A novel grey power-Markov model for the prediction of China's electricity consumption

Liqin Sun[1] · Youlong Yang[1] · Tong Ning[1] · Jiadi Zhu[1]

Abstract

Forecasting the electricity consumption has always played an important role in the management of power system management, which requires higher forecasting technology. Therefore, based on the principle of "new information priority", combined with rolling mechanism and Markov theory, a novel grey power-Markov prediction model with time-varying parameters (RGPMM(λ,1,1)) is designed, which overcomes the inherent defects of fixed structure and poor adaptability to the changes of original data. In addition, in order to prove the validity and applicability of the prediction model, we have used the model to predict China's total electricity consumption, and have compared it with the prediction results by a series of benchmark models. The result shows that the can better adapt to the characteristics of electricity consumption data, and it also shows the advantages of the proposed forecasting model. In this paper, the proposed forecasting model is used to predict China's total electricity consumption in the next six years from 2018 to 2023, so as to provide certain reference value for power system management and distribution.

Introduction

China is a developing country and is in a period of rapid development. Electricity consumption forecast is an important part of Power Economic Planning, energy investment, and environmental protection (Lin and Liu 2016). Electricity consumption forecast has become an important research area in the operation and management of modern power systems (Kavousi-Fard et al. 2014). The high and low accuracy of electricity consumption forecasting (Amber et al. 2018) is of great significance to economic development and power planning. Accurate electricity consumption forecast is affected by a series of factors, such as population (Hussain et al. 2016), economic growth (Lin and Liu 2016), power facilities (Khosravi et al. 2012), and climate factors (Hernández et al. 2013), making the prediction problem a

challenging and complex task. In order to solve these problems, in recent years, many domestic and foreign experts, scholars and related research institutions have done a lot of in-depth research on electricity consumption forecast models. The main methods are non-linear intelligent models (Bekiroglu et al. 2018; Hernandez et al. 2014), traditional statistical analysis models (Chui et al. 2009; Mohamed and Bodger 2005), and grey prediction models (Xiao et al. 2017).

Ghani and Ahmad (2010) used SPSS software to establish a linear regression model based on the multiple regression method to predict and analyze the fish landing to demonstrate the effectiveness and feasibility of the method. Wang J et al. (2018) used an autoregressive moving average (ARMA) model based on the time series algorithm to predict short-term wind power. Zhang et al. (2019) combined BP and RBF neural network methods to predict and analyze wind speed to verify the effectiveness and accuracy of the combined prediction model. The above models usually require a large amount of data, so it is difficult to get accurate results when dealing with limited data. Grey theory (Julong 1982) is used to solve the problem of uncertainty with limited data and poor information. It focuses on building a grey prediction model with a small amount of

---

✉ Liqin Sun
liqinsun@stu.xidian.edu.cn

[1] School of Mathematics and Statistics, Xidian University, Xi'an, 710071, People's Republic of China

information. The grey theory is not to find the statistical law of time series data, but to associate the random process with time, and use the cumulative generation operation to process the original data. It reduces the inherent randomness of the data, transforms the irregular data into an exponential form, generates a sequence with strong regularity, and can predict the future direction of the data according to the theory. Grey systems are often used to discover the laws hidden in chaotic data. Wang et al. (1819) used the grey management degree and grey theory to establish a grey system model based on the basic data of urban heating and forecast its demand. Guo et al. (2013) proposed a new comprehensive adaptive grey model, CAGM(1,N), which can be applied to any actual forecasting issues and can obtain higher fitting and prediction accuracy compared with the traditional GM(1,N) model (Pai et al. 2008; Tien 2008).

These methods make predictions based on raw data by maximizing the fitting accuracy, but do not take into account the complex diversity of things themselves. The single grey model has a large error in the process of forecasting, and it is difficult to achieve the expected accuracy. The Markov model is suitable for the prediction of random problems, and it can better describe the dynamic trend of randomly changing objects and can make up for the shortcomings of the grey model. Therefore, two prediction models are combined. The grey model is used to forecast, and the predicted values are corrected by the Markov chain to effectively improve the accuracy of the prediction, so as to achieve the purpose of scientific prediction and analysis. So far, the combination of grey theory and Markov model has applied many areas of prediction, such as Yong and Yidan (1992) put forward the grey Markov model for the first time by combining the advantages of grey model and Markov theory, which has since been widely used in the prediction of traffic, natural disasters, energy consumption, and other fields. Kumar and models (2010) combined the grey Markov and time series model to ground breakingly predict energy consumption in India, which provided a feasible scheme for the prediction of India's energy consumption. CAO Jian et al. (2019) explored the internal relationship among the accidents of road transport on hazardous chemical and the traffic accidents in China based on grey Markov model, and analyzed the grey Markov combined prediction model in the prediction of safety accident. The effectiveness and feasibility of the method had been verified by experiments.

The above research shows that the grey Markov model has better prediction accuracy and ability. However, the actual operation found that when faced with the prediction of small sample data, the model still has a certain degree of contingency. Therefore, according to the characteristics of China's electricity consumption data, this paper optimizes and improves the traditional grey Markov model, and proposes RGPMM(λ,1,1) to predict China's electricity consumption more accurately, so as to provide more accurate information for the rational distribution of energy.

objective

This article has carried out the work in the following aspects:

1) Grey power prediction model (Wang et al. 2011). It is a new type of nonlinear grey prediction model. Its power index can reflect the nonlinear development characteristics of the data and is used to describe the development of things. The nonlinear situation has good prediction results.

2) The power index λ is generally taken as integer in the traditional grey prediction models, such as $\lambda = 2$, which is called Verhulst model. In this paper, λ belongs to the real number R, that is, λ can be taken the fraction to establish the model and the value of can be estimated through optimization theory. Then a novel grey prediction model can be established. In the process of modeling, the robustness of λ is analyzed.

3) The introduction of rolling mechanism (Akay and Atak 2007). In the forecasting process, because the data from the far past has little effect on the forecast, the Rolling Mechanism is introduced to continuously update the input information, which breaks the constraint of constant initial value in the classic grey prediction model and complies with the principle of "new information priority" (Julong 1989).

4) In this paper, the relative error of the grey power model is used as the index, and the weighted Markov theory (Liu et al. 2018) is used to correct the grey power model, which further improves the prediction accuracy and adaptability of the model.

Organization

The rest of this article is organized as follows: the "Basic knowledge" section briefly introduces the historical background of the grey model and the traditional GM(1,1) model, Markov theory, rolling mechanism, and the grey development zone. The "Methodology of improved grey prediction model" section introduces how to build the RGPMM(λ,1,1). The "Case studies on forecasting the total electricity consumption in China" section illustrates the practicability of the RGPMM(λ,1,1) by experiment, and forecasts the total electricity consumption in the next few years by this model. The "Conclusion" section contains the conclusions and suggestions for future work.

Basic knowledge

Basic GM(1,1)

In 1982, Professor Deng Julong first proposed the concept of grey system and built the GM(1,1). The process of the GM(1,1) is as follows Julong (1982), Lin et al. (2012), and Zeng et al. (2020):

Step 1: Transforing the original data. Let a set of non-negative sequences is $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}, (n \geq 4)$. The 1-AGO sequence is given by

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

where, $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 1, 2, \cdots, n$.

Step 2: Based on the sequence $X^{(1)}$, the whitening form equation of the prediction model can be established:

$$\frac{dx^{(1)}}{dt} + a \cdot x^{(1)} = b. \tag{2}$$

In Formula (2), a and b are the parameters to be estimated. The grey differential equation is:

$$x^{(0)}(k) + a \cdot z^{(1)}(k) = b, \tag{3}$$

where, $z^{(1)}(k) = \frac{1}{2} \cdot \left[x^{(1)}(k) + x^{(1)}(k-1)\right]$ is the background value and
$Z^{(1)} = \left\{z^{(1)}(2), z^{(1)}(3), \cdots, z^{(1)}(n)\right\}$ is the mean sequence of $X^{(1)}$ (Xiong et al. 2014).

Step 3: Estimating the model parameters. Set the parameters vector to be estimated as $\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}$ and solve it according to the least squares method to obtain

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \left(B^T B\right)^{-1} B^T Y, \tag{4}$$

where $B = \begin{pmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \cdots & \cdots \\ -z^{(1)}(n) & 1 \end{pmatrix}$, $Y = (x^{(0)}(2), x^{(0)}(3),$
$\cdots, x^{(0)}(n))^T$.

Step 4: Obtaining the time response function. According to Eq. 4, solve (2), then the time corresponding equation is computed as:

$$\hat{x}^{(1)}(t+1) = \left(x^{(0)}(1) - \frac{b}{a}\right) \cdot e^{-at} + \frac{b}{a},$$
$$t = 1, 2, \cdots, n, n + 1, \cdots.$$

Step 5: Obtaining the fitted and predicted values in the original domain. The simplified predicted value of the first-order accumulation operator sequence is $\hat{x}^{(0)}(t) = \hat{x}^{(1)}(t) - \hat{x}^{(0)}(t-1)$, namely,

$$\hat{x}^{(0)}(t) = \left[x^{(0)}(1) - \frac{\hat{b}}{\hat{a}}\right] \cdot \left(1 - e^{\hat{a}}\right) \cdot e^{-\hat{a}(t-1)},$$
$$t = 2, 3, \cdots, n, n + 1, \cdots.$$

where $\hat{x}^{(0)}(t)(t \leq n)$ are called fitted values, and $\hat{x}^{(0)}(t)(t > n)$ are called predicted values.

The flow chart is shown in Fig. 1.

Markov process

Markov process (Zhao et al. 2014) is a theory that studies the state of things and their transition. A Markov process in which time and state are both discrete is called a Markov chain. Markov chain analysis is a statistical analysis method based on the probability theory and stochastic process theory, using stochastic mathematical models to analyze the quantitative relationship of objective objects in the development and change process. Its characteristic is no after-effect, that is, the current state of the system is only related to the previous state, and has nothing to do with the subsequent state.
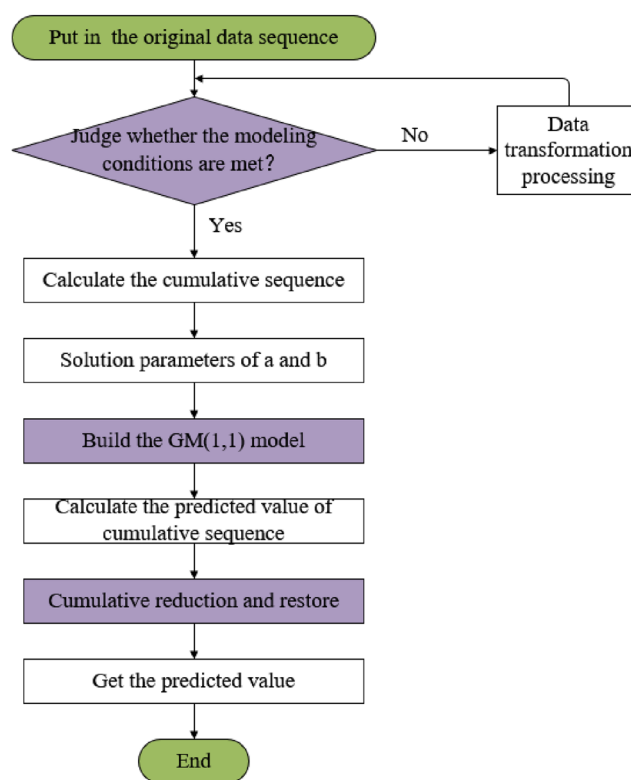

Fig. 1 The flow chart of GM(1,1)

Transition probability and transition probability matrix

In Markov process, the transition probability and the transition probability matrix of states need to be calculated, which are defined as follows:

Definition 1 Let $\{X_n, n \in T\}$ be a Markov chain, and call the conditional probability $p_{ij}(n, T) = P(X_{n+1}|X_n = i), i, j \in T$ the one-step transition probability of the Markov chain $\{X_n, n \in T\}$ at time n, which is referred to as the transition probability. That is, the conditional probability that the particle is in state i at time n and then is in state j after one step. The matrix composed of transition probability is the transition probability matrix. In a Markov chain, the system state transition can be represented by the transition probability matrix P as follows:

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

The steps of Markov process

Markov process is introduced to obtain the transition probability of residual state, so as to determine the state of the residual when $t > n$. The steps are as follows:

Step 1: Determine the residual state;

Step 2: Calculate the state transition probability matrix P according to the residual state;

Step 3: Determine the initial state vector;

Step 4: According to the state transition formula, calculate the result of the tth state transition, and take the one with higher probability of occurrence status.

