Heliyon 9 (2023) e18053

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

EMD-based gray combined forecasting model - Application to long-term forecasting of wind power generation

Minghao Ran, Jindi Huang, Wuyong Qian *, Tingting Zou, Chunyi Ji

School of Business, Jiangnan University, Wuxi, Jiangsu, 214122, China

ARTICLE INFO

Keywords:
Renewable energy
Energy forecasting
Carbon neutrality
Carbon peaking
Wind power

ABSTRACT

Wind power is the most promising renewable energy source after hydropower because of its mature technology and low price, and has great potential for carbon emission reduction. Long-term forecasts of its power generation can help power companies to develop operational plans, grid configuration and power dispatch, and can also provide a basis for the government to formulate energy and environmental policies. However, due to the characteristics of China's monsoon climate and wind power industry development, wind power generation data are characterized by nonlinear cycles and small data volume, which makes accurate prediction more difficult. To this end, this paper develops a new prediction model and applies it to the long-term prediction of wind power generation in China, and proposes some targeted policy recommendations based on the prediction results to promote the development of China's wind power industry.

1. Introduction

1.1. Background

According to the statistical bulletin on national economic and social development released by the National Bureau of Statistics of China, China's power generation in 2022 was 8534.25 billion kWh, which was 4.5% higher than that of the previous year in terms of growth rate, and steadily ranked as the world's top power generation country, as shown in Fig. 1, with huge power demand. However, for a long time, China's power generation is dominated by thermal power, with a high dependence on traditional energy sources such as coal and natural gas, and a relatively homogeneous structure, as shown in Fig. 2. Secondly, thermal power generation has caused great pollution to the natural environment. The emission of greenhouse gases, mainly carbon dioxide, has led to global warming, while the emission of acidic gases such as sulfur dioxide, as well as nitrogen oxides, has led to an increase in the amount of acid rain and dust pollution in many regions [1]. The power industry has become one of the largest polluting industries in China, and poses considerable challenges to the sustainable development of the economy and society [2]. Nowadays, human beings have reached a consensus on the above-mentioned energy tensions and environmental degradation problems, and it is urgent to actively seek and develop renewable energy sources, change the energy structure, alleviate the energy crisis, and improve the environmental problems from the source, and adjust and upgrade the energy structure.

The Chinese government is actively responding to the Paris Agreement's clear policy of "peak carbon by 2030″ and "carbon neutral

* Corresponding author.
E-mail address: wuyongqian@jiangnan.edu.cn (W. Qian).

https://doi.org/10.1016/j.heliyon.2023.e18053
Received 11 May 2023; Received in revised form 5 July 2023; Accepted 5 July 2023
Available online 7 July 2023

by 2060". To accelerate energy restructuring, China's power industry has resolutely implemented its "double carbon" policy and actively implemented clean energy substitution actions. According to the "Energy Production and Consumption Revolution Strategy (2016–2030)" published by China's National Development and Reform Commission and China's National Energy Administration, the share of non-fossil energy generation in all power generation will strive to reach 50% by 2030 [3].

Driven by Chinese government policies and real environmental problems, there is an urgent need to find clean renewable energy sources to reduce environmental pollution and meet the electricity consumption of society's production and life. As wind power is one of the more mature and competitively priced renewable energy sources, it has become the leading renewable energy source after hydropower [4]. The IEA estimates that by 2026, wind power is expected to be the leading renewable energy source. By 2026, the IEA expects renewable energy to be the largest source of electricity generation, accounting for about 37% of global electricity generation, with solar and wind power reaching nearly 18% of global electricity generation. Fig. 2 shows that among the various types of energy generation in China, wind power is growing strongly and rapidly compared to thermal power, hydropower, nuclear power, solar power, etc. In China's 14th Five-Year Plan for the Development of National Strategic Emerging Industries, the wind power industry has been listed as one of the national strategic emerging industries. The wind power industry has been listed as one of the national strategic emerging industries. As a result, under the combined effect of market demand and industrial policy, China's investment in the development and utilization of wind power projects has been increasing, and the wind power industry has shown unprecedented rapid momentum and has been able to leap forward in quality [5]. The wind power industry has shown unprecedented momentum and has been able to leapfrog in quality.

Accurate forecasting of wind power generation is important not only for the planning of production activities and power regulation, but also for the development of operational efforts, energy strategies, and energy policies of governments and power companies [6]. Specifically, due to the intermittent nature of wind power and the decentralized nature of the power stations, it is difficult to regulate the grid with wind power, while short-term forecasting can provide a basis for early scheduling by power companies to prevent grid fluctuations. The medium and long-term forecasts of wind power can provide a realistic basis for power companies to plan their power generation operations, staffing, and maintenance, and can also provide a basis for the Chinese government to formulate relevant policies to promote energy restructuring and achieve the "double carbon" goal.

This paper focuses on the long-term forecasting of wind power generation in China. However, due to the special climatic characteristics of China and the location of wind power plants, the wind power generation data in China exhibit non-linear characteristics characterized by cyclical seasonality and trends, and it is difficult to accurately grasp the potential characteristics of the data with a single model structure, so the forecasting results are often poor. Second, the development cycle of wind power in China is relatively short, the available data is relatively limited, and there are many factors affecting wind power generation. Based on the above

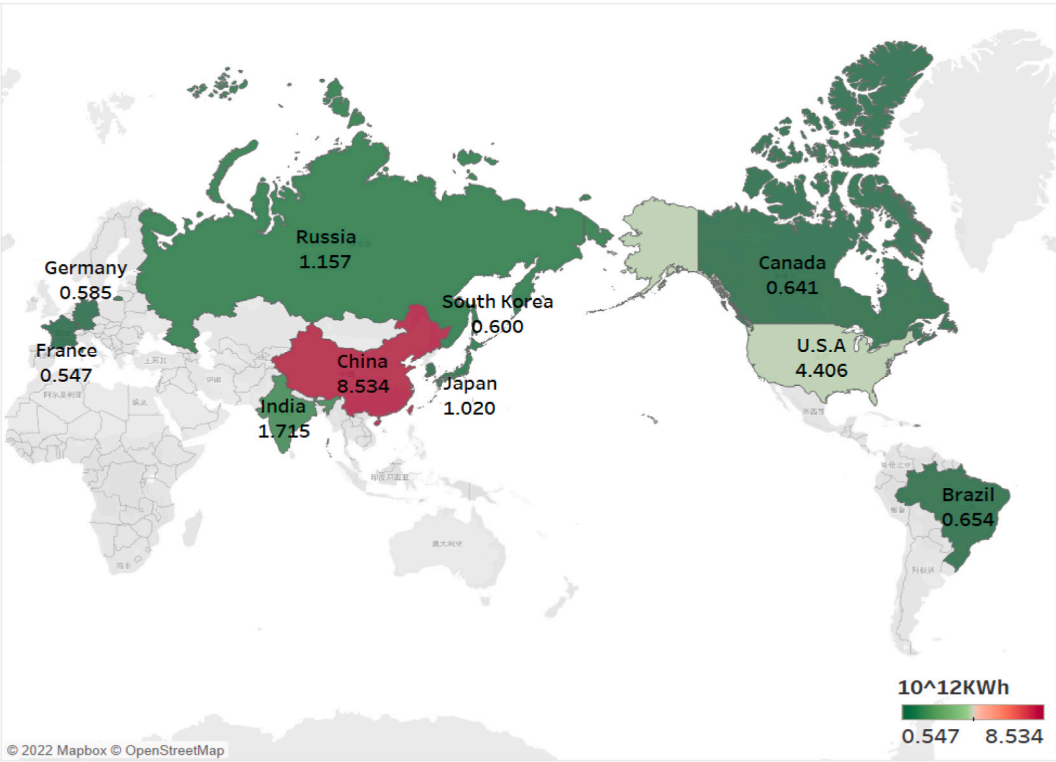

Fig. 1. Top 10 countries in global electricity generation in 2021
Data source: National Bureau of Statistics of China.

considerations, the prediction of wind power generation in China is essentially an uncertainty prediction problem based on small samples and poor information characteristics. Since gray forecasting methods have significant advantages in forecasting small-sample data, they have been widely used in various forecasting problems in recent years, especially in the fields of energy, transportation, and economy. However, the traditional gray forecasting model has a good fitting effect on the data series with exponential regularity, and it cannot obtain good forecasting accuracy for the dual attribute characteristics of periodic seasonality and trend of wind power generation in China. The modeling method based on the data decomposition method tends to obtain better prediction results than direct modeling using raw data [7]. In summary, to grasp the seasonal and trend characteristics of wind power data and obtain more accurate forecasting results, this paper proposes a gray DGM model based on EMD to optimize the original gray forecasting model and shows good fitting performance in both the training and test sets.

1.2. Literature review

1.2.1. Wind power forecasting study

Rapidly growing wind generation triggers random fluctuations in the electricity grid, which can compromise grid stability and increase grid imbalance costs. Therefore, accurate forecasting of wind generation is critical for capacity deployment, schedule improvement, energy restructuring, policy evaluation, and grid balancing of wind power high penetration systems. On the other hand, there is also a need to construct forecasting models for different periods to meet various needs, which can be generally classified into four categories, as shown in Table 1 [8].

The main methods of wind power forecasting are Statistical and econometric methods, machine learning methods, gray forecasting methods, and combined forecasting methods, as shown in Table 2.

