



# A hybrid prediction model based on improved multivariable grey model for long-term electricity consumption

Xiaohong Han<sup>1</sup> · Jun Chang<sup>1</sup>

Received: 8 April 2020 / Accepted: 6 October 2020 / Published online: 16 November 2020  
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

## Abstract

The accurate and stable prediction of electricity consumption is essential for intelligent power systems in rapidly developing countries. Grey prediction model is one of choices for prediction under the condition of limited historical data. Nonetheless, it seems rather sceptical using single-variable grey prediction model to predict the dynamics of a complex system. This paper presents a novel multivariable grey prediction model based on first-order linear difference equation for long-term electricity consumption prediction. The proposed model solves the problem of parameter estimation and variable prediction deriving from different approaches through rewriting the whitenization equation of multivariable grey model (MGM(1,  $m$ )). To validate the effectiveness of the proposed hybrid model, the electricity consumption is estimated and predicted over the data from Shanxi province and Beijing city in China from 1999 to 2018. The results show that the hybrid model provides a better estimation and prediction performance compared with other prediction model for predicting electricity consumption.

**Keywords** Difference equation · Electricity consumption forecasting · Multivariable grey model · Grey relational analysis

## List of symbols

### Grey relational analysis

$\chi_0(k)$	Reference data at time step $k$
$\chi_i(k)$	Comparative data of the $i$ th influencing factor at time step $k$
$\tilde{\chi}_0(k)$	Normalized reference data at time step $k$
$\tilde{\chi}_i(k)$	Normalized comparative data of the $i$ th influencing factor at time step $k$
$\zeta_{0i}(k)$	Grey relational coefficient between $\tilde{\chi}_0(k)$ and $\tilde{\chi}_i(k)$
$\rho$	Distinguishing coefficient
$\gamma_{0i}$	Grey relational grade between reference and comparative data sequences

### Multivariate grey model

$X^{(0)}$	Matrix of original data sequence
$X_i^{(0)}$	Original data sequence of the variable $i$

$x_i^{(0)}(j)$	Original data of the variable $i$ at time step $j$
$X^{(1)}$	Matrix of accumulated data sequence
$X_i^{(1)}$	Accumulated data sequence of the variable $i$
$x_i^{(1)}(j)$	Accumulated data of the variable $i$ at time step $j$
$Z^{(1)}(k)$	Vector of the background values at time step $k$
$\hat{x}_i^{(1)}(k)$	Estimated sequence of the variable $x_i^{(1)}$ at time step $k$
$\hat{X}^{(0)}(k)$	Estimated sequence of the variable $x_i^{(0)}$ at time step $k$

## 1 Introduction

### 1.1 Background

With the increase in population and large-scale industrialization, electricity consumption around the world is rising rapidly. Due to the non-storage of power resources and the uncertainty of coal, hydro, wind, and solar power, electricity consumption forecasting is very important for managing power resources successfully and using energy effectively. Excessive electricity supplying will lead to energy investment waste and energy dissipation, while insufficient electricity supplying will hinder economic development and social progress. Hence, accurate and reliable electricity

✉ Xiaohong Han  
hanxiaohong@tyut.edu.cn

<sup>1</sup> College of Data Science, Taiyuan University of Technology, Taiyuan 030024, Shanxi, China

consumption forecasting is critical to sustainable energy management [1]. Moreover, precise electricity consumption forecasting can guide government strategies for future energy usage and development.

## 1.2 Motivation and related work

There are various factors which affect electricity consumption, such as seasons, holidays, electricity policy, electricity price, human behaviours, social condition, political conditions, economy, technological development level, and so on [2], which makes the task of accurate prediction of electricity consumption more challenging. To tackle such tough problem, lots of forecasting techniques have been developed. These methods mainly include regression models [3–6], statistical models [7–10], machine learning methods [11–14], and so on. However, these widely utilized and relatively mature methods have a common limitation: requiring a great deal of training data. Practically, data for electricity consumption along with its influencing factors appear to be limited. Accordingly, the prediction model we use must be effective under this constraint.

It is worth noting that the grey forecasting model developed by Deng [15] is famous for finding the inner law based on incomplete information and few samples. Grey model (GM) with first order and one variable referred to GM (1, 1) has been utilized widely [16–23]. For example, Akay et al. [16] proposed a grey prediction model of single variable with rolling mechanism to predict the Turkey's total and industrial electricity consumption. Lee et al. [17] developed an improved grey forecasting model of one variable, which combines residual modification with genetic programming sign estimation. Xiong et al. [18] proposed a novel GM (1, 1) model based on optimizing initial condition according to the principle of new information priority. The optimized model and five other GM (1, 1) models are applied in the modelling of China's energy consumption and production. Hamzacebi et al. [19] proposed an optimized GM (1, 1) forecasting technique called Optimized GM (1, 1) to predict total electric energy demand of Turkey. The Optimized GM (1, 1) technique is implemented in direct and iterative manners. Zhu et al. [20] proposed to establish a self-adaptive grey fractional weighted model to predict Jiangsu's electricity consumption, which efficiently enhances the prediction quality of electricity consumption. This model introduced the fractional weighted coefficients to design a novel initial condition, which can capture the dynamic characteristics of the electricity consumption observations. Ding et al. [21] designed a novel optimized grey prediction model based on the principle of "new information priority", which combines a new initial condition and rolling mechanism. Wang et al. [22] proposed a novel hybrid forecasting model based on an improved grey forecasting mode optimized by multi-objective ant lion

optimization algorithm to obtain satisfactory forecasting results with high accuracy. Xu et al. [23] proposed a novel grey model with optimal time response function, in which particle swarm algorithm is used to obtain optimal values of unknown variables for shrinking simulation errors and improving adaptability to characteristics of raw data.

However, realistic and complex systems are often consisted of many variables which are not independent of each other and have mutual correlation among them. Therefore, GM (1, 1) with one variable is not suited to reflect more complex dynamics behaviours of a system. To deal with this issue, research on multivariable grey model (MGM(1,  $m$ )) ( $m$  represents the number of variables and  $m \geq 2$ ) has gained extensive attention [24–32]. For example, Wu et al. [24] proposed a novel multivariable grey forecasting model based on fractional-order accumulation. Senapati et al. [25] proposed a multivariable grey model based on convolution integral for solar energy generation prediction. Dang et al. [26] proposed a new electricity demand prediction model which is based on the development trend of multiple driving variables. Zhong et al. [27] proposed a novel multivariable grey theory model for short-term photovoltaic power generation volume forecasting in which particle swarm optimization algorithm is used for background value optimization. Bahrami et al. [28] proposed a new model based on wavelet transform and grey model for electric load forecasting in which the wavelet transform is used to eliminate the high-frequency components of the previous data. Rasheed et al. [29] developed an improved grey forecasting model through combining background value's interpolation optimization with original data sequence's data transformation to predict electricity consumption. Wang et al. [30] proposed to employ rolling mechanism and differential evolution algorithm to improve the prediction accuracy of the original grey model. Ayvaz et al. [31] proposed a nonhomogeneous discrete grey model for predicting yearly net electricity consumption in Turkey. Sun et al. [32] established a multivariable grey model by combining particle swarm optimization with improved grey theory to forecast electricity consumption in Bayannur region.

Although the above research on multivariable grey model has obtained great improvement in the performance of electricity consumption prediction, there are still inherent problems to be solved. One of the problems is that in these models the parameters estimation and variables prediction are solved through different approaches, which degrades the accuracy of grey prediction model for some data [33]. If the parameters estimation and variables prediction are derived from the same approach, the prediction error will be greatly reduced. In this paper, based on the difference equation [34], an improved MGM(1,  $m$ ) is proposed through rewriting the whitenization equation of MGM(1,  $m$ ) as a first-order linear difference equation with constant coefficients to

make parameters estimation and variables prediction derive the same approach; thus, the accuracy will be significantly improved.

### 1.3 Contribution

The main contributions and novelty of this paper can be summarized as follows: (a) an improved multivariable grey model (IMGGM(1,  $m$ )) is developed, which solves the problem of the parameter estimation and variable prediction using different approaches. (b) A hybrid forecasting model is developed based on IMGGM(1,  $m$ ) for electricity consumption.