Methodology of improved grey prediction model

The grey power model

The GPM(λ,1,1) is an extension of the traditional GM(1,1). In this paper, the power exponent of GPM(λ,1,1) is analyzed according to the information covering principle of grey system, and the following definitions are given.

Definition 2 (The grey power model) Assuming $X^{(0)}$ is a non-negative unimodal raw data sequence, $X^{(1)}$ is the $1 - AGO$ sequence of $X^{(0)}$. $Z^{(1)}$ is a sequence generated next to the mean of the $X^{(1)}$. Then, there is the following non-linear model which meets the three conditions of gray modeling, and the grey power model is

$$x^{(0)}(k) + a \cdot z^{(1)}(k) = b \cdot \left[z^{(1)}(k)\right]^{\lambda}$$

The whitening equation of the grey power model is

$$\frac{dx^{(1)}}{dt} + a \cdot x^{(1)} = b \cdot \left[x^{(1)}\right]^{\lambda} \quad (5)$$

Solving the above model by the solution method of GM(1,1), we can get the solution of the whitening equation is

$$x^{(t+1)} = \left\{e^{-(1-\lambda)at}\left[(1-\lambda)\int be^{(1-\lambda)at}dt + c\right]\right\}^{\frac{1}{1-\lambda}}. \quad (6)$$

Parameters analysis of *GPM*(λ, 1, 1)

Parameter λ estimation method

The parameter λ is an important coefficient in the GPM(λ,1,1). According to the above formulas, since $x^{(1)} \neq 0$, divide both sides of Eq. 5 by $\left[x^{(1)}\right]^{\lambda}$ and then take the deriation about t to get (7) as follows:

$$\frac{d^2x^{(1)}}{dt^2} \cdot \left[x^{(1)}\right]^{\lambda} - \lambda \cdot \left(\frac{dx^{(1)}}{dt}\right)^2\left[x^{(1)}\right]^{\lambda-1} = -a(1-\lambda)$$
$$\cdot \left[x^{(1)}\right]^{\lambda} \cdot \frac{dx^{(1)}}{dt} \quad (7)$$

According to the information coverage principle of grey derivative, we cover $\frac{dx^{(1)}}{dt}$ and $\frac{d^2x^{(1)}}{dt^2}$ in Eq. 7 with the first grey derivatives and the second grey derivatives of $x^{(1)}$, then we will get

$$\left[x^{(0)}(t) - x^{(0)}(t-1)\right] \cdot \left[z^{(1)}(t)\right]^{\lambda}$$
$$- \lambda \cdot \left[x^{(0)}(t)\right]^2 \cdot \left[z^{(1)}(t)\right]^{\lambda-1} \quad (8)$$
$$= -a(1-\lambda) \cdot \left[z^{(1)}(t)\right]^{\lambda} \cdot x^{(0)}(t)$$

Dividing the Eq. 8 with $t = k$ by the Eq. 8 with $t = k+1$, we can eliminate the unknown parameter a and get

$$\frac{\left[x^{(0)}(k) - x^{(0)}(k-1)\right] \cdot \left[z^{(1)}(k)\right]^{\lambda} - \lambda \cdot \left[x^{(0)}(k)\right]^2 \cdot \left[z^{(1)}(k)\right]^{\lambda-1}}{\left[x^{(0)}(k+1) - x^{(0)}(k)\right] \cdot \left[z^{(1)}(k+1)\right]^{\lambda} - \lambda \cdot \left[x^{(0)}(k+1)\right]^2 \cdot \left[z^{(1)}(k+1)\right]^{\lambda-1}}$$
$$= \frac{\left[z^{(1)}(k)\right]^{\lambda} \cdot x^{(0)}(k)}{\left[z^{(1)}(k+1)\right]^{\lambda} \cdot x^{(0)}(k+1)} \quad (9)$$

It follows from Eq. 9 that

$$
\begin{aligned}
\lambda = & \frac{\left[x^{(0)}(k+1) - x^{(0)}(k)\right] \cdot z^{(1)}(k+1) \cdot z^{(1)}(k) \cdot x^{(0)}(k)}{\left[x^{(0)}(k+1)\right]^2 \cdot z^{(1)}(k) \cdot x^{(0)}(k) - \left[x^{(0)}(k)\right]^2 \cdot z^{(1)}(k+1) \cdot x^{(0)}(k+1)} \\
& - \frac{\left[x^{(0)}(k) - x^{(0)}(k-1)\right] \cdot z^{(1)}(k) \cdot z^{(1)}(k+1) \cdot x^{(0)}(k+1)}{\left[x^{(0)}(k+1)\right]^2 \cdot z^{(1)}(k) \cdot x^{(0)}(k) - \left[x^{(0)}(k)\right]^2 \cdot z^{(1)}(k+1) \cdot x^{(0)}(k+1)}
\end{aligned}
\tag{10}
$$

From the expression of λ, we can see that it can not only reflect the grey derivative of the original data, but also reflect the role of grey integral. When $k = 2, 3, \cdots, n - 1$, the corresponding $(n - 2)$ values of λ can be computed, which is $\{\lambda_k\}$.

Let $g(\lambda) = \sum_{k=2}^{n-1} (\lambda - \lambda_k)^2$, the value of λ that makes $g(\lambda)$ take the minimum value is the constant value to be determined.

Since $g(\lambda)$ is a parabola with an opening upward, according to the first order condition of unconditional extremum, the optimal value of λ is

$$
\hat{\lambda} = \frac{1}{n-2} \sum_{k=2}^{n-1} \lambda_k.
\tag{11}
$$

In this case, $g\left(\hat{\lambda}\right)$ takes the minimum value.

Estimates of parameters *a* and *b*

After the optimal value of λ is determined, the parameters a and b can be estimated directly according to the least square method. Then, we can get the theorem 1.

Theorem 1 *Assuming $X^{(0)}$, $X^{(1)}$ and $Z^{(1)}$ are as defined in Definition 1, then the least squares estimate of the parameter sequence in $GPM(\lambda, 1, 1)$ is*

$$
\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \left(B^T B\right)^{-1} B^T Y,
\tag{12}
$$

where $B = \begin{pmatrix} -z^{(1)}(2) & \left[z^{(1)}(2)\right]^{\hat{\lambda}} \\ -z^{(1)}(3) & \left[z^{(1)}(3)\right]^{\hat{\lambda}} \\ \cdots & \cdots \\ -z^{(1)}(n) & \left[z^{(1)}(n)\right]^{\hat{\lambda}} \end{pmatrix}$, $Y = x^{(0)}(2), x^{(0)}$
$(3), \cdots, x^{(0)}(n)^T$.

Solution of *GPM*(λ, 1, 1)

According to Eq. 12 and the estimation results of parameters, we can simplify it to get $\hat{x}^{(1)}(t+1) =$

$\left[c \cdot e^{-(1-\hat{\lambda})\hat{a}t} + \hat{b}/\hat{a}\right]^{\frac{1}{1-\hat{\lambda}}}$. If the initial value $\hat{x}^{(1)}(1) = x^{(0)}(1)$, then the solution of GPM(λ,1,1) is

$$
\begin{aligned}
\hat{x}^{(1)}(t+1) = & \left\{\left[\left(x^{(0)}(1)\right)^{1-\hat{\lambda}} - \hat{b}/\hat{a}\right]\right. \\
& \left. \cdot e^{-(1-\hat{\lambda})\hat{a}t} + \hat{b}/\hat{a}\right\}^{\frac{1}{1-\hat{\lambda}}}.
\end{aligned}
\tag{13}
$$

Rolling modeling mechanism

It is easy to produce some unacceptable errors in practical applications. In order to reduce the errors, rolling mechanism is proposed. The length of training data set is set as c, and the predicted period of rolling modeling is set as d. The steps are as follows:

Step 1: The original sequence $\left\{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(c)\right\}$ is used to model and the d-period prediction value $\left\{\hat{x}^{(0)}(c+1), \hat{x}^{(0)}(c+2), \cdots, \hat{x}^{(0)}(c+d)\right\}$ is obtained;

Step 2: When predicting the sequence $\{\hat{x}^{(0)}(c+d+1), \hat{x}^{(0)}(c+d+2), \cdots, \hat{x}^{(0)}(c+2d)\}$, we use the latest c data points $\left\{\hat{x}^{(0)}(d+1), \hat{x}^{(0)}(d+2), \cdots, \hat{x}^{(0)}(d+c)\right\}$ to predict;

Step 3: Repeat step 2 and use the latest sequence to predict the next set of d data points until the required data points are predicted.

The flow chart is shown in Fig. 2.

Building the *RGPMM*(λ, 1, 1)

Due to the complexity of the real situation, there will always be a certain difference between the fitting value obtained by GPM(λ,1,1) and the real value. Then, the accuracy index of grey fitting is random and non-stationary. In order to correct the predictable result and improve the prediction accuracy of GPM(λ,1,1), the fluctuation of grey fitting accuracy index is analyzed and predicted by Markov theory in this paper. Combined with the rolling mechanism, the electricity consumption is forecasted in the future. The steps of RGPMM(λ,1,1) are as follows and the flow chart is shown in Fig. 3.