Statistical and econometric methods assume relevant statistical distributions for the original data and build predictive models based on the relevant assumptions. Statistical and econometric methods are more widely used in the field of wind power and many statistical and econometric models have been developed. Machine learning methods are all data-driven to build models, collecting a large amount of historical data or other exogenous data as input for prediction [9], compared with statistical and econometric methods, machine learning methods do not need to describe the model with the help of complex mathematical relationships and assumptions, but establish relationships with a large number of input and output processes that can simulate the relationship between historical data and target results, so in the case of big data, it is often able to make accurate predictions and has a stronger learning capability [10]. Since wind power generation data are affected by a variety of factors and have nonlinear and non-smooth characteristics, it is difficult for a single model structure to accurately capture its data characteristics, so the prediction effect is often poor. The combinatorial modeling approach based on data decomposition methods can achieve better prediction results than direct modeling using raw data, so combinatorial models are widely used in the field of wind power prediction [11], and some major studies are shown in Table 3.

1.2.2. Application of gray forecasting in the energy field

The gray prediction method takes the poor information uncertain system with partly known information and partly unknown information as the research object, and mines the system original information by generating and developing the known information to achieve the effective description of the system operation behavior and evolution law, as well as the quantitative prediction of the future state and change of the system [27]. The research on gray prediction theory is deepening. As the research of gray prediction theory deepens and expands, its application scope and field are also gradually expanded. In addition to being widely used in the economic and social fields, gray forecasting methods are also widely used in the energy field [28].

Xiao [29] and Wang [30] used a periodic truncated cumulative generating operator and a data grouping approach (DGGM(1,1)) to transform the periodic data series into a smoothed data series applicable to the GM(1,1) model for forecasting seasonal time series, respectively. Ding [31] forecasted the seasonal time series by constructing an adaptive gray forecasting model with modified initial values (NSGM(1,1) model) predicting the natural gas consumption in China. Qian [32] proposed a GM(1,1) model construction method based on HP filter decomposition for systems with periodic fluctuations, which can achieve effective prediction of evolutionary

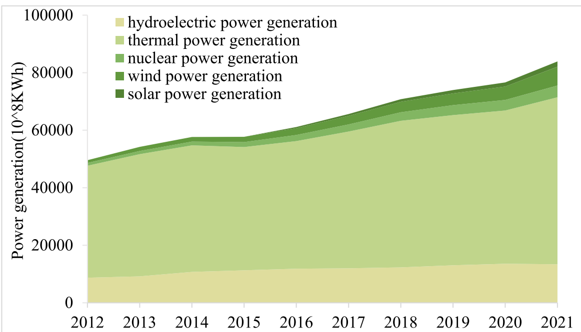

Fig. 2. Comparison of electricity generation capacity by power generation mode in China: 2012-2021
Data source: National Bureau of Statistics of China.

Table 1
Classification of wind power generation forecast content.

Prediction Scenarios	Period	Main role
Long-term forecast	Use "years" as the unit of prediction	It is mainly used for the feasibility study of wind farm design, operation plans, and other long-term strategy development.
Mid-term forecast	Use "weeks" as the unit of prediction	It is mainly used for wind farm scheduling and maintenance.
Short-term forecast	Use "hours" as the unit of prediction	In most cases, it is mainly used to predict the power in advance.
Ultra Short-term Forecast	Use "minutes" as the unit of prediction	It is mainly used for the control of wind turbines and the day's wind power market bidding.

Table 2
Common methods for wind power forecasting.

Method	Features
Statistical and econometric models	It is required for the raw data to obey a certain distribution. Involved in short, medium and long-term forecasting of wind power.
Machine learning models	Deep learning models are widely used to build predictive models with data-driven requirements for large amounts of data. Often used for short-term and ultra-short-term forecasting of wind power.
Gray prediction models	Good prediction for small data, poor information. Often used for long-term forecasting.
Combined predictive models	Two combined models exist, series and parallel, which tend to have higher prediction performance than a single model, but there is some flexibility in the choice of combined models for different scenarios. It is involved in all prediction scenarios of wind power.

Table 3
Some major studies in the field of wind power forecasting.

Method Type	Research Field	Author	Models
Statistical and econometric models	Short-term forecast	Yatiyana [12]	The autoregressive integrated moving average (ARIMA) method was used to develop an estimation model for wind power in Western Australia.
	Short-term forecast	Liu [13]	The autoregressive moving average-generalized autoregressive conditional heteroskedasticity (ARMA-GARCH) method is demonstrated to be advantageous in capturing the trend changes in mean wind speed variation and volatility.
	Short-term forecast	Wang [14]	A multi-step ahead wind speed prediction technique based on heteroskedasticity multicore learning is designed and its reliability is verified.
	Mid-term forecast	Dowell [15]	A sparse vector autoregressive model (SVAR) is developed in the parametric probability framework of log-normal distribution and its advantages over the traditional vector autoregressive model are demonstrated.
Machine learning models	Short-term forecast	Karakus [16]	An artificial neural network-adaptive neuro-fuzzy inference system (ANN-ANFIS) was developed using a polynomial linear regression (PAR) model.
	Short- and medium-term forecast	Cai [17]	The prediction results of the SVR were further enhanced by a multi-task Gaussian process (MTGP).
	Short-term forecast	Lahouar [18]	A quantile random forest model without row parameter tuning is constructed, and the accuracy of the prediction confidence interval explicitly constructed by it is significantly improved.
	Short-term forecast	Khosravi [19]	A model combining multilayer feedforward neural network (MLFFNN) and adaptive neuro-fuzzy inference system (ANFIS) with partial swarm optimization algorithm (ANFIS-PSO) is used.
Combination Forecast	Short-term forecast	Demolli [20]	A combination of five machine learning algorithms, minimum absolute shrinkage selection operator (LASSO), k nearest neighbor (kNN), extreme gradient boosting (XGBoost), random forest (RF), and support vector regression (SVR), is used.
	Short-term forecast	Luo [21]、LV [22] and Wu [23]	A combined prediction model is built based on the EMD data decomposition method to decompose the original series into different sub-series.
	Short-term forecast	Zhang [24]	A prediction model based on CEEMD-IGA–FNN–Markov is proposed to improve the accuracy of ultra-short-term wind speed prediction.
	Short- and medium-term forecast	Li [25] and Zhao [26]	Based on the TSD–FS–BM combinatorial model, a combinatorial model of VMD-GSO-ELM with simultaneous parameter optimization is proposed.

trends of systems with "periodic fluctuations" and achieve good results in wind power generation forecasting applications. Wu [33] used a nonlinear gray Bernoulli model (FANGBM(1,1) model) to forecast the total renewable energy consumption, hydroelectric power generation, wind energy consumption, solar energy consumption, and other renewable energy consumption, respectively. Guefano [34] combined a gray prediction model with a vector autoregressive model (VAR) to forecast electricity consumption. Qian [35] designed a new structural adaptive discrete gray forecasting model to capture the nonlinear, linear, periodic, and volatile characteristics present in the renewable energy generation series.

In summary, first, a large number of wind power forecasting studies have focused on short-term and ultra-short-term forecasting studies, while there are few Mid-term and long-term forecasting studies. Second, in terms of forecasting methods, statistical and econometric models require data to obey typical distributions. In the wind power application scenario, there are various stochasticity's in wind power generation, such as regional economy, seasonality, random effects, and outlier effects, and it is difficult for statistical models based on classical distributions to explain all stochastic components and influencing factors, and it is difficult to capture statistical distribution patterns when the amount of data or system information is limited. Machine learning requires a large amount of data for training, which is very time-consuming and weakly interpretable, and the application scenarios are limited to short-term and ultra-short-term forecasting studies. Third, from the perspective of combinatorial models, a large number of forecasting models built by data decomposition are Parallel concatenated, while there are few studies on types of tandem connections combinatorial forecasting. In this paper, we focus on the long-term forecasting of Chinese wind power generation, decompose the data into periodic and trend terms through EMD data decomposition, build a tandem-type combined forecasting model through seasonal factor and DGM forecasting model, and make some targeted suggestions for Chinese wind power industry based on the accurate forecasting model.

2. Research methodology

2.1. DGM

In the traditional GM (1.1) model, the jump from the discrete form of the model to the continuous form of the whitening equation has always troubled researchers in gray system theory. DGM (1.1) takes this as the starting point of research to solve this theoretical problem from the perspective of going from discrete to discrete, and establishes a discrete gray prediction model [36].

Definition 1 [36]: Assuming a non-negative sequence $x^{(0)}$ and its 1-AGO (cumulative generating operator) sequences $x^{(1)}$ respectively.

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$$
$$X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right)$$

which $x^{(0)}(k) \geq 0, x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 1, 2, \cdots, n$, called

$$x^{(1)}(k+1) = \beta_1 x^{(1)}(k) + \beta_2 \tag{1}$$

GM (1.1) model, or call it a discrete form of the GM (1.1) model.