### 1.4 Article structure

The rest of the paper is organized as follows: Section 2 provides a brief description of existing techniques related to our proposed method, including grey relational analysis, statistical correlation analysis, and MGM(1,  $m$ ). Section 3 presents the proposed improved multivariable grey model. In Sect. 4, two cases of electricity consumption prediction including Shanxi province and Beijing city are adopted to demonstrate the performance of the proposed approach. Some conclusions and discussion of this study are drawn in Sect. 5. Figure 1 displays the schematic overview of the paper.

## 2 Materials and methods

### 2.1 Grey relational analysis

Before performing the estimation of variables, correlation analysis is required to do in order to quantitatively determine the degree of correlation among factors. Grey relational analysis introduced by Professor Deng [15] is one of the most widely used methods in grey system theory, which is appropriate for solving complicated interrelationships between multiple factors and variables. In this paper, we use grey relational analysis to quantitatively analyse the degree of association among factors before performing estimation of variables. The steps of grey relational analysis are as follows:

**Step 1:** Determine the reference sequence and the comparative sequence.

Suppose  $\chi_0(k)$  is the reference sequence;  $\chi_i(k)$  is the comparative sequence,  $i = 1, 2, \dots, m$ ;  $k = 1, 2, \dots, n$ ;  $m$  is the number of factors influencing on the reference variable;  $n$  is the number of samples.

**Step 2:** Normalize reference sequence and comparative sequence.

Since different factors have different physical meanings, the scale of the data is not necessarily the same, which makes later analysis and comparison more difficult. Therefore, data

must be normalized before relational analysis. The normalization is calculated as follows:

$$\tilde{\chi}_i(k) = \frac{\chi_i(k) - \min_{1 \leq i \leq m} \chi_i(k)}{\max_{1 \leq i \leq m} \chi_i(k) - \min_{1 \leq i \leq m} \chi_i(k)} \tag{1}$$

in which  $k = 1, 2, \dots, n$ ;  $i = 0, 1, 2, \dots, m$ .

**Step 3:** Determine the grey relational coefficient of reference sequence and comparative sequence.

The grey relational coefficient  $\zeta_{0i}(k)$  is computed by means of Eqs. (2)–(5):

$$\zeta_{0i}(k) = \frac{\Delta_{\min} + \rho \Delta_{\max}}{\Delta_{0i}(k) + \rho \Delta_{\max}} \tag{2}$$

$$\Delta_{0i}(k) = |\tilde{\chi}_0 - \tilde{\chi}_i|, \quad i = 1, 2, \dots, m \tag{3}$$

$$\Delta_{\max} = \max_{1 \leq i \leq m} \max_{1 \leq k \leq n} \Delta_{0i}(k) \tag{4}$$

$$\Delta_{\min} = \min_{1 \leq i \leq m} \min_{1 \leq k \leq n} \Delta_{0i}(k) \tag{5}$$

in which  $\rho \in [0, 1]$  is the distinguishing coefficient which is normally set to 0.5;  $k = 1, 2, \dots, n$ ;  $i = 0, 1, 2, \dots, m$ ;  $\tilde{\chi}_0$  is the normalized reference sequence;  $\tilde{\chi}_i$  is the normalized comparative sequence.

**Step 4:** Calculate grey relational degree.

Grey relational degree indicates the numerical measure of similarity between reference sequence and comparative sequence, which is done by averaging the grey relational coefficients through Eq. (6):

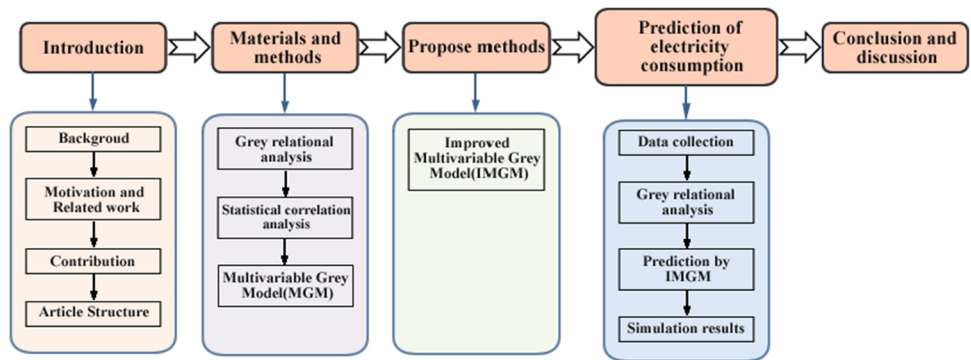
$$\gamma_{0i} = \frac{1}{n} \sum_{k=1}^n \zeta_{0i}(k) \tag{6}$$

in which  $k = 1, 2, \dots, n$ ;  $i = 0, 1, 2, \dots, m$ .

### 2.2 Statistical correlation analysis

To further ascertain the level of association between the reference data and the compatibility data, the statistical correlation analysis is performed, in which correlation coefficient is used to examine the degree of correlation between two variables. The greater the absolute value of correlation coefficients, the stronger the correlation degree is; the closer the correlation coefficient is to 1 or  $-1$ , the stronger the correlation degree is; the closer the correlation coefficient is to 0, the weaker the correlation degree is. Generally, the correlation coefficient was within the range of [0.8, 1.0], and the correlation degree is considered as very strong. In this research, we use Pearson correlation coefficient (PCC) to perform statistical correlation analysis. The two variables can be defined

**Fig. 1** The schematic overview of the paper



as  $X = [x_1, x_2, \dots, x_n]$  and  $Y = [y_1, y_2, \dots, y_n]$ , and their PCC is calculated as follows:

$$\rho(x, y) = \frac{X \cdot Y - (\sum_{i=1}^n x_i \sum_{i=1}^n y_i)/n}{\sqrt{[\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2/n] * [\sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2/n]}} \tag{7}$$

where  $\rho(x, y)$  is the PCC of variables  $x$  and  $y$ , and  $-1 \leq \rho(x, y) \leq 1$ .  $\rho(x, y) = -1$  or  $1$  indicates completely negative and positive correlation, respectively, and  $\rho(x, y) = 0$  indicates no correlation at all. Besides, the larger the value of  $\rho(x, y)$  is, the stronger the correlation degree between  $x$  and  $y$  is.

### 2.3 Multivariable grey model

Grey prediction theory is often used to handle with a system which has sparse data and incomplete information. The modelling procedures of the multivariable grey model are presented as follows:

**Step 1:** Obtain the original and transformed sequences for modelling.

Assume that  $X^{(0)} = (X_1^{(0)}, X_2^{(0)}, \dots, X_m^{(0)})$  is an original nonnegative sequence, where  $X_i^{(0)} = (x_i^{(0)}(1), x_i^{(0)}(2), \dots, x_i^{(0)}(n))^T$  is the observation sequence of the  $i$ th variable at times  $1, 2, \dots, n$ .  $i = 1, 2, \dots, m$ ;  $m$  is the number of variables.

Then, the original sequence is processed using the first-order accumulated generation operator (AGO). The matrix  $X^{(1)} = (X_1^{(1)}, X_2^{(1)}, \dots, X_m^{(1)})$  is the first-order accumulating generation matrix of  $X^{(0)}$ .  $X_i^{(1)} = (x_i^{(1)}(1), x_i^{(1)}(2), \dots, x_i^{(1)}(n))^T$  is the first-order accumulating generation sequence of  $X_i^{(0)}$ , where

$$x_i^{(1)}(j) = \sum_{k=1}^j x_i^{(0)}(k), \quad j = 1, 2, \dots, n, \quad i = 1, 2, \dots, m \tag{8}$$

**Step 2:** Establish the multivariable grey model.

The multivariable grey model (MGM(1,  $m$ )), which is sometimes called the whitening equations of MGM(1,  $m$ ), is established as:

$$\begin{cases} \frac{dx_1^{(1)}(t)}{dt} = a_{11}x_1^{(1)}(t) + a_{12}x_2^{(1)}(t) + \dots + a_{1m}x_m^{(1)}(t) + b_1 \\ \frac{dx_2^{(1)}(t)}{dt} = a_{21}x_1^{(1)}(t) + a_{22}x_2^{(1)}(t) + \dots + a_{2m}x_m^{(1)}(t) + b_2 \\ \dots \\ \frac{dx_m^{(1)}(t)}{dt} = a_{m1}x_1^{(1)}(t) + a_{m2}x_2^{(1)}(t) + \dots + a_{mm}x_m^{(1)}(t) + b_m \end{cases} \tag{9}$$

Let

$$X^{(1)}(t) = \{x_1^{(1)}(t), x_2^{(1)}(t), \dots, x_m^{(1)}(t)\}^T \tag{10}$$

$$A = \begin{pmatrix} a_{11} & a_{11} & \dots & a_{11} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \vdots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_{21} \\ \vdots \\ b_m \end{pmatrix} \tag{11}$$

$A$  is a developing grey matrix, and  $B$  is an endogenous control grey matrix.

