Forecasting electricity consumption in Shanghai using a nonlinear interval grey model

Zhiqiang WU

University of Shanghai for Science and Technology

Yan GAO (✉ gaoyan@usst.edu.cn)

University of Shanghai for Science and Technology

Bo WANG

University of Shanghai for Science and Technology

Research Article

Keywords: Electricity forecasting, Electricity consumption, Interval grey model, Genetic algorithm, BP neural network

Posted Date: September 8th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-831308/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Forecasting electricity consumption in Shanghai using a nonlinear interval grey model

Zhiqiang Wu¹, Yan Gao^{1,*}, Bo Wang¹

University of Shanghai for Science and Technology, Shanghai, 200093, China

*Corresponding author.

E-mail: gaoyan@usst.edu.cn (Yan Gao)

Abstract: Carbon emissions from fossil energy not only cause a lot of extreme weathers, but also global warming. Accurately forecasting of electricity demand can promote the development of the renewable energy, which is vital to achieving the goal of carbon peak and carbon neutrality. In this paper, a nonlinear interval grey model based on genetic algorithm and BP neural network optimization (BPGA-IGM (1,1)) is proposed to predict electricity consumption. Firstly, based on the forecast of China's energy consumption and China's coal consumption, the reliability and superiority of the BPGA-IGM (1,1) model have been verified. Then, the model and other competing models are applied to forecast Shanghai's electricity consumption. The empirical results show that the model designed in this paper could obtain more accurate and reliable prediction results. According to the empirical results, Shanghai's electricity consumption continues to rise to a higher level of no less than 1978.19 million kWh by 2025. On the basis of this issue, several suggestions have been offered.

Keywords: Electricity forecasting; Electricity consumption; Interval grey model; Genetic algorithm; BP neural network;

1 Introduction

With the increasing energy demand for economic development, the large amount of carbon emissions generated by fossil energy not only causes a lot of extreme weathers, but also makes global warming (Naqvi et al. 2020). Energy-saving and emission-reduction technologies to solve energy shortages and ecological imbalances have become the most concerned topic in the world (Chen et al. 2021). As an important part of the world economy, China's economic development is inseparable from energy consumption, which means the importance of energy forecasting technology to China's economic development. In particular, with the implementation of "carbon peak" and "carbon neutral" target in China, the realization of the dual carbon goals requires more reliable and accurate energy forecasting technology.

Electricity consumption is closely connected with people's daily lives as well as the industrial production. Although the reliability of the power industry is of great significance to social development, it is also the main source of carbon dioxide emissions (Chen et al. 2020). Therefore, as the hub of the energy chain, the power industry is not only an important support for the economy and society, but also an important participant in the emission chain. According to relevant literature, 80% of China's carbon dioxide emissions come from electricity production activities. And as the level of global energy consumption continues to rise, power energy as one of the

most important secondary energy sources, occupies a relatively high proportion in global energy consumption. By the end of 2020, China's electricity consumption was 5,200 kWh/person, OECD countries 8,000 kWh/person and the United States 128,000 kWh/person. As a developing country, China's total consumption of electricity energy will continue to grow. And China's carbon peak and carbon neutrality are also implemented under this circumstance. On the other hand, the level of electricity consumption will not only affect the cost of electricity production, but also affect social welfare, which is an important research object of demand side management (Yuan et al. 2020). Therefore, accurately electricity consumption forecasting is beneficial to achieving a balance between supply and demand, and has a guiding role in the supply management of power resources, which to a certain extent alleviates the tightness of energy supply (Tao et al. 2020).

According to the China Statistical Yearbook, Shanghai's electricity consumption has risen to a high level since 1995. According to the Granger causality test, there is a one-way causal relationship between Shanghai's electricity consumption and economic development. At the same time, the test results also show that there is a corresponding equilibrium relationship among power consumption, economic development and energy consumption structure in Shanghai. In addition, according to Shanghai's Municipal Administration, Shanghai has set a clear goal to achieve carbon peaks by 2025. Therefore, accurate forecasting of electricity energy demand is vital to promote Shanghai's economic development, as well as to achieve the goal of carbon peak and carbon neutrality. However, due to the influences of various external factors such as political conditions, population, economy, etc. electricity consumption often featured by uncertainty and complex characteristics, which consequently makes electricity consumption forecasting even difficult and challenging.

Currently, there are a variety of methods that can support the electricity demand forecasting, on the basis of the performance of the forecast results, they can be divided into two categories: certainty forecasting methods and uncertainty forecasting methods. Furthermore, certainty prediction methods mainly include time series method, artificial intelligence method and grey forecasting method. And uncertainty forecasting methods include the following categories: interval prediction methods and probability prediction methods (Khosravi et al. 2010).

In the certainty forecasting methods, time series method is the simplest forecasting method in the field of prediction due to its explanatory ability and the simplicity (Jiang et al. 2020). Among many time series models, auto-regressive moving average model (Sen et al. 2016; Ma et al. 2020), Holt-Winters method (Trull et al. 2020), regression model (Liao et al. 2017) and exponential smoothing model (Li et al. 2019)) are commonly applied by most researchers to predict energy consumption due to their good prediction accuracy and easy implementation ways. Although the effectiveness of the time series method has been widely tested by above researchers, the traditional time series method can only come up with the overall trend of electricity consumption, and it cannot accurately predict the non-linear trend (Lin et al. 2016; He et al. 2019). Artificial intelligence methods based on computer algorithms are the most commonly used method in power demand forecasting (Wang

et al. 2018; Guo et al. 2018). Due to its great advantages compare with the traditional models in solving the large-scale and complex nonlinear problems, it has been adopted by many scholars to conduct the energy consumption prediction (Khan et al. 2020). In order to ensure the robustness of modeling for the nonlinear electricity demand model, the neural network model has been widely adopted in many intelligent algorithms due to its operability (Bedi et al. 2019; Wen et al.2020). In addition to the various neural network models, the support vector machine (SVM) has been favored by relevant scholars for it could generate excellent prediction results in processing the nonlinear time series prediction (Cao et al. 2016) .

However, above methods' prediction accuracy depends on the sufficiency of the initial data. When the scale of the data source is relatively low or insufficient, their prediction accuracy cannot be guaranteed. On the other hand, the grey prediction method requires only a small number of samples to reach a high prediction accuracy. Therefore, when the sample size is low or insufficient, the prediction performance of the grey model is often more optimal compared with other model, and thus it is widely used in the energy consumption prediction (Wu et al. 2019; Duan et al. 2020; Xu et al. 2020). Ikram et al. (2021) used an optimized grey model to predict future trends in carbon dioxide emissions and renewable electricity production, and then assess the sustainability of China and the United States. Ding et al. (2018) adopted a modified grey prediction model to accurately predict the overall electricity consumption as well as the industrial electricity consumption in China. Xie et al. (2020) proposed a reverse-fitting fractional grey scale model to predict China's annual electricity consumption as well as verified the reliability of the model through examples. Targeting at the complexity and uncertainty of power system, Zhu et al. (2020) established an adaptive grey fractional weighting model to predict the electricity consumption in Jiangsu, which could effectively improve the accuracy of electricity consumption prediction. Wu et al. (2018) adopted the grey convex correlation analysis method to explore the relationship between electricity consumption and its related factors, and then established a multi-variable grey prediction model for the purpose of effectively predicting electricity consumption in Shandong Province. In order to accurately predict the seasonal fluctuations of electricity consumption in major economic sectors, Wang et al. (2017) proposed a seasonal grey model (SGM (1,1)) on the basis of the accumulation operator generated by seasonal factors.

Unfortunately, certainty forecasting methods only output the expected value of a single point at a specific moment, which are difficult to reflect the uncertainty of electricity consumption. Due to the ability to describe the variation range, probability distribution, and possible scenarios of electricity consumption, the uncertainty prediction method can obtain more comprehensive prediction information (Shrivastava et al. 2015). In the short-term load forecast, interval prediction methods represented by lower upper bound estimation (LUBE) method (Liu et al. 2020; Khosravi et al. 2011), Bootstrap method (Beyaztas et al. 2020; Sheng et al. 2013), and Delta method (Tang et al.2020), as well as probability prediction methods represented by Bayesian method (Brusaferri et al.2019), have been widely used. However, in the

medium and long-term electricity demand forecast, due to the insufficiency of the electricity demand data, the accuracy and reliability of the above uncertain forecasting method is difficult to guarantee.