Theorem 1 [36]: Assume that the sequence $x^{(0)}$ and the sequence $x^{(1)}$ as defined in Definition 1, the $\widehat{\beta} = [\beta_1, \beta_2]^{\mathrm{T}}$ is a parameter column, and

$$Y = \begin{bmatrix} x^{(1)}(2) \\ x^{(1)}(3) \\ \vdots \\ x^{(1)}(n) \end{bmatrix}, B = \begin{bmatrix} x^{(1)}(1) & 1 \\ x^{(1)}(2) & 1 \\ \vdots & \vdots \\ x^{(1)}(n-1) & 1 \end{bmatrix}$$

Then, the least squares estimated parameter column $x^{(1)}(k+1) = \beta_1 x^{(1)}(k) + \beta_2$ satisfying

$$\widehat{\beta} = [\beta_1, \beta_2]^{\mathrm{T}} = \left(B^{\mathrm{T}} B\right)^{-1} B^{\mathrm{T}} Y \tag{2}$$

Theorem 2 [36]: Assume Y and B are as defined in Theorem 1 $\widehat{\beta}$ that $\widehat{\beta} = (B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}Y$,

Let $x^{(1)}(1) = x^{(0)}(1)$, then $x^{(1)}(k+1) = \beta_1 x^{(1)}(k) + \beta_2$ the time response equation of

$$\widehat{x}^{(1)}(k+1) = {\beta_1}^k \left(x^{(0)}(1) - \frac{\beta_2}{1-\beta_1}\right) + \frac{\beta_2}{1-\beta_1}, k = 1, 2, \cdots, n-1.$$

Proof: Shaped like

$$x^{(1)}(k+1) = Ax^{(1)}(k) + B \tag{3}$$

of the difference equation is solved identically as

$$x^{(1)}(k) = CA^k + \frac{B}{1-A} \tag{4}$$

where C is an arbitrary constant that can be determined according to the initial conditions given by the problem

The above equation is the same as $x^{(1)}(k+1) = \beta_1 x^{(1)}(k) + \beta_2$ the difference equation that is exactly equivalent to $A = \beta_1, B = \beta_2$ and therefore has

$$x^{(1)}(k) = C{\beta_1}^k + \frac{\beta_2}{1-\beta_1} \tag{5}$$

When $k = 0$ When, take $x^{(1)}(0) = x^{(0)}(1)$, substitute into the equation, we get $C = \left[x^{(0)}(1) - \frac{\beta_2}{1-\beta_1}\right]$, substitute C back into Equation to obtain the proof.

The reduction equation is given by the following equation [36]:

$$\widehat{x}^{(0)}(k+1) = \widehat{x}^{(1)}(k+1) - \widehat{x}^{(1)}(k)$$

$$= (\beta_1 - 1)\left(x^{(0)}(1) - \frac{\beta_2}{1-\beta_1}\right)\beta_1^k, k = 1, 2, \cdots, n-1 \tag{6}$$

2.2. EMD

Data-driven empirical modal decomposition (EMD) based on data is powerful and adaptive in analyzing nonlinear and non-stationary data sets. It essentially acts as a binary filter, separating complex signals with wide frequency bands into relatively simple components with different time scales, IMF, which includes information about the local characteristics of the trend and fluctuations of the original signal at different scales, and helps to analyze the true physical meaning of the signal to some extent [37]. Empirical modal decomposition decomposes the signal based on the time-scale characteristics of the data itself and therefore does not require any prior setting of the characteristics of the basic functions. Huang [38], the proposer of EMD, argued that any signal can be split into the sum of several implicit modal components. And the implicit modal components have two constraints.

(1) In the whole data segment, the number of extreme value points and the number of over zero points must be equal or must not differ by more than one at most.
(2) At any moment, the average value of the upper envelope formed by the local maxima and the lower envelope formed by the local minima is zero, i.e., the upper and lower envelopes are locally symmetric concerning the time axis.

The decomposition steps are as follows.

(1) Draw the upper and lower envelopes according to the upper and lower extreme points of the original signal, respectively.
(2) Find the mean value of the upper and lower envelopes and draw the mean envelope.
(3) The original signal minus the mean envelope to obtain the intermediate signal.
(4) determine whether the intermediate signal satisfies the two conditions of IMF, if so, the signal is an IMF component; if not, use the signal as the basis and redo the analysis of 1∼4. The acquisition of IMF components usually requires several iterations.

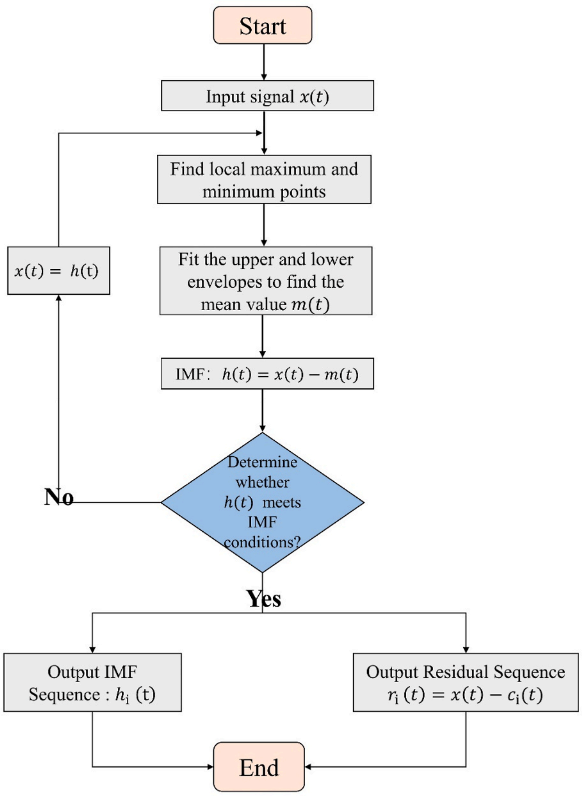

Fig. 3. EMD algorithm flow.

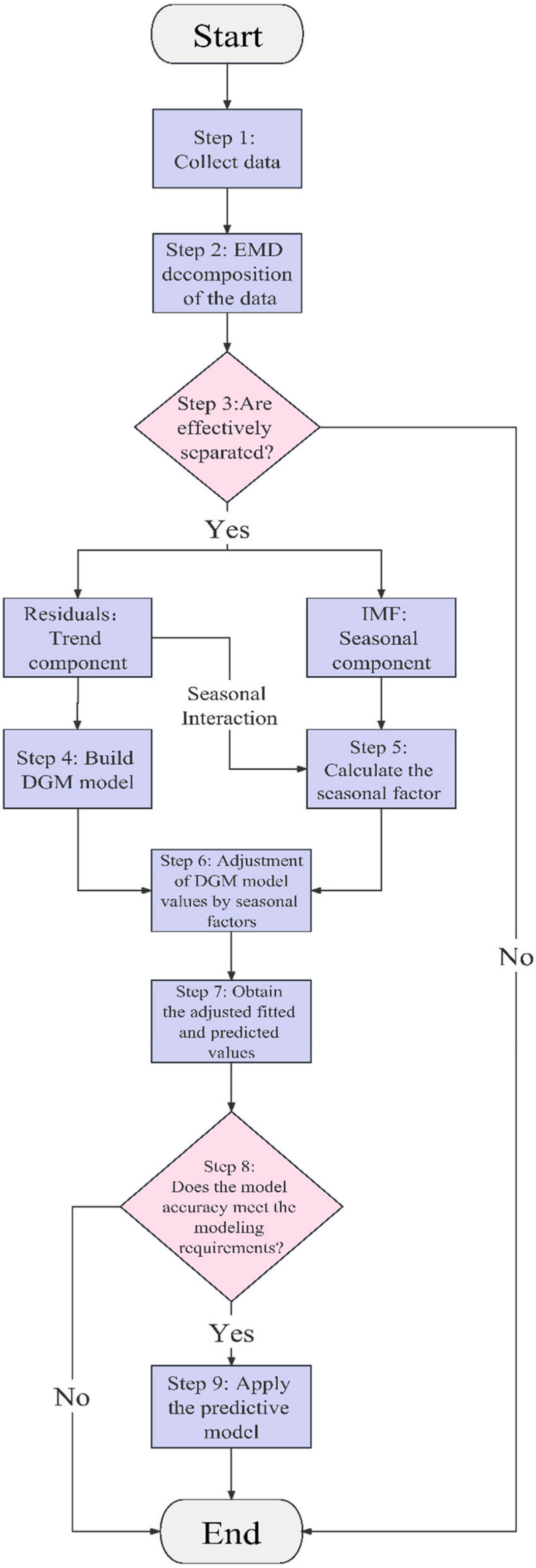

Fig. 4. Flow chart of EMD-DGM modeling.

The EMD method enables the decomposition of any type of signal, especially non-stationary and non-linear data [39]. Since wind power generation data are nonlinear and non-stationary, empirical modal decomposition helps to extract its features and thus improves model prediction capability. The specific algorithm flow chart is shown in Fig. 3.

2.3. Combined prediction method of EMD-DGM

Since complex time series are often composed of many interacting components, it is often difficult to build corresponding forecasting models directly based on the original data. To address this problem, this paper adopts a decomposition method for complex time series, and adopts a "divide and conquer" approach to build the corresponding prediction models separately [40]. In this paper, we adopt a "divide and conquer" approach to building the corresponding forecasting models.

Time series are usually assumed to be composed of four components: a trend component, a cyclic component, a periodic component, and an irregular component. The long-term trend component reflects the long-term pattern of change in a complex data series, while the cyclical component reflects the non-fixed-cycle change in the data series [41]. Since the long-term trend component and the cyclic component are difficult to distinguish in practical complex time series identification, the trend and cyclic components are often used together as the trend-cyclic component [42]. The cyclic component, compared to the cyclic component, reflects the fixed-period variation, while the irregular component reflects the irregular fluctuation of the complex time series, and the component is usually considered as the smooth component with bounded variance and zero means. The combination of complex data series components can be generally classified into two categories, which are additive form and multiplicative form [43]. In this paper, the multiplicative model is used to obtain the seasonal adjustment factor with the following equation

$$Y_t = S_t Y_t^T, t = 1, 2, 3, 4 \tag{7}$$

Where Y_t^T is the decomposition trend component, and S_t is the seasonal adjustment factor.

Let the series $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$ be a set of periodic fluctuation series, for which the EMD decomposition is applied to obtain the corresponding trend term series $X^{(0)T} = (x^{(0)T}(1), x^{(0)T}(2), \cdots, x^{(0)T}(n))$.