Then, the dynamical model (8) can be rewritten in matrix form as Eq. (12):

$$\frac{dX^{(1)}(t)}{dt} = AX^{(1)}(t) + B \tag{12}$$

**Step 3:** Estimate the model parameters.

To identify parameter matrix  $A$  and vector  $B$ , Eq. (8) is discretized as

$$X^{(0)}(k) = AZ^{(1)}(k) + B \tag{13}$$

in which  $X^{(0)}(k) = \{x_1^{(0)}(k), x_2^{(0)}(k), \dots, x_m^{(0)}(k)\}^T$ , and  $Z^{(1)}(k) = \{z_1^{(1)}(k), z_2^{(1)}(k), \dots, z_m^{(1)}(k)\}^T$  is the mean generation matrix of  $X^{(1)}$ , and the sequence  $Z_i^{(1)} =$

$(z_i^{(1)}(2), z_i^{(1)}(3), \dots, z_i^{(1)}(n))^T$  is the generated mean sequence of  $X_i^{(1)}$ , in which

$$z_i^{(1)}(k) = \frac{1}{2} (x_i^{(1)}(k) + x_i^{(1)}(k - 1)), \quad i = 2, 3, \dots, m, \quad k = 2, 3, \dots, n \tag{14}$$

By the least square method, the parameter matrix  $A$  and  $B$  can be obtained as

$$(A, B)^T = (L^T L)^{-1} L^T M \tag{15}$$

where

$$L = \begin{bmatrix} z_1^{(1)}(2) & z_2^{(1)}(2) & \dots & z_m^{(1)}(2) & 1 \\ z_1^{(1)}(3) & z_2^{(1)}(3) & \dots & z_m^{(1)}(3) & 1 \\ \dots & \dots & \dots & \dots & \dots \\ z_1^{(1)}(n) & z_2^{(1)}(n) & \dots & z_m^{(1)}(n) & 1 \end{bmatrix} \quad M = \begin{bmatrix} X^{(0)}(2) \\ X^{(0)}(3) \\ \vdots \\ X^{(0)}(n) \end{bmatrix}$$

**Step 4:** Obtain the time response function for predicting in the transformed domain.

Using the estimated parameters and  $x_i^{(1)}(1) = x_i^{(0)}(0), i = 1, 2, \dots, m$ , as the initial condition, the time response function can be achieved:

$$\hat{x}_i^{(1)}(k) = e^{-\hat{A} \cdot (k-1)} (x_i^{(1)}(1) - \hat{A}^{-1} \hat{B}) + \hat{A}^{-1} \hat{B}, \tag{16}$$

where  $\hat{X}^{(1)}(k)$  represents the prediction vector of  $X^{(1)}(k)$ .

**Step 5:** Obtain the fitted and forecasted values in the original domain.

By using the first-level inverse accumulating generation operator (I-AGO), the fitted and predicted values in the original domain can be calculated by Eq. (17):

$$\hat{x}_i^{(0)}(k) = \hat{x}_i^{(1)}(k) - \hat{x}_i^{(1)}(k - 1) \tag{17}$$

in which  $\hat{x}_i^{(0)}(k)$  represents the prediction vector of  $x^{(0)}(k)$ .

### 3 Proposed methods

#### 3.1 Improved multivariable grey model

Although the grey prediction model has enjoyed high popularity in many predicting applications, there are still inherent limitations to be improved. One of the drawbacks is that the estimation of parameters and the prediction of variables are deriving from different approaches, which degrades the accuracy of grey prediction model for some data. If the parameter estimation and variable prediction use the same approach, the prediction error will be greatly reduced.

In the original multivariable grey model, the whitenization equation of MGM(1,  $m$ ) (Eq. (9)) is used to predict variable values, while the estimation of parameters uses the discretized equation (Eq. (13)). The estimated parameters by Eq. (13) can only be regarded as approximate values of the parameters in Eq. (9); thus, the accuracy of prediction would be reduced due to adopting different methods to estimate parameters and predict variables.

In fact, Eq. (13) can be written as follows:

$$(X^{(1)}(k) - X^{(1)}(k - 1)) = A'(X^{(1)}(k) + X^{(1)}(k - 1)) + B \tag{18}$$

Namely,

$$X^{(1)}(k) = \Gamma X^{(1)}(k - 1) + \psi \tag{19}$$

in which  $\Gamma = \frac{1-A'}{1+A'}$ ,  $\psi = \frac{B}{1+A'}$ ,  $A' = \frac{A}{2}$ ; Eq. (19) is a first-order linear difference equation with constant coefficients. If both parameter estimation and variable prediction are calculated using the same difference equation, the accuracy will be significantly improved.

**Definition 1**  $X^{(1)}(k + 1) = AX^{(1)}(k) + C(k)$  is the first-order discrete grey model of constant coefficient nonhomogeneous difference equation.

**Theorem 1** If  $X^{(1)}(t)$  meets  $\hat{X}^{(1)}(k + 1) = A\hat{X}^{(1)}(k) + C$ , then

$$\hat{X}^{(1)}(k) = A\hat{X}^{(1)}(0)A^k + \Omega \tag{20}$$

Here,  $\Omega = E - EA^k$ ,  $E = \frac{C}{1-A}$ ,  $(A, C)^T = (Q^T Q)^{-1} Q^T P$

$$Q = \begin{bmatrix} x_1^{(1)}(1) & x_2^{(1)}(1) & \dots & x_m^{(1)}(1) & 1 \\ x_1^{(1)}(2) & x_2^{(1)}(2) & \dots & x_m^{(1)}(2) & 1 \\ \dots & \dots & \dots & \dots & \dots \\ x_1^{(1)}(n) & x_2^{(1)}(n) & \dots & x_m^{(1)}(n) & 1 \end{bmatrix} \tag{21}$$

$$P = \begin{bmatrix} x_1^{(1)}(1) & x_2^{(1)}(1) & \dots & x_m^{(1)}(1) \\ x_1^{(1)}(2) & x_2^{(1)}(2) & \dots & x_m^{(1)}(2) \\ \dots & \dots & \dots & \dots \\ x_1^{(0)}(n) & x_2^{(1)}(n) & \dots & x_m^{(1)}(n) \end{bmatrix} \tag{22}$$

The fitted and forecasted values can be obtained by Eq. (23) in the original domain:

$$\hat{X}^{(0)}(k) = \hat{X}^{(1)}(k) - \hat{X}^{(1)}(k - 1) \tag{23}$$

**Theorem 2** If  $X^{(1)}(t)$  meets  $\hat{X}^{(1)}(k + 1) = A\hat{X}^{(1)}(k) + C(k)$ , then

$$\hat{X}^{(1)}(k) = A^k \hat{X}^{(1)}(0) + \Omega' \tag{24}$$



in which  $\Omega' = \sum_{r=1}^k A^{r-1} C(k-r)$ .