A reliable uncertainty forecasting method for predicting medium and long-term electricity consumption is important for the power sector and low-carbon development. Considering the uncertainty and nonlinearity of Shanghai's electricity consumption, a nonlinear interval grey model is proposed in this paper. The main contributions of this paper can be described in the following aspects:

Firstly, considering the uncertainty of electricity consumption, an interval grey prediction model (IGM (1,1) model) is established via integrating the uncertainty of grey action into the traditional GM (1,1) model.

Then, considering the difference among action mechanism of various elements within the system at different moments, we integrate the uncertainty of development coefficient into the IGM (1,1) model. Subsequently, in order to improve the prediction accuracy of the model, genetic algorithm is applied for the purpose of optimizing the IGM (1,1) model.

Furthermore, in order to meet the nonlinear nature of electricity consumption, BP neural network is applied to the interval modification of the model, and a nonlinear interval grey model is constructed.

Finally, the nonlinear interval grey model is applied to the prediction of Shanghai's electricity consumption.

The rest part of the paper is organized as follows: In the section 2, we introduce the operation mechanism and modeling process of the nonlinear interval grey model in detail. In the section 3, two cases are used to verify the reliability and validity of the model. In the section 5, we forecast the electricity consumption of Shanghai city based on the established model, and make comparison with several typical certainty forecasting models as well as uncertainty forecasting models. Last but not least, the basic conclusion and relevant suggestions are given in the section 5.

2 Methodology

In this section, a detailed introduction to the nonlinear interval grey model are given. Firstly, the operation mechanism of the interval grey model is explained in detail. Secondly, in order to improve the prediction performance of the interval grey model, genetic algorithm is applied to the parameter optimization. Subsequently, in order to obtain more reliable prediction results, BP neural network is applied to optimize the residual error of the prediction results. Finally, several evaluation metrics and modeling steps is explained.

2.1 The interval grey model

Definition 1. Let $X^{(0)} = [x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)]$ is a real data sequence, and

$X^{(1)} = [x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)]$ is the AGO generation sequence, where the calculation of

$x^{(1)}(k)$ is expressed as follows:

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), i = 1, \dots, n. \quad (1)$$

The background value sequence $z^{(1)} = [z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n)]$ is obtained based on the AGO generation sequence $X^{(1)}$, where the calculation of $z^{(1)}(k)$ is shown below:

$$z^{(1)}(k) = 0.5 \times (x^{(1)}(k) + x^{(1)}(k-1)), k = 2, \dots, n. \quad (2)$$

Definition 2. The grey differential equation of GM (1,1) model is defined as follows:

$$x^{(0)}(k) + az^{(1)}(k) = b, \quad (3)$$

where a represents development coefficient and b represents grey action.

Theorem 1 (Ding et al. 2018). Assuming $Y = [x^{(0)}(2), \dots, x^{(0)}(n)]^T$ and

$$B = \begin{bmatrix} -z^{(1)}(2) & \cdots & -z^{(1)}(n) \\ 1 & \cdots & 1 \end{bmatrix}^T, \text{ the values of development coefficient } a \text{ and grey action } b$$

can be computed by the least square method, as shown below:

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (B^T B)^{-1} B^T Y. \quad (4)$$

Theorem 2 (Ding et al. 2018). The continuous time response function of $X^{(1)}$ and the continuous time response function of $X^{(0)}$ are shown below:

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{\hat{b}}{\hat{a}} \right) e^{-\hat{a}k} + \frac{\hat{b}}{\hat{a}}, k = 1, \dots, n-1, \quad (5)$$

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), k = 1, \dots, n-1, \quad (6)$$

where $\hat{x}^{(1)}(k+1)$ represents the response value of $X^{(1)}$ at time $k+1$, and $\hat{x}^{(0)}(k+1)$ means the response value of $X^{(0)}$ at time $k+1$.

As explained in the previous study, the grey action quantity b represents the influence of system external factors on the development trend of the system (Zeng et al. 2020; Ye et al. 2019). In order to fully tap the uncertain information which included in the grey action, the essence of the grey action b should be the grey number.

Theorem 3 (Zeng et al. 2020). The grey action quantity b at each moment is shown below:

$$\hat{b}(k) = x^{(0)}(k) + \hat{a}z^{(1)}(k), k = 2, \dots, n, \quad (7)$$

where $\hat{b}(k)$ represents the grey action quantity b at the k -th moment.

Theorem 4 (Zeng et al. 2020). Assuming $\hat{b}_{\max}$ is the maximum grey action quantity at

all times, and $\hat{b}_{\min}$ is the minimum grey action quantity, the mathematical model definition of the interval grey model (IGM (1,1)) can be obtained as follows:

$$\hat{x}_{\max}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{\hat{b}_{\max}}{\hat{a}} \right) e^{-\hat{a}k} + \frac{\hat{b}_{\max}}{\hat{a}}, \quad (8)$$

$$\hat{x}_{\min}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{\hat{b}_{\min}}{\hat{a}} \right) e^{-\hat{a}k} + \frac{\hat{b}_{\min}}{\hat{a}}, \quad (9)$$

where $k=1, \dots, n-1$, $\hat{x}_{\max}^{(1)}(k+1)$ represent the maximum response of $X^{(1)}$ at the $k+1$ time, and $\hat{x}_{\min}^{(1)}(k+1)$ is the minimum response of $X^{(1)}$ at the $k+1$ time. According to Eqs. (6), the upper bound and lower bound of the response value of $X^{(0)}$ at time $k+1$ are $[\hat{x}_{\min}^{(0)}(k+1), \hat{x}_{\max}^{(0)}(k+1)]$.

2.2 Optimization of IGM (1,1) based on genetic algorithm

It can be observed from the IGM (1,1) model that it fully explores the uncertain information contained in the grey action by changing the form of grey action into the form of interval number. However, due to the differences in the mechanism of the internal elements of the system at each moment, the system presents to have different development trends at each moment. It indicates that the development coefficient which affect the speed of system development should also be grey.

As shown in Theorem 1, we can find that both the development coefficient a and the grey action b are directly related to the background value sequence $Z^{(1)}$. Based on the mathematical mechanism of the grey model, $z^{(1)}(k)$ is the integral form of $x^{(1)}(k)$ under the Δk time period, that is, $z^{(1)}(k) = \int_{k-1}^k x^{(1)}(t) dt$. In the traditional GM (1,1) model, $Z^{(1)}$ is generated by the mean value of $X^{(1)}$'s continuous neighbor sequence, which is only an approximately representation of its integral form. Based on this mechanism, we set the background value $z^{(1)}(k)$ into an interval form, and obtains the interval value of the development coefficient a .

From Definition 1, it is easy to prove that the curve of the first-order cumulative sequence $X^{(1)}$ in the time period Δk is monotonic and non-decreasing. Therefore, the integral form of $z^{(1)}(k)$ under the mathematical mechanism of the grey model could satisfy the inequality: $x^{(1)}(k-1) \leq \int_{k-1}^k x^{(1)}(t) dt$.

Definition 3. The interval bound of the background value $z^{(1)}(k)$ is computed as follows:

$$z_{\min}^{(i)}(k) = x^{(i)}(k-1), \quad (10)$$

$$z_{\max}^{(i)}(k) = 0.5 \times (x^{(i)}(k-1) + x^{(i)}(k)), \quad (11)$$

where $z_{\min}^{(i)}(k)$ represent the lower bound of $z^{(i)}(k)$, $z_{\max}^{(i)}(k)$ means the upper bound of $z^{(i)}(k)$, $k = 2, \dots, n$.

Theorem 5. Let the lower bound of the development coefficient a is $a_{\min}$ while the upper bound is $a_{\max}$, so the interval values of grey action quantity b corresponding to the lower bound $a_{\min}$ and the upper bound $a_{\max}$ can be obtained, as shown follows:

$$b \in [b_{\min}^L, b_{\max}^L] \quad a \in a_{\min}, \quad (12)$$

$$b \in [b_{\min}^U, b_{\max}^U] \quad a \in a_{\max}, \quad (13)$$

where L refers to the lower bound of the development coefficient a , and U represents the upper bound of the development coefficient a .

