

## The Extended Grey GM(2, 1, $\sum \exp(ct)$ ) Model and Its Application in the Predictions

Maolin CHENG\*

*School of Mathematical Sciences, Suzhou University of Science and Technology, Suzhou 215009, China*  
*E-mail: cml@mail.usts.edu.cn*

Bin LIU

*School of Business, Suzhou University of Science and Technology, Suzhou 215009, China*  
*E-mail: ustsliubin@yeah.net*

**Abstract** The conventional grey GM(2, 1) model built for the fast growing time sequence generally has big errors. To improve the modeling precision, the paper improves from the following two aspects: First, the paper transforms the accumulated generating sequence of original time sequence quantitatively to make the transformed time sequence have the better adaptability to the model; second, the paper extends the conventional grey GM(2, 1) model's structure to make the extended model meet the variation law of fast growing sequence better. The extended grey model is called the GM(2, 1,  $\sum \exp(ct)$ ) model. The paper offers the parameter optimization method and the solving method of time response sequence of GM(2, 1,  $\sum \exp(ct)$ ) model. Using the model and methods proposed, the paper builds the GM(2, 1,  $\sum \exp(ct)$ ) models for the natural gas consumption of China and Chongqing City, China, respectively. Results show that the models built have high simulation precision and prediction precision.

**Keywords** GM(2, 1,  $\sum \exp(ct)$ ) model; parameter optimization method; time response equation; natural gas consumption; prediction

### 1 Introduction

The common models used for grey predictions include the GM(1, 1) model<sup>[1,2]</sup>, the GM(1, 1) power model<sup>[3,4]</sup>, the GM(1,  $N$ ) model<sup>[5,6]</sup>, the GM( $N$ , 1) model<sup>[7]</sup>, and so on, in which the GM( $N$ , 1) is an important type. In the GM( $N$ , 1) models, the GM(2, 1) model is used the most frequently. It has two characteristic roots and its dynamic process can reflect various sequences including the monotonic, the nonmonotonic and the oscillating sequences.

The GM( $N$ , 1) model is complicated. Currently the research on the GM(2, 1) model is getting somewhere. Domestic and foreign scholars have made studies from the following aspects. 1) The research on the transformation of accumulated generating sequence. To solve the poor prediction precision of conventional second-order univariate grey GM(2, 1) model for some sequences, Zeng and Luo<sup>[8]</sup> introduced the fractional order accumulated operator on the basis of conventional GM(2, 1) model and proposed a discrete GM(2, 1) model based on the fractional

---

Received November 12, 2022, accepted February 6, 2023

Supported by Natural Science Foundation of China (11401418)

\*Corresponding author

order accumulation. The model used the same grey difference equation for both parameter identification and simulation & prediction, and thus eliminated the errors in the transformation from the difference equation (used for parameter identification) to the differential equation (used for simulation & prediction). In addition, the fractional order accumulation way could reduce the randomness and uncertainty of original sequence better relative to the conventional first-order accumulation way, and could improve the model's prediction precision significantly by selecting a proper accumulation order. 2) The research on the model's parameter estimation and optimization. Tang and Lu<sup>[9]</sup> inferred a range format equation based on the time response function of GM(2, 1) model and gave a parameter estimation method of least square method for the equation. The method avoided the complicated construction of background value and the errors generated in the transformation from the difference equation to the differential equation. Su and Shao<sup>[10]</sup> proposed an optimization method for time response coefficients  $c_1$  and  $c_2$  of GM(2, 1) model. The method avoided the common method's defect of determining the simulation value equals to the true value at the first time point and the last time point for the initial value without considering the influence of the middle time point on the determination of response coefficient. They proved the response coefficient determined with the method proposed could improve the model's prediction precision significantly. 3) The research on the model's initial condition determination. For the GM(2, 1) model, Zhao and Chen<sup>[11]</sup> substituted the secondary parameter estimate of least squares for the constant of whitening differential equation determined by boundary conditions, and proved that the new method had a better result with an calculation example. To solve the poor precision of GM(2, 1) model generally using two initial values to determine parameters  $c_1$  and  $c_2$ , Su, et al.<sup>[12]</sup> considered the influence of value at each time point and got parameters  $c_1$  and  $c_2$  based on the new objective function. The method improved the model's precision overall. An<sup>[13]</sup> solved and obtained parameters  $c_1$  and  $c_2$  through the objective function which was the minimum of sum of squares of errors between the actual values and theoretical solution of time response equation. Considering the GM(2, 1) model required two initial values and might have some problems when the values were given, Shen and Zhao<sup>[14]</sup> proposed a new method to determine the system's parameters with the least square idea, and proved that the improved method was very effective and practical through a numerical experiment. 4) The research on the model's nature. Zeng and Xiao<sup>[15]</sup> introduced the accumulation method for the parameter estimation of GM(2, 1) model, solved the ill-conditioned problem existing in the GM(2, 1) model, and proved the multiply transformation didn't change the model's development coefficient and precision. Xiao and Guo<sup>[16]</sup> discussed the ill-conditioned problem existing in the parameter identification process of GM(2, 1) model using the vector multiply transformation. 5) The research on the model's form change and the extended model. Many researchers have made related studies. Liu, et al.<sup>[17]</sup> extended the conventional GM(2, 1) model's grey action into a linear function to construct a GM(2, 1) model with the optimized grey action. They have proved that the GM(2, 1) model with the improved grey action had high simulation precision with an example. Shui, et al.<sup>[18]</sup> introduced the time term into the grey action to optimize the GM(2, 1) model. Yong and Wei<sup>[19]</sup> optimized the grey derivative and background value of unequal-interval DGM(2, 1) model. Shao and Su<sup>[20]</sup> constructed a grey GM(2, 1) model with the new second-order grey derivative. Xu and Dang<sup>[21]</sup>

proposed a GM(2, 1) model with the optimized structure, i.e., the SOGM(2, 1) model. The new model first constructed a new grey equation with the optimized structure using the background value sequence and the inverse accumulated generating sequence, and then derived the parameter estimate based on the minimum error. Comparing with other models in terms of results, the SOGM(2, 1) was effective and feasible. Cheng and Shi<sup>[22]</sup> transformed the grey GM(2, 1) model's basic form into a second-order difference equation form based on the structure of linear difference equation, and then solved parameters and made predictions with the same difference equation, in which case the model's grey differential equation (used for parameter solving) highly consisted with the whitening equation (used for predictions) structurally. It has been proved theoretically and practically that the improved grey GM(2, 1) model had significantly improved prediction precision.