Assuming that the effect of the seasonal adjustment factor is fixed for the time series [43], the seasonal adjustment factor will be obtained as

$$S_i = \sum \frac{x^{(0)}(ki)}{x^{(0)}(ki)^T}, i = 1, 2, \cdots, T \tag{8}$$

Where i denotes the period T for each point in time within $x^{(0)}(ki)$ and $x^{(0)}(ki)^T$ denote the true values of the data in the series at the same time point as $x^{(0)}(k)$ the true value of the data at the same time point of each cycle and the corresponding value of the trend term after EMD decomposition.

Due to the late start and short development time of China's wind industry, less data can be collected, while China's wind power generation data shows complex characteristics such as cyclical seasonality, trend, and randomness. Traditional forecasting models are not effective in predicting them. In response to the above situation, this paper decomposes Chinese wind power generation data based on the EMD algorithm, and finds that the EMD algorithm can effectively strip cycle seasonality and trend, and its stripped trend term has an obvious quasi-exponential law, as the gray forecasting model can accurately predict the data with quasi-exponential law in the case of small data. Therefore, this paper constructs a discrete DGM gray prediction model for the trend term series and decomposes the seasonal index S_i according to Eqs. (7) and (8), and then uses the seasonal index to correct the predicted trend to establish the EMD-DGM model, as shown in Fig. 4. The algorithm flow profile is shown in Table 4.

Table 4
Algorithm flow.

Steps	Specific operation
Start	Initializing the software.
Step 1	Collection of cyclical seasonal data.
Step 2	Perform EMD decomposition of the data.
Step 3	Determine whether the period component and trend component are effectively separated, if they are effectively separated then go to the next step, if not, the algorithm ends.
Step 4	DGM modeling of trend components for trend prediction.
Step 5	The periodic components were separated from the raw data and the seasonal factors were calculated using the formula.
Step 6	Adjustment of the fitted and predicted values of the DGM model using seasonal factors.
Step 7	Obtain adjusted fitted and predicted values.
Step 8	Judge whether the fitting accuracy and prediction accuracy of the model meet the requirements, if they do, proceed to the next step, if they don't, the algorithm ends.
Step 9	Application of the model for specific case studies.
End	Exporting specific data.

2.4. Model evaluation indicators

To measure the prediction accuracy of a model scientifically and reasonably, it is often necessary to introduce corresponding evaluation indexes for analysis. The commonly used evaluation indexes for model prediction effectiveness mainly include mean absolute error (MAE), root mean square error (RMSE), absolute.

Pairwise percentage error (APE), mean absolute percentage error (MAPE).

1. Mean absolute error

The mean absolute error (MAE) is the average of the absolute errors between the predicted and true values, and is a linear score in which all individual differences are equally weighted on the mean.

$$\mathrm{MAE}=\frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}\left|\mathrm{e}(\mathrm{i})\right| \tag{9}$$

With $x^{(0)}(i)$ is the true value, and $\widehat{x}^{(0)}(i)$ is the predicted value, and $e(i)=x^{(0)}(i)-\widehat{x}^{(0)}(i)$ is the error value, then the mean absolute error (MAE) is calculated as

2. Root mean square error

Root mean square error (RMSE) is the average of the sum of squares of errors between the predicted and true values and then the square root, which enhances the role of errors with large values in the overall evaluation system and improves the sensitivity of the indicator.

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}e^{(2)}(i)} \tag{10}$$

$e(i)$ as defined above, then the root mean square error (RMSE) is calculated as

3. Absolute percentage error and average absolute percentage error

Absolute percentage error (APE) is a percentage that represents the percentage value between a single error and the true value, i.e., the percentage of a single error. Mean absolute percentage error (MAPE), on the other hand, is the average of the absolute percentage errors, and it is one of the most common metrics used to assess prediction accuracy.

$x^{(0)}(i), e(i)$ As shown above, the absolute percentage error (APE) and the mean absolute percentage error (MAPE) are calculated as

$$APE=\left|\frac{e(i)}{x^{(0)}(i)}\right|\times 100\% \tag{11}$$

$$MAPE=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{e(i)}{x^{(0)}(i)}\right|\times 100\% \tag{12}$$

The use of mean absolute percentage error (MAPE) to measure model prediction effectiveness usually follows the grading criteria shown in Table 5.

3. Case studies

3.1. Data sources and data processing methods

In the context of the goal of “carbon peaking and carbon neutrality”, wind power is a highly efficient and clean energy source that gradually replaces inefficient and polluting fossil energy sources. However, wind energy is an unstable renewable energy source, as shown in Fig. 5, the wind power generation in China is characterized by obvious trends and cyclical fluctuations, so this study investigates the changes in wind power generation in China from the perspective of cyclical seasonality and trends, which is of strategic significance to improve the energy structure and rational allocation of power resources [1]. This study is of strategic importance to improve the energy mix and rationalize the allocation of power resources.

Table 5
Prediction accuracy grading table.

MAPE (%)	Predicted Effect	MAPE (%)	Predicted Effect
<10	Good	20–50	General
10–20	Better	>50	Poor

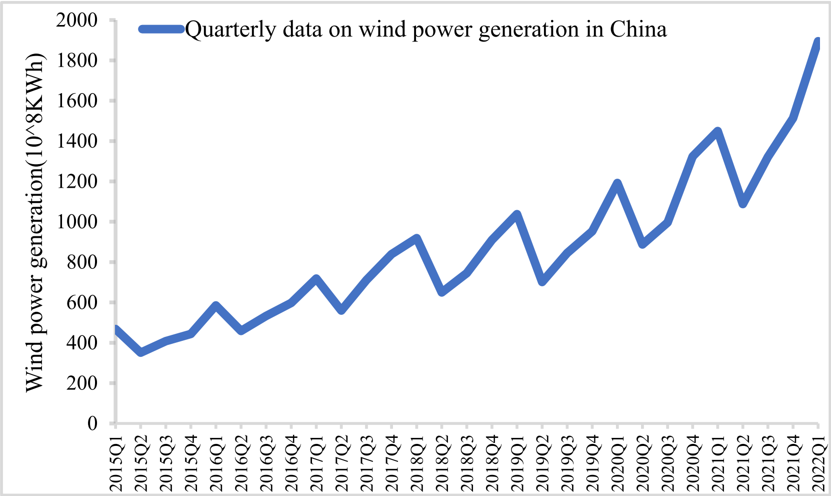

Fig. 5. China Wind Power Quarterly Data: 2015Q1-2022Q1
Data source: National Bureau of Statistics of China.

Such as the appeal, in order to compare the accuracy of the four models, this study selected the cyclical seasonal wind power generation data for the spring of China from 2015 to 2022 for the case study, and the data required for modeling were obtained from the National Bureau of Statistics of China, as shown in Table 6. Considering the latitudinal location of China, the months of March–May in a year are therefore classified as spring, June–August as summer, September–November as autumn, and December–February as winter. Spring, summer, autumn and winter are defined as Q1-Q4 in that order, abbreviated as 2015Q1-2022Q1, as shown in the specific data. It is obvious from the table that the seasonal variation of wind power generation is great, showing a general seasonal pattern of Q1>Q4>Q3>Q2 by and large. However, starting from the spring of 2020, the data of wind power generation gradually deviates from the overall trend. The reason behind this is mainly the impact of the new crown pneumonia epidemic, which led to the closure of a large number of factories and enterprises and brought some hindrance to the normal development of wind power. Among them, the data from 2015Q1-2020Q4 is used as the model training set, and the data from 2021Q1-2022Q1 is used as the model testing set.

3.2. Model construction

In Fig. 6 below, the raw data are decomposed by EMD to present obvious trend and period terms, where IMF is the period term and residue is the trend term.

In this paper, the traditional DGM(1,1) model, the DGGM(1,1) model based on seasonal grouping, and the Holt-Winters model are proposed as comparative models. Among them, the DGGM model with seasonal grouping is constructed as the GM(1.1) model, respectively, and the corresponding model time response equation is.

$$\hat{x}^{(1)}(1,t) = 3324.132e^{0.170(t-1)} - 2856.20$$

$$\hat{x}^{(1)}(2,t) = 2768.184e^{0.160(t-1)} - 2417.28$$

$$\hat{x}^{(1)}(3,t) = 3852.391e^{0.138(t-1)} - 3444.39$$

$$\hat{x}^{(1)}(4,t) = 3370.203e^{0.172(t-1)} - 2926.20$$

Which $\hat{x}^{(1)}(i,t), i = 1, 2, 3, 4$ stands for $Q1, Q2, Q3, Q4$.

Table 6
Wind power generation in China from 2015 to 2022 to spring: 10^8 KW h

	Q1	Q2	Q3	Q4
2015	468	350.9	408	444
2016	585.5	458.8	533.2	597.5
2017	718.3	559.5	711.2	839
2018	918.9	649.2	745.4	909.9
2019	1037.8	701.3	845.1	952.4
2020	1192.4	887.2	996	1323.7
2021	1449.5	1087.7	1321.4	1513.3
2022	1895			

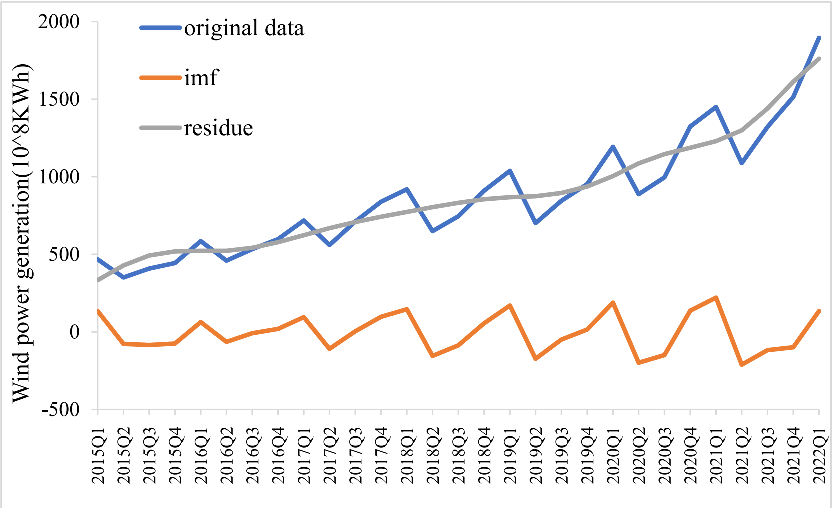

Fig. 6. EMD decomposition results for wind power in China (spring 2015 to spring 2022)
Data source: National Bureau of Statistics of China.