Based on **Definition 3** and **Theorem 5**, it can be summarized that when $z^{(i)}(k)$ takes the upper bound of the interval, the development coefficient a and the grey action b are the solution of the IGM(1,1) model. But from Theorem 5, two upper and lower bounds can be observed for the solution of the IGM (1,1) model based on the interval-type development coefficient a . In order to optimize this model, genetic algorithm is applied to optimize the upper and lower bounds of the IGM (1,1) model. The optimization model of the upper bound is shown in Equations (14), and the optimization model of the lower bound is shown below:

$$\begin{aligned} \text{Min} \quad & \sum_{k=1}^n |\hat{x}^{(0)}(k) - x^{(0)}(k)| \\ \text{s.t.} \quad & \begin{cases} a \in [a_{\min}, a_{\max}] \\ b \in [b_{\min}^L, b_{\max}^U] \end{cases}, \end{aligned} \quad (14)$$

$$\begin{aligned} \text{Min} \quad & \sum_{k=1}^n |\hat{x}^{(0)}(k) - x^{(0)}(k)| \\ \text{s.t.} \quad & \begin{cases} a \in [a_{\min}, a_{\max}] \\ b \in [b_{\max}^L, b_{\min}^U] \end{cases}, \end{aligned} \quad (15)$$

where $\hat{x}^{(0)}(k)$ represents the predicted value, which is obtained through the grey model under each parameter's setting, and $x^{(0)}(k)$ means the original value.

The upper and lower bounds of the IGM (1,1) model which are optimized by solving the optimization model are shown in (14) and (15) with genetic algorithm. We named the optimized IGM (1,1) model as the GA-IGM (1,1) model. Its mathematical

model can be described as follows:

$$\hat{x}_{\max}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{\hat{b}_{\max}}{\hat{a}_{\max}} \right) e^{-\hat{a}_{\max}k} + \frac{\hat{b}_{\max}}{\hat{a}_{\max}}, \quad (16)$$

$$\hat{x}_{\min}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{\hat{b}_{\min}}{\hat{a}_{\min}} \right) e^{-\hat{a}_{\min}k} + \frac{\hat{b}_{\min}}{\hat{a}_{\min}}, \quad (17)$$

where $k=1, \dots, n-1$, $\hat{b}_{\max}$ represent the upper bound of the grey action b of the GA-IGM (1,1) model; $\hat{b}_{\min}$ means the lower bound of the grey action b of the GA-IGM (1,1) model; $\hat{a}_{\max}$ refers to the upper bound of the model development coefficient a of the GA-IGM (1,1) model; $\hat{a}_{\min}$ is the lower bound of the model development coefficient a of the GA-IGM (1,1) model. To sum up, based on Eqs. (6), the upper and lower bounds of the response value of $X^{(0)}$ at $k+1$ time can be obtained and described as $[\hat{x}_{\min}^{(0)}(k+1), \hat{x}_{\max}^{(0)}(k+1)]$.

2.3 The nonlinear interval grey model

Apparently, the GA-IGM (1,1) model present a more reliable prediction result compare with the traditional GM (1,1) model with the setting of the prediction result as an interval number. At the same time, as explained in the introduction, electricity consumption is susceptible to external factors and shows a nonlinear growth trend. Regrettably, both of is the traditional GM (1,1) model and the GA-IGM (1,1) model would generate prediction results featured by exponential growth trend, which is unable to reflect the non-linear trend of the electricity consumption.

As the most widely used technical method in the field of artificial intelligence, BP neural network is equipped with strong nonlinear mapping ability, so it is widely used for the purpose of optimizing nonlinear problems. Therefore, in order to solve the problem that the GA-IGM (1,1) model fails in showing a non-linear trend, we use the BP neural network to optimize the GA-IGM (1,1) model. And the newly established model is named as the nonlinear interval grey model (BPGA-IGM (1,1) model). The relevant definition of the BPGA-IGM (1,1) model is as follows.

Definition 4. Let $E_U = [\varepsilon_U^2, \varepsilon_U^3, \dots, \varepsilon_U^n]$ is the upper bound residual sequence of the GA-IGM(1,1) model, and $E_L = [\varepsilon_L^2, \varepsilon_L^3, \dots, \varepsilon_L^n]$ is the lower bound residual sequence of the GA-IGM(1,1) model, where ε_U^k and ε_L^k are computed as follows:

$$\varepsilon_U^k = \hat{x}_{\max}^{(0)}(k) - x^{(0)}(k), \quad (18)$$

$$\varepsilon_L^k = x^{(0)}(k) - \hat{x}_{\min}^{(0)}(k), \quad (19)$$

where $k=2,\dots,n$, $\hat{x}_{\max}^{(0)}(k)$ is the upper bound of the response value of $X^{(0)}$ at the time k , and $\hat{x}_{\min}^{(0)}(k)$ represents the lower bound of the response value of $X^{(0)}$ at the time k .

Definition 5. Let $\Delta E_U = [\Delta e_U^2, \Delta e_U^3, \dots, \Delta e_U^n]$ is the upper bound growth rate sequence of the GA-IGM(1,1) model, and $\Delta E_L = [\Delta e_L^2, \Delta e_L^3, \dots, \Delta e_L^n]$ is the lower bound growth rate sequence of the GA-IGM(1,1) model, where the formula of Δe_U^k and Δe_L^k are shown below:

$$\Delta e_U^k = \hat{x}_{\max}^{(0)}(k) - \hat{x}_{\max}^{(0)}(k-1), \quad (20)$$

$$\Delta e_L^k = \hat{x}_{\min}^{(0)}(k) - \hat{x}_{\min}^{(0)}(k-1), \quad (21)$$

where $k=2,\dots,n$, Δe_U^k is the upper bound growth rate of the GA-IGM (1,1) model at the time k , and Δe_L^k is the lower bound growth rate of the GA-IGM(1,1) model at the time k .

Definition 6. Let $D = [d^2, d^3, \dots, d^n]$ is the interval width sequence of the GA-IGM (1,1) model, where the formula of d^k is shown below:

$$d^k = \hat{x}_{\max}^{(0)}(k) - \hat{x}_{\min}^{(0)}(k), \quad (22)$$

where $k=2,\dots,n$, d^k means the interval difference of the GA-IGM (1,1) model at the k -th moment.

In accordance with the time series, the results generated by the GA-IGM (1,1) model, the residual sequence, the growth rate sequence, and the interval width sequence, the residual prediction model based on BP neural network can be constructed. For the purpose of effective optimizing the solution of the GA-IGM (1,1) model, we take 0.75 times the residual sequence as the target output of the BP neural network model, and other sequence data as the input of the model. Among them, the relevant parameters of the BP neural network can be described as follows: the network structure is defined as 6x10x1, the training function is "traingdx", the maximum number of iterations is 5000, the learning rate is 0.05, the momentum factor is 0.9, the learning rate growth ratio is 1.05, and the learning rate reduction ratio 0.75.

Theorem 6. Suppose that $E_{BP}^U = [\varepsilon_{BP}^U(2), \dots, \varepsilon_{BP}^U(n)]$ and $E_{BP}^L = [\varepsilon_{BP}^L(2), \dots, \varepsilon_{BP}^L(n)]$ is obtained by the BP neural network model, so the solution of the BPGA-IGM (1,1) can be obtained, as shown follows:

$$\hat{x}_U^{(0)}(k) = \hat{x}_{\max}^{(0)}(k) - \varepsilon_{BP}^U(k), \quad (23)$$

$$\hat{x}_L^{(0)}(k) = \hat{x}_{\min}^{(0)}(k) + \varepsilon_{BP}^L(k), \quad (24)$$

where $k = 2, \dots, n$, $\hat{x}_U^{(0)}(k)$ represents the upper bound of the BPGA-IGM(1,1) solution at the k -th time, and $\hat{x}_L^{(0)}(k)$ means the lower bound of the BPGA-IGM(1,1) solution at the k -th time, $\hat{x}_{\max}^{(0)}(k)$ refers to the upper bound of the GA-IGM(1,1) solution at the k -th time, and $\hat{x}_{\min}^{(0)}(k)$ is the lower bound of the GA-IGM(1,1) solution at the k -th time.

2.4 Evaluation metrics

In order to evaluate the reliability and accuracy of models, five evaluation methods are selected as evaluation metrics, including mean absolute percentage error (MAPE), prediction intervals coverage probability (PICP), mean prediction interval width (MPIW), normalized mean prediction interval width (NMPIW) and coverage width-based criterion (CWC). Those metrics can be expressed as follows:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad (25)$$

$$\text{PICP} = \frac{1}{n} \sum_{i=1}^n c_i, \quad (26)$$

$$\text{MPIW} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_U^i - \hat{y}_L^i), \quad (27)$$

$$\text{NMPIW} = \frac{\text{MPIW}}{R}, \quad (28)$$

$$\text{CWC} = \text{NMPIW}(1 + \gamma e^{-\eta(\text{PICP}-u)}), \quad (29)$$

where y_i is the raw data, $\hat{y}_i$ is the predicted value of the certainty forecasting methods, while in the uncertain methods, $\hat{y}_i$ is the median value of the predicted interval. $\hat{y}_U^i$ is the upper bound of the prediction interval, and $\hat{y}_L^i$ is the lower bound of the prediction interval. $c_i = 1$ if $y_i \in [\hat{y}_L^i, \hat{y}_U^i]$, otherwise $c_i = 0$. R is the range of the underlying target. η is a constant, $u = 0.9$ is a reliability parameter, $\gamma = 0$ if $\text{PICP} \geq u$, otherwise $\gamma = 0$.