Fig. 2 The rolling mechanism
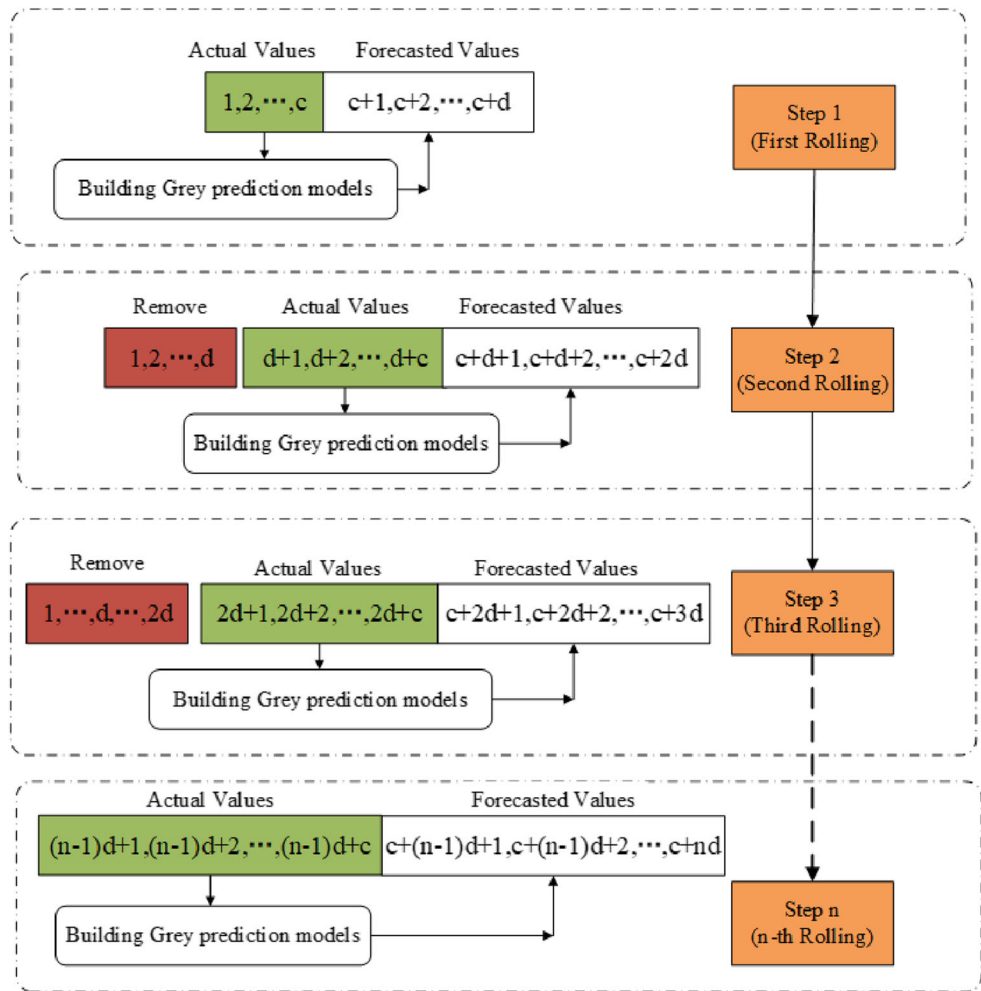
Step 1: Calculate fitted values and predicted values According to the time series $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}$, GPM(1,1) is established to obtain

$$
\begin{aligned}
&\hat{x}^{(0)}(t+1) \\
&= \left\{ \left[\left(x^{(0)}(1) \right)^{1-\hat{\lambda}} - \frac{\hat{b}}{\hat{a}} \right] \cdot e^{-\left(1-\hat{\lambda}\right)at} + \frac{\hat{b}}{\hat{a}} \right\}^{\frac{1}{1-\hat{\lambda}}} \\
&\quad - \left\{ \left[\left(x^{(0)}(1) \right)^{1-\hat{\lambda}} - \frac{\hat{b}}{\hat{a}} \right] \cdot e^{-\left(1-\hat{\lambda}\right)a(t-1)} + \frac{\hat{b}}{\hat{a}} \right\}^{\frac{1}{1-\hat{\lambda}}}
\end{aligned}
\tag{14}
$$

where $\hat{x}^{(0)}(t)(t \leq n)$ are called fitted values, and $\hat{x}^{(0)}(t)(t > n)$ are called predicted values.

Step 2: Calculate the grey fitting accuracy index The grey fitting accuracy index is set to $Y(t) = x^{(0)}(t)/\hat{x}^{(0)}(t)$, which reflects the deviation degree of the data fitted by model from the original data.

Step 3: Division of state interval Considering that $Y(t)$ is divided into m state $E_i = [\otimes_{1i}, \otimes_{2i}], i = 1, 2, \cdots, m$.

The grey elements $\otimes_{1i}$ and $\otimes_{2i}$ are the lower bound and the upper bound of the ith state, where $\otimes_{1i} = Y(t) + a_i \overline{Y}$, $\otimes_{2i} = Y(t) + b_i \times \overline{Y}$, $\overline{Y} = \frac{1}{n} \cdot \sum_{i=1}^{n} Y(i)$. The a_i and b_i are constants that need to be determined based on experience and data.

Considering the limited amount of electricity consumption data in this paper, it is more appropriate to use the cluster analysis to determine the number of classification classes and the classification interval.

Step 4: Establish a state transition matrix The transition probability of state E_i to state E_j is

$$
p_{ij}(\omega) = \frac{M_{ij}(\omega)}{M_i}.
$$

where, $M_{ij}(\omega)$ is the number of samples $Y(t)$ transferred from the state of E_i to the state of E_j through ω steps; M_i is the total number of occurrences of E_i, and satisfies $\sum_{j=1}^{m} p_{ij}(\omega) = 1, i, j = 1, 2, \cdots, m$. Therefore, $p_{ij}(\omega)$ reflects the probability of transition from E_i to E_j through ω steps.
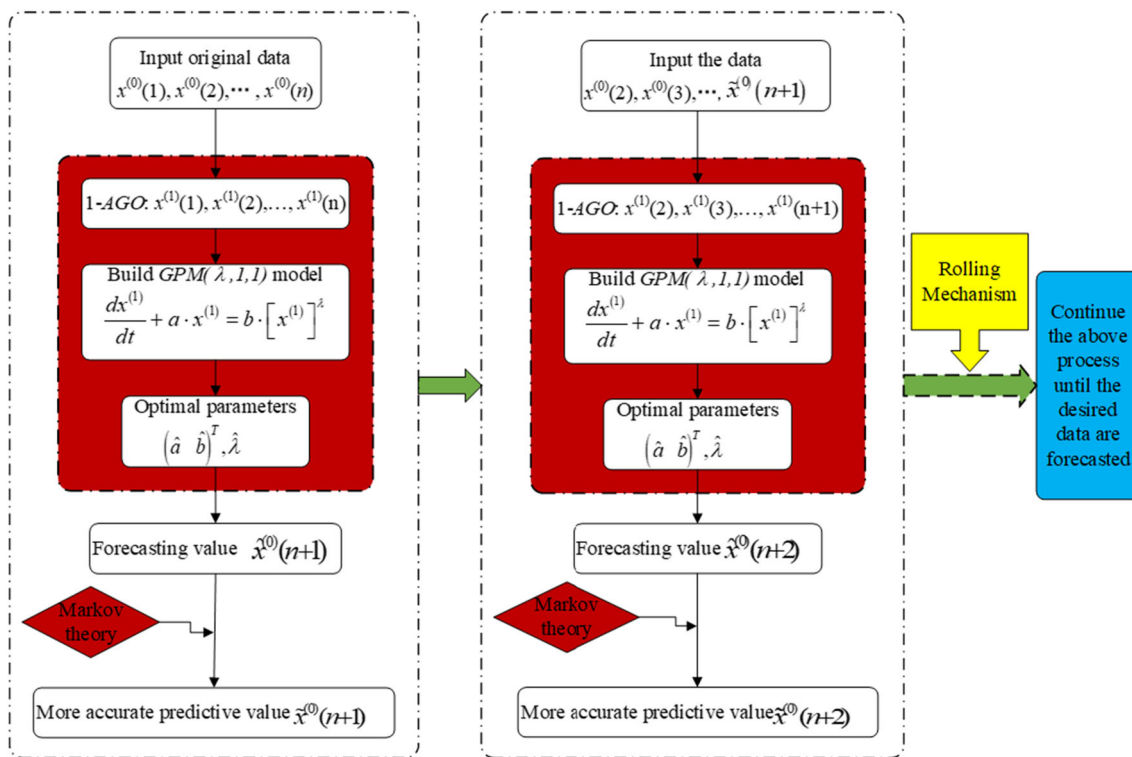
Fig. 3 The flow chart of RGPMM(λ,1,1)

The state transition probability matrix is

$$R(\omega) = \begin{pmatrix} p_{11}(\omega) & p_{12}(\omega) & \cdots & p_{1m}(\omega) \\ p_{21}(\omega) & p_{22}(\omega) & \cdots & p_{2m}(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1}(\omega) & p_{m2}(\omega) & \cdots & p_{mm}(\omega) \end{pmatrix}$$

The $R(\omega)$ reflects the transfer law between the various states of the system. By examining $R(\omega)$ and the current state, we can predict the future development and change of the system.

The autocorrelation coefficient of each order is

$$r_\omega = \frac{\sum\limits_{l=1}^{n-\omega} \left[Y(l) - \overline{Y} \right] \cdot \left[Y(l+\omega) - \overline{Y} \right]}{\sum\limits_{l=1}^{n} \left[Y(l) - \overline{Y} \right]^2}$$

By normalizing r_ω, the Markov weight of each order is

$$\theta_\omega = \frac{|r_\omega|}{\sum\limits_{\omega=1}^{m} |r_\omega|}, \omega \leq m.$$

where, θ_ω is the Markov weight of the ωth order, and the ωth order generally is the maximum order when $|r_\omega| \geq 0.3$.

Step 5: Calculate more accurate the predicted value The transition probability matrix is used to predict the state interval E_i of the grey fitting precision index. The interval interpolation is used to determine the predicted value. Therefore, $\tilde{x}^{(0)}(n+1) = \hat{x}^{(0)}(n+1) * \widehat{Y}(n+1)$, where

$$\widehat{Y}(n+1) = \otimes_{1i} \times \frac{p_{i-1}}{p_{i-1} * p_{i+1}} + \otimes_{2i} \times \frac{p_{i+1}}{p_{i-1} * p_{i+1}} \quad (15)$$

Step 6: Calculate more accurate predicted values in the next few years Through a rolling mechanism, the input data is updated and the RGMM(λ,1,1) is established to forecast the next year's value. Continue the above process until the desired data forecasted.

Evaluation metrics

Here are three kinds of evaluation metrics to evaluate the prediction accuracy. Only when the three metrics are all passed, the RGPMM(λ,1,1) can be used to predict, and its predicted values have reference significance. There are three kinds of evaluation metrics as follows:

A: The residual test The three statistical indicators are determined, namely MAE (Mean Absolute Error) (Hamzaçebi 2007), MAPE (Mean Absolute Percentage Error) (Azadeh et al. 2008), and RMSE (The Root Mean Squared Error)

(Geem and Roper 2009). The formulas for MAE,MAPE and RMSE are as follows:

$$MAE = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| x^{(0)}(i) - \hat{x}^{(0)}(i) \right|,$$

$$MAPE = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| \frac{x^{(0)}(i) - \hat{x}^{(0)}(i)}{x^{(0)}(i)} \right|$$

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} \left(x^{(0)}(i) - \hat{x}^{(0)}(i) \right)^2}$$

where $x^{(0)}(i)$ is the original value at time i, and the $\hat{x}^{(0)}(i)$ is the fitted value at time i. Table 1 shows criteria of forecasting performance.

B: The correlation degree The test method of correlation degree is γ_g:

$$\gamma_g = \frac{1}{n-1} \cdot \sum_{i=1}^{n} \gamma_g(i) = \frac{1}{n-1}$$

$$\cdot \sum_{i=1}^{n} \frac{\min\{\Delta(i)\} + \rho \cdot \max\{\Delta(i)\}}{\Delta(i) + \rho \cdot \max\{\Delta(i)\}} x$$

where, $\rho = 0.5$, $\Delta(i) = \left| \hat{x}^{(0)}(i) - x^{(0)}(i) \right|$, $i = 1, 2, \cdots, n$, n is the number of samples.