In particular, (1) when  $C_k = B_0 + B_1 k$ , then

$$\hat{X}^{(1)}(k) = A^k \hat{X}^{(1)}(0) + \Omega' \tag{25}$$

in which  $\Omega'' = \sum_{r=1}^k A^{r-1} [B_0 + B_1(k-r)]$ ,  
 $(A, B_0, B_1)^T = (Q^T Q)^{-1} Q^T P$

$$Q = \begin{bmatrix} x_1^{(1)}(1) & x_2^{(1)}(1) & \dots & x_m^{(1)}(1) & 1 & 1 \\ x_1^{(1)}(2) & x_2^{(1)}(2) & \dots & x_m^{(1)}(2) & 1 & 2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_1^{(1)}(n) & x_2^{(1)}(n) & \dots & x_m^{(1)}(n) & 1 & n \end{bmatrix} \tag{26}$$

$$P = \begin{bmatrix} x_1^{(1)}(1) & x_2^{(1)}(1) & \dots & x_m^{(1)}(1) \\ x_1^{(1)}(2) & x_2^{(1)}(2) & \dots & x_m^{(1)}(2) \\ \dots & \dots & \dots & \dots \\ x_1^{(0)}(n) & x_2^{(1)}(n) & \dots & x_m^{(1)}(n) \end{bmatrix} \tag{27}$$

(2) When  $C_k = B_0 + B_1 k + B_2 k^2$ , then

$$\hat{X}^{(1)}(k) = A^k \hat{X}^{(1)}(0) + \Omega''' \tag{28}$$

Here,  $\Omega''' = \sum_{r=1}^k A^{r-1} [B_0 + B_1(k-r) + B_2(k-r)^2]$ ,

$(A, B_0, B_1, B_2)^T = (Q^T Q)^{-1} Q^T P$

$$Q = \begin{bmatrix} x_1^{(1)}(1) & x_2^{(1)}(1) & \dots & x_m^{(1)}(1) & 1 & 1 & 1 \\ x_1^{(1)}(2) & x_2^{(1)}(2) & \dots & x_m^{(1)}(2) & 1 & 2 & 2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_1^{(1)}(n) & x_2^{(1)}(n) & \dots & x_m^{(1)}(n) & 1 & n & n^2 \end{bmatrix} \tag{29}$$

$$P = \begin{bmatrix} x_1^{(1)}(1) & x_2^{(1)}(1) & \dots & x_m^{(1)}(1) \\ x_1^{(1)}(2) & x_2^{(1)}(2) & \dots & x_m^{(1)}(2) \\ \dots & \dots & \dots & \dots \\ x_1^{(0)}(n) & x_2^{(1)}(n) & \dots & x_m^{(1)}(n) \end{bmatrix} \tag{30}$$

The fitted and forecasted values can be obtained by Eq. (31) in the original domain:

$$\hat{X}^{(0)}(k) = \hat{X}^{(1)}(k) - \hat{X}^{(1)}(k-1) \tag{31}$$

### 3.2 Hybrid prediction model

In the part, we integrate correlation analysis test and improved multivariable grey model (IMGGM) to construct a hybrid prediction model of electricity consumption. First, correlation analysis test is done using grey relational analysis and statistical correlation analysis before performing the estimation of variables of interest. Then, the values got from correlation analysis test are processed using the first-order accumulated generation operator (AGO). Finally, we input the transformed sequence by AGO into the IMGGM to get the fitted and forecasted values. The detailed steps of this hybrid model are shown in Fig. 2.

## 4 Experimental analysis

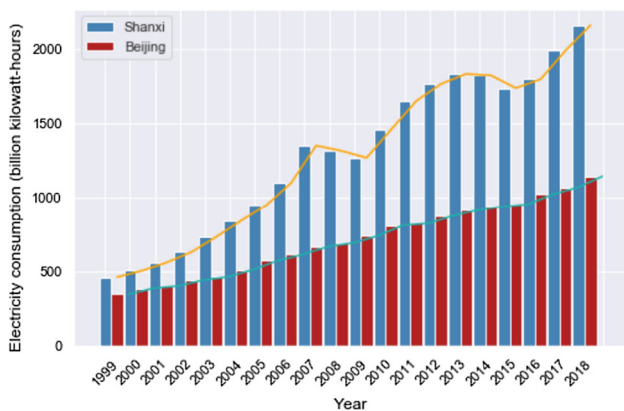
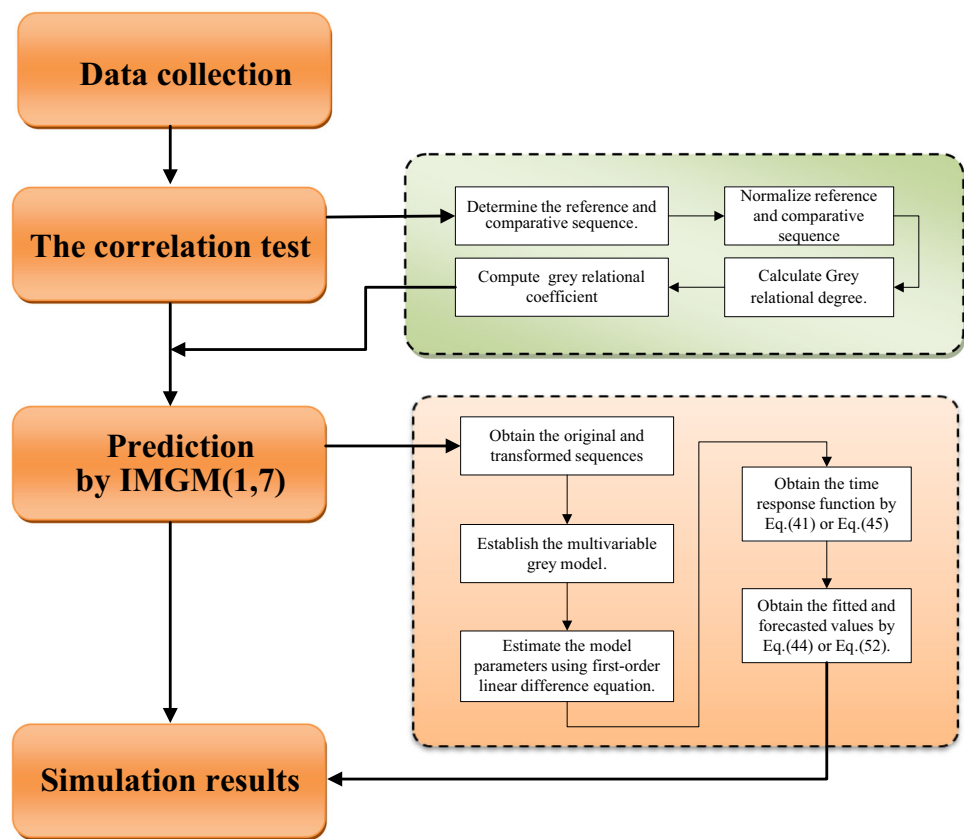
In this part, we predict the electricity consumption using our hybrid IMGGM approach to test the effectiveness and performance of our approach. All the experimental data sets are performed in the MATLAB R2016b environment on the Microsoft Windows 7 Pro operating system. The hardware environment is on a computer with Intel(R) Core(TM) i5-4440 3.10 GHz CPU and 8-GB RAM.

The electricity consumption data are driving from Shanxi province and Beijing city of China from 1999 to 2018 shown in Fig. 3, which is available from National Bureau of Statistics of China (<http://data.stats.gov.cn/>). It is noteworthy that the electricity consumption can be affected by many influencing factors such as gross domestic product (GDP), third industry proportion, urbanization rate, mean annual temperature, consumption level of urban residents, second industry proportion, third industry proportion, total volume of retail sales, permanent residents, fixed assets investment, urbanization rate, and so on [35]. In this research, we consider six influencing factors including GDP, consumption level of urban residents, second industry proportion, total volume of retail sales, permanent residents, fixed assets investment based on the grey relational grade and the statistical correlations. The grey relational grade is to measure the degree of correlation between the two factors according to the similarity or difference of the two factors. The values of grey relational grade within the range of [0.5, 0.9] are considered moderate link among factors [39]. The statistical correlation analyses the degree of connection between the two factors. The degrees of the statistical correlations above 0.9 are marked a strong inter-correlation among the factors of interest [40]. Here, we define the variable  $\chi_0$  as the electricity consumption, and  $\chi_1, \chi_2, \chi_3, \chi_4, \chi_5$ , and  $\chi_6$  denote consumption level of urban residents, second industry proportion, total volume of retail sales, GDP, permanent residents, and fixed assets investment, respectively, as shown in Table 1. The data of electricity consumption and its influencing factors are shown in Tables 2 and 3.

The data set is divided into two groups: the data from 1999 to 2013 are utilized as the training data set and 2014–2018 as the testing data set. To verify the presented hybrid forecasting model, we compare our proposed improved multivariable grey model (IMGGM) with the original multivariable grey model (MGM), generalized regression neural network model with fly optimization algorithm (FOAGRNN) [36], grey prediction model with convolution integral (GMC) [24], and self-adaptive fractional weighted grey model (SFOGM) [20].