Among the above evaluation metrics, MAPE is the accuracy evaluation metrics of certain prediction methods. If the MAPE value is smaller, the model result is more accurate. PICP is an interval coverage evaluation metrics for uncertain prediction methods. If the PICP value is higher, the prediction result of the model is more

reliable (Shrivastava et al.2015). MPIW and NMPIW are the interval width evaluation metrics of uncertain forecasting methods. Since the narrow prediction interval width has more reference information and actual value than the wide interval width, lower MPIW and NMPIW means more effective prediction results (Khosravi et al.2014). The CWC comprehensively considers the model's prediction width and interval coverage. If the CWC indicator value is lower, it means that the quality of the value prediction interval is higher (Shrivastava et al.2015).

2.5 Modeling steps

Based on the definition of the BPGA-IGM (1,1) model, the process of forecasting with the BPGA-IGM (1,1) model is proposed in this paper, as shown in Fig.1.

Step 1: Get the raw data set $X^{(0)}$.

Step 2: Build the IGM (1,1) model and solve it. Firstly, compute the AGO generation sequence $X^{(1)}$ and the background value sequence $Z^{(1)}$. Then compute the development coefficient a and $[\hat{b}_{\min}, \hat{b}_{\max}]$. Finally, obtained the solution of the IGM (1,1) model.

Step 3: Solve the GA-IGM (1,1) model. Firstly, compute the interval value of the development coefficient a as $[a_{\min}, a_{\max}]$, the interval value of the lower bound of the grey action b as $[b_{\min}^L, b_{\min}^U]$ and the upper bound of the grey action b as $[b_{\max}^L, b_{\max}^U]$. Secondly, apply the genetic algorithm to optimize the development coefficient and the grey action.

Step 4: Build and solve the BPGA-IGM (1,1) model. Firstly, BP neural network is used to predict the residual of the upper and lower bound. Then, modify the upper and lower bound with the residual.

Step 5: Use the BPGA-IGM (1,1) model to forecast Shanghai's electricity consumption.

Fig.1 Flowchart of the BPGA-IGM (1,1) model

3 Validation of the BPGA-IGM (1,1) model

In this section, we use two cases to verify the reliability and validity of the BPGA-IGM (1,1) model. One case is to forecast China's total energy consumption. Another is to forecast China's coal consumption. In addition, the LUBE method, Bootstrap method, Delta method and Bayesian method are used for model comparison. PICP, MPIW, NMPIW and CWC are selected as evaluation metrics.

3.1 Validation case 1: China's total energy consumption

In this case, China's energy consumption data from 2005 to 2019 is used to verify the validity of the model. The raw data is shown in Table 1. Based on the model solving process as shown in Fig.1, the parameter of the IGM (1,1) model is $\hat{a} = -0.0372$,

$\hat{b}_{\max} = 310449.629$, $\hat{b}_{\min} = 271415.994$. The parameter of the IGM(1,1) model is

$\hat{a}_{\max} = -0.0372$, $\hat{b}_{\max} = 310449.629$, $\hat{a}_{\min} = -0.0378$, $\hat{b}_{\min} = 276568.009$. Finally, the

prediction results of each model are shown in Table 1. Table 2 shows evaluation metrics of various models. The comparison of the evaluation metrics of each model is shown in Fig.2.

As shown in Table 2 and Fig.2, although the PICP of the IGM (1,1) model is 100%, the other metrics of this model are inferior to other models. Especially the interval width metrics and coverage metrics of the IGM (1,1) model have the worst rank among all models. In particular, the interval width metrics and coverage metrics of the IGM (1,1) model have the worst rank among all models. In addition, from Table 2 and Fig.2, the PICP of the BPGA-IGM (1,1) model is also 100%. Meanwhile, its interval width metrics (MPIW, NMPIW) and coverage metrics (CWC) also have optimal values compared with other models, among which their values are 10288.02, 4.56% and 4.56% respectively. The superiority of these metrics indicates the reliability and validity of the prediction results of the BPGA-IGM (1,1) model.

Table 1 The prediction results of case 1

Year	Raw data	LUBE		Bootstrap		Delta		Bayesian		IGM (1,1)		BPGA-IGM (1,1)	
		Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower
2005	2613	27443	24830	27447	24830	27443	24830	27443	24830	26136	26136	26136	26136
	69	7.45	0.55	3.19	0.79	7.45	0.65	3.49	4.51	9.00	9.00	9.00	9.00
2006	2864	30079	27214	30078	28147	30079	26281	30058	27235	32620	28643	29640	28645
	67	0.35	3.65	8.65	7.73	0.35	7.12	1.82	2.18	2.08	3.34	0.77	8.59
2007	3114	31841	29140	34654	29503	32277	29649	32701	29586	33856	29728	31730	30938
	42	6.67	9.34	9.42	7.27	3.38	0.90	4.67	9.33	5.16	9.18	1.05	9.08
2008	3206	33678	30777	33517	30500	34471	32029	33693	30428	35139	30855	32881	31949
	11	1.46	4.43	2.86	0.04	0.79	3.30	5.96	6.04	6.79	6.46	8.53	2.15
2009	3361	35734	32489	34333	31931	36008	33721	35325	31900	36471	32025	34365	33336
	26	2.33	5.63	0.68	9.70	0.15	2.92	1.07	0.93	4.75	0.76	3.80	1.44
2010	3606	36868	34261	37399	33841	37700	35034	37867	34262	37853	33238	36466	35541
	48	0.40	5.60	4.11	9.38	0.94	9.19	5.16	0.84	7.46	8.28	2.28	5.38
2011	3870	39922	36820	40630	36037	39705	36176	40601	36806	39288	34498	38724	37714
	43	3.65	1.62	3.27	6.40	6.78	8.78	6.59	9.41	4.05	5.82	8.25	7.34
2012	4021	41804	38703	43065	38203	41826	37266	42203	38223	40777	35806	40354	39248
	38	9.49	7.20	4.96	1.10	2.20	5.53	8.72	7.28	4.37	0.79	6.78	3.07
2013	4169	42486	39269	44631	40043	43775	38371	43767	39615	42322	37163	41717	40569
	13	1.38	6.08	2.09	4.24	8.65	9.96	3.58	2.42	9.04	1.32	4.78	0.55
2014	4283	44975	40691	45554	41419	45359	39530	44985	40680	43926	38571	42927	41899
	34	0.70	7.30	2.91	5.00	3.69	5.61	8.38	9.62	9.44	6.16	9.34	6.52
2015	4341	46297	41957	46072	42352	46527	40759	45629	41193	45591	40033	44098	42895
	13	9.67	4.37	9.52	9.16	1.04	1.75	3.03	2.97	7.78	4.82	8.00	4.98
2016	4414	47484	43064	46356	42944	47331	42059	46421	41876	47319	41550	45318	43864
	92	2.59	7.08	8.17	9.08	5.01	8.99	9.80	4.20	7.08	7.53	9.84	1.70
2017	4558	48559	44019	46509	43304	47861	43423	47900	43265	49113	43125	46649	45371
	27	6.84	4.39	9.99	4.95	8.35	5.99	2.72	1.28	1.28	5.29	2.03	8.34
2018	4719	49543	44832	47592	43517	48202	44832	49546	44838	50974	44759	48122	46892

18	25	9.41	8.75	0.45	2.11	7.92	8.99	0.84	9.16	5.18	9.89	9.59	9.59
20	4870	50450	46519	46635	43641	48419	46264	51082	46317	52906	46456	49751	48450
19	00	7.80	3.35	8.17	0.79	5.04	9.13	1.57	8.43	4.54	3.95	6.14	2.20

Table 2 Evaluation metrics of each model in case 1

Model	PICP (%)	MPIW	NMPIW (%)	CWC (%)
LUBE	100.00	35051.39	15.53	15.53
Bootstrap	93.33	36173.38	16.03	16.03
Delta	86.67	37837.53	16.77	105.41
Bayesian	100.00	38777.26	17.19	17.19
IGM (1,1)	100.00	47803.69	21.19	21.19
BPGA-IGM (1,1)	100.00	10288.02	4.56	4.56

Fig.2 Comparison of model metrics in case 1

3.2 Validation case 2: China's coal consumption

In case 2, we use the model to predict China's coal consumption. Furthermore, in order to facilitate the verification of the reliability of the BPGA-IGM (1,1) model, the LUBE method, Bootstrap method, Delta method and Bayesian method are also used to predict coal consumption. Based on the model solving process, the parameter of the

IGM (1,1) model is $\hat{a} = -0.0192$, $\hat{b}_{\max} = 243292.890$, $\hat{b}_{\min} = 201770.187$. The parameter of

the IGM(1,1) model is $\hat{a}_{\max} = -0.0192$, $\hat{b}_{\max} = 243292.890$, $\hat{a}_{\min} = -0.0193$,

$\hat{b}_{\min} = 203742.337$. The prediction results of each model are shown in Table 3.