C: The posterior error test The process of posterior error test is as follows:

Step 1: Calculate $S_0 = \sqrt{\frac{\sum_{i=1}^{n}(X^{(0)}(i) - \overline{X}^{(0)})^2}{n-1}}$ of $\{X^{(0)}(i)\}$ and $S_1 = \sqrt{\frac{\sum_{i=1}^{n}(\varepsilon^{(0)}(i) - \overline{\varepsilon}^{(0)})^2}{n-1}}$ of $\{\varepsilon^{(0)}(i)\}$, where $\overline{X}^{(0)} = \frac{1}{n} \cdot \sum_{i=1}^{n} X^{(0)}(i)$, $\varepsilon^{(0)}(i) = x^{(0)}(i) - \hat{x}^{(0)}(i)$, $\overline{\varepsilon}^{(0)} = \frac{1}{n} \cdot \sum_{i=1}^{n} \varepsilon^{(0)}(i)$, $i = 1, 2, \cdots, n$

Step 2: The standard deviation ratio $C = \frac{S_1}{S_0}$;

Step 3: The error probability $P = \{|\varepsilon^{(0)}(i) - \overline{\varepsilon}^{(0)}(i)| < 0.6745 \times S_0\}$;

Step 4: The discrimination rules are shown in Table 2.

Case studies on forecasting the total electricity consumption in China

Since forecasting electricity consumption is important for the dispatch and operation of the power system, the total electricity consumption forecast at the national level is put forward. To prove the prediction accuracy of RGPMM(λ,1,1) in the "Methodology of improved grey prediction model

Table 1 MAPE criteria for model evaluation

MAPE(%)	Forecasting ability
<10	Highly accurate forecasting
10–20	Good forecasting
20–50	Reasonable forecasring
>50	Weak and inaccurate forecasting

section", it is compared with GM(1,1), GM(1,1,X_n) (Bahrami et al. 2014; Dang et al. 2004), OICGM(1,1) (Akdi et al. 2020), and NOGM(1,1) models (Ding et al. 2018).

Experimental data

The experimental data is from the China Statistical Yearbook published by the National statistics Bureau of China (Bureau, 2019). The experimental data set includes the total amount of electricity consumption from 2008 to 2017, as shown in Table 3.

Selecting the input data set and determining the length of the input data set are the prerequisites to affect the accuracy of the prediction model. In this paper, the optimized subset method proposed by Wang (Wang et al. 2011) is used to determine the optimal length of the input data set, and $c = 8$ can be obtained. Eight data points are used as input data points, and the data of the last two years are used to test the prediction effect of the model.

The prediction results of experiment

In Table 3, the original sequence is $X^{(0)} = (3.45414, 3.70322, 4.19345, 4.70009, v4.97626, 5.42034, 5.63837, 5.80200, 6.12971, 6.48210)$. The flow chart of using the RGPM(λ,1,1) model to predict China's electricity consumption is shown in Fig. 4.

According to Theorem 1 and Formula (11) in the "Methodology of improved grey prediction model" section, the parameters of λ and $\begin{pmatrix} a \\ b \end{pmatrix}$ are estimated twice by MATLAB software programming to predict the total electricity consumption in 2016, 2017. Its calculation and results are as follows:

In 2016, the estimates calculated by MATLAB are $\hat{\lambda} = 0.2539$ and $\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} 0.0041 \\ 2.4323 \end{pmatrix}$.

Substitute the parameter estimates into Eq. 13, the time response sequence function by RGPM(λ,1,1) model in 2016 is

$$\hat{x}^{(1)}(t+1) = \left\{ \left[\left(x^{(0)}(1) \right)^{(1-0.2539)} - \frac{2.4323}{0.0041} \right] \right.$$

$$\left. * e^{-(1-0.2539)*0.0041*t} + \frac{2.4323}{0.0041} \right\}^{\frac{1}{(1-0.2539)}}$$

Table 2 the gradation of prediction accuracy

P-value	C-value	Gradation
>0.95	<0.35	Good
>0.85	<0.50	Qualified
>0.70	<0.65	Barely qualified
≤ 0.70	≥ 0.65	Unqualified

Table 3 Total electricity energy consumption in China from 2008 to 2017($10^{12}kwh$)

Year	2008	2009	2010	2011	2012
Actual value	3.45414	3.70322	4.19345	4.70009	4.97626
Year	2013	2014	2015	2016	2017
Actual value	5.42034	5.63837	5.80200	6.12971	6.48210

Similarly, by entering the new initial value, the estimated values of parameters in 2017 are $\hat{\lambda} = -0.0035$ and $\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} -0.0587 \\ 4.0794 \end{pmatrix}$.

Then, the time response sequence function of 2017 is

$$\hat{x}^{(1)}(t+1) = \left\{ \left[\left(x^{(0)}(1) \right)^{(1+0.0035)} - \frac{4.0794}{-0.0587} \right] *e^{(1+0.0035)*0.0587*t} + \frac{4.0794}{-0.0587} \right\}^{\frac{1}{(1+0.0035)}}$$

The different parameter values of each step indicate that the RGPM(λ,1,1) can make the prediction result dynamic according to the characteristics of the input data. Figure 5 intuitively shows that the forecast results by RGPM(λ,1,1) are in good agreement with the real values.

The prediction values of each grey model are shown in Table 4 and Figs. 6 and 7.

The related parameters and run time of each grey model are shown in Table 5.

Through Table 4 and Figs. 6 and 7, it can be intuitively shown that the fitting degree between the predicted results of RGPM(λ,1,1) and the real values is better than that

of GM(1,1), GM(1,1,X_n), OICGM(1,1), and NOGM(1,1). This result can also be verified by using the experimental test indicators mentioned in the "Evaluation metrics" section. The prediction effect and the detection results of each grey model are shown in Table 6.

Through Table 6, the values of 6 detection indexes by GM(1,1) are $MAE = 0.6041$, $MAPE = 10.9339\%$, $RMSE = 0.7006$, $r_g = 0.4991$, $P = 0.9000$, and $C = 0.3870$. Since $MAPE = 10.9339\% \in [10\%, 20\%]$, it indicates that the prediction ability of GM(1,1) is weaker than other 4 grey prediction models. Moreover, its $P = 0.9000 < 1.0000$, which further indicates that the predictive ability of GM(1,1), is weaker than that of GM(1,1,X_n), OICGM(1,1), NOGM(1,1), and RGPM (λ,1,1). This result further shows that the grey model with non-integer power is more suitable for the prediction of total electricity consumption, and the intervention of the rolling mechanism further improves the accuracy of the prediction results.

In summary, RGPM(λ,1,1) is superior to other four grey models in all detection indices. That is, its prediction performance is better than other four models. MAE and MAPE are increased by several percentage points respectively, indicating that the prediction accuracy of RGPM(λ,1,1) is greatly improved compared with other four grey models. So, RGPM(λ,1,1) is a better choice to

Fig. 4 The rolling mechanism
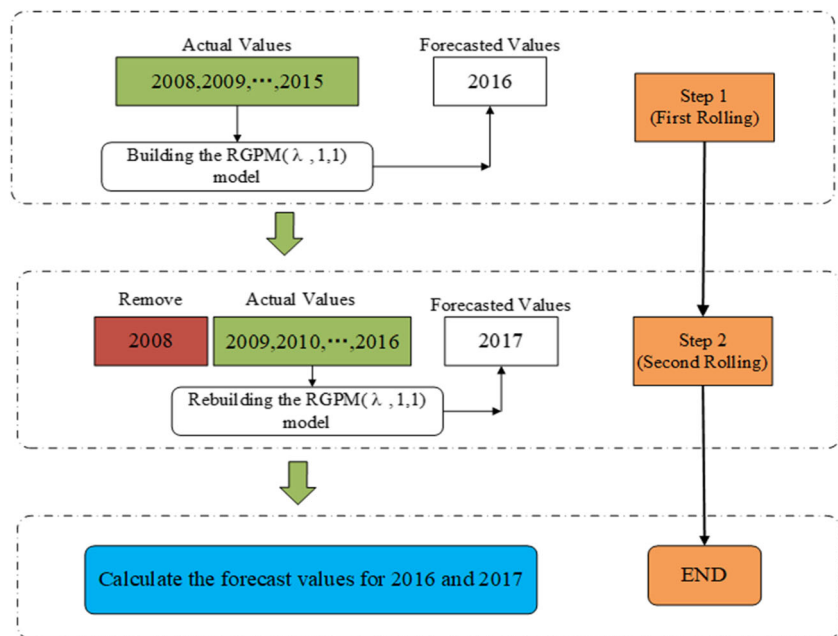
**Fig. 5** The predicted value of RGPM(λ,1,1)
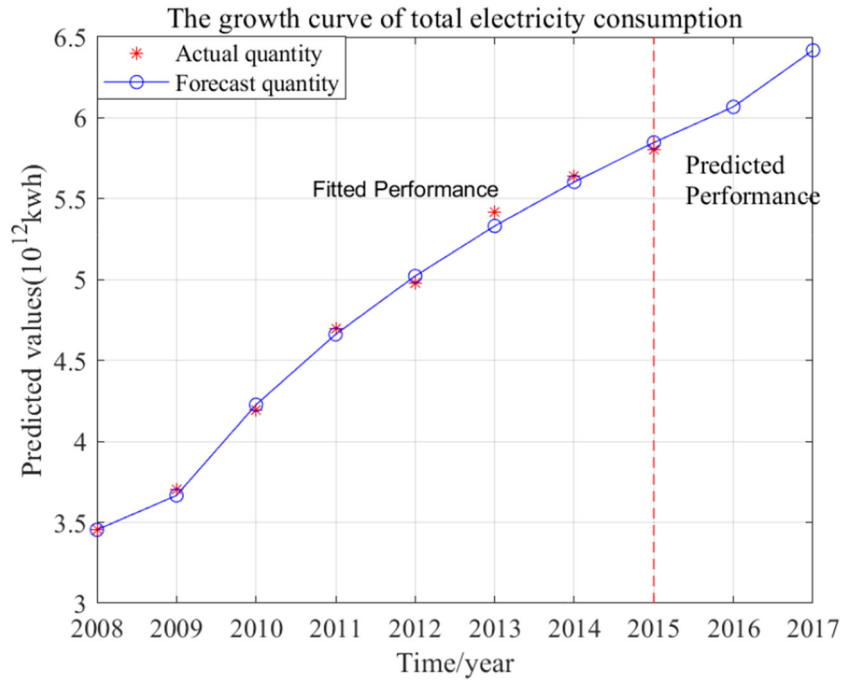
Table 4 Comparison of the original data and the predicted values of each model

Year	Actual value($*10^{12}$)	GM(1,1)	GM(1,1,x_n)	OICGM(1,1)	NOGM(1,1)	RGPM(λ,1,1)
2008	3.45414	3.45414	3.45414	3.45414	3.45414	3.45414
2009	3.70322	3.77109	3.94588	3.92903	4.0711	3.6663
2010	4.19345	3.9498	4.23301	4.21493	4.3673	4.2270
2011	4.70009	4.1371	4.54102	4.52163	4.6851	4.6620
2012	4.97626	4.3332	4.87145	4.85064	5.0260	5.0223
2013	5.42034	4.5386	5.22593	5.20360	5.317	5.3318
2014	5.63837	4.7537	5.60619	5.58224	5.7840	5.6041
2015	5.80200	4.97905	6.01413	5.988437	6.2049	5.8477
2016	6.12971	5.21507	6.45175	6.424188	6.6564	6.0682
2017	6.48210	5.46228	6.92121	6.891646	7.1408	6.4173

