametric probabilistic forecasting of wind power generation. IEEE Transactions on Power Systems 2017;32(4):2767-78.

[40] Lu H, Chang G. Wind power forecast by using improved radial basis function neural network. In 2018 IEEE Power & Energy Society General Meeting. IEEE; 2018.

[41] Viet D, Phuong V, Duong M, Kies A, Schyska B. A short-term wind power forecasting tool for vietnamese wind farms and electricity market. In 2018 4th International Conference on Green Technology and Sustainable Development. IEEE; 2019.

[42] Shahid F, Zameer A, Mehmood A, Raja M. A novel wavenets long short term memory paradigm for wind power prediction. Applied Energy 2020;269.

[43] Hu T, Wu W, Guo Q, Sun H, Shi L, Shen X. Very short-term spatial and temporal wind power forecasting: A deep learning approach. CSEE Journal of Power and Energy Systems 2019;6(2):434-43.

[44] Chen Q, Folly K. Short-term wind power forecasting based on spatial correlation and artificial neural network. In 2020 International SAUPEC/RobMech/PRASA Conference. IEEE; 2020.

[45] Jiao R, Huang X, Ma X, Han L, Tian W. A model combining stacked auto encoder and back propagation algorithm for short-term wind power forecasting. IEEE Access 2018;6:17851-8.

[46] Shen W, Jiang N, Li N. An EMD-RF based short-term wind power forecasting method. in 2018 IEEE 7th data driven control and learning systems conference. IEEE; 2018.

[47] Li C, Lin S, Xu F, Liu D, Liu J. Short-term wind power prediction based on data mining technology and improved support vector machine method: A case study in Northwest China. Journal of Cleaner Production 2018;205:909-22.

[48] Ju Y, Sun G, Chen Q, Zhang M, Zhu H, Rehman M. A model combining convolutional neural network and lightGBM algorithm for ultra-short-term wind power forecasting. IEEE Access 2019;7:28309-18.

[49] Zhang Y, Sun H, Guo Y. Wind power prediction based on PSO-SVR and grey combination model. IEEE Access 2019;7:136254-67.

[50] Scheuerer M, Möller D. Probabilistic wind speed forecasting on a grid based on ensemble model output statistics. Annals of Applied Statistics 2015; 9(3): 1328-1349.

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

[52] Gel Y, Raftery A, Gneiting T, Claudia Tebaldi, Doug Nychka, William Briggs et al. Calibrated probabilistic mesoscale weather field forecasting: The geostatistical output perturbation method. Journal of the American Statistical Association 2004;99(467):575-583.

[53] Yan J, Zhang H, Liu Y, Han S, Li L, Lu Z. Forecasting the High Penetration of Wind Power on Multiple Scales Using Multi-to-Multi Mapping. IEEE Transactions on Power Systems 2018;33(3):3276-3284.

[54] Hoolohan V, Tomlin A, Cockerill T. Improved near surface wind speed predictions using Gaussian process regression combined with numerical weather predictions and observed meteorological data. Renewable Energy 2018;126:1043-54.

[55] Nadja H, Juš K. Design of a hybrid mechanistic/Gaussian process model to predict full-scale wastewater treatment plant effluent. Computers and Chemical Engineering 2020;140.

[56] Fang S, Chiang HD. A high-accuracy wind power forecasting model. IEEE Transactions on Power Systems 2016;32(2):1589-90.

[57] Yan J, Li K, Bai EW, Deng J, Foley A. Hybrid probabilistic wind power forecasting using temporally local Gaussian process. IEEE Transactions on Sustainable Energy 2016;7(1):87-95.

[58] Pandit R, Infield D. Comparative analysis of Gaussian Process power curve models based on different stationary covariance functions for the purpose of improving model accuracy. Renewable Energy 2019;140:190-202.

[59] Lei L, Yin X, Jia X, Sobhani B. Day ahead powerful probabilistic wind power forecast using combined intelligent structure and fuzzy clustering algorithm. Energy 2020;192.

[60] National meteorological information center of China. http://data.cma.cn/data/ cdcindex/cid/6d1b5efbdcbf9a58.html, Accessed date: 28 July 2020.

[61] Michael A, Mohammed K, Alex E. Location selection using heat maps: Relative advantage, task-technology fit, and decision-making performance. Computers in Human Behavior 2019;101:151-62.