The evaluation metrics of each model in case 2 are shown in Table 4. Fig.3 shows the comparison of the evaluation metrics of each model. As shown in Table 4 and Fig.3, the PICP of the BPGA-IGM (1,1) model is 100%, which shows that the model has reliable prediction results. Meanwhile, since the interval width metrics of the BPGA-IGM (1,1) model has the lowest value compared with other models, in which its MPIW and NMPIW are 10915.61 and 11.89% respectively, the model has more effective prediction results. According to CWC, the BPGA-IGM (1,1) model has

the optimal CWC, which shows that the model has high quality prediction results.

Table 3 The prediction results of case 2

Year	Raw data	LUBE		Bootstrap		Delta		Bayesian		IGM (1,1)		BPGA-IGM (1,1)	
		Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower
2005	1892	19869	17975	19869	17684	19494	17981	19910	17935	18923	18923	18923	18923
	31	2.71	3.28	0.81	4.18	1.40	0.80	7.42	4.90	1.16	1.16	1.16	1.16
2006	2074	21777	19703	23932	22834	22353	18744	21576	19904	24932	20739	21789	20740
	02	2.21	2.00	9.56	6.66	7.27	2.76	0.48	3.74	0.09	5.66	3.91	0.50
2007	2257	22852	22494	23707	21451	23707	21450	23709	21450	25415	21142	23188	22107
	95	8.72	4.79	5.14	0.43	8.99	5.68	0.73	0.17	9.90	1.63	2.89	1.97
2008	2292	24069	21823	24679	22181	24864	23338	24255	21591	25909	21552	23748	22636
	37	8.71	2.98	2.63	9.91	1.60	8.35	9.22	4.51	3.66	5.75	9.97	7.22
2009	2406	25544	22488	25861	23141	25845	24459	25435	22697	26412	21970	24671	23598
	66	1.18	1.64	7.95	7.60	8.22	7.28	4.39	8.04	3.19	9.54	3.62	2.37
2010	2495	26941	24246	26880	23960	26662	23176	26330	23583	26925	22397	25525	24418
	68	5.21	3.32	1.23	8.80	7.26	5.55	6.10	0.73	0.36	4.55	7.20	9.68
2011	2717	27973	25572	27656	24615	27331	25667	28546	25794	27447	22832	27241	26108
	04	4.93	5.60	3.56	7.92	8.51	3.20	2.80	5.57	7.06	2.35	1.81	1.87
2012	2754	28605	26167	28211	25130	27872	26006	28923	26169	27980	23275	27656	26539
	65	6.20	6.70	8.49	1.74	8.14	4.31	8.60	0.46	5.21	4.55	8.02	5.47
2013	2809	28947	26357	28591	25530	28305	26233	29478	26721	28523	23727	28192	26905
	99	5.75	8.83	6.58	6.61	1.13	7.95	3.35	5.37	6.80	2.79	7.42	2.77
2014	2818	29117	26396	28841	25840	28646	26376	29563	26805	29077	24187	28317	27037
	44	4.68	4.60	8.52	6.08	7.08	8.62	2.90	4.64	3.83	8.74	5.27	0.52
2015	2769	29195	26392	29005	26079	28913	26456	29075	26317	29641	24657	28216	27091
	64	5.38	1.57	5.28	4.06	4.02	3.34	4.94	3.25	8.34	4.10	4.74	0.02
2016	2746	29227	26380	29121	26262	29118	26387	28839	26081	30217	25136	28318	26853
	08	3.63	9.20	2.71	7.55	6.95	9.15	8.32	7.73	2.42	0.60	3.32	8.63
2017	2762	29236	26371	29220	26403	29273	26483	29001	26244	30803	25624	28375	27292
	31	7.78	9.14	2.96	1.58	8.81	3.55	9.49	2.84	8.20	0.02	4.68	4.04
2018	2784	29235	26366	29323	26510	29388	26351	29222	26465	31401	26121	28767	27485
	36	7.54	1.67	9.18	4.55	2.81	3.96	1.22	0.28	7.84	4.16	8.52	9.59
2019	2809	29230	26362	29443	26592	29469	26398	29478	26721	32011	26628	28914	27736
	99	1.68	8.44	2.36	3.27	5.17	6.29	1.10	6.90	3.57	4.86	5.40	7.89

Table 4 Evaluation metrics of each model in case 2

Model	PICP (%)	MPIW	NMPIW (%)	CWC (%)
LUBE	100.00	24483.50	26.68	26.68
Bootstrap	93.33	26751.07	29.15	29.15
Delta	86.67	23823.77	25.96	163.18
Bayesian	100.00	25909.46	28.23	28.23
IGM (1,1)	100.00	44471.41	48.46	48.46
BPGA-IGM (1,1)	100.00	10915.61	11.89	11.89

Fig.3 Comparison of model metrics in case 2

4 Shanghai's electricity consumption forecast

As shown in the above cases, the BPGA-IGM (1,1) model can obtain more reliable prediction results than other models. In this section, we use the BPGA-IGM (1,1) model to predict the electricity consumption in Shanghai. At the same time, in order to verify the superiority of the model, several typical certain and uncertain forecasting models are established for the purpose of conducting comparative analysis of model performance.

4.1 Data description

The data of Shanghai's electricity consumption ranges from 1999 to 2018 is selected as the raw data, as shown in Fig.4. The data comes from the "China Statistical Yearbook" (<http://www.stats.gov.cn/tjsj/ndsj/>). As presented in Fig.4, Shanghai's electricity consumption has risen to a high level since 1999. The level of electricity consumption in 2018 has reached 1567 hundred million kilowatt-hours, which is three times of that in 1999 (501.2 hundred million kilowatt-hours). However, Shanghai's electricity consumption shows a continually rising trend, which means that Shanghai's electricity consumption will reach to an even higher level. In that case, a reliable forecast of Shanghai's electricity consumption is of important reference significance for power management as well as economic development planning of Shanghai's power resource.

It is also worth noting that the growth trend of electricity consumption in Shanghai presents a typical nonlinear growth trend. Taking the year of 2010 and 2014 as example, Shanghai's electricity consumption shows strong nonlinear fluctuations in the two years. In 2010, Shanghai's electricity consumption increased significantly compared with the previous years. In 2014, due to some factors, Shanghai's electricity consumption in 2014 dropped slightly compared to 2013. Considering the existence of a non-linear growth trend in the electricity consumption in Shanghai, we construct the BPGA-IGM (1,1) model on the basis of the grey model and applies it to Shanghai's electricity consumption forecast for the next few years.

Fig.4 Electricity consumption of Shanghai

4.2 Construction and application of the BPGA-IGM (1,1) model

As shown in Fig.1, the IGM (1,1) model and the GA-IGM (1,1) model are the basic models for the BPGA-IGM (1,1) model. Firstly, on the basis of the theoretical description of the IGM (1,1) model, the development coefficient a and the interval of the grey action b can be obtained. Ultimately, we could obtain the results as

$\hat{a} = -0.0484$, $\hat{b}_{\min} = 521.6532$, $\hat{b}_{\max} = 821.996$. The forecasting results of the IGM (1,1)

model is shown in Table 5, and the forecast trend is shown in Fig.5.