The studies above all improved the prediction precision of GM(2, 1) model from different perspectives, but the modeling precision was still poor in some cases, especially for the modeling prediction of fast growing time sequence in which case the model shows big errors. To improve the modeling precision for fast growing time sequence, the paper first makes a transformation of the accumulated generating sequence of original time sequence quantitatively to make the transformed time sequence have the better adaptability to the model. Next, the paper extends the grey action of conventional grey GM(2, 1) model into a superimposed exponential form to make the extended model meet the variation law of fast growing sequence better. The extended grey model is called the GM(2, 1,  $\sum \exp(ct)$ ) model. The paper gives the parameter optimization methods and the solving methods of time response equation of two forms of GM(2, 1,  $\sum \exp(ct)$ ) models. The paper builds the GM(2, 1,  $\sum \exp(ct)$ ) models for the natural gas consumption of China and Chongqing city, China, respectively. Results show that the models built have high precision which is significantly higher than that of the conventional GM(2, 1) model and superior to that of the models studied in referencing documents.

## 2 The Parameter Estimation Method and Modeling of Conventional Grey GM(2, 1) Model

The paper supposes the original time sequence is  $x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$  and the once accumulated generating sequence is  $x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$  in which  $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)$ ,  $k = 1, 2, \dots, n$ . The once degressive generating sequence is  $\alpha^{(1)}x^{(0)} = (\alpha^{(1)}x^{(0)}(2), \alpha^{(1)}x^{(0)}(3), \dots, \alpha^{(1)}x^{(0)}(n))$  in which  $\alpha^{(1)}x^{(0)}(k) = x^{(0)}(k) - x^{(0)}(k-1)$ ,  $k = 2, \dots, n$ .

**Definition 1**<sup>[23]</sup> We call  $z^{(1)} = \{z^{(1)}(1), z^{(1)}(2), \dots, z^{(1)}(n)\}$  the background value sequence of  $x^{(1)}(k)$  in which  $z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k) + x^{(1)}(k-1))$ ,  $k = 2, 3, \dots, n$ .

**Definition 2**<sup>[23]</sup> We call  $\alpha^{(1)}x^{(0)}(k) + a_1x^{(0)}(k) + a_2z^{(1)}(k) = b_0$  the grey differential equation of grey GM(2, 1) model.

**Definition 3**<sup>[23]</sup> We call  $\frac{d^2x^{(1)}(t)}{dt^2} + a_1\frac{dx^{(1)}(t)}{dt} + a_2x^{(1)}(t) = b_0$  the whitening equation of grey GM(2, 1) model.

**Theorem 1**<sup>[23]</sup> The grey GM(2, 1) model's parameter estimates are

$$\hat{B} = \begin{pmatrix} a_1 \\ a_2 \\ b_0 \end{pmatrix} = (X'X)^{-1}X'Y, \quad (1)$$

where

$$X = \begin{pmatrix} -x^{(0)}(2) & -z^{(1)}(2) & 1 \\ -x^{(0)}(3) & -z^{(1)}(3) & 1 \\ \vdots & \vdots & \vdots \\ -x^{(0)}(n) & -z^{(1)}(N) & 1 \end{pmatrix}, \quad (2)$$

$$Y = \begin{pmatrix} x^{(0)}(2) - x^{(0)}(1) \\ x^{(0)}(3) - x^{(0)}(2) \\ \vdots \\ x^{(0)}(n) - x^{(0)}(n-1) \end{pmatrix}. \quad (3)$$

**Theorem 2**<sup>[23]</sup> The grey GM(2,1) model's time response function is

$$x^{(1)}(t) = \frac{b_0}{a_2} + \frac{b_0g_4 - b_0g_2 - a_2c_1g_4 + a_2c_2g_2}{g_1} e^{-(\frac{a_1}{2} - g_6)t} - \frac{b_0g_5 - b_0g_3 + a_2c_2g_3 - a_2c_1g_5}{g_1} e^{-(\frac{a_1}{2} + g_6)t}, \quad (4)$$

where  $g_1 = a_2(g_2g_5 - g_4g_3)$ ,  $g_2 = e^{-\frac{a_1}{2} - g_6}$ ,  $g_3 = e^{g_6 - \frac{a_1}{2}}$ ,  $g_4 = e^{-a_1 - \sqrt{a_1^2 - 4a_2}}$ ,  $g_5 = e^{-a_1 + \sqrt{a_1^2 - 4a_2}}$ ,  $g_6 = \frac{\sqrt{a_1^2 - 4a_2}}{2}$ ,  $c_1 = x^{(1)}(1) = x^{(0)}(1)$ ,  $c_2 = x^{(1)}(2) = x^{(0)}(1) + x^{(0)}(2)$ .

In Theorem 2, we can get the time response equation of GM(2,1) model easily using command dsolve of software MATLAB.

According to the time response function sequence, we can get the simulation and prediction values of original sequence with  $\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$ .

Suppose there are  $N$  years of observation data of which the data from year 1 to year  $n$  are used for the modeling and the data from year  $n+1$  to year  $N$  are used for the prediction. The simulation value of  $x^{(0)}$  is  $\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$  ( $k = 2, 3, \dots, n$ ), and the prediction value is  $\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$  ( $k = n+1, \dots, N$ ), and then the average simulation relative error is  $\text{MAPE}_1 = \frac{1}{n-1} \sum_{k=2}^n \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%$  and the average prediction relative error is  $\text{MAPE}_2 = \frac{1}{N-n} \sum_{k=n+1}^N \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%$ .

To improve the model's simulation and prediction precision, we use  $z^{(1)}(k) = \alpha x^{(1)}(k-1) + (1-\alpha)x^{(1)}(k)$  ( $0 \leq \alpha \leq 1$ ) to substitute for background value  $z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k) + x^{(1)}(k-1))$ . We determine  $\alpha$  by making  $\max(\text{MAPE}_1, \text{MAPE}_2)$  have the minimum value.

### 3 The Form of Extended Grey GM(2,1, $\sum \exp(ct)$ )

The paper mainly explores the grey modeling of fast growing sequence. To make the data have the better adaptability to the model, the paper first improves the accumulated generating sequence of  $x^{(0)}(t)$  and then has the following Definition 4.

**Definition 4** Suppose the original time sequence is  $x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$  and the new accumulated generating sequence is defined as  $x^{(r)} = \{x^{(r)}(1), x^{(r)}(2), \dots, x^{(r)}(n)\}$  in which  $x^{(r)}(k) = \sum_{i=1}^k \frac{x^{(0)}(i)}{v+g \cdot r^i}$ ,  $v > 0, g > 0, 0 < r \leq 1, k = 1, 2, \dots, n$ .