Fig. 6 Comparison of the original data and the predicted values of each model
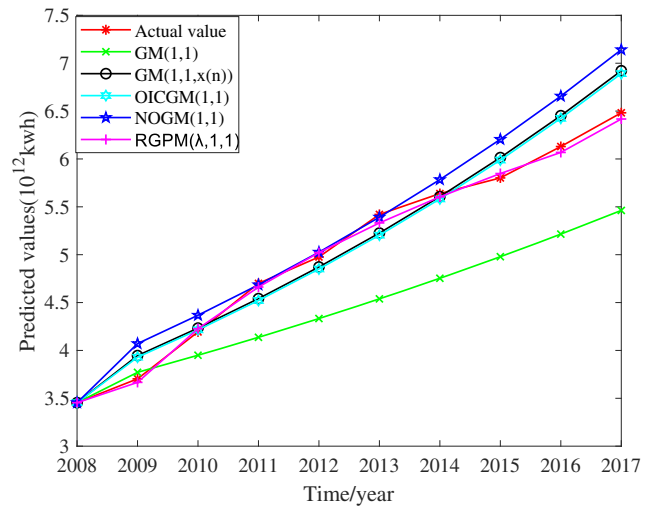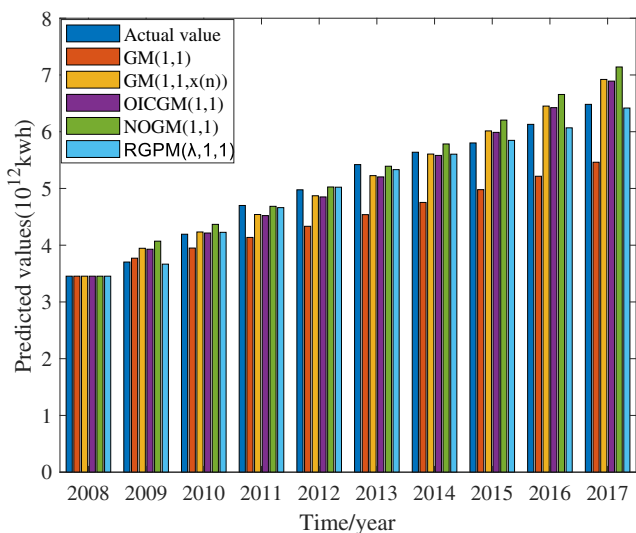
Fig. 7 The histogram of actual values and predicted values for each model

analyze the forecast value of the country's total electricity consumption. However, the more accurate the prediction of national electricity consumption, the better it can provide more accurate forecast information for power workers. In order to improve the prediction performance of the grey power model, the weighted Markov model is used to modify it.

The prediction results of *RGPMM*(λ, 1, 1)

The RGPM(λ,1,1) is used to get the fitted value $\hat{x}^{(0)}(t)$, $t = 1, 2, \cdots, 8$ in each year, where 1-8 denote the 2008-2015, respectively. The accuracy index of grey fitting is obtained according to $Y(t) = \frac{x^{(0)}(t)}{\hat{x}^{(0)}(t)}$ presented in the "Methodology of improved grey prediction model" section, as shown in Table 7.

From Table 7, it can be found that the grey fitting accuracy index has a strong volatility, and the weighted Markov model can be used to predict the state of the grey fitting accuracy.

The Q-type clustering (Narasimhan et al. 2005) is used to divide the state. The clustering result is shown in Fig. 8. According to the cluster diagram, the accuracy indices of grey fitting can be divided into 5 states, which are denoted as $\{E_1, E_2, E_3, E_4, E_5\}$. The clustering results are shown in

Fig. 8 and Table 8. Then, the Markov state interval can be obtained as follows:

$$E_1 \in [95.7350\%, 96.98035\%),$$

$$E_2 \in [96.98035\%, 99.5000\%),$$

$$E_3 \in [99.5000\%, 101.0000\%),$$

$$E_4 \in [101.0000\%, 101.7210\%),$$

$$E_5 \in [101.7210\%, 102.2500\%).$$

According to the grey fitting accuracy index sequence, the autocorrelation coefficients of each order can be calculated as $r_0 = 1.0000$, $r_1 = 0.2306$, $r_2 = -0.0694$, $r_3 = -0.0412$, $r_4 = -0.4134$, $r_5 = -0.1910$, $r_6 = -0.0172$, $r_7 = 0.0016$. The autocorrelation diagram is shown in Fig. 9 and Table 8.

According to Fig. 10, it is found that when m is taken as 4, the condition of $|r_\omega| \geq 0.3$ is satisfied, and there are $r_1 = 0.2306$, $r_2 = -0.0694$, $r_3 = -0.0412$, $r_4 = -0.4134$. Then the Markov weights of each order are obtained as follows:

$$\theta_1 = \frac{0.2306}{0.2306+0.0694+0.0412+0.4134} = 0.3056, \theta_2$$

$$= \frac{0.0694}{0.2306+0.0694+0.0412+0.4134} = 0.0920,$$

$$\theta_3 = \frac{0.0412}{0.2306+0.0694+0.0412+0.4134} = 0.0546, \theta_4$$

$$= \frac{0.4134}{0.2306+0.0694+0.0412+0.4134} = 0.5478.$$

The Markov transition probability matrixes of each step are calculated as follows:

$$p(1) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0.5 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad p(2) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix},$$

$$p(3) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad p(4) = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The accuracy index of grey fitting is in the state of E_3 in 2015, so we can get:

The one-step transition probability vector is (0.5, 0, 0, 0, 0.5), and the corresponding Markov weight is $\theta_1 = 0.3056$;

Table 5 Related parameters and run time of each model

| Parameter	GM(1,1)	GM(1,1,x_n)	OICGM(1,1)	NOGM(1,1)	RGPM(λ,1,1)
a	−0.07024	−0.07024	−0.07024	−0.0702	−0.0587
b	3.56658	3.56658	3.56658	3.56658	4.0794
Run time	0.044531s	0.0439114s	0.0439495s	0.0530092s	0.046502s

Table 6 The absolute error and the inspection result of grey models

Year	GM(1,1)	GM(1,1,x_n)	OICGM(1,1)	NOGM(1,1)	RGPM(λ,1,1)
2008	0	0	0	0	0
2009	0.0679	0.2427	0.2258	0.3679	0.0369
2010	0.2437	0.0396	0.0215	0.1738	0.0335
2011	0.5630	0.1591	0.1785	0.0150	0.0381
2012	0.6431	0.1048	0.1256	0.0497	0.0460
2013	0.8817	0.1944	0.2167	0.0286	0.0885
2014	0.8847	0.0322	0.0561	0.1456	0.0343
2015	0.8229	0.2121	0.1864	0.4029	0.0457
2016	0.9146	0.3220	0.2945	0.5267	0.0615
2017	1.0198	0.4391	0.4095	0.6587	0.0648
MAE	0.6041	0.1746	0.1715	0.2369	**0.0449**
MAPE	10.9339%	3.2828%	3.2260%	4.4208%	**0.8566%**
RMSE	0.7006	0.2184	0.2094	0.3262	**0.0501**
r_g	0.4991	0.7487	0.7490	0.7207	**0.9093**
P	0.9000	**1.0000**	**1.0000**	**1.0000**	**1.0000**
C	0.3870	0.2100	0.2071	0.2487	**0.1282**

Table 7 Grey fitting accuracy index and state division

Year	2008	2009	2010	2011	2012	2013	2014	2015
Y(t)/%	100	95.7350	98.2257	101.1512	99.6787	102.2436	101.1984	99.9776
State	E_3	E_1	E_2	E_4	E_3	E_5	E_4	E_3

Fig. 8 The figure of Q-cluster for 2008–2015
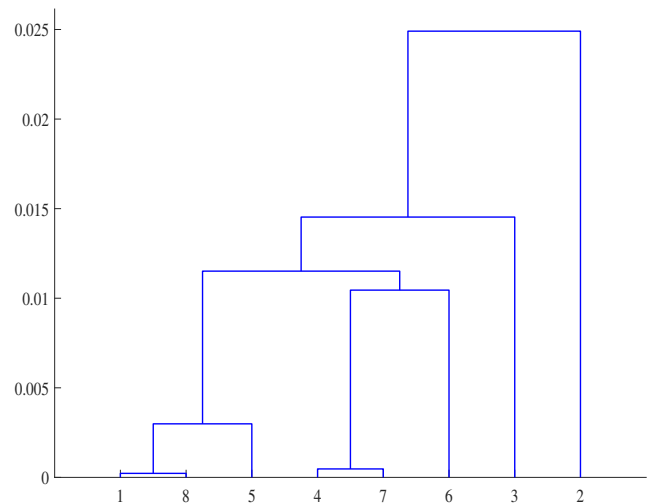
Table 8 The correlation coefficients (r) and Q-cluster results during 2008-2015

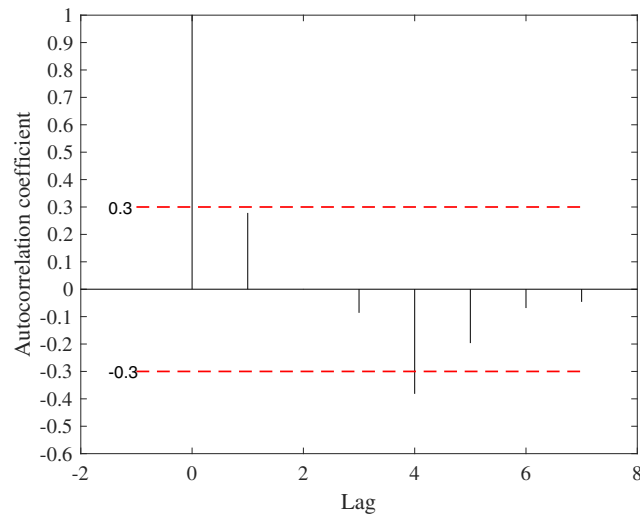| Label	1	2	3	4	5	6	7	8
Year	2008	2009	2010	2011	2012	2013	2014	2015
r	1.0000	0.2306	−0.0694	−0.0412	−0.4134	−0.1910	−0.0172	0.0016
Q-cluster:	Class 1: 2013							
	Class 2: 2011,2014							
	Class 3: 2008,2012,2015							
	Class 4: 2010							
	Class 5: 2009							

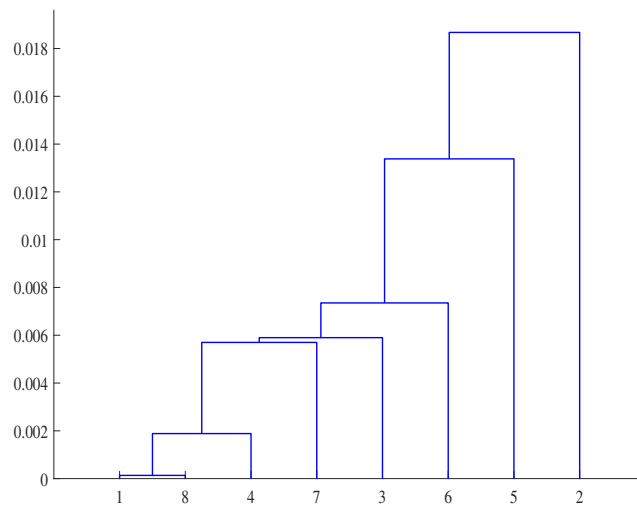**Fig. 9** The each order of autocorrelation coefficient

Fig. 10 The figure of Q-cluster for 2008-2015

**Table 9** Calculation results of weighted Markov transition probability

Start year	state	weight	E_1	E_2	E_3	E_4	E_5
2015	E_3	0.3056	0.5	0	0	0	0.5
2014	E_4	0.0920	0	0.5	0	0.5	0
2013	E_5	0.0546	0	0	0.5	0.5	0
2012	E_3	0.5478	0	0	1	0	0
Weighted transition probability			0.1528	0.0460	0.5751	0.0733	0.1528