Table 5 The prediction results of the IGM (1,1) model (100 million kilowatt-hours)

Year	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Original value	501.20	559.45	592.98	645.71	745.97	821.44	921.97	990.15	1072.38	1138.22
Upper bound	501.20	867.07	910.07	955.20	1002.57	1052.28	1104.47	1159.24	1216.72	1277.06
Lower bound	501.20	559.34	587.08	616.19	646.75	678.82	712.48	747.81	784.90	823.82
Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Original value	1153.38	1295.87	1339.62	1353.45	1410.60	1369.03	1406.00	1486.00	1527.00	1567.00
Upper bound	1340.39	1406.86	1476.63	1549.86	1626.71	1707.38	1792.05	1880.92	1974.20	2072.10
Lower bound	864.68	907.55	952.56	999.80	1049.38	1101.42	1156.04	1213.37	1273.54	1336.69

Fig.5 The prediction results of the IGM (1,1) model

According to the above-mentioned principle explanation of the GA-IGM (1,1) model, the GA-IGM (1,1) model functions to modify the development coefficient a on the basis of the IGM (1,1) model. By using genetic algorithm for parameter optimization, the optimal parameters of the GA-IGM(1,1) model are obtained as

$\hat{a}_{\max} = -0.0484$, $\hat{b}_{\max} = 821.996$, $\hat{a}_{\min} = -0.0494$, $\hat{b}_{\min} = 534.6907$. The forecast results of the GA-IGM (1,1) model can be obtained and showed in Table 6, and the forecast development trend of electricity consumption in Shanghai is also showed in Fig.6.

Table 6 Prediction results of the GA-IGM (1,1) model (100 million kilowatt-hours)

Year	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Original value	501.20	559.45	592.98	645.71	745.97	821.44	921.97	990.15	1072.38	1138.22
Upper bound	501.20	867.07	910.07	955.20	1002.57	1052.28	1104.47	1159.24	1216.72	1277.06
Lower bound	501.20	573.50	602.54	633.05	665.11	698.79	734.18	771.36	810.42	851.46

Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Original value	1153.38	1295.87	1339.62	1353.45	1410.60	1369.03	1406.00	1486.00	1527.00	1567.00
Upper bound	1340.39	1406.86	1476.63	1549.86	1626.71	1707.38	1792.05	1880.92	1974.20	2072.10
Lower bound	894.58	939.88	987.48	1037.48	1090.02	1145.22	1203.22	1264.15	1328.17	1395.43

Fig.6 Forecast trend of the GA-IGM (1,1) model

In light of the principle of the BPGA-IGM (1,1) model, the BPGA-IGM (1,1) model applies the BP neural network to modify the GA-IGM (1,1) model, so that the GA-IGM (1,1) model would equip with nonlinear characteristics. In accordance with the residual prediction results of the BP neural network, the solution of the GA-IGM (1,1) model is modified. The correction results are shown in Table 7, and the prediction trend is described in Fig.7.

Table 7 Prediction results of the BPGA-IGM (1,1) model (100 million kilowatt-hours)

Year	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Original value	501.20	559.45	592.98	645.71	745.97	821.44	921.97	990.15	1072.38	1138.22
Upper bound	501.20	636.82	659.95	730.90	808.46	889.27	967.78	1051.63	1121.59	1186.29
Lower bound	501.20	558.39	592.23	643.90	717.81	788.82	863.36	937.81	1008.89	1074.16

Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Original value	1153.38	1295.87	1339.62	1353.45	1410.60	1369.03	1406.00	1486.00	1527.00	1567.00
Upper bound	1212.23	1360.98	1412.14	1440.22	1476.10	1430.79	1525.41	1571.15	1619.24	1670.87
Lower bound	1095.30	1183.13	1227.44	1273.63	1335.42	1324.64	1371.03	1417.19	1456.50	1500.14

Fig.7 Forecast trend of the BPGA-IGM (1,1) model

As presented in Fig.5, a huge difference exists between the predicted results of the traditional GM (1,1) model and the actual results. The IGM (1,1) model takes the comprehensive influence of uncertain and complex factors on the system development into account. By setting grey action b as interval grey order, its prediction results present an interval form. In comparison with GM (1,1) model, it can provide more effective referential information to decision-makers. As shown in Fig.6, taking the difference of action mechanism of various elements in the system at different moments into consideration, the GA-IGM (1,1) model optimizes the development coefficient a as well as the grey action amount b by means of genetic algorithm, hence effectively optimizing the lower bound of the prediction of the IGM (1,1) model. However, both of the IGM (1,1) model and the GA-IGM (1,1) model show an exponential growth trend, which on the other hand, is inconsistent with the non-linear trend of the original data (as shown in Fig.6). The BPGA-IGM (1,1) model applies BP neural network to modify residual error of the GA-IGM (1,1) model, so that the prediction results would have nonlinear characteristics and will be closer to the original data, as shown in Fig.7.

4.3 Comparative analysis with certainty forecasting methods

In this section, in order to verify the rationality of the BPGA-IGM (1,1) model in power consumption prediction of Shanghai, several certainty forecasting methods including GM (1,1) model, Verhulst model, linear regression model and neural network model are selected for comparative analysis. The prediction results and prediction errors of each model are shown in Table 8. Fig.8 shows the forecast trends of each model.

From Table , the MAPE of the median interval of the BPGA-IGM model, the neural network model and Verhulst model are all lower than 3%, which are 1.79%,

2.02% and 2.39%, respectively. The MAPE of GM (1,1) model and linear regression model are both higher than 5%, and they are 8.57% and 5.44% respectively. From the comparison results of MAPE, the BPGA-IGM model, neural network and Verhulst model are significantly superior to the GM (1,1) model and linear regression model.

As presented in Fig.8, compared with the other three linear models, the prediction results of the neural network model are closer to the non-linear trend of the original data and it has better prediction results. However, the electricity consumption forecast results of various model are unique. We take 5% of the original data as the criterion for the accuracy of the model. The prediction results of the above models all have some deviation points, as shown by the red stars in Fig.8. When the deterministic forecast results are biased, the wrong decision information can easily make decision-makers make wrong decisions, which in turn leads to unreasonable power production, planning and scheduling. The BPGA-IGM (1,1) model starts from the "non-uniqueness" nature of the grey model, and comprehensively considers all aspects of system development, and has better fault tolerance.

Table 8 The prediction results of certainty forecasting methods (100 million kilowatt-hours)

Year	Original data	GM (1,1)		Verhulst		Linear regression		Neural network		BPGA-IGM	
		Value	Error (%)	Value	Error (%)	Value	Error (%)	Value	Error (%)	Median	Error (%)
1999	501.2	501.20	0.00	501.20	0.00	540.41	7.82	501.20	0.00	501.20	0.00
2000	559.45	709.41	26.80	563.42	0.71	598.74	7.02	558.47	0.18	597.61	6.82
2001	592.98	744.59	25.57	629.35	6.13	657.08	10.81	603.14	1.71	626.09	5.58
2002	645.71	781.51	21.03	698.31	8.15	715.41	10.80	641.08	0.72	687.40	6.46
2003	745.97	820.27	9.96	769.46	3.15	773.75	3.72	734.34	1.56	763.14	2.30
2004	821.44	860.95	4.81	841.84	2.48	832.09	1.30	817.52	0.48	839.04	2.14
2005	921.97	903.64	1.99	914.43	0.82	890.42	3.42	904.63	1.88	915.57	0.69
2006	990.15	948.45	4.21	986.19	0.40	948.76	4.18	990.76	0.06	994.72	0.46
2007	1072.38	995.49	7.17	1056.14	1.51	1007.10	6.09	1012.38	5.60	1065.24	0.67
2008	1138.22	1044.85	8.20	1123.37	1.30	1065.43	6.39	1087.25	4.48	1130.23	0.70
2009	1153.38	1096.67	4.91	1187.15	2.93	1123.77	2.57	1213.24	5.19	1153.76	0.03
2010	1295.87	1151.05	11.18	1246.89	3.78	1182.10	8.78	1303.07	0.56	1272.05	1.84
2011	1339.62	1208.13	9.82	1302.19	2.79	1240.44	7.40	1324.00	1.17	1319.79	1.48
2012	1353.45	1268.05	6.31	1352.82	0.05	1298.77	4.04	1367.66	1.05	1356.92	0.26
2013	1410.6	1330.93	5.65	1398.71	0.84	1357.11	3.79	1444.86	2.43	1405.76	0.34
2014	1369.03	1396.93	2.04	1439.93	5.18	1415.45	3.39	1436.44	4.92	1377.72	0.63
2015	1406	1466.20	4.29	1476.66	5.03	1473.78	4.82	1463.24	4.07	1448.22	3.00
2016	1486	1538.91	3.56	1509.14	1.56	1532.12	3.10	1486.00	0.00	1494.17	0.55
2017	1527	1615.23	5.78	1537.68	0.70	1590.46	4.16	1505.40	1.41	1537.87	0.71
2018	1567	1695.33	8.19	1562.63	0.28	1648.79	5.22	1522.00	2.87	1585.50	1.18
MAPE (%)			8.57		2.39		5.44		2.02		1.79

Fig.8The forecast trends of several certainty forecasting methods

4.4 Comparative analysis with uncertainty forecasting methods

In this section, the LUBE method, Bootstrap method, Delta method and Bayesian method are used to predict electricity consumption for the verification of the reliability of the BPGA-IGM (1,1) model. According to the modeling principles of different models, we construct the above-mentioned four uncertain prediction models and use them to predict Shanghai's electricity consumption. The prediction results of uncertainty forecasting methods are shown in Table 9. And Fig.9 shows the forecast trends of each uncertainty forecasting model.