Next, the paper extends the conventional grey GM(2,1) model's structure to meet the requirements of this type of modeling, and then has the following Definitions 5 and 6.

**Definition 5** We record  $\alpha^{(1)}x^{(r)}(k) = x^{(r)}(k) - x^{(r)}(k-1)$  and the new background value is  $z^{(r)}(k) = \int_{k-1}^k x^{(r)}(t)dt$ . We call  $\alpha^{(1)}x^{(r)}(k) + a_1x^{(r)}(k) + a_2z^{(r)}(k) = b_0 + \sum_{i=1}^p f_i e^{c_i t}$  the grey differential equation of extended grey model GM(2, 1,  $\sum \exp(ct)$ ).

**Definition 6** We call  $\frac{d^2x^{(r)}}{dt^2} + a_1 \frac{dx^{(r)}}{dt} + a_2x^{(r)} = b_0 + \sum_{i=1}^p f_i e^{c_i t}$  the whitening equation of extended grey model GM(2, 1,  $\sum \exp(ct)$ ).

Especially, when  $f_i = 0$ , the equation above becomes the whitening equation of conventional grey GM(2, 1) model.

Because background value  $z^{(r)}(k)$ 's generation coefficient can be an equal-weighted or unequal-weighted constant, there will be two different extended models.

**Definition 7** If  $z^{(r)}(k)$ 's generation coefficient is an equal-weighted constant, i.e.,  $z^{(r)}(k) = 0.5x^{(r)}(k-1) + 0.5x^{(r)}(k)$ , we call  $\alpha^{(1)}x^{(r)}(k) + a_1x^{(r)}(k) + a_2z^{(r)}(k) = b_0 + \sum_{i=1}^p f_i e^{c_i t}$  the extended model GM(2, 1, 0.5,  $\sum \exp(ct)$ ).

**Definition 8** If  $z^{(r)}(k)$ 's generation coefficient is an unequal-weighted constant, i.e.,  $z^{(r)}(k) = \alpha x^{(r)}(k-1) + (1-\alpha)x^{(r)}(k)$ , we call  $\alpha^{(1)}x^{(r)}(k) + a_1x^{(r)}(k) + a_2z^{(r)}(k) = b_0 + \sum_{i=1}^p f_i e^{c_i t}$  the extended model GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ).

## 4 The Parameter Optimization and Modeling of Extended Grey Model GM(2, 1, $\sum \exp(ct)$ )

### 4.1 The Parameter Optimization Method and Model Building of Extended Grey Model GM(2, 1, 0.5, $\sum \exp(ct)$ )

**Theorem 3** The grey differential equation of extended grey model GM(2, 1, 0.5,  $\sum \exp(ct)$ ) is

$$\alpha^{(1)}x^{(r)}(k) + a_1x^{(r)}(k) + a_2z^{(r)}(k) = b_0 + \sum_{i=1}^p f_i e^{c_i t}. \quad (5)$$

If the generation coefficient is an equal-weighted constant, i.e.,  $z^{(r)}(k) = 0.5x^{(r)}(k-1) + 0.5x^{(r)}(k)$ , then for the given  $v, g, r, c_1, c_2, \dots, c_p$ , with the least square method, we can get

$$\hat{B} = \begin{pmatrix} a_1 \\ a_2 \\ b_0 \\ f_1 \\ \vdots \\ f_p \end{pmatrix} = (X'X)^{-1}X'Y, \quad (6)$$

where

$$X = \begin{pmatrix} -\sum_{i=1}^2 \frac{x^{(0)}(i)}{v+g \cdot r^i} & -0.5 \sum_{i=1}^1 \frac{x^{(0)}(i)}{v+g \cdot r^i} - 0.5 \sum_{i=1}^2 \frac{x^{(0)}(i)}{v+g \cdot r^i} & 1 & b_1(\frac{e^{c_1}-1}{c_1})e^{c_1} & \dots & b_p(\frac{e^{c_p}-1}{c_p})e^{c_p} \\ -\sum_{i=1}^3 \frac{x^{(0)}(i)}{v+g \cdot r^i} & -0.5 \sum_{i=1}^2 \frac{x^{(0)}(i)}{v+g \cdot r^i} - 0.5 \sum_{i=1}^3 \frac{x^{(0)}(i)}{v+g \cdot r^i} & 1 & b_1(\frac{e^{c_1}-1}{c_1})e^{2c_1} & \dots & b_p(\frac{e^{c_p}-1}{c_p})e^{2c_p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -\sum_{i=1}^n \frac{x^{(0)}(i)}{v+g \cdot r^i} & -0.5 \sum_{i=1}^{n-1} \frac{x^{(0)}(i)}{v+g \cdot r^i} - 0.5 \sum_{i=1}^n \frac{x^{(0)}(i)}{v+g \cdot r^i} & 1 & b_1(\frac{e^{c_1}-1}{c_1})e^{(n-1)c_1} & \dots & b_p(\frac{e^{c_p}-1}{c_p})e^{(n-1)c_p} \end{pmatrix}, \quad (7)$$

$$Y = \begin{pmatrix} \frac{x^{(0)}(2)}{v+g \cdot r^2} \\ \frac{x^{(0)}(3)}{v+g \cdot r^3} \\ \dots \\ \frac{x^{(0)}(n)}{v+g \cdot r^n} \end{pmatrix}. \quad (8)$$

*Proof* The extended grey model GM(2, 1, 0.5,  $\sum \exp(ct)$ )'s whitening equation is

$$\frac{d^2 x^{(r)}}{dt} + a_1 \frac{dx^{(r)}}{dt} + a_2 x^{(r)} = b_0 + \sum_{i=1}^p f_i e^{c_i t}, \quad (9)$$

i.e.,

$$\int_{k-1}^k \frac{d^2 x^{(r)}(t)}{dt^2} dt = -a_1 \int_{k-1}^k \frac{dx^{(r)}(t)}{dt} - a_2 \int_{k-1}^k x^{(r)}(t) dt + \int_{k-1}^k (b_0 + f_1 e^{c_1 t} + \dots + f_p e^{c_p t}) dt. \quad (10)$$