Table 10 Grey fitting accuracy index of 2017

Year	2009	2010	2011	2012	2013	2014	2015	2016
Y(t)/%	100	97.3600	100.5900	99.8000	102.6600	101.3300	99.2300	99.9900

Table 11 The correlation coefficients (r) and Q-cluster results during 2009–2016

Label	1	2	3	4	5	6	7	8
Year	2009	2010	2011	2012	2013	2014	2015	2016
r	1.0000	0.0104	−0.0471	−0.3873	−0.2401	0.1349	0.0282	0.0009
Q-cluster:	Class 1: 2015							
	Class 2: 2009,2012,2016							
	Class 3: 2011							
	Class 4: 2014							
	Class 5: 2013							
	Class 6: 2010							
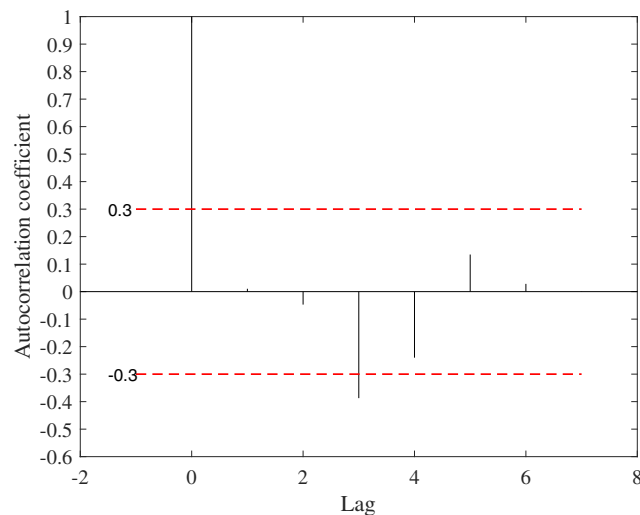
Fig. 11 The each order of autocorrelation coefficient

Table 12 Comparison of prediction effects from the models

Model	2016	2017	MSE	MAPSE
True value(10^{12})	6.12971	6.48210	/	/
RGPM(1,1)	6.0682	6.4173	0.0040	1.00%
RGPMM(1,1)	6.0943	6.5258	**0.0016**	**0.63%**

Table 13 The correlation coefficients (r) and Q-cluster results during 2010-2017

Label	1	2	3	4	5	6	7	8
Year	2010	2011	2012	2013	2014	2015	2016	2017
r	1.0000	0.2228	−0.5766	−0.3175	0.0345	0.0842	0.0527	0.0000
Q-cluster:	Class 1: 2013							
	Class 2: 2014							
	Class 3: 2011							
	Class 4: 2010,2012,2015,2016,2017							

The two-step transition probability vector is $(0, 0.5, 0, 0.5, 0)$, and the Markov weight is $\theta_2 = 0.0920$;

The three-step transition probability vector is $(0, 0, 0.5, 0.5, 0)$, and the Markov weight is $\theta_3 = 0.0546$;

The four-step transition probability vector is $(0, 0, 1, 0, 0)$, and the Markov weight is $\theta_4 = 0.5478$.

According to $p_i = \sum\limits_{\omega=1}^{m} \theta_\omega p_i(\omega)$, the weighted Markov prediction probability is calculated, as shown in Table 9.

From Table 9, the weighted probability value of the transition probability can be obtained. Based on results, $max\{p_i\} = p_3 = 0.5751$, that is, the probability of being in the state of E_3 in 2016 is the largest. And its adjacent states are E_2 and E_4. According to Formula (15), perform linear interpolation on the interval of state E_3, and the predicted value of grey fitting accuracy is

$$\widehat{Y}(9) = \otimes_{1i} \times \frac{0.0460}{0.0460 + 0.0733} + \otimes_{2i} \times \frac{0.0733}{0.0460 + 0.0733}$$
$$= 100.4300\%.$$

Through the correction of the grey fitting accuracy index, the more accurate predicted value in 2016 can be obtained as follows:

$$\widetilde{x}^{(0)}(9) = \hat{x}^{(0)}(9) * \widehat{Y}(9) = 6.0682 \times 100.4300\% = 6.0943$$

In modeling process, the introduction of rolling mechanism makes the established model make full use of the latest information to forecast the electricity consumption in 2017. Then using the latest data from 2009 to 2016, the proposed RGPMM(λ,1,1) model can be built to obtain the predicted value in 2017. The results are as follows:

The grey fitting accuracy index in 2017 is in Table 10.

The clustering result and the autocorrelation diagram are shown in Table 11 and Figs. 10 and 11.

After the correction by Markov blanket, the more accurate prediction value in 2017 is 6.5258.

Comparing the forecasting performance of *RGPM*(λ, 1, 1) and *RGPMM*(λ, 1, 1)

The predicted performance of RGPMM(λ,1,1) model is compared with the RGPM(λ,1,1) model. MSE and MAPE are selected to evaluate the predictive performance of the models, and their expressions are respectively as follows:

$$MSE = \tfrac{1}{2} * \sum_{i=9}^{10} (x_i - \hat{x}_i)^2, \, MAPE$$
$$= \tfrac{1}{2} * \sum_{i=9}^{10} \left| \tfrac{x_i - \hat{x}_i}{x_i} \right| \times 100\%$$

MSE and MAPE of each model are calculated. The result is in Table 12.

In Table 12, it is found that the values of MSE and MAPE by RGPMM(λ,1,1) are smaller than RGPM(λ,1,1). The result shows that the RGPMM(λ,1,1) further improves the accuracy.

Forecast the total electricity consumption in the next 6 years

Because that RGPMM(λ,1,1) has been shown to provide accurate prediction, we use it to forecast China's electricity consumption from 2018 to 2023. During the process, the results by Q-cluster and the correlation coefficients of each

Table 14 The correlation coefficients (r) and Q-cluster results during 2011-2018

Label	1	2	3	4	5	6	7	8
Year	2011	2012	2013	2014	2015	2016	2017	2018
r	1.0000	−0.1076	−0.3212	−0.0884	−0.0879	0.0363	0.0686	0.0003
Q-cluster:	Class 1: 2013,2014							
	Class 2: 2011,2015,2016,2017,2018							
	Class 3: 2012							

Table 15 The correlation coefficients (r) and Q-cluster results during 2012–2019

Label	1	2	3	4	5	6	7	8
Year	2012	2013	2014	2015	2016	2017	2018	2019
r	1.0000	−0.4445	0.0280	−0.1651	0.0121	0.1111	−0.0428	0.0012
Q-cluster:	Class 1: 2017,2018							
	Class 2: 2012,2015,2016,2019							
	Class 3: 2014							
	Class 4: 2013							

Table 16 The correlation coefficients (r) and Q-cluster results during 2013–2020

Label	1	2	3	4	5	6	7	8
Year	2013	2014	2015	2016	2017	2018	2019	2020
r	1.0000	−0.2805	−0.4265	0.2937	0.0960	−0.1968	−0.0145	0.0286
Q-cluster:	Class 1: 2013							
	Class 2: 2019,2020							
	Class 3: 2014							
	Class 4: 2015,2017,2018							
	Class 5: 2016							

Table 17 The correlation coefficients (r) and Q-cluster results during 2014–2021

Label	1	2	3	4	5	6	7	8
Year	2014	2015	2016	2017	2018	2019	2020	2021
r	1.0000	−0.4712	−0.0961	0.2323	−0.0760	−0.1330	0.0445	−0.0005
Q-cluster:	Class 1: 2019,2020							
	Class 2: 2014,2017,2018,2021							
	Class 3: 2016							
	Class 4: 2015							

Table 18 The correlation coefficients (r) and Q-cluster results during 2015–2022

Label	1	2	3	4	5	6	7	8
Year	2015	2016	2017	2018	2019	2020	2021	2022
r	1.0000	0.2038	−0.4346	−0.1974	−0.0552	−0.0086	−0.0074	−0.0006
Q-cluster:	Class 1: 2016,2021							
	Class 2: 2015,2022							
	Class 3: 2019,2020							
	Class 4: 2017,2018							
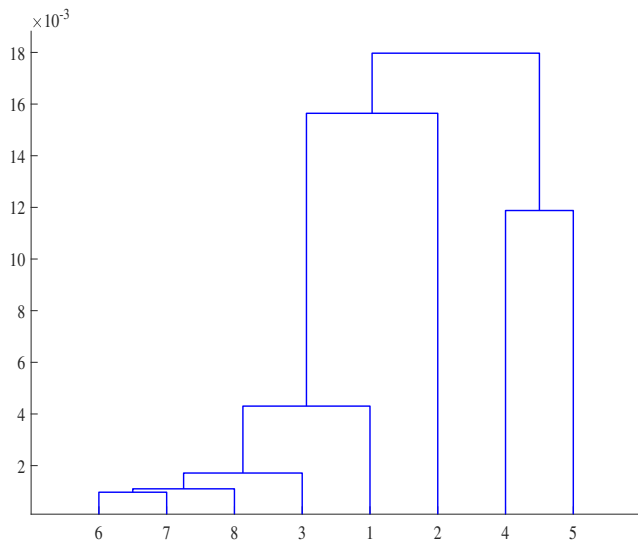
Fig. 12 The figure of Q-cluster during 2010-2017

Fig. 13 Each order of autocorrelation coefficient for 2018

Fig. 14 The figure of Q-cluster during 2011–2018

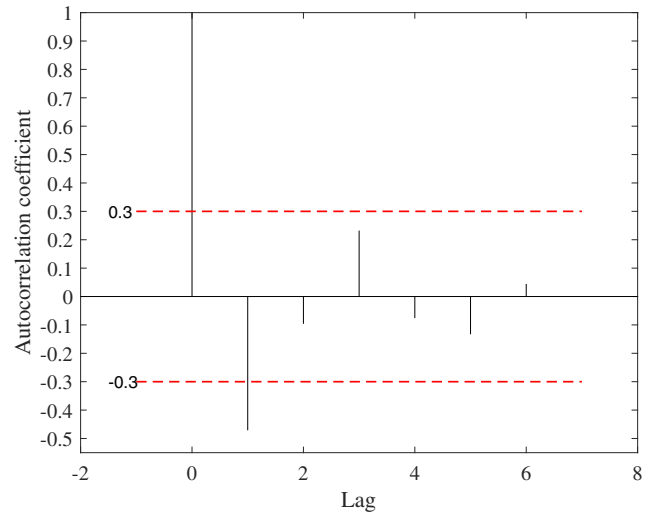**Fig. 15** The each order of autocorrelation coefficient for 2019