The parameters of MGM model are calculated using the least square method. The parameters of IMGGM model are estimated by Eqs. (20)–(31). FOAGRNN model uses fly optimization algorithm (FOA) for the parameter selection of the generalized regression neural network (GRNN) model. The

**Fig. 2** The proposed hybrid prediction model



**Fig. 3** The electricity consumption in Shanxi province and Beijing city from 1999 to 2019

main parameters of FOA are the maximum iteration number (*maxgen*), the population size (*sizepop*), the initial fruit fly swarm location (*X\_axis*, *Y\_axis*), and the random flight distance range (*FR*). In this case, suppose *maxgen* = 100, *sizepop* = 10, (*X\_axis*, *Y\_axis*) ∈ [0, 1], *FR* ∈ [− 10, 10]. SFOGM model utilizes particle swarm optimization (PSO) to determine the parameters of the self-adaptive fractional parameter and the input time parameter through minimizing the mean absolute percentage error between the simulation value and the actual value. The main parameters of PSO are

**Table 1** Nomenclature of the reference variable and its influence factors

Variable names	Description
$X_0$	Electricity consumption
$X_1$	Consumption level of urban residents
$X_2$	Second industry proportion
$X_3$	Total volume of retail sales
$X_4$	GDP
$X_5$	Permanent residents
$X_6$	Fixed assets investment

the acceleration factors ( $c_1$  and  $c_2$ ), population size (*popsiz*), iteration times (*itetimes*), the dimensions (*Dim*), and inertia weight (*w*). In this experiment, set  $c_1 = c_2 = 2$ , *popsiz* = 30, *itetimes* = 1000, *Dim* = 2.

In multivariable grey model, the parameter estimation and variable prediction are crucial; otherwise, this can lead to performance degradation. The original multivariable grey model uses the whitenization equation of MGM(1, *m*) (Eq. 9) and discretized equation (Eq. 13)) to estimate variable values and parameter values. The parameters estimated through Eq. (13) are approximate values of the parameters; thus, the accuracy of prediction will be reduced. In this study, a first-order linear difference equation with constant coefficients is introduced to compute both parameter estimation and variable predic-

**Table 2** Data of all factors from 1999 to 2018 for Shanxi province

Year	$X_0$	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
1999	459.34	4200	785.47	587.1	1667.10	3247	477.57
2000	501.99	4311	858.37	629.1	1845.72	3272	548.16
2001	557.58	3552	956.01	679.9	2029.53	3294	663.58
2002	628.82	4764	1134.31	755.4	2324.80	3314	813.36
2003	731.77	5172	1463.38	729.3	2855.23	3335	1100.86
2004	841.55	6441	1919.40	884.8	3571.37	3355	1443.88
2005	946.33	7104	2357.04	1410.7	4230.53	3375	1826.58
2006	1097.68	8143	2755.66	1635.4	4878.61	3393	2255.74
2007	1348.81	9252	3454.49	1953.3	6024.45	3411	2861.46
2008	1314.33	10,262	4242.36	2421.1	7315.40	3427	3531.16
2009	1267.54	10,617	3993.80	2809.0	7358.31	3574	4943.16
2010	1460.00	12,680	5234.00	3318.2	9200.86	3593	6040.54
2011	1650.41	14,055	6635.26	3903.4	11,237.55	3611	6063.17
2012	1765.80	15,091	6731.56	4506.8	12,112.83	3630	7073.06
2013	1832.30	16,341	6613.06	5139.3	12,665.25	3648	8863.30
2014	1822.63	17,189	6293.91	5717.9	12,761.49	3664	11,031.89
2015	1737.21	19,018	5194.27	6033.7	12,766.49	3682	12,354.53
2016	1797.18	19,724	5028.99	6480.5	13,050.41	3702	14,074.15
2017	1990.61	23,345	6778.89	6918.1	15,528.42	3718	14,197.98
2018	2160.53	24,000	7089.19	7338.5	16,818.11	3730	14,305.25

**Table 3** Data of all factors from 1999 to 2018 for Beijing city

Year	$X_0$	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
1999	344.13	10,007	907.34	1313.3	2678.82	1350	1652.2
2000	384.43	10,145	1033.29	1443.3	3161.66	1364	1678.6
2001	399.94	10,301	1142.35	1593.5	3707.96	1385	1689.3
2002	439.96	12,505	1249.99	1744.8	4315	1423	1693.6
2003	461.24	13,826	1487.15	1916.7	5007.21	1456	1999.9
2004	510.11	15,438	1853.58	2191.8	6033.21	1493	2333
2005	570.54	16,478	2026.51	2911.7	6969.52	1538	2595.41
2006	611.57	18,185	2191.43	3295.3	8117.78	1601	3012.45
2007	667.01	20,320	2509.4	3835.2	9846.81	1676	3597.29
2008	689.72	21,872	2626.41	4645.5	11,115	1771	3520.95
2009	739.15	23,812	2855.55	5309.9	12,153.03	1860	4149.63
2010	809.9	26,949	3388.38	6340.3	14,113.58	1962	4916.53
2011	821.71	30,037	3752.48	7222.2	16,251.93	2019	5519.84
2012	874.3	32,857	4059.27	8123.5	17,879.4	2069	6064.86
2013	913.1	35,836	4292.56	8872.1	19,800.81	2115	6797.54
2014	937.05	38,515	4544.8	9638	21,330.83	2152	6873.44
2015	952.72	41,846	4542.64	10,338	23,014.59	2171	7446.02
2016	1020.27	52,721	4944.44	11,005.1	25,669.13	2173	7888.69
2017	1066.89	57,100	5326.76	11,575.4	28,014.94	2171	8307.33
2018	1142.38	59,325	5477.35	11,747.7	33,105.97	2154	8933.56

tion, as the parameter estimation and variable prediction use the same approach, and the accuracy will be significantly improved.

To verify the effectiveness of our method, we use two evaluation criteria including the mean absolute error (MAE) [37] and the mean absolute percent error (MAPE) [38]. MAE is used to reflect the overall level of errors; MAPE is regarded as



**Table 4** Results of grey relational grade for Shanxi province

Variable	$\chi_0$	$\chi_1$	$\chi_2$	$\chi_3$	$\chi_4$	$\chi_5$	$\chi_6$
$\chi_0$	1.0000	0.7089	0.8101	0.6819	0.7103	0.7575	0.6755
$\chi_1$	0.6787	1.0000	0.7286	0.8341	0.8171	0.6457	0.8160
$\chi_2$	0.7484	0.6900	1.0000	0.6609	0.6806	0.7301	0.6674
$\chi_3$	0.6400	0.8271	0.6928	1.0000	0.7526	0.6161	0.8979
$\chi_4$	0.7365	0.8558	0.7685	0.8092	1.0000	0.6757	0.7832
$\chi_5$	0.7560	0.6757	0.7894	0.6583	0.6447	1.0000	0.6867
$\chi_6$	0.5972	0.7815	0.6674	0.8821	0.6934	0.6125	1.0000

a measure of the prediction accuracy of a forecasting method in statistics. The evaluation criteria are given as follows:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \tag{32}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \tag{33}$$

in which  $N$  is the number of forecasting data points, which is 20 in our simulation,  $y_i$  and  $\hat{y}_i$  represent the actual and predicted values of electricity consumption, respectively, at time  $i$ .

**4.1 The correlation test**

Before conducting the prediction of electricity consumption, the correlation test is required to be done to determine the variables for modelling. Tables 4 and 6 display the calculation results of the grey relational grade of two cases when the distinguishing coefficient  $\rho$  equals 0.5. The overall values change between 0.5 and 0.9 with the minimum and maximum values being 0.5572 and 0.8979, respectively. This indicates moderate link among factors [39]. To further ascertain the level of association between the reference data and the compatibility data, the statistical correlations [40] are computed, as shown in Tables 5 and 7. It is apparent that most correlation degrees are above 0.9 which marks a strong inter-correlation among the factors of interest. Both measures agree well with each other which can therefore suggest that the factors considered in our model are reasonably correlated. Therefore, we select six influencing factors including GDP, consumption level of urban residents, second industry proportion, total volume of retail sales, permanent residents, and fixed assets investment in our prediction model.