The indexes of the fitted Holt-Winters model are shown in Table 7. Using Equations (1)–(6), the time response of the original DGM model is

$$\widehat{x}^{(1)}(k+1) = 1.05^{k} * 8449 - 7981, k = 1, 2, \cdots, n-1$$

After decomposing the raw data into EMD data, the cycle impact factor was calculated using Eqs. (7) and (8), and the calculated seasonal factors are shown in Table 8 below. Using Equations (1)–(5), a DGM model was developed for the trend term, and its time response equation was

$$\widehat{x}^{(1)}(k+1) = 1.05^{k} * 8708.2 - 8240.2, k = 1, 2, \cdots, n-1$$

The data are reduced using equation (6), and then the DGM model is modified using seasonal factors to finally complete the decomposition, integration, and prediction of the data.

3.3. Model comparison

To verify the superiority of the model, we compared the EMD-DGM model with other related benchmark models Holt-Winters model, DGM model and DGGM model, and selected three indicators, MAE, RMSE and MAPE, as the evaluation criteria to compare the accuracy of the models, where MAE, RMSE and MAP are calculated by Equations (9)–(12). The actual, predicted and error values are shown in Table 9, and the evaluation results of the indicators are shown in Table 10. The distribution of the predicted and actual values of the four models is shown in Fig. 7.

The traditional DGM gray prediction model tends to have more satisfactory simulation and prediction effects for data series with near-exponential growth, but it is less applicable to the periodic fluctuation data similar to this study. From the model results, the simulation errors of the DGM gray prediction model for periodic fluctuation data are relatively large in both the training and test sets, and its MAE, RMSE, and MAPE indicators are higher than the corresponding indicators of other models. The value of its RMSE is more than twice that of the newly proposed EMD-DGM model, and the value of MAPE is as high as 13%, which is between 10% and 20%, with unsatisfactory simulation accuracy. If the DMG model is not adjusted and optimized, and the model is still used for subsequent forecasting, it will be difficult to achieve better forecasting results for cyclical seasonal data.

The DGGM(1,1) model divides the seasonal time series into several groups, constructs a GM(1,1) model for each group separately, and finally integrates them uniformly. This model has the best fitting accuracy in the training set, and its MAE and MAPE are 33.15 and 4%, respectively, with the smallest error, but the RMSE value is larger than that of EMD-DGM, which proves that its fitted data do not

Table 7
Fitting results of Holt-Winters model: 2015Q1-2022Q1.

Fitting statistics	
Stable R-side	0.346
R-side	0.959
RMSE	80.393

Table 8
Seasonal indices (spring 2015 to spring 2022).

	Q1	Q2	Q3	Q4
Seasonal Index	1.16201	0.8261	0.91846	1.02331

Table 9
Forecast values and errors generated using four different models: 2015Q1-2022Q1.

Time	Actual value	DGM		DGGM		Holt-Winters		EMD-DGM	
		Forecasted value	Error (%)	Forecasted value	Error (%)	Forecasted value	Error (%)	Forecasted value	Error (%)
Training Stage									
2015Q1	468			468	0.00	457.53	2.24	397.89	14.98
2015Q2	350.9			454.32	29.47	358.32	2.11	382.90	9.12
2015Q3	408			474.26	16.24	426.77	4.60	445.74	9.25
2015Q4	444			495.08	11.50	476.18	7.25	532.18	19.86
2016Q1	585.5			516.82	11.73	592.14	1.13	629.61	7.53
2016Q2	458.8			539.51	17.59	443.98	3.23	452.68	1.33
2016Q3	533.2			563.19	5.62	528.36	0.91	526.98	1.17
2016Q4	597.5			587.91	1.61	589.1	1.41	629.17	5.30
2017Q1	718.3			613.72	14.56	759.22	5.70	744.36	3.63
2017Q2	559.5	582	4.02	640.66	14.51	559.12	0.07	535.19	4.35
2017Q3	711.2	651.2	8.44	668.79	5.96	644.89	9.32	623.03	12.40
2017Q4	839	803.8	4.20	698.15	16.79	739.25	11.89	743.84	11.34
2018Q1	918.9	939.1	2.20	728.8	20.69	967.69	5.31	880.02	4.23
2018Q2	649.2	728	12.14	760.79	17.19	717.41	10.51	632.73	2.54
2018Q3	745.4	878.2	17.82	794.19	6.55	810.49	8.73	736.58	1.18
2018Q4	909.9	936.7	2.95	829.05	8.89	864.54	4.99	879.41	3.35
2019Q1	1037.8	948.9	8.57	865.45	16.61	1036.02	0.17	1040.41	0.25
2019Q2	701.3	657.2	6.29	903.44	28.82	770.64	9.89	748.05	6.67
2019Q3	845.1	742	12.20	943.1	11.60	873.24	3.33	870.83	3.04
2019Q4	952.4	927.43	2.62	984.5	3.37	984.6	3.38	1039.69	9.17
2020Q1	1192.4	1247.17	4.59	1027.72	13.81	1119.69	6.10	1230.04	3.16
2020Q2	887.2	789.1	11.06	1072.84	20.92	827.69	6.71	884.38	0.32
2020Q3	996	1078.6	8.29	1119.93	12.44	1015.72	1.98	1029.54	3.37
2020Q4	1323.7	1283.83	3.01	1169.1	11.68	1145.25	13.48	1229.18	7.14
Test Stage									
2021Q1	1449.5	1438.07	0.79	1220.42	15.80	1033.53	0.99	1454.22	0.33
2021Q2	1087.7	1164.17	7.03	1273.99	17.13	1194.66	4.98	1045.57	3.87
2021Q3	1321.4	1237.97	6.31	1329.92	0.64	1417.96	9.59	1217.18	7.89
2021Q4	1513.3	1667.2	10.17	1388.3	8.26	1601.15	6.30	1453.21	3.97
2022Q1	1895	1774.8	6.34	1449.25	23.52	1033.53	15.51	1719.26	9.27

Table 10
Prediction accuracy and performance evaluation of the four models (2015Q1-2022Q1).

Model	MAE	RMSE	MAPE
Training Stage			
DGM	97.65	113.39	13%
DGGM	33.15	86.12	4%
Holt-Winters	40.58	60	5.02%
EMD-DGM	40.64	50.42	6%
Test Stage			
DGM	198.92	245.58	13%
DGGM	96.3	124.76	6.6%
Holt-Winters	142.53	169.14	9.09%
EMD-DGM	77.38	97	5%

capture the characteristics of periodic seasonal data as well as the EMD-DGM model. And the model has a MAPE value of 6.6% in the test set, which is not the best fit compared with other models, although the difference between the before and after fitting accuracy is not significant. The reason for this is that the DGGM(1,1) model is relatively susceptible to the influence of random disturbances. The prediction accuracy of the DGGM(1,1) model decreases significantly when the amount of modeled data is small and abnormal data appear. Combined with the current events, this study speculates that the presence of the epidemic has caused the wind power data series to deviate from the overall trend to some extent. And due to the exponential form of the model, the subsequent error will increase rapidly if the model continues to be used for forecasting.

The Holt-Winters model introduces a seasonal term based on the Holt model, which can be used to deal with the behavior of

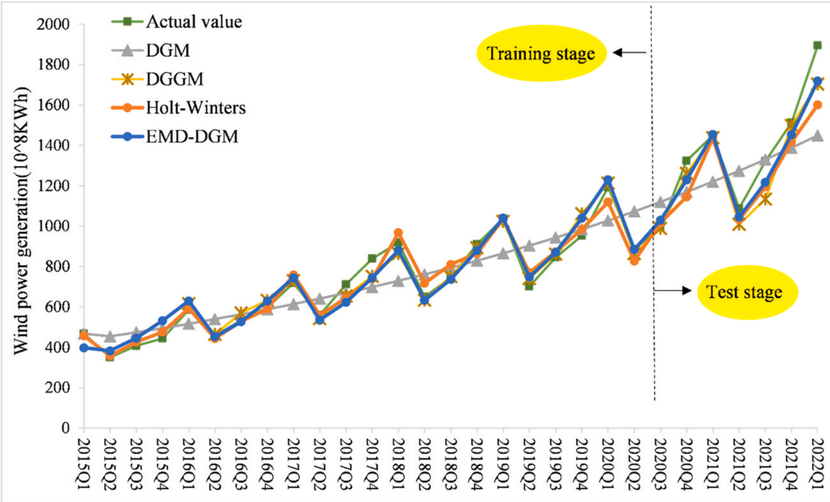

Fig. 7. Distribution of actual and forecast quarterly values conducted by the four models: 2015Q1-2022Q1.

periodic fluctuations in the time series, and is suitable for non-stationary series containing linear trends and periodic fluctuations. The model fits relatively well in the training set, and the prediction accuracy is basically the same compared with the EMD-DGM model, but the fitting effect in the test set is far inferior to the latter, with MAE value of 142.53, RMSE value of 169.14, and MAPE value of 9.09%, and the values of each index are approximately twice that of the EMD-DGM model, which shows that the Holt-Winters model does not have an advantage in the trend portrayal of the test set.