We record  $\alpha^{(1)} x^{(r)}(k) = x^{(r)}(k) - x^{(r)}(k-1)$  and the background value is  $z^{(r)}(k) = \int_{k-1}^k x^{(r)}(t) dt$ , and then

$$\begin{aligned} x^{(r)}(k) - x^{(r)}(k-1) &= -a_1 x^{(r)}(k) - a_2 z^{(r)}(k) + b_0 + f_1 \left( \frac{e^{c_1} - 1}{c_1} \right) e^{c_1(k-1)} + \\ &\quad \dots + f_p \left( \frac{e^{c_p} - 1}{c_p} \right) e^{c_p(k-1)}, \end{aligned} \quad (11)$$

$$\begin{aligned} \frac{x^{(0)}(k)}{v+g \cdot r^k} &= -a_1 \sum_{i=1}^k \frac{x^{(0)}(i)}{v+g \cdot r^i} - a_2 z^{(r)}(k) + b_0 + f_1 \left( \frac{e^{c_1} - 1}{c_1} \right) e^{c_1(k-1)} + \\ &\quad \dots + f_p \left( \frac{e^{c_p} - 1}{c_p} \right) e^{c_p(k-1)}. \end{aligned} \quad (12)$$

Because  $z^{(r)}(k) = \alpha_k x^{(r)}(k-1) + (1-\alpha_k) x^{(r)}(k) = \alpha_k \sum_{i=1}^{k-1} \frac{x^{(0)}(i)}{v+g \cdot r^i} + (1-\alpha_k) \sum_{i=1}^k \frac{x^{(0)}(i)}{v+g \cdot r^i}$ , if the generation coefficient is an equal-weighted constant, i.e.,  $\alpha_k = 0.5$ , for the given  $v, g, r, c_1, c_2, \dots, c_p$ , with the least square method, we can get

$$\hat{B} = \begin{pmatrix} a_1 \\ a_2 \\ b_0 \\ f_1 \\ \vdots \\ f_p \end{pmatrix} = (X'X)^{-1} X'Y, \quad (13)$$

where

$$X = \begin{pmatrix} -\sum_{i=1}^2 \frac{x^{(0)}(i)}{v+g \cdot r^i} & -0.5 \sum_{i=1}^1 \frac{x^{(0)}(i)}{v+g \cdot r^i} & -0.5 \sum_{i=1}^2 \frac{x^{(0)}(i)}{v+g \cdot r^i} & 1 & b_1 \left( \frac{e^{c_1}-1}{c_1} \right) e^{c_1} & \dots & b_p \left( \frac{e^{c_p}-1}{c_p} \right) e^{c_p} \\ -\sum_{i=1}^3 \frac{x^{(0)}(i)}{v+g \cdot r^i} & -0.5 \sum_{i=1}^2 \frac{x^{(0)}(i)}{v+g \cdot r^i} & -0.5 \sum_{i=1}^3 \frac{x^{(0)}(i)}{v+g \cdot r^i} & 1 & b_1 \left( \frac{e^{c_1}-1}{c_1} \right) e^{2c_1} & \dots & b_p \left( \frac{e^{c_p}-1}{c_p} \right) e^{2c_p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -\sum_{i=1}^n \frac{x^{(0)}(i)}{v+g \cdot r^i} & -0.5 \sum_{i=1}^{n-1} \frac{x^{(0)}(i)}{v+g \cdot r^i} & -0.5 \sum_{i=1}^n \frac{x^{(0)}(i)}{v+g \cdot r^i} & 1 & b_1 \left( \frac{e^{c_1}-1}{c_1} \right) e^{(n-1)c_1} & \dots & b_p \left( \frac{e^{c_p}-1}{c_p} \right) e^{(n-1)c_p} \end{pmatrix}, \quad (14)$$

$$Y = \begin{pmatrix} \frac{x^{(0)}(2)}{v+g \cdot r^2} \\ \frac{x^{(0)}(3)}{v+g \cdot r^3} \\ \vdots \\ \frac{x^{(0)}(n)}{v+g \cdot r^n} \end{pmatrix}. \quad (15)$$

Using the command `dsolve` of software MATLAB, we can get the extended GM(2, 1) model's time response equation  $\hat{x}^{(r)}(t)$  easily.

After estimating the parameters, we can calculate and get the simulation and prediction values of  $x^{(0)}$  with the following Theorem 4.

**Theorem 4**  $x^{(0)}$ 's simulation value is

$$\hat{x}^{(0)}(k) = (\hat{x}^{(r)}(k) - \hat{x}^{(r)}(k-1))(v + gr^k), \quad k = 1, 2, \dots, n, \quad (16)$$

and prediction value is

$$\hat{x}^{(0)}(k) = (\hat{x}^{(r)}(k) - \hat{x}^{(r)}(k-1))(v + gr^k), \quad k = n+1, \dots, N. \quad (17)$$

*Proof* We suppose the time response equation is  $\hat{x}^{(r)}(t)$ . Because  $\hat{x}^{(r)}(k) = \sum_{i=1}^k \frac{\hat{x}^{(0)}(i)}{v+g \cdot r^i}$ , then we get

$$\hat{x}^{(r)}(k) - \hat{x}^{(r)}(k-1) = \sum_{i=1}^k \frac{\hat{x}^{(0)}(i)}{v+g \cdot r^i} - \sum_{i=1}^{k-1} \frac{\hat{x}^{(0)}(i)}{v+g \cdot r^i} = \frac{\hat{x}^{(0)}(k)}{v+g \cdot r^k}, \quad (18)$$

and then  $\hat{x}^{(0)}(k) = (\hat{x}^{(r)}(k) - \hat{x}^{(r)}(k-1))(v + gr^k)$ .

When  $k = 1, 2, \dots, n$ , we get simulation value  $\hat{x}^{(0)}(t)$ , and when  $k = n+1, n+2, \dots, N$ , we get prediction value  $\hat{x}^{(0)}(t)$ .

In fact, in Theorems 3 and 4, the values of  $v, g, r, c_1, c_2, \dots, c_p$  are required, which can be determined using an optimization method. To avoid big simulation or prediction errors, we define the objective function as the minimum of  $\max(\text{MAPE}_1, \text{MAPE}_2)$ , i.e.,

$$\begin{aligned} & \min_{v, g, r, c_1, c_2, \dots, c_p} (\max(\text{MAPE}_1, \text{MAPE}_2)), \quad (19) \\ \text{s.t.} \quad & \begin{cases} \text{MAPE}_1 = \frac{1}{n-1} \sum_{k=2}^n \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%, \\ \text{MAPE}_2 = \frac{1}{N-n} \sum_{k=n+1}^N \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%, \\ \hat{x}^{(0)}(t) = (\hat{x}^{(r)}(t) - \hat{x}^{(r)}(t-1))(v + gr^t), \\ \hat{x}^{(r)}(t) = \text{dsolve}('D2x + a_1 Dx + a_2 x = \\ b_0 + b_1 \exp(c_1 t) + \dots + b_p \exp(c_p t)', x(1) = c_1, x(2) = c_2'), \\ (a_1, a_2, b_0, f_1, \dots, f_p)^T = (X'X)^{-1} X'Y, \\ 0 \leq c_1, c_2, \dots, c_p \leq 1, \quad v > 0, \quad g > 0, \quad 0 < r \leq 1. \end{cases} \quad (20) \end{aligned}$$

We can use MATLAB to solve the optimization problem, get all parameters and the specific time response equation, and then calculate the simulation and prediction values of  $x^{(0)}$  and the relative errors and average relative errors in the periods.