**4.2 Simulation results**

According to the proposed IMGGM, we first use the first-order accumulated generation operator to process the original

**Table 5** Results of statistical correlation for Shanxi province

Variable	$\chi_0$	$\chi_1$	$\chi_2$	$\chi_3$	$\chi_4$	$\chi_5$	$\chi_6$
$\chi_0$	1.0000	0.9610	0.9738	0.9501	0.9832	0.9670	0.9134
$\chi_1$	0.9610	1.0000	0.9126	0.9924	0.9877	0.9612	0.9822
$\chi_2$	0.9738	0.9126	1.0000	0.9031	0.9618	0.9478	0.8464
$\chi_3$	0.9501	0.9924	0.9031	1.0000	0.9842	0.9644	0.9912
$\chi_4$	0.9832	0.9877	0.9618	0.9842	1.0000	0.9776	0.9565
$\chi_5$	0.9670	0.9612	0.9478	0.9644	0.9776	1.0000	0.9430
$\chi_6$	0.9134	0.9822	0.8464	0.9912	0.9565	0.9430	1.0000

**Table 6** Results of grey relational grade for Beijing city

Variable	$\chi_0$	$\chi_1$	$\chi_2$	$\chi_3$	$\chi_4$	$\chi_5$	$\chi_6$
$\chi_0$	1.0000	0.6650	0.7837	0.6734	0.6626	0.7240	0.7037
$\chi_1$	0.5572	1.0000	0.5923	0.7050	0.8773	0.6090	0.7743
$\chi_2$	0.7927	0.7018	1.0000	0.7599	0.6855	0.7486	0.7216
$\chi_3$	0.6567	0.7732	0.7358	1.0000	0.7729	0.6840	0.8317
$\chi_4$	0.5706	0.8860	0.5877	0.7217	1.0000	0.6101	0.7703
$\chi_5$	0.6291	0.6090	0.6536	0.5994	0.5931	1.0000	0.5640
$\chi_6$	0.6670	0.8173	0.6720	0.8159	0.8046	0.6308	1.0000

**Table 7** Results of statistical correlation for Beijing city

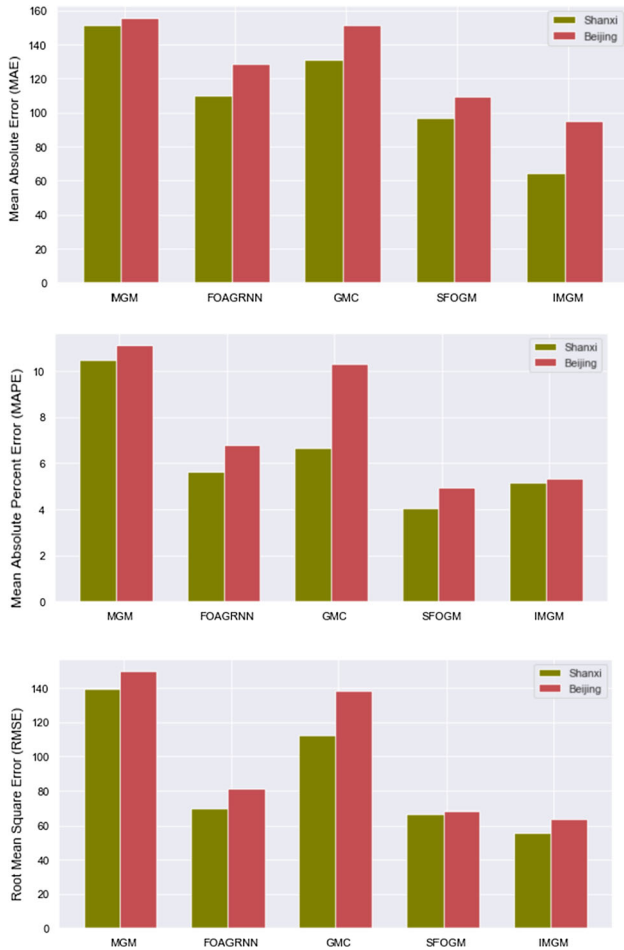
Variable	$\chi_0$	$\chi_1$	$\chi_2$	$\chi_3$	$\chi_4$	$\chi_5$	$\chi_6$
$\chi_0$	1.0000	0.9615	0.9952	0.9829	0.9801	0.9775	0.9851
$\chi_1$	0.9615	1.0000	0.9698	0.9762	0.9916	0.9112	0.9792
$\chi_2$	0.9952	0.9698	1.0000	0.9930	0.9848	0.9793	0.9939
$\chi_3$	0.9829	0.9762	0.9930	1.0000	0.9874	0.9731	0.9975
$\chi_4$	0.9801	0.9916	0.9848	0.9874	1.0000	0.9400	0.9918
$\chi_5$	0.9775	0.9112	0.9793	0.9731	0.9400	1.0000	0.9676
$\chi_6$	0.9851	0.9792	0.9939	0.9975	0.9918	0.9676	1.0000

sequence. And then based on the multivariable grey model, the first-order linear difference equation with constant coefficients (Eq. 19) is used to calculate the parameters of developing grey matrix and endogenous control grey matrix. Finally, using the estimated parameters and the first-order accumulating generation sequence, the fitted and forecasted values can be obtained by Eq. (31) in the original domain.

Table 8 lists results of MAPE and MAE for IMGGM, SFOGM, GMC, FOAGRNN, and MGM models in the example of Shanxi province and Beijing city from 2014 to 2018, where the values in bold represent the smallest values of MAE and MAPE. According to the results of the above-mentioned forecasting models listed in Table 8 and Fig. 4, it can be observed that the proposed hybrid forecasting model obtains the highest prediction accuracy (via the MAE and MAPE criteria). We can see that the developed model significantly outperforms than the four compared models in most

**Table 8** Results of MAPE and MAE for IMGGM, SFOGM, GMC, FOAGRNN, MGM models in the example of Shanxi province and Beijing city

Region	Index	MGM	FOAGRNN	GMC	SFOGM	IMGGM
Shanxi	MAE	151.5081	110.1308	131.295	97.0248	<b>64.2708</b>
	MAPE	10.4578	5.6373	6.6547	<b>4.0485</b>	5.1667
Beijing	MAE	155.3449	128.3347	151.4587	109.2398	<b>95.0955</b>
	MAPE	11.0951	6.7878	10.2873	<b>4.9212</b>	5.3468



**Fig. 4** The illustration of MAE, MAPE, RMSE of different models for Shanxi province and Beijing city

indices, except for SFOGM model in the MAPEs index. The MAPEs index of the developed model is 5.1667 and 5.3468% for Shanxi province and Beijing city, respectively. For the original grey model of MGM, the prediction accuracy is poor because this model uses the different approaches to estimate the parameters and variables; thus, the estimated parameters are the approximate values of developing grey matrix and endogenous control grey matrix, which degrades the accuracy of grey prediction model. Among the four compared forecasting methods of SFOGM, GMC, FOAGRNN, and MGM, SFOGM gets higher prediction accuracy when compared with GMC, FOAGRNN, and MGM, because

SFOGM uses the particle swarm algorithm to estimate the adjustable fractional weighted coefficients and corresponding time parameter of the initial condition, which promotes the forecasting precision. In addition, this model introduces the fractional weighted coefficients to design the optimized initial condition that captures the dynamic characteristics of the electricity consumption observations. SFOGM and GMC perform better than MGM in both cases. However, when comes to the forecasting models using optimization algorithm, such as FOAGRNN and SFOGM, GMC still shows large forecasting errors. The MAPEs of the SFOGM, FOAGRNN, and GMC are 4.0485, 5.6373, and 6.6547% for Shanxi province and 4.9212, 6.7878, and 10.2873% for Beijing city, respectively. The grey forecasting model using optimization algorithm, i.e. SFOGM, obtained more reliable and accurate electricity consumption forecasting results than GMC and MGM, both of which are not using optimization algorithm. The forecasting results of SFOGM are more consistent with the actual values; however, SFOGM estimates the parameters and variables using different approaches, which degrades the accuracy of grey prediction model. All in all, the proposed model is more suitable than the others to forecast annual electricity consumption.