The EMD-DGM model fits well on the training and test sets. Its MAE, and MAPE values ranked second among the comparison models in the training set and first in the test set, indicating the high fitting accuracy of the model. From the distribution plots of the predicted and actual values of the EMD-DGM model, the model portrays the seasonal trends and key turning points of the series data more accurately, and its RMSE values are the smallest in both the training and test sets. The EMD-DGM model uses the EMD algorithm to decompose the seasonal time series and trend characteristics, constructs a discrete DGM gray prediction for the trend characteristics of the data series model, decompose the seasonal indices, and then use the seasonal indices to correct the predicted trends. Overall, the EMD-DGM model can better simulate the seasonal variation pattern of wind power generation in China, and has better adaptability compared with the DGM, DGGM, and Holt-Winters forecasting models.

3.4. China wind power forecast 2022–2027

Based on the superiority of the EMD-DGM model, this study will use the model to forecast the wind power generation in China from summer 2022 to winter 2027, and the forecast results are important for power companies to develop operational strategies, ensure the stable development of the wind power industry, and the government to develop policies to adjust the energy structure and promote carbon emission reduction based on the forecast results. In order to make the forecast results more ideal, all data from spring 2015 to spring 2022 are included in the modeling process in this study. The trend term of EMD decomposition is modeled as DGM, whose time response equation is.

$$\widehat{x}^{(1)}(k+1) = 1.05^k * 8573.64 - 8240.2, k = 1, 2, \cdots, n-1.$$

The data were reduced using Equation (6), and then the DGM model predicted data were corrected using the seasonal indices in Table 4 to complete the modeling of EMD-DGM, and the predicted values in Table 11 were obtained, and the correlation distribution is shown in Fig. 8.

According to the model forecast results in Table 11, China's wind power generation will reach 485,643 million kWh by the winter

Table 11
EMD-DGM model predicted electricity consumption results from 2022Q2-2023Q4: 10ˆ8 KW h

Year	EMD-DGM model prediction results			
	Q1	Q2	Q3	Q4
2022		1360.810	1587.494	1855.875
2023	2211.259	1649.490	2019.087	2249.583
2024	2680.361	1999.418	2332.493	2726.814
2025	3248.979	2423.587	2827.313	3305.291
2026	3938.226	2937.735	3427.114	4006.483
2027	4773.699	3560.953	4154.148	4856.434

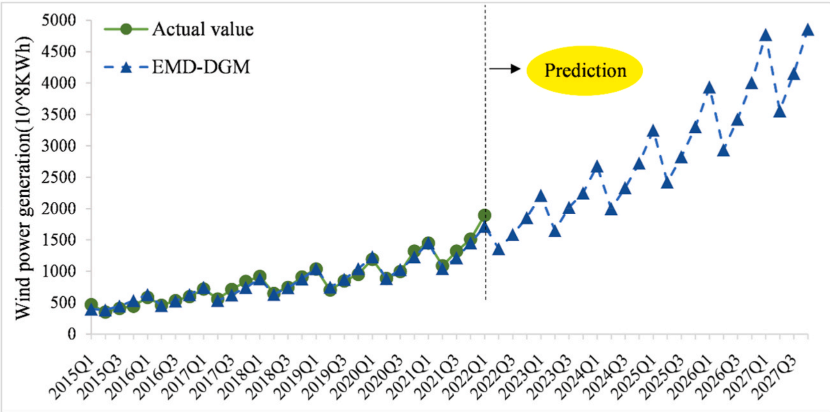

Fig. 8. Distribution of predicted and actual values: 2015Q1-2023Q4.

of 2027. This study finds that China's wind power generation will continue to exhibit trend and seasonal characteristics, with steady growth while still featuring seasonal fluctuations. Unlike historical data, wind power generation is higher in winter relative to spring over the next five years, while it remains at its lowest level in summer, and the gap between the maximum and minimum values is gradually increasing. The seasonal fluctuations of wind power cannot be ignored, and the relevant power departments need to allocate and regulate the power resources reasonably to ensure the smooth implementation of various activities.

3.5. Retesting the prediction accuracy of forecasting models

In order to study the effectiveness of the prediction model proposed in this paper, the wind power generation data from 2022Q2 to 2022Q4 have been updated at the National Bureau of Statistics of China at the time of this paper becoming a manuscript, and for this reason, in order to verify the ability of the model to predict the future, it is compared to further verify the prediction ability of the model. The results are shown in Table 12. Overall, the forecasting models are all less than 10%, which proves their excellent forecasting ability and further validates the excellent performance of the model proposed in this paper in dealing with cyclical seasonal data in the case of small data.

4. Policy recommendations and conclusion

4.1. Policy recommendations

The established EMD-DGM model has made an accurate forecast of wind power generation in China. In this regard, through the analysis of the forecast results, this paper makes some targeted suggestions on the seasonal characteristics of wind power generation in China, the relationship between new power generation and social demand for electricity, the geographical distribution and types of wind power generation, and the improvement of power generation efficiency.

(1) Synergistically develop multiple power generation modes, and quickly adjust energy structure

According to the "Opinions on Comprehensively and Accurately Implementing the New Development Concept and Doing a Good Job in Carbon Emission and Carbon Neutral Work" issued by the Chinese State Council in October 2021, the proportion of non-fossil energy consumption in China is expected to reach about 20% by 2025 and more than 80% by 2060. And according to the Energy Research Institute of the National Development and Reform Commission, it is expected that wind power will account for 38.5% of energy consumption by 2050, a predicted result that is still a long way from the goal of reaching 80% of non-fossil energy consumption. At the same time, Fig. 9 points out that between 2014 and 2021, the electricity demand of the whole society far exceeds the amount of wind power generation, and according to the forecast results, the current development rate of wind power generation is far from enough to supply the whole society with electricity, and as the proportion of fossil energy generation decreases, the contradiction of electricity consumption in China will be further highlighted, which requires China's policy to increase policy support, so that a variety

Table 12
Retesting of forecast accuracy: 2022Q2 - 2022Q4.

Quarterly	True Value	Predicted value	Prediction error (%)
2022Q2	1443.3	1360.81	5.1%
2022Q3	1691.4	1587.494	6.1%
2022Q4	1989.8	2029.8	8.4%

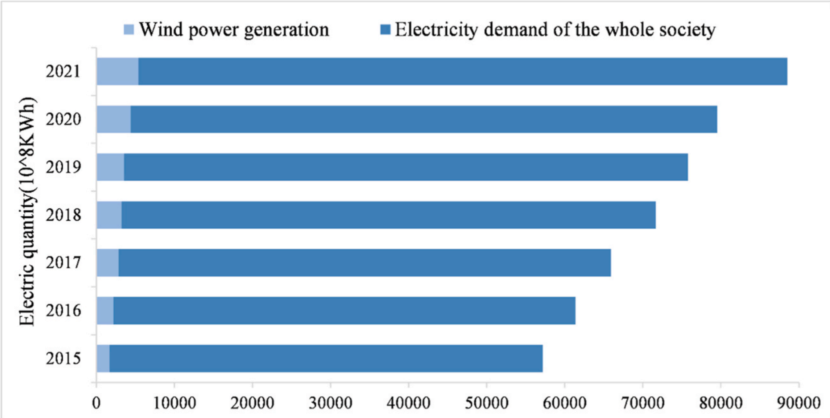

Fig. 9. Electricity demand for the whole society: 2014-2021
Data source: National Bureau of Statistics of China.

of clean energy generation methods synergistic development, to accelerate the adjustment of energy structure. This will require China's policy to increase policy support, so that multiple clean energy generation methods can be developed in a coordinated manner to accelerate the adjustment of the energy structure.

According to the above analysis and forecast, the seasonal characteristics of China's wind power generation will continue, with an uneven pattern of low power generation in summer and autumn and high-power generation in spring and winter. However, the hot weather in summer and autumn has always been the peak of electricity consumption, so we should adopt various clean energy generation methods such as hydroelectric power, nuclear power and tidal power to make up for the electricity demand gap. For the current situation in China, the most effective way is to adopt wind power and hydropower to complement each other. Based on China's climate characteristics, summer and autumn are China's rainy seasons with high precipitation, and China is also a mountainous country with a high potential energy of water, which has a good foundation for hydroelectric power generation. Adopting the complementary power generation method of wind power and hydropower can solve the dilemma of high electricity demand but low wind power generation in summer and autumn in China.

(2) Vigorously develop onshore wind power, and steadily promote offshore wind power

With ample wind and land resources, low costs and huge market potential, onshore wind is currently one of the most competitive sources of new power generation. Currently, China is the largest onshore wind power market, accounting for about one-third of the world's installed capacity. The vigorous development of onshore wind power is the best path to achieve the dual carbon goal and drive economic development. Compared to onshore wind power, offshore wind power has higher construction costs and technical difficulties, and longer construction periods. However, offshore wind power also has its unique advantages, with higher energy efficiency, power generation efficiency and average service life of wind resources. And offshore wind power does not occupy land resources, generally built in the coastal area, and the coastal area of electricity demand is large, so it can also significantly reduce transmission costs. With the continuous upgrading of technology policy boost, the cost of offshore wind power cost gradually reduced, and the scale effect also emerged, which will usher in a booming wave of development.