#### 4.2 The Parameter Optimization Method and Model Building of Extended Grey Model GM(2, 1, $\alpha$ , $\sum \exp(ct)$ )

**Theorem 5** The grey differential equation of extended grey model GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) is

$$\alpha^{(1)}x^{(r)}(k) + a_1x^{(r)}(k) + a_2z^{(r)}(k) = b_0 + \sum_{i=1}^p f_i e^{c_i t}. \quad (21)$$

If the generation coefficient of  $z^{(r)}(k)$  is an unequal-weighted constant, i.e.,  $z^{(r)}(k) = \alpha x^{(r)}(k-1) + (1-\alpha)x^{(r)}(k)$ , then for the given  $\alpha, v, g, r, c_1, c_2, \dots, c_p$ , with the least square method, we can get

$$\hat{B} = \begin{pmatrix} a_1 \\ a_2 \\ b_0 \\ f_1 \\ \vdots \\ f_p \end{pmatrix} = (U'V)^{-1}U'V, \quad (22)$$

where

$$U = \begin{pmatrix} -\sum_{i=1}^2 \frac{x^{(0)}(i)}{v+g \cdot r^i} - \alpha \sum_{i=1}^1 \frac{x^{(0)}(i)}{v+g \cdot r^i} - (1-\alpha) \sum_{i=1}^2 \frac{x^{(0)}(i)}{v+g \cdot r^i} & 1 & b_1(\frac{e^{c_1}-1}{c_1})e^{c_1} & \dots & b_p(\frac{e^{c_p}-1}{c_p})e^{c_p} \\ -\sum_{i=1}^3 \frac{x^{(0)}(i)}{v+g \cdot r^i} - \alpha \sum_{i=1}^2 \frac{x^{(0)}(i)}{v+g \cdot r^i} - (1-\alpha) \sum_{i=1}^3 \frac{x^{(0)}(i)}{v+g \cdot r^i} & 1 & b_1(\frac{e^{c_1}-1}{c_1})e^{2c_1} & \dots & b_p(\frac{e^{c_p}-1}{c_p})e^{2c_p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\sum_{i=1}^n \frac{x^{(0)}(i)}{v+g \cdot r^i} - \alpha \sum_{i=1}^{n-1} \frac{x^{(0)}(i)}{v+g \cdot r^i} - (1-\alpha) \sum_{i=1}^n \frac{x^{(0)}(i)}{v+g \cdot r^i} & 1 & b_1(\frac{e^{c_1}-1}{c_1})e^{(n-1)c_1} & \dots & b_p(\frac{e^{c_p}-1}{c_p})e^{(n-1)c_p} \end{pmatrix}, \quad (23)$$

$$V = \begin{pmatrix} \frac{x^{(0)}(2)}{v+g \cdot r^2} \\ \frac{x^{(0)}(3)}{v+g \cdot r^3} \\ \dots \\ \frac{x^{(0)}(n)}{v+g \cdot r^n} \end{pmatrix}. \quad (24)$$

*Proof* The proof is similar to that of Theorem 3, which is omitted here.

In fact, in Theorems 5, the values of  $\alpha, v, g, r, c_1, c_2, \dots, c_p$  are required, which can be determined using an optimization method. To avoid big simulation or prediction errors, we define the objective function as the minimum of  $\max(\text{MAPE}_1, \text{MAPE}_2)$ , i.e.,

$$\min_{\alpha, v, g, r, c_1, c_2, \dots, c_p} (\max(\text{MAPE}_1, \text{MAPE}_2)), \quad (25)$$



$$\text{s.t.} \begin{cases} \text{MAPE}_1 = \frac{1}{n-1} \sum_{k=2}^n \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%, \\ \text{MAPE}_2 = \frac{1}{N-n} \sum_{k=n+1}^N \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%, \\ \hat{x}^{(0)}(t) = (\hat{x}^{(r)}(t) - \hat{x}^{(r)}(t-1))(v + gr^t), \\ \hat{x}^{(r)}(t) = \text{dsolve}('D2x + a_1 Dx + a_2 x = \\ b_0 + b_1 \exp(c_1 t) + \dots + b_p \exp(c_p t)', x(1) = c_1, x(2) = c_2'), \\ (a_1, a_2, b_0, f_1, \dots, f_p)^T = (U'V)^{-1}U'V, \\ 0 \leq \alpha, c_1, c_2, \dots, c_p \leq 1, \quad v > 0, \quad g > 0, \quad 0 < r \leq 1. \end{cases} \quad (26)$$

We can use software MATLAB to solve the optimization problem, get all parameters and the specific time response equation. In this way, we get the original sequence's simulation value which is  $\hat{x}^{(0)}(k) = (\hat{x}^{(r)}(k) - \hat{x}^{(r)}(k-1))(v + gr^k)$  ( $k = 1, 2, \dots, n$ ) and prediction value which is  $\hat{x}^{(0)}(k) = (\hat{x}^{(r)}(k) - \hat{x}^{(r)}(k-1))(v + gr^k)$  ( $k = n+1, n+2, \dots, N$ ).

## 5 Application Examples of Extended Grey GM(2, 1, $\sum \exp(ct)$ ) Model

### 5.1 Example 1

In 2021, China's natural gas consumption reached 4663.6 billion  $\text{m}^3$ , increasing by 12.7% on a year-on-year basis, with growing rate greater than those of raw coal and crude oil, in a reasonable range of medium-high rate of growth. Data show that in recent years, the growing rate of China's natural gas consumption has been kept in the range of 25 billion  $\text{m}^3 \sim 30$  billion  $\text{m}^3$ . However, under the impact of COVID-19 pandemic, the growing rate of China's natural gas consumption fell below 20 billion  $\text{m}^3$  in 2020, increasing by 5.5% on a year-on-year basis. In 2021, the growth of China's natural gas consumption was mainly because of the strong economic recovery in China. The industry and the manufacturing industry showed the sound momentum of recovery and the COVID-19 pandemic caused a low base effect of natural gas consumption in 2020, which both pulled the consumption of natural gas in 2021. Currently, the COVID-19 pandemic is still affecting various aspects of economy and life, coupled with the conflicts between Russia and Ukraine, which may cause a sharp decline of natural gas consumption globally. Everyone is very concerned how China's natural gas consumption will change in the future. The paper builds the grey prediction models using the model and method proposed. The raw data of China's natural gas consumption are shown in Table 1.