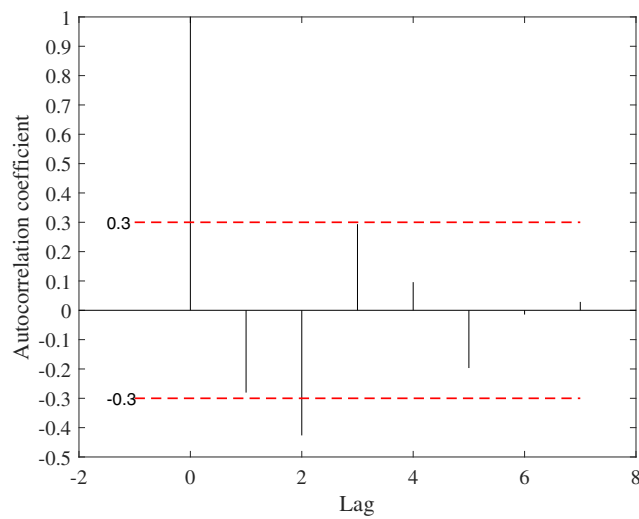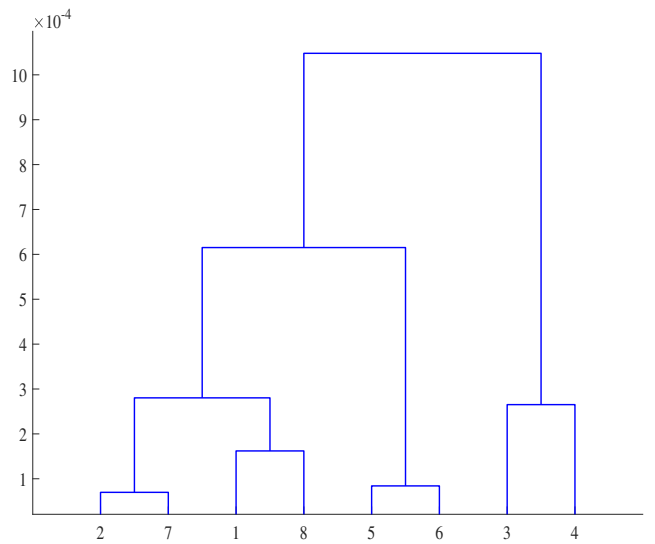**Fig. 16** The figure of Q-cluster during 2012–2019
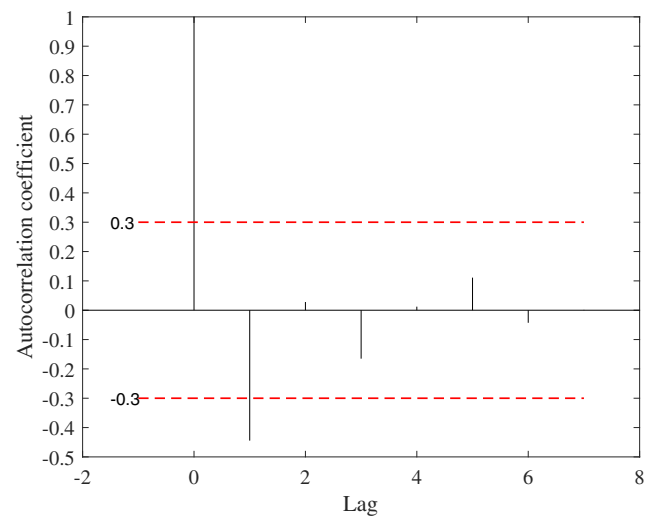
Fig. 17 The each order of autocorrelation coefficient of 2020
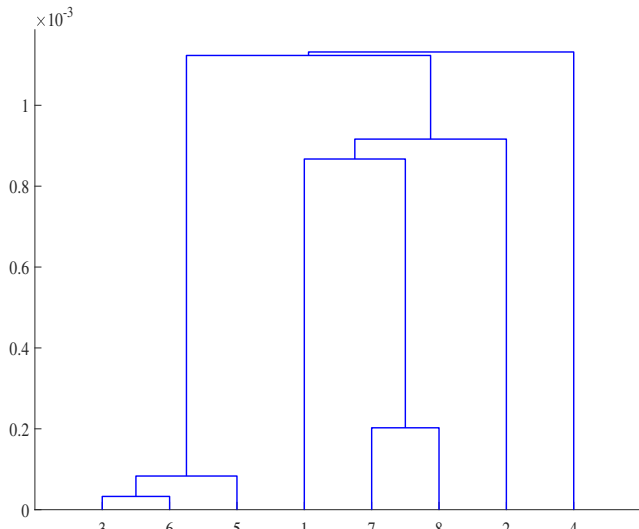
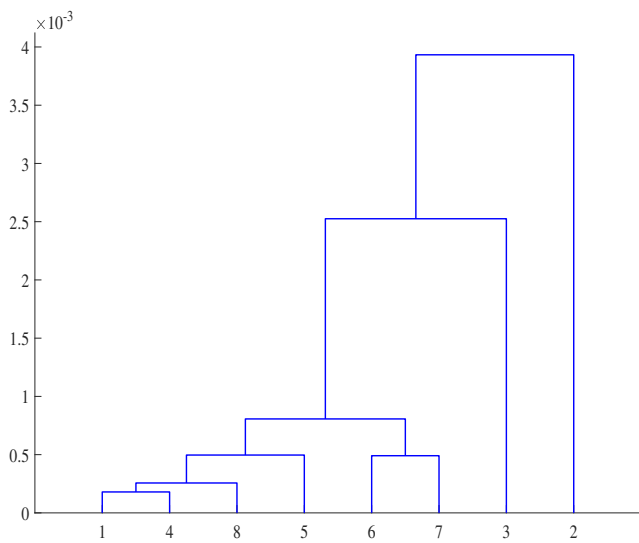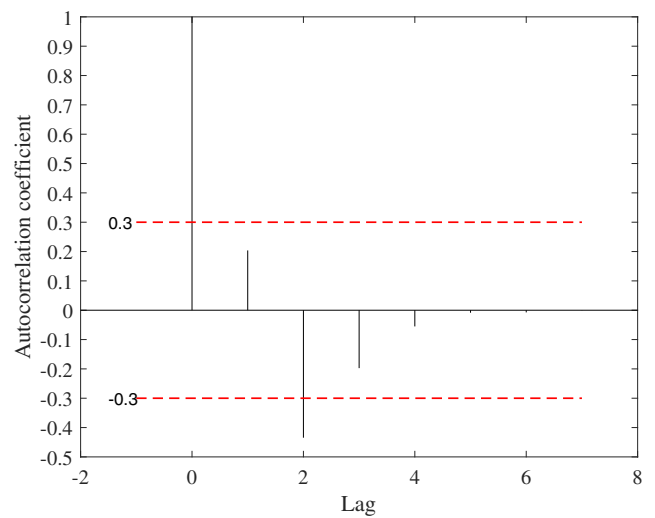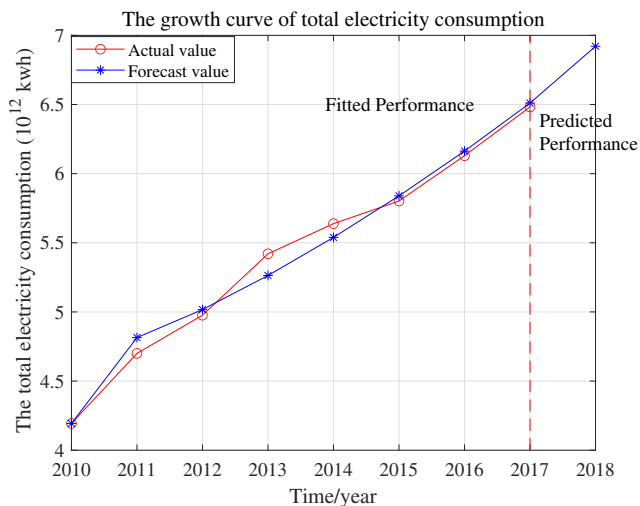**Fig. 18** The figure of Q-cluster during 2013–2020

Fig. 19 The each order of autocorrelation coefficient for 2021

Fig. 20 The figure of Q-cluster during 2014-2021

Fig. 21 The each order of autocorrelation coefficient for 2022

Fig. 22 The figure of Q-cluster during 2015–2022

Fig. 23 The each order of autocorrelation coefficient for 2023

Fig. 24 The predicted value of RGPMM(λ,1,1) for 2018

year are in Tables 13, 14, 15, 16, 17, and 18 and Figs. 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, and 23.

The prediction results are modified by Markov theory during 2018-2023 are shown in Figs. 24, 25, 26, 27, 28, and 29.

The parameters and forecasted results are presented in Table 19.

It can be seen from experimental results in Table 19 that the total electricity consumption will maintain a growth trend. It is estimated that by 2023, electricity consumption will increase from $6.48210 * 10^{12} kwh$ in 2017 to $9.9826 * 10^{12} kwh$, which is nearly $10^{13} kwh$, almost twice that of 2012. This prediction result is of great significance to energy planning and policy-making. In order to ensure
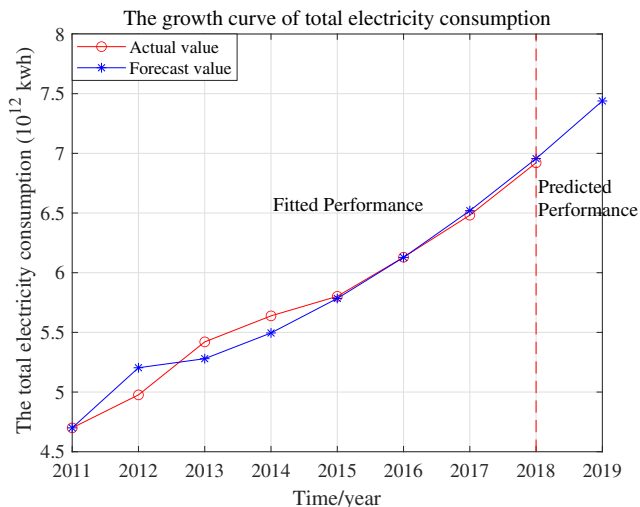
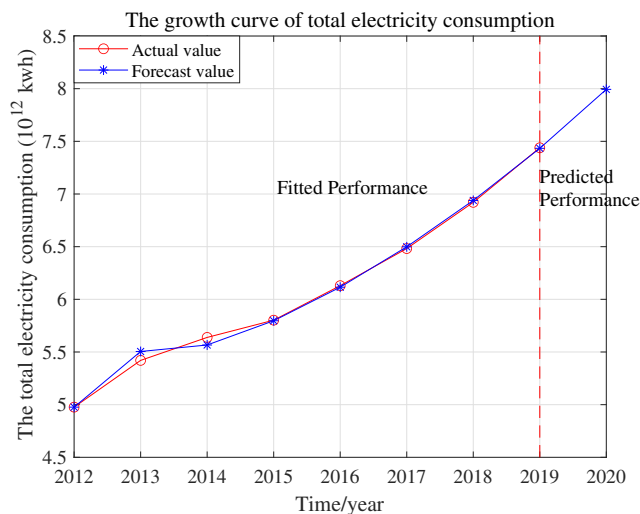**Fig. 26** The predicted value of RGPMM(λ,1,1) for 2020

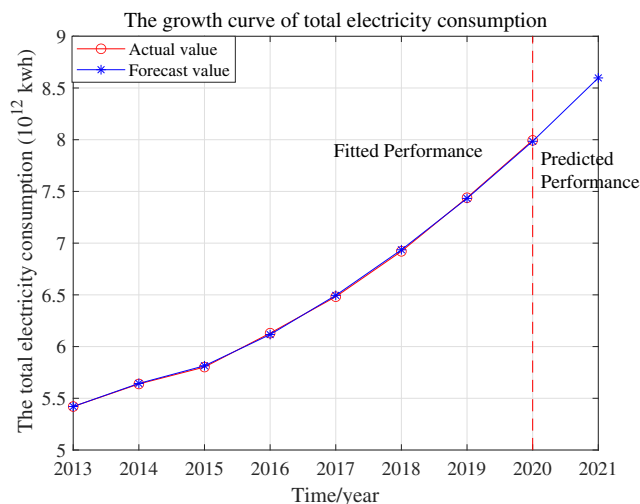**Fig. 27** The predicted value of RGPMM(λ,1,1) for 2021