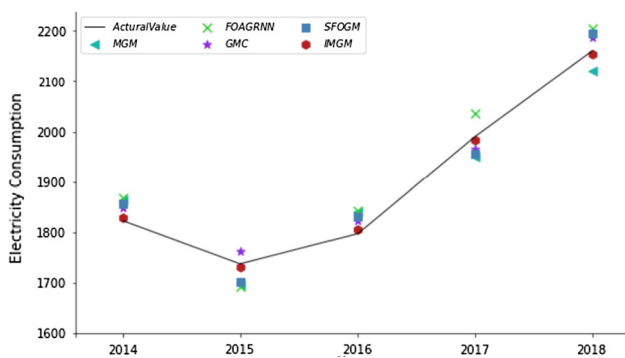
Table 9 and Fig. 5 give the forecasting results of electricity consumption with the IMGGM, MGM, FOAGRNN, GMC, SFOGM models. Table 9 also lists the relative errors of the five forecasting models. According to Table 9 and Fig. 5, it can be clearly seen that all of the five forecasting models capture the changing trend, but the performance of IMGGM, SFOGM, GMC, and FOAGRNN is better than the MGM model.

Table 10 gives the number of the relative errors falling in the scope of  $[-3, +3\%]$  and the number of the relative errors falling in the scope of  $[-1, +1\%]$ . Table 10 also presents the maximum and minimum relative errors. The error range  $[-3, +3\%]$  is always considered as a standard to measure the performance of the forecasting model [41]. Therefore, this paper uses this range to compare the five forecasting models.

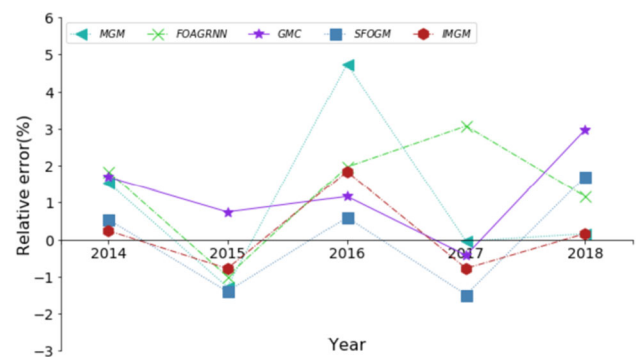
From Table 10, it can be seen that the relative errors of IMGGM, SFOGM, GMC, and FOAGRNN models are in the scope of  $[-3, +3\%]$ . The number of the relative error in  $[-1, +1\%]$  using IMGGM, SFOGM, GMC, FOAGRNN model is 3, which means 60% of forecasting points are in  $[-1, +1\%]$ . Using MGM models, the number of the relative error in  $[-1, +1\%]$  is 2, which means 40% of forecasting points

**Table 9** Forecasting results of electricity consumption using IMGGM, SFOGM, GMC, FOAGRNN, MGM models

Region	Year	Actual value	MGM		FOAGRNN		GMC		SFOGM		IMGGM	
			Result	Error (%)	Result	Error (%)	Result	Error (%)	Result	Error (%)	Result	Error (%)
Shanxi	2014	1822.63	1823.58	1.536	1823.80	1.838	1823.69	1.687	1822.87	0.536	1822.66	1.235
	2015	1737.21	1736.15	-1.281	1736.36	-0.996	1737.60	0.744	1736.07	-0.403	1736.51	-0.786
	2016	1797.18	1800.42	4.736	1798.44	0.964	1797.87	0.163	1797.46	0.591	1798.35	1.834
	2017	1990.61	1990.43	-0.048	1992.66	2.972	1990.16	-0.428	1989.40	-1.489	1989.91	-0.781
	2018	2160.53	2160.50	0.152	2161.22	0.165	2162.51	2.967	2161.60	1.693	2160.50	0.157
Beijing	2014	937.05	937.93	1.614	938.05	0.807	937.96	1.649	937.27	0.552	937.01	0.138
	2015	952.72	951.74	-1.369	951.99	-0.961	953.12	0.835	951.72	-0.394	952.08	-0.820
	2016	1020.27	1023.16	4.824	1021.39	0.991	1020.85	0.134	1020.57	0.684	1021.25	1.731
	2017	1066.89	1066.77	0.002	1068.73	2.938	1066.44	-0.516	1065.84	-1.480	1066.3	-0.737
	2018	1142.38	1142.34	0.139	1143.04	1.263	1144.06	2.882	1143.34	1.735	1142.38	1.195



**Fig. 5** Forecasting results of Shanxi’s annual electricity consumption from 2014 to 2018



**Fig. 6** Error analysis for Shanxi province

are in  $[-1, +1\%]$ . The maximum error using IMGGM model is 1.834%, smaller than that of 3 models containing MGM, FOAGRNN, and GMC model. The minimum relative error using IMGGM is  $-0.157\%$ , better than that of three models containing FOAGRNN, GMC, and SFOGM model.

Figure 6 describes the error analysis of the five forecasting models. From Fig. 6, the deviation between the forecasting results and the actual value can be captured, which shows the performance of IMGGM, SFOGM, GMC, FOAGRNN is better than the MGM model.

Table 11 lists the electricity consumption of Shanxi province and Beijing city in 2019 and 2020 forecasted by the five models.

### 5 Conclusion and discussion

This paper primarily focuses on the prediction of electricity consumption. We use the grey forecasting model since it can achieve prediction based on restricted number of available

**Table 10** The maximum and minimum relative errors and the number of the relative error in  $[-3, +3\%]$  and  $[-1, +1\%]$ , in which  
 $rate\_1 = \frac{\text{Points in } [-1, +1\%]}{\text{Total Points}}$ ,  
 $rate\_3 = \frac{\text{Points in } [-3, +3\%]}{\text{Total Points}}$

Region	Index	MGM	FOAGRNN	GMC	SFOGM	IMGGM
Shanxi	Maximum relative errors	4.736%	2.972%	2.967%	1.693%	1.834%
	Minimum relative errors	-0.048%	0.165%	0.163%	0.403%	0.157%
	rate_1	2/5 (40%)	3/5 (60%)	3/5 (60%)	3/5 (60%)	3/5 (60%)
	rate_3	4/5 (80%)	5/5 (100%)	5/5 (100%)	5/5 (100%)	5/5 (100%)
Beijing	Maximum relative errors	4.824	2.938	2.882	1.735	1.731
	Minimum relative errors	0.002	0.807	0.134	-0.394	0.138
	rate_1	2/5 (40%)	3/5 (60%)	3/5 (60%)	3/5 (60%)	3/5 (60%)
	rate_3	4/5 (80%)	5/5 (100%)	5/5 (100%)	5/5 (100%)	5/5 (100%)

**Table 11** Forecasting results of electricity consumption of Shanxi province and Beijing city in 2019 and 2020

Region	Year	MGM	FOAGRNGMC	SFOGM	IMGM	
Shanxi	2019	2267.64	2267.38	2271.12	2271.27	2271.11
	2020	2345.29	2348.42	2345.62	2345.00	2349.20
Beijing	2019	1171.79	1173.13	1172.23	1169.72	1174.54
	2020	1206.67	1205.9	1201.14	1204.32	1201.01

data. Since realistic and complex systems are often consisted of many variables which are not independent of each other and have mutual correlation among them, we focus on research on multivariable grey model. Different from the original multivariable grey model, which uses different approaches for parameters estimation and variables prediction thus degrading the prediction accuracy, the proposed improved multivariable grey model used the same first-order linear difference equation with constant coefficients for both parameters estimation and variables prediction so that the accuracy is significantly improved.

To investigate the performance of the proposed methods in the prediction and estimation of electricity consumption, other driving factors of the system are first acquired. These include consumption level of urban residents, second industry proportion, total volume of retail sales, GDP, permanent residents, and fixed assets investment. Grey relational analysis and statistical correlation analysis are used to justify their strong relations. Using 1999–2018 historical data, it is found that IMGGM, SFOGM, GMC, and FOAGRNN models have superior estimation performance over the MGM model. For training data, the MAPE of traditional MGM model is 10.4578, that of GMC model is 6.6547, that of FOAGRNN is 5.6373, and that of IMGGM is 5.1667. IMGGM is best. For the out-of-sample data in 2014, the relative error of MGM model is 1.614, that of SFOGM model is 0.807, that of GMC is 1.649, that of SFOGM is 0.552, and that of IMGGM is 0.138. IMGGM is the smallest. Therefore, it can be used for other real cases for electricity consumption forecasting. In theory, the grey forecasting model is suitable for addressing the limited sample forecasting problems [42]. Limited sample is suitable for short-term projection. In practice, the trends of these relative factors may change or the relationship between the reference series and comparison series may vary in the long term, so the IMGGM is also applicable for short-term projection.