The paper first builds the extended grey GM(2, 1, 0.5,  $\sum \exp(ct)$ ) model proposed, i.e., the following model:

$$\frac{d^2 x^{(r)}}{dt^2} + a_1 \frac{dx^{(r)}}{dt} + a_2 x^{(r)} = b_0 + \sum_{i=1}^p f_i e^{c_i t}. \quad (27)$$

The grey model is adapted to the small-sample modeling, so building an extended grey GM(2, 1, 0.5,  $\sum \exp(ct)$ ) model only requires the number of sample point is more than the number of estimated parameters to allow the parameter estimated have the unique solution. Besides, there should not be too many sample points which may cause the reduction of fitting precision of model. In Example 1, the sample point  $N=15$ . To improve the model's precision

and avoid overfitting, after the calculation and comparison, the paper chooses a proper value  $p = 2$ . In this case, the model's precision meets requirements.

Using the method proposed, we calculate and get the parameter estimates:

$$(c_1, c_2, v, g, r) = (0.1476, 0.9671, 0.6210, 1.4399, 0.5060), \quad (28)$$

$$(a_1, a_2, b_0, f_1, f_2) = (-0.16276577, 0.37105884, -30541.757, 0.029735426, 0.029735435). \quad (29)$$

In this way, we get the time response equation

$$\begin{aligned} \hat{x}^{(r)}(t) = & 0.08062362e^{0.1475805t} + 0.02588276e^{0.9670552t} + \\ & (25124.2 + 36016.71i)e^{(0.08138289 - 0.6036851i)t} + \\ & (25124.2 - 36016.71i)e^{(0.08138289 + 0.6036851i)t} - 82309.74. \end{aligned} \quad (30)$$

**Table 1** Related calculation results of grey modeling for China's natural gas consumption

Year	No.	$x^{(0)}(t)$	GM(2, 1, 0.5, $\sum \exp(ct)$ ) Model		GM(2, 1, $\alpha$ , $\sum \exp(ct)$ ) Model	
			Simulation Value	Relative Error / %	Simulation Value	Relative Error / %
2007	1	9343.26	-	-	-	-
2008	2	10900.77	10900.77	0	10900.77	0
2009	3	11764.41	12024.29	2.2091	11766.79	0.0203
2010	4	14425.92	14427.48	0.0108	13907.62	3.5928
2011	5	17803.98	17435.45	2.0699	16715.74	6.1123
2012	6	19302.62	20393.12	5.6495	19601.3	1.5473
2013	7	22096.39	22761.12	3.0083	22073.57	0.1033
2014	8	23986.7	24321.87	1.3973	23919.96	0.2783
2015	9	25178.55	25332.73	0.6123	25339.36	0.6387
2016	10	26931.01	26514.5	1.5466	26930.95	0.0002
2017	11	31452.06	28827.85	8.3435	29486.83	6.2483
2018	12	35866.3	33046.64	7.8616	33593.37	6.3372
			Prediction Value	Relative Error / %	Prediction Value	Relative Error / %
2019	13	38999.04	39145.13	0.3746	39099.89	0.2586
2020	14	41858.38	45397.91	8.4560	44542.29	6.4119
2021	15	46636.0	46678.0	0.0901	46582.55	0.1146
Average simulation relative error (2007–2018)			-	2.97	-	2.26
Average prediction relative error (2019–2021)			-	2.96	-	2.25
Average overall relative error (2007–2021)			-	2.97	-	2.26

With  $\hat{x}^{(0)}(k) = (\hat{x}^{(r)}(k) - \hat{x}^{(r)}(k-1))(v + gr^k)$ , we calculate and get the simulation and prediction values of original sequence in these periods. With  $RE(t) = \left| \frac{x^{(0)}(t) - \hat{x}^{(0)}(t)}{x^{(0)}(t)} \right| \times 100\%$ , we calculate the relative errors of simulation and prediction values of original sequence in the periods. With  $MAPE_1 = \frac{1}{n-1} \sum_{k=2}^n \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%$ , we calculate the average simulation relative error. With  $MAPE_2 = \frac{1}{N-n} \sum_{k=n+1}^N \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%$ , we calculate and get the average prediction relative error. Table 1 shows the results.

Next, the paper builds the extended grey GM(2, 1,  $\alpha, \sum \exp(ct)$ ) model proposed, i.e., the following model

$$\frac{d^2x^{(r)}}{dt^2} + a_1 \frac{dx^{(r)}}{dt} + a_2 x^{(r)} = b_0 + \sum_{i=1}^p f_i e^{c_i t}. \quad (31)$$

Similarly, we choose  $p = 2$ , and calculate and get the parameter estimates using the method proposed

$$(\alpha, c_1, c_2, v, g, r) = (0.5739, 0.1732, 0.5416, 0.5640, 1.6600, 0.4587), \quad (32)$$

$$(a_1, a_2, b_0, f_1, f_2) = (-0.2291295, 0.366367, -24292.59, 0.02261703, 0.02261698). \quad (33)$$

In this way, we get the time response equation

$$\begin{aligned} \hat{x}^{(r)}(t) = & 0.08062362e^{0.1475805t} + 0.02588276e^{0.9670552t} + \\ & (25124.2 + 36016.71i)e^{(0.08138289-0.6036851i)t} + \\ & (25124.2 - 36016.71i)e^{(0.08138289+0.6036851i)t} - 82309.74. \end{aligned} \quad (34)$$

With  $\hat{x}^{(0)}(k) = (\hat{x}^{(r)}(k) - \hat{x}^{(r)}(k-1))(v + gr^k)$ , we calculate and get the simulation and prediction values of original sequence in the periods. Table 1 shows the results. Table 1 gives the relative errors and average relative errors in the periods.

Then, we build the conventional grey GM(2, 1) model for original sequence, i.e., the following model

$$\frac{d^2x^{(1)}}{dt^2} + a_1 \frac{dx^{(1)}}{dt} + a_2 x^{(1)} = b_0. \quad (35)$$

We calculate and get the parameter estimates

$$(a_1, a_2, b_0) = (-0.4636994, 0.03901764, -3262.972). \quad (36)$$

In this case, the time response equation is

$$\hat{x}^{(1)}(t) = 98.40717e^{0.3532442t} + 83123.59e^{0.1104551t} - 83628.12. \quad (37)$$

With  $\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$ , we calculate and get the simulation and prediction values of original sequence in the periods. Table 2 shows the relative errors and average relative errors in the periods.