(3) Enhance long-distance transmission technology, and accelerate the construction of extra-high voltage backbone channels

The main wind power stations in China are located in the northwest, northeast, north China, and coastal areas, as shown in Fig. 10, and the overall pattern of "west-to-east" and "north-to-south" power transmission requires consideration of long-distance transmission losses and grid construction. Therefore, enhancing long-distance transmission technology and improving the construction of transmission networks is the basic condition to ensure the reasonable distribution of power resources. Accelerating the construction of extra-high voltage backbone channels not only enables high-capacity, low-loss and high-efficiency power transmission, but also enhances cross-region and cross-province power exchange capacity, improves the flexibility and reliability of power grid operation, and is of great significance to the rational allocation of power resources.

4.2. Conclusion

In this paper, we propose a combined EMD-DGM forecasting model based on EMD data decomposition, which can effectively capture the trend and periodicity of cyclical seasonal data. Comparing the fitting results of the SARIMA model, DGM model, and DDGM model, the EMD-DGM model possesses high fitting accuracy in both training and prediction sets and has good adaptability to cyclical seasonal data. The main conclusions of this paper are as follows.

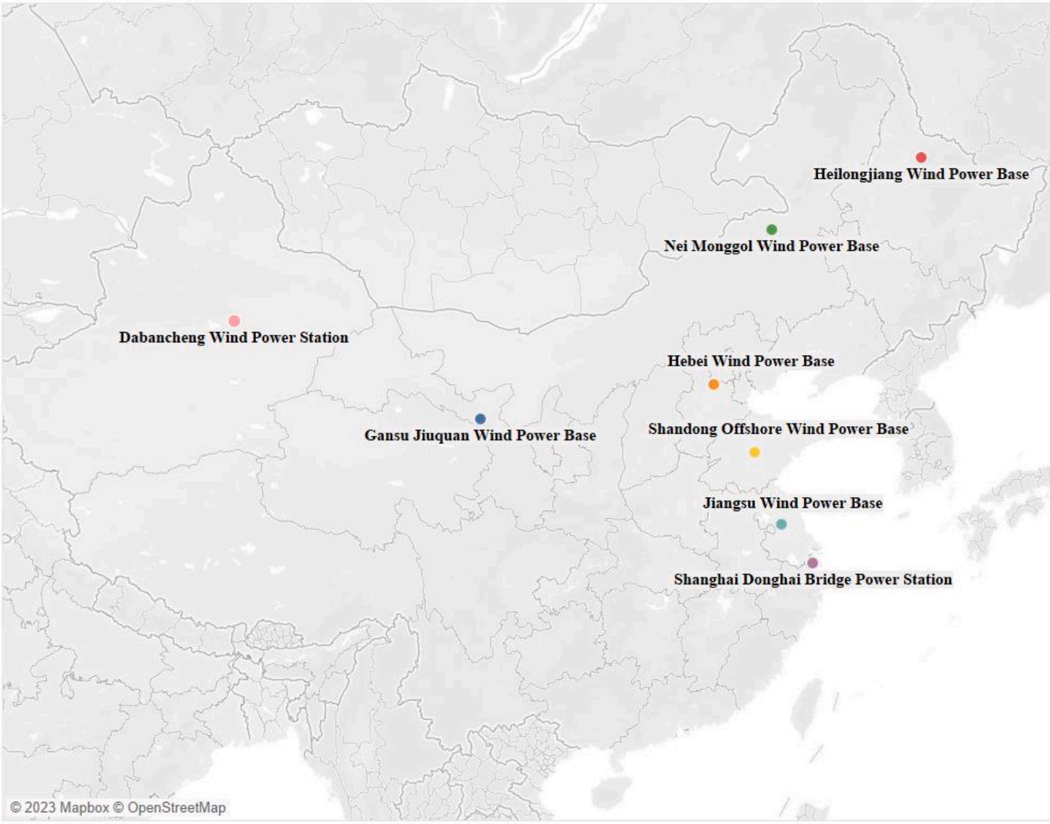

Fig. 10. Distribution of major wind power stations in China
Data source: National Energy Administration of China.

(1) In this paper, a parallel combined forecasting model is established by EMD decomposition and seasonal factors, which enhances the interpretability of the decomposed data and the interpretability of the forecasting model.
(2) The traditional DGM gray forecasting model has a good forecasting effect for data with exponential patterns, but it does not apply to data with trend and periodicity characteristics. In this paper, the EMD-DGM forecasting model is proposed by decomposing the data into trend and periodic terms through EMD and calculating the seasonal index based on the multiplicative seasonal model. The model achieves good forecasting accuracy on the Chinese 2015–2022 wind power dataset.
(3) After forecasting the Chinese wind power data set, Chinese wind power will continue to show seasonal cyclical characteristics. And the development rate of wind power is not able to keep up with the development rate of China's social electricity consumption. Therefore, based on the current situation in China, this paper proposes the coordinated development of multiple power generation modes, accelerating energy structure adjustment, vigorously developing onshore wind power, steadily promoting offshore wind power, upgrading long-distance transmission technology, and accelerating the construction of extra-high voltage backbone channels.
(4) EMD, as an adaptive data decomposition algorithm, has excellent data decomposition capability and can effectively extract the characteristics of cyclical seasonal data, but the "endpoint effect" of EMD is still obvious, and in this case, there is an obvious "endpoint anomaly" in 2015Q1. For this reason, trying a better data decomposition algorithm may significantly improve the prediction accuracy of the model. Second, this paper assumes that the seasonal impact factor is constant, however, in real life, sudden climate changes and weather anomalies occur, and the seasonal impact factor is constantly changing. In order to further improve the prediction accuracy of the model, how to construct seasonal factors with variable weights by using each year's data set for weighting will be the key to further research. So far, we have some ideas, but they are not comprehensive, and we share them here.

Let the periodic seasonal series be X with period n, $\mathrm{n} = 1, 2, 3, 4, \ldots n$, and m be the components of its period, $m = 1, 2, 3, 4$. Decompose it by EMD, and the trend component of the decomposed series is $\mathrm{TR(nm)}, n = 1, 2, 3, 4, \ldots n,$ and the periodic component is $PE(nm)$, $n = 1, 2, 3, 4, \ldots n$. Let the seasonal factor be Q_{ns}, $s = 1, 2, 3, 4$. Let $X = \mathrm{TR(nm)} * Q$, at this time, $Q_s = \frac{X}{\mathrm{TR(nm)}}$. Arrange X into a matrix form T according to the period n as follows.

$$T=\begin{array}{c} \\ Q1\\ Q2\\ Q3\\ Q4 \end{array}\overbrace{\begin{bmatrix} 1 & 2 & 3 & \dots & n\\ \mathrm{TR}(11)*Q_{11} & \mathrm{TR}(21)*Q_{21} & \mathrm{TR}(31)*Q_{31} & \dots & \mathrm{TR}(\mathrm{n}1)*Q_{n1}\\ \mathrm{TR}(12)*Q_{12} & \mathrm{TR}(22)*Q_{22} & \mathrm{TR}(32)*Q_{32} & \dots & \mathrm{TR}(\mathrm{n}1)*Q_{n2}\\ \mathrm{TR}(13)*Q_{13} & \mathrm{TR}(23)*Q_{23} & \mathrm{TR}(33)*Q_{33} & \dots & \mathrm{TR}(\mathrm{n}1)*Q_{n3}\\ \mathrm{TR}(14)*Q_{14} & \mathrm{TR}(23)*Q_{23} & \mathrm{TR}(33)*Q_{33} & \dots & \mathrm{TR}(\mathrm{n}1)*Q_{n4} \end{bmatrix}}^{\text{Periodicity}}$$

At this time, the seasonal factor of each cycle can be obtained, and in fact, the seasonal factor proposed in this paper is the average of its seasonal factor of each cycle. The difficulty of the current study is that (1) sudden weather is uncertain and its impact cycle after the occurrence of sudden weather is also uncertain. This means that although dynamic seasonal factors can be constructed, it is not possible to choose which cycle length seasonal factor fragment to predict. (2) Even if we artificially determine a seasonal factor fragment, it is still static in the forecasting process, so its dynamization in forecasting still needs to be solved.

Author contribution statement

Minghao Ran: conceived and designed the experiments; performed the experiments; analyzed and interpreted the data; wrote the paper.

Jindi Huang: conceived and designed the experiments; performed the experiments; wrote the paper.

Wuyong Qian: conceived and designed the experiments; contributed reagents, materials, analysis tools or data.

Tingting Zou: analyzed and interpreted the data.

Chunyi Ji: analyzed and interpreted the data.

Data availability statement

Data included in article/supplementary material/referenced in article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Humanities and Social Science Foundation of Ministry of Education (22YJAZH033), the major project of philosophy and social science research in colleges and universities in Jiangsu Province (2021SJZDA134), the National Natural Science Foundation of China (71871106) and the Fundamental Research Funds for the Central Universities (JUSRP321016).

References

[1] Project Team on the Strategy and Pathway for Peaking Carbon Emissions and Carbon Neutrality, Analysis of a peaked carbon emission pathway in China toward carbon neutrality, Eng. Times 7 (2021) 1673–1677.

[2] G. Msigwa, J.O. Ighalo, P.-S. Yap, Considerations on environmental, economic, and energy impacts of wind energy generation: projections towards sustainability initiatives, Sci. Total Environ. 849 (2022), 157755, https://doi.org/10.1016/J.SCITOTENV.2022.157755.

[3] X. Zhang, X. Zhao, Z. Jiang, S. Shao, How to achieve the 2030 CO2 emission-reduction targets for China's industrial sector: retrospective decomposition and prospective trajectories, Global Environ. Change 44 (2017) 83–97.