Table 9 The prediction results of uncertainty forecasting methods (100 million kilowatt-hours)

Year	Original data	LUBE		Bootstrap		Delta		Bayesian	
		Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower
1999	501.2	536.27	456.19	551.32	451.08	551.37	450.38	526.26	476.14
2000	559.45	587.41	481.48	612.48	506.42	488.07	382.09	597.42	533.41
2001	592.98	582.63	500.70	621.50	511.10	650.24	536.30	632.43	561.30
2002	645.71	713.26	558.73	707.57	592.61	669.61	554.65	724.56	601.73
2003	745.97	803.27	711.25	803.15	681.34	783.53	669.72	789.05	657.33
2004	821.44	875.85	785.07	893.30	763.98	897.36	776.30	876.55	721.03
2005	921.97	974.54	835.87	978.87	841.91	994.72	861.85	967.44	787.50
2006	990.15	1057.58	923.93	1062.45	917.85	1075.99	931.58	950.38	854.10
2007	1072.38	1144.64	1001.52	1144.98	992.79	1146.21	992.06	1126.00	919.66
2008	1138.22	1205.79	1059.22	1225.34	1065.69	1210.76	1048.53	1195.18	983.70
2009	1153.38	1281.17	1125.77	1301.11	1134.18	1273.04	1103.89	1258.60	1045.97
2010	1295.87	1351.00	1200.35	1369.63	1195.64	1333.89	1158.52	1266.79	1106.28
2011	1339.62	1415.53	1253.96	1428.94	1248.18	1392.35	1211.16	1370.11	1164.38
2012	1353.45	1425.02	1300.65	1478.24	1291.02	1446.86	1260.04	1418.87	1119.91
2013	1410.6	1519.79	1327.90	1517.85	1324.53	1496.07	1303.73	1463.35	1272.49
2014	1369.03	1525.16	1356.31	1498.85	1399.78	1539.29	1341.47	1503.78	1281.70
2015	1406	1566.44	1357.11	1572.65	1368.22	1576.42	1373.12	1540.43	1307.20
2016	1486	1598.96	1401.70	1590.70	1381.30	1607.82	1399.05	1573.54	1388.75
2017	1527	1638.04	1411.42	1604.32	1390.31	1634.12	1419.87	1653.35	1426.27

Fig.9 The forecast results of several uncertainty forecasting methods

As presented in Table 9 and Fig.9, although several typical uncertainty prediction methods can effectively deal with the uncertain characteristics of electricity consumption, their prediction intervals may not contain the original data at some nodes. Although this situation does not exist in the IGM (1,1) model, as mentioned above, the model shows an exponential growth trend, which is inconsistent with the non-linear trend of electricity consumption. In addition, the prediction interval width of the IGM (1,1) model is significantly larger than that of other uncertainty prediction models. The BPGA-IGM (1,1) model completely predicts the uncertainty of power consumption, and its prediction interval width is much lower than the IGM (1,1) model.

To further illustrate the reliability of the BPGA-IGM (1,1) model, the evaluation metrics of several uncertain prediction models are shown in Table 10. Fig.10 shows the comparison of several models. As shown in Table 10 and Fig.10, compared with several other models, the BPGA-IGM (1,1) model has the most reliable and optimal prediction results. According to Table 10, the PICP of the BPGA-IGM (1,1) model is 100% while the MPIW and NMPIW of the BPGA-IGM (1,1) model are 120.10 and 11.27% respectively. Compared with other models, the BPGA-IGM (1,1) model has the best metrics value. These metrics mean that the BPGA-IGM (1,1) model has the most reliable prediction results. In addition, the CWC of the BPGA-IGM (1,1) model is lowest compared with other models, in which the CWC of the BPGA-IGM (1,1) model is 11.27% while the CWC of other models are 14.20%, 14.65%, 15.07%, 16.49% and 43.91% respectively. The lowest CWC indicates that the quality of the prediction interval of the BPGA-IGM (1,1) model is the best.

Table 10 Evaluation metrics of several uncertainty forecasting models

Model	PICP (%)	MPIW	NMPIW (%)	CWC (%)
LUBE	95.00	151.34	14.20	14.20
Bootstrap	95.00	156.18	14.65	14.65
Delta	95.00	160.66	15.07	15.07
Bayesian	95.00	175.78	16.49	16.49

IGM (1,1)	100.00	467.98	43.91	43.91
BPGA-IGM (1,1)	100.00	120.10	11.27	11.27

Fig.10 Comparison of several uncertainty forecasting models

4.5 Electricity consumption forecast in Shanghai

Accurate and reliable electricity consumption forecast plays a guiding role in the supply management planning of electricity resource. Based on the above analysis, the BPGA-IGM (1,1) model shows a better prediction result compare with other models. Therefore, we apply the BPGA-IGM (1,1) model to perform the electricity consumption prediction of Shanghai city from the year of 2019 to 2025. Fig.11 shows the forecast trend of electricity consumption in Shanghai from 2019 to 2025. It is easy to find that the overall electricity consumption of Shanghai shows a rising trend for the next few years. In 2025, the electricity consumption of Shanghai will reach to a figure no less than 1978.19 hundred million kWh, which is an increase of 26.24% compared with the consumption level in 2018. As the predicted results are within the form of interval, it is possible that the electricity consumption in the adjacent years shows nonlinear development characteristics.

Fig.11 Electricity consumption forecast of Shanghai from 2019 to 2025

5 Conclusions and recommendations

5.1 Conclusions

In this paper, a nonlinear interval grey model based on genetic algorithm and BP neural network optimization is proposed. Using Shanghai's panel data from 1999 to 2018, We can obtain a more reliable forecast of Shanghai's electricity consumption from 2019 to 2025. The following conclusions can be drawn from this study.

(1) Considering the uncertainty and nonlinear characteristics of electricity consumption, an interval grey prediction model (IGM (1,1) model) which integrates the uncertainty of grey action into the traditional GM (1,1) model is established. On this basis, genetic algorithm is used to optimize the parameters of the model. In addition, the BP neural network is used for interval modification of the model. Finally, a nonlinear interval grey model is established.

(2) The BPGA-IGM (1,1) model and several other uncertainty prediction models are used to predict China's total energy consumption and China's coal consumption. In both cases, the interval coverage evaluation metrics (PICP) of the BPGA-IGM (1,1) model are both 100%, which shows the reliability of the model's prediction results. In addition, in both cases, the BPGA-IGM (1,1) model has the best interval width evaluation metrics (MPIW, NMPIW). Moreover, its CWC also shows the best solution. These excellent metrics show that the BPGA-IGM (1,1) model has the best prediction results.

(3) The proposed BPGA-IGM (1,1) model is applied to forecast the electricity consumption in Shanghai. Compared with several classic certainty model, the BPGA-IGM (1,1) model comprehensively considers all aspects of system development and has better prediction accuracy. At the same time, the BPGA-IGM (1,1) model has better fault tolerance than certainty forecasting methods by setting the results in interval form. Compared with several uncertainty model, the BPGA-IGM (1,1) model has the most reliable and optimal prediction results.

(4) Based on the BPGA-IGM (1,1) model, Shanghai's electricity consumption from 2019 to 2025 is predicted. According to the forecast results in Fig.11, Shanghai's electricity consumption will continue to rise in the next few years. By the year of 2025, Shanghai's electricity consumption will reach to a figure no less than 1978.19 million kWh, indicating that more reasonable management on the electricity supply is needed.

5.2 Recommendations

The sustained and stable growth of the national economy and the continuous improvement of people's living standards have greatly increased the demand for electricity. According to the research results obtained from this paper, in the next five years, the electricity consumption in Shanghai will continue to rise to a high consumption level of no less than 1978.19 hundred million kWh. Making a reasonable electricity supply plan is a new challenge which poses to all managers. To this end, we put forward several suggestions as followings:

(1) As far as the forecast results are concerned, although Shanghai's electricity consumption will continue to grow in the next few years, the growth rate will remain

stable. In addition to ensuring stable power planning, the power sector can appropriately increase the input of renewable energy to reduce the consumption of fossil energy, which has direct significance for the realization of the "dual carbon" goal.