**Table 2** Related calculation results of the grey models for China's natural gas consumption built with the other methods

Year	No.	$x^{(0)}(t)$	Conventional model	GM(2, 1)	Grey model proposed by Cheng and Liu <sup>[3]</sup>	
			Simulation Value	Relative Error / %	Simulation Value	Relative Error / %
2007	1	9343.26	-	-	-	-
2008	2	10900.77	10900.77	$5.01 \times 10^{-14}$	10900.78	$2.35 \times 10^{-5}$
2009	3	11764.41	12192.05	3.64	12975.05	10.3
2010	4	14425.92	13641.85	5.44	14983.62	3.87
2011	5	17803.98	15271.95	14.2	17003.06	4.5
2012	6	19302.62	17108.07	11.4	19071.5	1.2
2013	7	22096.39	19180.89	13.2	21219.42	3.97
2014	8	23986.7	21527.5	10.3	23480.16	2.11
2015	9	25178.55	24193.3	3.91	25896.38	2.85
2016	10	26931.01	27234.69	1.13	28526.41	5.92
2017	11	31452.06	30722.76	2.32	31452.17	$3.53 \times 10^{-4}$
2018	12	35866.3	34748.45	3.12	34789.97	3.0
			Prediction Value	Relative Error / %	Prediction Value	Relative Error / %
2019	13	38999.04	39429.73	1.1	38705.7	0.752
2020	14	41858.38	44921.72	7.32	43436.68	3.77
2021	15	46636.0	51430.97	10.3	49323.01	5.76
Average Simulation Relative Error (2007–2018)			-	6.23	-	3.43
Average Prediction Relative Error (2019–2021)			-	6.23	-	3.42
Average Overall Relative Error (2007–2021)			-	6.23	-	3.43

Then, we build a model using the extended grey GM(1,1) power model and modeling method proposed by Cheng and Liu<sup>[3]</sup> and get the parameter estimates:

$$(\alpha, a, b_0, b_1, b_2) = (0.1244701, -0.3914404, 2647.282, -622.7652, -56.80383). \quad (38)$$

In this case, the time response equation is

$$\begin{aligned} \hat{x}^{(1)}(t) &= \left\{ e^{-a(1-\alpha)(t-1)} \left[ (x^{(1)}(1))^{(1-\alpha)} + (1-\alpha)e^{-a(1-\alpha)}g(t) \right] \right\}^{\frac{1}{1-\alpha}} \\ &= \left\{ e^{0.3427(t-1)} [2994.2631 + 1.2334g(t)] \right\}^{1.1422}, \end{aligned} \quad (39)$$

where

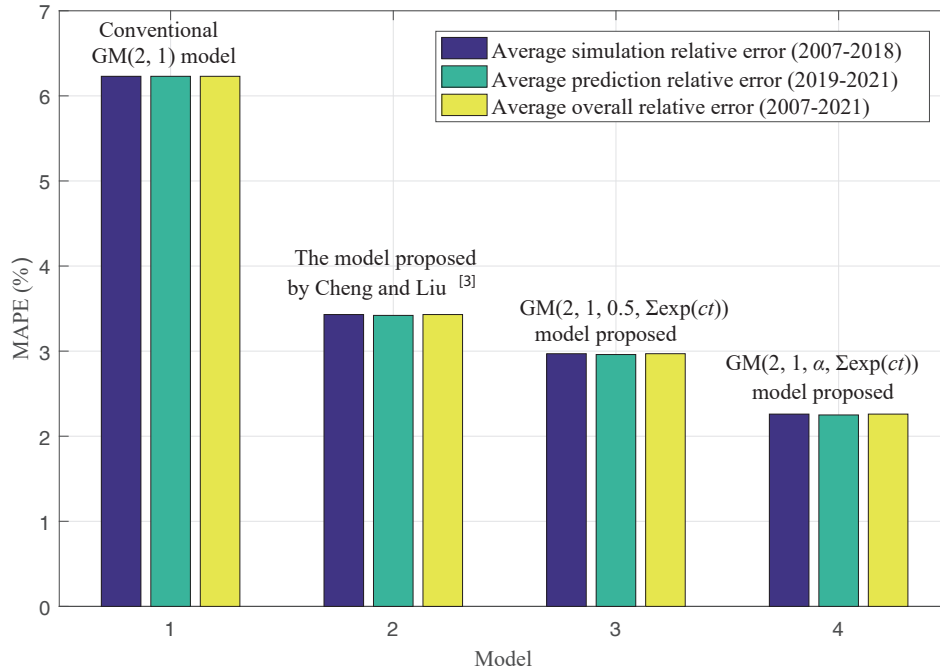
$$\begin{aligned}
g(t) &= \frac{b_2 t^2 e^{a(1-\alpha)t}}{a(1-\alpha)} - \frac{2b_2 t e^{a(1-\alpha)t}}{a^2(1-\alpha)^2} + \frac{b_1 t e^{a(1-\alpha)t}}{a(1-\alpha)} + \frac{2b_2 e^{a(1-\alpha)t}}{a^3(1-\alpha)^3} - \frac{b_1 e^{a(1-\alpha)t}}{a^2(1-\alpha)^2} + \frac{b_0 e^{a(1-\alpha)t}}{a(1-\alpha)} - \\
&\quad \frac{2b_2 e^{a(1-\alpha)t}}{a^3(1-\alpha)^3} + \frac{2b_2 e^{a(1-\alpha)t}}{a^2(1-\alpha)^2} - \frac{b_2 e^{a(1-\alpha)t}}{a(1-\alpha)} + \frac{b_1 e^{a(1-\alpha)t}}{a^2(1-\alpha)^2} - \frac{e^{a(1-\alpha)t}}{a(1-\alpha)} - \frac{b_0 e^{a(1-\alpha)t}}{a(1-\alpha)} \\
&= 165.7452t^2 e^{-0.3427t} + 2784.3774t e^{-0.3427t} + 400.0235e^{-0.3427t} - 2378.0627.
\end{aligned} \tag{40}$$

With  $\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$ , we calculate and get the original sequence's simulation and prediction values which are shown in Table 2. Table 2 gives the relative errors and average relative errors in the periods.

Table 2 shows related calculating results of grey models for China's natural gas consumption built with the other methods.