[4] G. Sun, C. Jiang, P. Cheng, Y. Liu, X. Wang, Y. Fu, Y. He, Short-term wind power forecasts by a synthetical similar time series data mining method, Renew. Energy 115 (2018) 575–584, https://doi.org/10.1016/j.renene.2017.08.071.

[5] Z. Wu, X. Huang, R. Chen, X. Mao, X. Qi, The United States and China on the paths and policies to carbon neutrality, Environ. Manag. 320 (2022), 115785, https://doi.org/10.1016/J.JENVMAN.2022.115785.

[6] N. Khasanzoda, I. Zicmane, S. Beryozkina, M. Safaraliev, S. Sultonov, A. Kirgizov, Regression model for predicting the speed of wind flows for energy needs based on fuzzy logic, Renew. Energy 191 (2022) 723–731, https://doi.org/10.1016/J.RENENE.2022.04.017.

[7] M. Ghofrani, R. Azimi, Ghayekhloo, A Hybrid Wind Power Forecasting Model Based on Data Mining and Wavelets Analysis, Energy Conversion and Management, 2016.

[8] J. Jung, R.P. Broadwater, Current status and future advances for wind speed and power forecasting, Renew. Sustain. Energy Rev. 31 (2014) 762–777, https://doi.org/10.1016/j.rser.2013.12.054.

[9] H. Liu, C. Chen, X. Lv, X. Wu, M. Liu, Deterministic wind energy forecasting: a review of intelligent predictors and auxiliary methods, Energy Convers. Manag. 195 (2019) 328–345, https://doi.org/10.1016/j.enconman.2019.05.020.

[10] A. Aghajani, R. Kazemzadeh, A. Ebrahimi, A novel hybrid approach for predicting wind farm power production based on wavelet transform, hybrid neural networks and imperialist competitive algorithm, Energy Convers. Manag. 121 (2016) 232–240, https://doi.org/10.1016/j.enconman.2016.05.024.

[11] A. Aghajani, R. Kazemzadeh, A. Ebrahimi, A novel hybrid approach for predicting wind farm power production based on wavelet transform, hybrid neural networks and imperialist competitive algorithm, Energy Convers. Manag. 121 (2016) 232–240, https://doi.org/10.1016/j.enconman.2016.05.024.

[12] E. Yatiyana, S. Rajakaruna, A. Ghosh, Wind speed and direction forecasting for wind power generation using ARIMA model, 2017 Australasian Universities Power Engineering Conference, AUPEC 2017. 2017-November, 1–6, https://doi.org/10.1109/AUPEC.2017.8282494, 2018.

[13] H. Liu, E. Erdem, J. Shi, Comprehensive evaluation of ARMA-GARCH(-M) approaches for modeling the mean and volatility of wind speed, Appl. Energy 88 (2011) 724–732, https://doi.org/10.1016/j.apenergy.2010.09.028.

[14] Y. Wang, Z. Xie, Q. Hu, S. Xiong, Correlation aware multi-step ahead wind speed forecasting with heteroscedastic multi-kernel learning, Energy Convers. Manag. 163 (2018) 384–406, https://doi.org/10.1016/j.enconman.2018.02.034.

[15] J. Dowell, P. Pinson, Very-short-term probabilistic wind power forecasts by sparse vector autoregression, IEEE Trans. Smart Grid 7 (2016) 763–770, https://doi.org/10.1109/TSG.2015.2424078.
[16] O. Karakuş, E.E. Kuruoglu, M.A. Altinkaya, One-day ahead wind speed/power prediction based on polynomial autoregressive model, IET Renew. Power Gener. 11 (2017) 1430–1439, https://doi.org/10.1049/iet-rpg.2016.0972.
[17] H. Cai, X. Jia, J. Feng, W. Li, Y.M. Hsu, J. Lee, Gaussian Process Regression for numerical wind speed prediction enhancement, Renew. Energy 146 (2020) 2112–2123, https://doi.org/10.1016/j.renene.2019.08.018.
[18] A. Lahouar, J. Ben Hadj Slama, Hour-ahead wind power forecast based on random forests, Renew. Energy 109 (2017) 529–541, https://doi.org/10.1016/j.renene.2017.03.064.
[19] A. Khosravi, R.N.N. Koury, L. Machado, J.J.G. Pabon, Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive neuro-fuzzy inference system, Sustain. Energy Technol. Assessments 25 (2018) 146–160, https://doi.org/10.1016/j.seta.2018.01.001.
[20] H. Demolli, A.S. Dokuz, A. Ecemis, M. Gokcek, Wind power forecasting based on daily wind speed data using machine learning algorithms, Energy Convers. Manag. 198 (2019), 111823, https://doi.org/10.1016/j.enconman.2019.111823.
[21] L. Luo, H. Li, J. Wang, J. Hu, Design of a combined wind speed forecasting system based on decomposition-ensemble and multi-objective optimization approach, Applied Mathemat. Models 89 (2021) 49–72, https://doi.org/10.1016/J.APM.2020.07.019.
[22] S.X. Lv, L. Wang, Deep learning combined wind speed forecasting with hybrid time series decomposition and multi-objective parameter optimization, Appl. Energy 311 (2022), 118674, https://doi.org/10.1016/J.APENERGY.2022.118674.
[23] Z. Wu, L. Xiao, A secondary decomposition based hybrid structure with meteorological analysis for deterministic and probabilistic wind speed forecasting, Appl. Soft Comput. 85 (2019), 105799, https://doi.org/10.1016/J.ASOC.2019.105799.
[24] Y. Zhang, J. Han, G. Pan, Y. Xu, F. Wang, A multi-stage predicting methodology based on data decomposition and error correction for ultra-short-term wind energy prediction, J. Clean. Prod. 292 (2021), 125981, https://doi.org/10.1016/J.JCLEPRO.2021.125981.
[25] C. Li, Z. Xiao, X. Xia, W. Zou, C. Zhang, A hybrid model based on synchronous optimisation for multi-step short-term wind speed forecasting, Appl. Energy 215 (2018) 131–144, https://doi.org/10.1016/J.APENERGY.2018.01.094.
[26] E. Zhao, S. Sun, S. Wang, New developments in wind energy forecasting with artificial intelligence and big data: a scientometric insight, Data Sci. Manag. 5 (2022) 84–95, https://doi.org/10.1016/j.dsm.2022.05.002.
[27] B. Zeng, S.F. Liu, Z.G. Fang, N.M. Xie, Grey combination prediction model and its application, China Manag. Sci. 17 (2009) 150–155.
[28] Z.X. Wang, Y.Q. Jv, A non-linear systematic grey model for forecasting the industrial economy-energy-environment system, Technol. Forecast. Soc. Change 167 (2021), https://doi.org/10.1016/j.techfore.2021.120707.
[29] X. Xiao, J. Yang, S. Mao, J. Wen, An improved seasonal rolling grey forecasting model using a cycle truncation accumulated generating operation for traffic flow, Appl. Math. Model. 51 (2017) 386–404, https://doi.org/10.1016/j.apm.2017.07.010.
[30] Z.X. Wang, Q. Li, L.L. Pei, Grey forecasting method of quarterly hydropower production in China based on a data grouping approach, Appl. Math. Model. 51 (2017) 302–316, https://doi.org/10.1016/j.apm.2017.07.003.
[31] S. Ding, A novel self-adapting intelligent grey model for forecasting China's natural-gas demand, Energy 162 (2018) 393–407, https://doi.org/10.1016/j.energy.2018.08.040.
[32] W. Qian, J. Wang, An improved seasonal GM(1,1) model based on the HP filter for forecasting wind power generation in China, Energy 209 (2020), https://doi.org/10.1016/j.energy.2020.118499.
[33] W. Wu, X. Ma, B. Zeng, Y. Wang, W. Cai, Forecasting short-term renewable energy consumption of China using a novel fractional nonlinear grey Bernoulli model, Renew. Energy 140 (2019) 70–87, https://doi.org/10.1016/j.renene.2019.03.006.
[34] S. Guefano, J.G. Tamba, T.E.W. Azong, L. Monkam, Forecast of electricity consumption in the Cameroonian residential sector by Grey and vector autoregressive models, Energy 214 (2021), https://doi.org/10.1016/j.energy.2020.118791.
[35] W. Qian, A. Sui, A novel structural adaptive discrete grey prediction model and its application in forecasting renewable energy generation, Expert Syst. Appl. (2021) 186, https://doi.org/10.1016/j.eswa.2021.115761.
[36] N.M. Xie, S.F. Liu, Modeling Mechanisms of Discrete GM(1,1) Models and Gray Prediction Models, Systems Engineering Theory and Practice, 2005.
[37] D. Kim, K.O. Kim, Extending the scope of empirical mode decomposition by smoothing, J. Adv. Signal Proces. (2012) 168.
[38] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, H.H. Liu, The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis, vol. 454, Proceedings of the Royal Society of London Series A, 1998.
[39] N.E. Huang, Z. Wu, A review on Hilbert-Huang transform: method and its applications to geophysical studies, Rev. Geophys. 46 (2008).
[40] M. Ghofrani, R. Azimi, Ghayekhloo, A Hybrid Wind Power Forecasting Model Based on Data Mining and Wavelets Analysis, Energy Conversion and Management, 2016.
[41] S. Hylleberg, Seasonality in regression, 221-226, Academic Press, 1986, https://doi.org/10.1016/B978-0-12-363455-9.50014-X.
[42] I.J. Lu, C. Lewis, S.J. Lin, The forecast of motor vehicle, energy demand and CO2 emission from Taiwan's road transportation sector, Energy Pol. 37 (2009) 2952–2961, https://doi.org/10.1016/J.ENPOL.2009.03.039.
[43] B. Abraham, J. Ledolter, Statistical Methods for Forecasting, Wiley, 1983.