(2) In order to cope with the impact of uncertain factors on power demand, the power sector can adjust the capacity of generators according to the forecast upper bound to deal with future uncertainties. At the same time, low-carbon or zero-carbon electricity production methods should be vigorously promoted.

From the perspective of the government, the government should vigorously support renewable energy power generation companies to provide support for environmental protection. Secondly, ensuring capital investment to improve the level of power services and increase the efficiency of power generation is also an aspect that cannot be ignored.

Ethical approval and consent to participate

Not applicable.

Consent to publish

Not applicable.

Authors contributions

All authors contributed to the study conception and design. Data preparation, data collection, methodology, software and analysis were performed by [Zhiqiang Wu]. Formal analysis, funding acquisition, project administration, and supervision were performed by [Yan Gao]. Supervision, validation and review were performed by [Bo Wang]. The first draft of the manuscript was written by [Zhiqiang Wu] and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

This work is supported in part by National Nature Science Foundation of China under Grant 72071130.

Competing Interests

The authors declare that they have no competing interests.

Availability of data and materials

All data generated or analyzed during this study are included in this published article

References

- Bedi J, Toshniwal D (2019) Deep learning framework to forecast electricity demand. *Appl Energ* 238:1312-1326.
- Beyaztas U, Shang HL (2020) Robust bootstrap prediction intervals for univariate and multivariate autoregressive time series models. *J Appl Stat* 10.1080/02664763.2020.1856351.
- Brusaferri A, Matteucci M, Portolani P (2019) Bayesian deep learning based method for probabilistic forecast of day-ahead electricity prices. *Appl Energ* 250:1158-1175

- Cao GH, Wu LJ (2016) Support vector regression with fruit fly optimization algorithm for seasonal electricity consumption forecasting. *Energy* 115:734-745.
- Chen GJ, Hou FJ, Li JQ, Chang KL (2020) Decoupling analysis between carbon dioxide emissions and the corresponding driving forces by Chinese power industry. *Environ Sci Pollut R* 28(2):2369-2378.
- Chen JQ, Guo YS, Su HN, Ma XJ, Zhang ZY, Chang BS (2021) Empirical analysis of energy consumption transfer in China's national economy from the perspective of production and demand. *Environ Sci Pollut R* 28(15):19202-19221.
- Ding S, Hipel KW, Dang YG (2018) Forecasting China's electricity consumption using a new grey prediction model. *Energy* 149:314-328.
- Duan HM, Luo XL (2020) Grey optimization Verhulst model and its application in forecasting coal-related CO₂ emissions. *Environ Sci Pollut R* 27(35): 43884-43905.
- Guo ZF, Zhou KL, Zhang XL, Yang SL (2018) A deep learning model for short-term power load and probability density forecasting. *Energy* 160:1186-1200.
- He YY, Qin Y, Wang S, Wang X, Wang C (2019) Electricity consumption probability density forecasting method based on LASSO-Quantile Regression Neural Network. *Appl Energ* 233:565-575.
- Ikram M, Sroufe R, Zhang QY, Ferasso M (2021) Assessment and prediction of environmental sustainability: novel grey models comparative analysis of China vs. the USA. *Environ Sci Pollut R* 28(14):17891-17912.
- Jiang WH, Wu XG, Gong Y, Yu WX, Zhong XH (2020) Holt-Winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity consumption. *Energy* 193:807-814.
- Khan AN, Nadeem MA, Hussain MS, Aslam M, Bazmi AA (2020) A forecasting model approach of sustainable electricity management by developing adaptive neuro-fuzzy inference system. *Environ Sci Pollut R* 27(15):17607-17618.
- Khosravi A, Nahavandi S, Creighton D (2010) Construction of optimal prediction intervals for load forecasting problems. *IEEE T Power Syst* 25(3):1496-1503.
- Khosravi A, Mazloumi E, Nahavandi S (2011) Prediction intervals to account for uncertainties in travel time prediction. *IEEE T Intell Transp* 12(2):537-547.
- Khosravi A, Nahavandi S (2014) An optimized mean variance estimation method for uncertainty quantification of wind power forecasts. *Int J Elec Power* 61:446-454.
- Li P, Zhang JS (2019) Is China's energy supply sustainable? New research model based on the exponential smoothing and GM(1,1) methods. *Energies* 12(2): 236.
- Liao H, Liu YA, Gao YX, Hao Y, Ma XW, Wang K (2017) Forecasting residential electricity demand in provincial China. *Environ Sci Pollut R* 24(7): 6414-6425.
- Lin BQ, Liu C (2016) Why is electricity consumption inconsistent with economic growth in China? *Energ Policy* 88:310-316.
- Liu FJ, Li CS, Xu YH, et al (2020) A new lower and upper bound estimation model using gradient descend training method for wind speed interval prediction. *Wind Energy* 24(3): 290-304.
- Ma ML, Wang ZZ (2020) Prediction of the energy consumption variation trend in South Africa based on ARIMA, NGM and NGM-ARIMA models. *Energies* 13(1):1-15.
- Naqvi SAA, Shah SAR, Mehdi MA (2020) Revealing empirical association among ecological footprints, renewable energy consumption, real income, and financial development: a global perspective. *Environ Sci Pollut R* 27(34):42830-42849.

- Sen P, Roy M, Pal P (2016) Application of ARIMA for forecasting energy consumption and GHG emission: A case study of an Indian pig iron manufacturing organization. *Energy* 116:1031-1038.
- Sheng CY, Zhao J, Wang W, et al (2013) Prediction intervals for a noisy nonlinear time series based on a bootstrapping reservoir computing network ensemble. *IEEE T Neur Net Lear* 24(7):1036-1048.
- Shrivastava NA, Khosravi A, Panigrahi BK (2015) Prediction interval estimation of electricity prices using PSO-Tuned support vector machines. *IEEE T Ind Inform* 11(2):322-331
- Tang G, Wu YF, Li CS, et al (2020) A novel wind speed interval prediction based on error prediction method. *IEEE T Ind Inform* 16(11):6806-6815.
- Tao L, Gao Y (2020) Real-time pricing for smart grid with distributed energy and storage: A noncooperative game method considering spatially and temporally coupled constraints. *Int J Elec Power* 115:105487.
- Trull O, Garcia-Diaz JC, Troncoso A (2020) Stability of multiple seasonal Holt-Winters models applied to hourly electricity demand in Spain. *Appl Sci-Basel* 10(7):2630.
- Wang ZX, Li Q, Pei LL (2017) Grey forecasting method of quarterly hydropower production in China based on a data grouping approach. *Appl Math Model* 51:302-316.
- Wang L, Hu HL, Ai XY, Liu H (2018) Effective electricity energy consumption forecasting using echo state network improved by differential evolution algorithm. *Energy* 153:801-815.
- Wen L, Yuan XY (2020) Forecasting the annual household electricity consumption of Chinese residents using the DPSO-BP prediction model. *Environ Sci Pollut R* 27(17):22014-22032.
- Wu LF, Gao XH, Xiao YL, Yang YG, Chen XN (2018) Using a novel multi-variable grey model to forecast the electricity consumption of Shandong Province in China. *Energy* 157:327-335.
- Wu WQ, Ma X, Zeng B, Wang Y, Cai W (2019) Forecasting short-term renewable energy consumption of China using a novel fractional nonlinear grey Bernoulli model. *Renew Energ* 140:70-87.
- Xie WL, Wu WZ, Liu C, Zhao JJ (2020) Forecasting annual electricity consumption in China by employing a conformable fractional grey model in opposite direction. *Energy* 202:117682.
- Xu ZC, Liu LY, Wu LF (2020) Forecasting the carbon dioxide emissions in 53 countries and regions using a non-equipart grey model. *Environ Sci Pollut R* 28(13):15659-15672.
- Ye J, Dang YG, Ding S, Yang YG (2019) A novel energy consumption forecasting model combining an optimized DGM (1, 1) model with interval grey numbers. *J Clean Prod* 229:256-267.
- Yuan GX, Gao Y, Ye B, Huang RP (2020) Real-time pricing for smart grid with multi-energy microgrids and uncertain loads: a bilevel programming method. *Int J Elec Power* 123:106206.
- Zeng B, Ma X, Shi JJ (2020) Modeling method of the grey GM(1,1) model with interval grey action quantity and its application. *Complexity* 3:1-10.
- Zhu XY, Dang YG, Ding S (2020) Using a self-adaptive grey fractional weighted model to forecast Jiangsu's electricity consumption in China. *Energy* 190:116417.