Although the proposed model got the competitive performance, one of limitations is that this model cannot deal with the conflict between prediction accuracy and generalization ability well. The first-order linear difference equation with constant coefficients used for both parameters estimation and variables prediction lacks generalization ability. In future studies, we will devote to studying first-order non-

homogeneous difference equation with variable coefficients and second-order nonhomogeneous difference equation with variable coefficients, which can be used to study more complex models.

**Acknowledgements** We acknowledge research project supported by the Natural Science Foundation of Shanxi Province, China (Grant No. 201801D121136), and research project supported by Shanxi Scholarship Council of China (Grant No. HGKY2019024).

## Compliance with ethical standards

**Conflict of interest** No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work.

## References

- Kaytez F, Taplamacioglu MC, Cam E et al (2015) Forecasting electricity consumption: a comparison of regression analysis, neural networks and least squares support vector machines. *Int J Electr Power Energy Syst* 67:431–438
- He Y, Qin Y, Wang S et al (2019) Electricity consumption probability density forecasting method based on LASSO-quantile regression neural network. *Appl Energy* 233:565–575
- Kewo A, Munir R, Lapu AK (2015) IntelligEnSia based electricity consumption prediction analytics using regression method. In: 2015 international conference on electrical engineering and informatics (ICEEI). IEEE, pp 523–528
- Bianco V, Manca O, Nardini S (2013) Linear regression models to forecast electricity consumption in Italy. *Energy Sources Part B* 8(1):86–93
- Torrini FC, Souza RC, Oliveira FLC et al (2016) Long term electricity consumption forecast in Brazil: a fuzzy logic approach. *Socio-Econ Plan Sci* 54:18–27
- Tian C, Hao Y (2018) A novel nonlinear combined forecasting system for short-term load forecasting. *Energies* 11(4):712
- Chujai P, Kerdprasop N, Kerdprasop K (2013) Time series analysis of household electric consumption with ARIMA and ARMA models. In: Proceedings of the international multi conference of engineers and computer scientists, vol 1, pp 295–300
- Pappas SS, Ekonomou L, Karamousantas DC et al (2008) Electricity demand loads modeling using auto regressive moving average (ARMA) models. *Energy* 33(9):1353–1360
- Pappas SS, Ekonomou L, Karampelas P et al (2010) Electricity demand load forecasting of the Hellenic power system using an ARMA model. *Electr Power Syst Res* 80(3):256–264
- Marwala L, Twala B (2014) Forecasting electricity consumption in South Africa: ARMA, neural networks and neuro-fuzzy systems. In: 2014 international joint conference on neural networks (IJCNN). IEEE, pp 3049–3055
- Chen YH, Hong WC, Shen W et al (2016) Electric load forecasting based on a least squares support vector machine with fuzzy time series and global harmony search algorithm. *Energies* 9(2):70
- Zhang X, Wang J, Zhang K (2017) Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm. *Electr Power Syst Res* 146:270–285
- Li HZ, Guo S, Li CJ et al (2013) A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm. *Knowl-Based Syst* 37:378–387

14. Marwala L, Twala B (2013) Univariate modelling of electricity consumption in South Africa: neural networks and neuro-fuzzy systems. In: 2013 IEEE international conference on systems, man, and cybernetics. IEEE, pp 2238–2243
15. Ju-Long D (1982) Control problems of grey systems. *Syst Control Lett* 1(5):288–294
16. Akay D, Atak M (2007) Grey prediction with rolling mechanism for electricity demand forecasting of Turkey. *Energy* 32(9):1670–1675
17. Lee YS, Tong LI (2011) Forecasting energy consumption using a grey model improved by incorporating genetic programming. *Energy Convers Manag* 52(1):147–152
18. Xiong P, Dang Y, Yao T et al (2014) Optimal modeling and forecasting of the energy consumption and production in China. *Energy* 77:623–634
19. Hamzacebi C, Es HA (2014) Forecasting the annual electricity consumption of Turkey using an optimized grey model. *Energy* 70:165–171
20. Zhu X, Dang Y, Ding S (2020) Using a self-adaptive grey fractional weighted model to forecast Jiangsu's electricity consumption in China. *Energy* 190:116417
21. Ding S, Hipel KW, Dang Y (2018) Forecasting China's electricity consumption using a new grey prediction model. *Energy* 149:314–328
22. Wang J, Du P, Lu H et al (2018) An improved grey model optimized by multi-objective ant lion optimization algorithm for annual electricity consumption forecasting. *Appl Soft Comput* 72:321–337
23. Xu N, Dang Y, Gong Y (2017) Novel grey prediction model with nonlinear optimized time response method for forecasting of electricity consumption in China. *Energy* 118:473–480
24. Wu L, Gao X, Xiao Y et al (2018) Using a novel multi-variable grey model to forecast the electricity consumption of Shandong Province in China. *Energy* 157:327–335
25. Senapati RN, Sahoo NC, Mishra S (2016) Convolution integral based multivariable grey prediction model for solar energy generation forecasting. In: 2016 IEEE international conference on power and energy (PECon). IEEE, pp 663–667
26. Dang Y, Ding S, Zhao K (2017) Modelling and forecasting of Jiangsu's total electricity consumption using the novel grey multi-variable model. In: 2017 international conference on grey systems and intelligent services (GSIS). IEEE, pp 193–199
27. Zhong Z, Yang C, Cao W et al (2017) Short-term photovoltaic power generation forecasting based on multivariable grey theory model with parameter optimization. *Mathematical Problems in Engineering*, 2017
28. 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. *Energy* 72:434–442
29. Rasheed MK, Majeed AH (2019) Developing an improved grey prediction model for application to electricity consumption prediction: toward enhanced model accuracy. *J Phys Conf Ser* 1362(1):012137
30. Wang M, Luo Q, Kuang L et al (2019) Optimized rolling grey model for electricity consumption and power generation prediction of China. *IAENG Int J Appl Math* 49(4):1–11
31. Ayvaz B, Kusakci AO (2017) Electricity consumption forecasting for Turkey with nonhomogeneous discrete grey model. *Energy Sources Part B* 12(3):260–267
32. Sun W, Yan Y (2008) The model based on grey theory and PSO for electricity consumption forecasting. In: 2008 international conference on intelligent computation technology and automation (ICICTA), vol 1. IEEE, pp 152–156
33. Moonchai S, Chutsagulprom N (2020) Short-term forecasting of renewable energy consumption: augmentation of a modified grey model with a Kalman filter. *Appl Soft Comput* 87:105994
34. Villadsen J, Michelsen ML (1978) Solution of differential equation models by polynomial approximation. Prentice-Hall, Englewood Cliffs
35. Xu Y, Yang W, Wang J (2017) Air quality early-warning system for cities in China. *Atmos Environ* 148:239–257
36. Li HZ, Guo S, Li CJ, Sun JQ (2013) A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm. *Knowl Based Syst* 37:378–387
37. Li R, Jin Y (2018) The early-warning system based on hybrid optimization algorithm and fuzzy synthetic evaluation model. *Inf Sci* 435:296–319
38. Du P, Wang J, Yang W et al (2018) Multi-step ahead forecasting in electrical power system using a hybrid forecasting system. *Renew Energy* 122:533–550
39. Guo X, Liu S, Wu L et al (2015) A multi-variable grey model with a self-memory component and its application on engineering prediction. *Eng Appl Artif Intell* 42:82–93
40. LeBlanc DC (2004) Statistics: concepts and applications for science. Jones and Bartlett Learning, Sudbury
41. Niu D, Wang Y, Wu DD (2010) Power load forecasting using support vector machine and ant colony optimization. *Expert Syst Appl* 37(3):2531–2539
42. Wu LF, Liu SF, Yao LG et al (2013) The effect of sample size on the grey system model. *Appl Math Model* 37(9):6577e83

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