Fig. 25 The predicted value of RGPMM(λ,1,1) for 2019
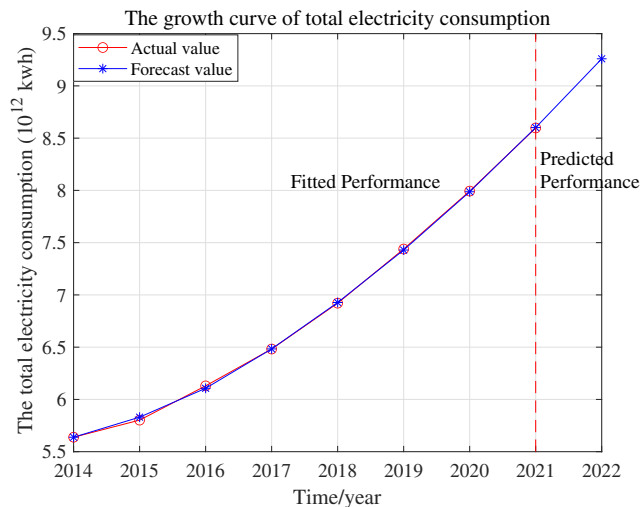

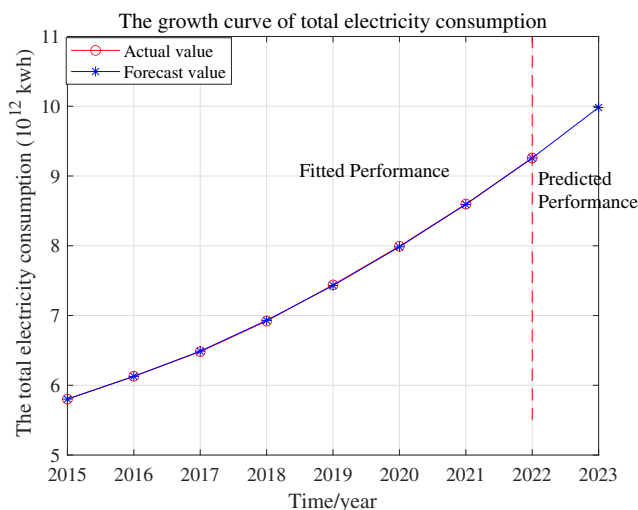**Fig. 28** The predicted value of RGPMM(λ,1,1) for 2022

Fig. 29 The predicted value of RGPMM(λ,1,1) for 2023

secure and stable energy and power supply, it is necessary to accelerate the establishment and improvement of auxiliary services for the power allocation, continuously improve the capacity of power system, and ensure the perfection of power infrastructure in the supply chain. Meanwhile, it is worth noting that huge consumption will bring a heavy burden to energy planning and environmental protection. Therefore, the government needs to take appropriate actions and plans to meet the high energy demand in the future, so as to avoid the phenomenon of resource waste caused by unreasonable resource arrangement.

Conclusion

Electricity consumption forecast is of great significance to the economic development and the guarantee of people's life. However, due to the increase in electricity demand and in the complexity of the power system, it is becoming more and more difficult to accurately predict the power

consumption. Therefore, it is meaningful to design a prediction method which is suitable for the limited data.

Aiming at the problem of power demand under limited data, a grey power-Markov forecasting model based on rolling mechanism is proposed. It predicts the electricity consumption on the basis of grey theory by introducing rolling mechanism and Markov state prediction. According to the research results, the following conclusions can be drawn:

1) In view of the problems faced by power system load forecasting, RGPMM(λ,1,1) is proposed. The example shows that it gets a better prediction effect than the traditional GM(1,1) models, and it improves the prediction accuracy to a certain extent. This model provides a new way to forecast electricity consumption. The structure of the new grey prediction model is simple, the modeling process is easy to operate, and it is easy to be applied in other fields.

2) Total electricity consumption is not only an important indicator of economic development, but also an important indicator of formulating the energy strategy and related environmental protection policies. The forecast results of the total electricity consumption indicate that China's total power consumption will continue to maintain a strong growth trend in the next few years. It has certain reference value for formulating the energy strategy and related environmental protection policies.

However, the grey prediction model proposed in this paper also has some defects. It only solves the univariate prediction problem, and no corresponding solution is proposed for the multivariate prediction problem. Therefore, there is still a lot of work to be done in the future. Future work: the forecasting model proposed in this article is for univariate, and multivariate gray models can be considered in the future. In addition, the new model proposed in this paper can also be used to forecast the industrial electricity consumption, the agricultural electricity consumption, the residential electricity consumption and other fields.

Table 19 The predicted values(10^{12} kwh) the model parameters of RGPMM(λ,1,1) during 2018–2023

Year	2018	2019	2020	2021	2022	2023
a	−0.0587	−0.0774	−0.0827	−0.0824	−0.0816	−0.0792
b	4.7181	6.0109	6.6204	6.3135	6.044	5.9985
λ	−0.0341	−0.132	−0.1525	−0.1152	−0.0769	−0.0464
Predicted value	6.9209	7.4390	7.9934	8.5973	9.2592	9.9826

Data availability The data in this article are from: "'http://www.stats.gov.cn/tjsj/ndsj/2019/indexch.htm".

Declarations

Ethical approval Not applicable

Consent to participate Not applicable

Consent to publish Not applicable

Competing interests The authors declare no competing interests.

References

Azadeh A, Ghaderi SF, Sohrabkhani S (2008) A simulated-based neural network algorithm for forecasting electrical energy consumption in Iran[J]. Energy policy 36(7):2637–2644

Akay D, Atak M (2007) Grey prediction with rolling mechanism for electricity demand forecasting of Turkey[J]. Energy 32(9):1670–1675

Amber KP, Ahmad R, Aslam MW et al (2018) Intelligent techniques for forecasting electricity consumption of buildings[J]. Energy 157:886–893

Akdi Y, Gölveren E, Okkaoğlu Y (2020) Daily electrical energy consumption: periodicity, harmonic regression method and forecasting[J]. Energy 191:116524

Bekiroglu K, Duru O, Gulay E et al (2018) Predictive analytics of crude oil prices by utilizing the intelligent model search engine[J]. Appl Energy 228:2387–2397

Bahrami S, Hooshmand RA, Parastegari M (2014) Short term electric load forecasting by wavelet transform and grey model improved by PSO (particle swarm optimization) algorithm[J]. Energy 72:434–442

Chui F, Elkamel A, Surit R et al (2009) Long-term electricity demand forecasting for power system planning using economic, demographic and climatic variables[J]. Eur J Indu Eng 3(3):277–304

CAO Jian SHI, Shiliang CAO et al (2019) Huajuan Study on prediction and relationship of road transportation accidents of dangerous chemicals and traffic accidents based on GM (1, 1)-Markov model[J]. J Safety Sci Technol 15(1):26–31

China Statistical Yearbook of 2019

Ding S, Hipel KW, Dang Y (2018) Forecasting China's electricity consumption using a new grey prediction model[J]. Energy 149:314–328

Dang Y, Liu SF, Liu B (2004) The GM models that x(n) be taken as initial value[J]. Kybernetes 33(2):247–254

Guo H, Xiao X, Forrest JA (2013) Research on a comprehensive adaptive grey prediction model CAGM (1, N)[J]. Appl Math Comput 225:216–227

Ghani IMM, Ahmad S (2010) Stepwise multiple regression method to forecast fish landing[J]. Procedia-Soc Behav Sci 8:549–554

Geem ZW, Roper WE (2009) Energy demand estimation of South Korea using artificial neural network[J]. Energy policy 37(10):4049–4054

Hussain A, Rahman M, Memon JA (2016) Forecasting electricity consumption in Pakistan: the way forward[J]. Energy Policy 90:73–80

Hamzaçebi C (2007) Forecasting of Turkey's net electricity energy consumption on sectoral bases[J]. Energy policy 35(3):2009–2016

Hernández L, Baladrón C, Aguiar JM et al (2013) Experimental analysis of the input variables' relevance to forecast next day's aggregated electric demand using neural networks[J]. Energies 6(6):2927–2948

Hernandez L, Baladron C, Aguiar JM et al (2014) A survey on electric power demand forecasting: future trends in smart grids, microgrids and smart buildings[J]. IEEE Commun Surv Tutorials 16(3):1460–1495

Yong HE, Yidan BAO (1992) Grey-markov forecasting model and its application[J]. Syst Eng-Theory Practice 12(4):59–63

Julong D (1982) Grey control system[J]. J Huazhong Univ Sci Technol 3(9):18

Julong D (1989) Introduction to grey system theory[J]. J Grey Syst 1(1):1–24

Khosravi A, Nahavandi S, Creighton D et al (2012) Interval type-2 fuzzy logic systems for load forecasting: a comparative study[J]. IEEE Trans Power Syst 27(3):1274–1282

Kavousi-Fard A, Samet H, Marzbani F (2014) A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting[J]. Expert Syst Appl 41(13):6047–6056

Kumar U, models JainVK. (2010) Time series (Grey-Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India[J]. Energy 35(4):1709–1716

Lin B, Liu C (2016) Why is electricity consumption inconsistent with economic growth in China?[J]. Energy Policy 88:310–316

Liu T, Wang B, Yang C (2018) Online Markov Chain-based energy management for a hybrid tracked vehicle with speedy Q-learning[J]. Energy 160:544–555

Lin Z, Zhang Q, Liu H (2012) Parameters optimization of GM (1, 1) model based on artificial fish swarm algorithm[J] Grey Systems, Theory and Application

Mohamed Z, Bodger P (2005) Forecasting electricity consumption in New Zealand using economic and demographic variables[J]. Energy 30(10):1833–1843

Narasimhan M, Jojic N, Bilmes JA (2005) Q-clustering[j]. Advances in Neural Information Processing Systems 18:979–986

Pai TY, Chiou RJ, Wen HH (2008) Evaluating impact level of different factors in environmental impact assessment for incinerator plants using GM (1, N) model[J]. Waste Manag 28(10):1915–1922

Tien TL (2008) The indirect measurement of tensile strength for a higher temperature by the new model IGDMC (1, n)[J]. Measurement 41(6):662–675

Wang J, Zhu S, Zhao W et al (2011) Optimal parameters estimation and input subset for grey model based on chaotic particle swarm optimization algorithm[J]. Expert Syst Appl 38(7):8151–8158

Wang J, Zhou Q, Zhang X (2018) Wind power forecasting based on time series ARMA model[C] \\ IOP Conference Series: Earth and Environmental Science. IOP Publish 199(2):022015

Wang S, Wang P, Zhang Y (1819) A prediction method for urban heat supply based on grey system theory[J]. Sustain Cit Soc 52(10):2020

Wang ZX, Dang Y, Pei LL (2011) Modeling approach for oscillatory sequences based on GM(1, 1) power model[J]. Syst Eng Electron 33(11):2440–2444

Xiong P, Dang Y, Yao T et al (2014) Optimal modeling and forecasting of the energy consumption and production in China[J]. Energy 77:623–634

Xiao X, Yang J, Mao S et al (2017) An improved seasonal rolling grey forecasting model using a cycle truncation accumulated generating operation for traffic flow[J]. Appl Math Model 51:386–404

Zeng B, Ma X, Shi J (2020) Modeling method of the grey GM(1, 1) model with interval grey action quantity and its application[J]. Complexity 2020

Zhang Y, Chen B, Pan G et al (2019) A novel hybrid model based on VMD-WT and PCA-BP-RBF neural network for short-term wind speed forecasting[J]. Energy Convers Manag 195:180–197

Zhao W, Wang J, Lu H (2014) Combining forecasts of electricity consumption in China with time-varying weights updated by a high-order Markov chain model[J]. Omega 45:80–91

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.