Figure 1 is the histogram of average simulation relative errors, average prediction relative errors and average overall relative errors of four models.

We can see from Table 1, Table 2 and Figure 1 that the models built with the method proposed have high simulation and prediction precision which is much higher than that of conventional grey GM(2, 1) model and superior to that of the grey power model proposed by Cheng and Liu<sup>[3]</sup>. It indicates the extended model and method proposed have high reliability and effectiveness.



**Figure 1** Histogram of average relative errors of four models

## 5.2 Example 2

To verify the effectiveness and practicability of the novel model proposed, the paper chooses the natural gas consumption of Chongqing City, China, from year 2001 to year 2016 in [8] as

the original data to build the  $GM(2, 1, \alpha, \sum \exp(ct))$  model. We write natural gas consumption as in the unit of 0.1 billion cubic meters. Table 3 shows the data divided into two parts in which the data from year 2001 to year 2012 are used for the modeling and the data from year 2013 to year 2016 are used for predictions.

**Table 3** Related calculation results of grey modeling for natural gas consumption of Chongqing city

Year	No.	$x^{(0)}(t)$	GM(2, 1, $\alpha, \sum \exp(ct)$ ) Model		Grey GM(2, 1) Model Proposed by Zeng and Luo <sup>[8]</sup>		Grey GM(2, 1) Model Proposed by Cheng and Shi <sup>[22]</sup>	
			Simulation Value	Relative Error %	Simulation Value	Relative Error %	Simulation Value	Relative Error %
2001	1	322.51	-	-	-	-	-	-
2002	2	331.87	331.87	1.82e-10	336.16	1.29	331.87	4.25e-37
2003	3	349.11	343.23	1.68	355.96	1.96	374.843	7.37
2004	4	403.52	425.928	5.55	408.31	1.19	416.283	3.16
2005	5	472.15	485.01	2.72	459.5	2.68	460.721	2.42
2006	6	532.67	532.637	0.00611	514.41	3.43	509.543	4.34
2007	7	578.95	577.741	0.209	572.15	1.17	563.456	2.68
2008	8	648.38	626.778	3.33	633.31	2.32	623.056	3.91
2009	9	657.91	684.179	3.99	698.23	6.13	688.954	4.72
2010	10	752.49	752.511	0.00284	767.37	1.98	761.822	1.24
2011	11	821.82	832.359	1.28	841.15	2.35	842.397	2.5
2012	12	943.86	921.88	2.33	920.04	2.52	931.494	1.31
			Prediction Value	Relative Error %	Prediction Value	Relative Error %	Prediction Value	Relative Error %
2013	13	959.97	1015.98	5.83	1004.52	4.64	1030.01	7.3
2014	14	1092.25	1104.99	1.17	1095.08	0.26	1138.95	4.28
2015	15	1175.32	1172.68	0.225	1192.26	1.44	1259.42	7.16
2016	16	1187.96	1193.33	0.452	1296.62	9.15	1392.62	17.2
Average Simulation Error (2001–2012)			-	1.92	-	2.25	-	3.06
Average Prediction Error (2013–2016)			-	1.92	-	3.87	-	8.98
Average Overall Relative Error (2001–2016)			-	1.92	-	2.68	-	4.64

The paper first builds the extended grey  $GM(2, 1, \alpha, \sum \exp(ct))$  model proposed, i.e., the following model

$$\frac{d^2 x^{(r)}}{dt^2} + a_1 \frac{dx^{(r)}}{dt} + a_2 x^{(r)} = b_0 + \sum_{i=1}^p f_i e^{c_i t}. \quad (41)$$

Similarly, we choose  $p=2$ . Using the method proposed, we calculate and get the following parameter estimates

$$(\alpha, c_1, c_2, v, g, r) = (0.3695, 0.1670, 0.0970, -5.5287, 4.9427, 1.1409), \quad (42)$$

$$(a_1, a_2, b_0, f_1, f_2) = (21.07227, 6.124531, 22122.44, 0.01919518, 0.01919509). \quad (43)$$

In this way, we get the time response equation

$$\begin{aligned} \hat{x}^{(r)}(t) = & 0.002029439e^{0.166998t} - 51.98184e^{-1.421085e-16t} - \\ & 2.704999e11 \cdot e^{-20.7775t} - 581.5639e^{-0.2947674t} - 0.00004474394e^{0.1670054t} + \\ & 0.002392029e^{0.09699935t} - 0.00004489377e^{0.09700675t} + 3664.085e^{-0.0000074t}. \end{aligned} \quad (44)$$

With  $\hat{x}^{(0)}(k) = (\hat{x}^{(r)}(k) - \hat{x}^{(r)}(k-1))(v + gr^k)$ , we calculate and get the simulation and prediction values of original sequence in the periods. Table 3 shows the relative errors and average relative errors in the periods.

Next, we compare the model with the improved grey GM(2, 1) models proposed in other documents in terms of precision. First, we build the grey FDGM(2, 1) model proposed in [8] and get the following time response equation

$$\hat{x}^{(0.64)}(k) = 2236.086989 \times 1.078417^k + 130.502500 \times (-0.198018)^k - 2063.081509. \quad (45)$$

In this case, according to the reduction formula, we get the reducing value  $\hat{x}^{(0)}(k) = \sum_{i=1}^k C_{-0.64+k-j-1}^{k-j} \hat{x}^{(0.64)}(j)$  and get the simulation and prediction values of original sequence in the periods which are shown in Table 3. Table 3 gives the relative errors and average relative errors in the periods.

Then, we build the grey DEDGM(2, 1) model proposed in [22] and get the following time response equation

$$\hat{x}^{(1)}(k) = 2965.3028 \frac{z_2 \cdot z_1^k - z_1 \cdot z_2^k}{z_2 - z_1} + 3234.21 \frac{z_2^k - z_1^k}{z_2 - z_1} - 2911.6751, \quad (46)$$

where  $z_1(k) = 0.22804147$ ,  $z_2(k) = 1.1057655265$ .

With  $\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$ , we calculate the simulation and prediction values of original sequence in the periods which are shown in Table 3. See Table 3 for the relative errors and average relative errors in the periods.

## 6 Quality Evaluation on Models

### 6.1 Robustness Evaluation

The robustness evaluation refers to a statistical evaluation on the stability of the model built. We suppose that, for the model, the parameter estimate is  $\eta$ , the residua sum of squares is  $s$ , the sub-sample variance is  $m_s^2$ , the degree of freedom is  $n_s$  and the population variance is  $\delta_s^2$ . We give  $\eta$  a small perturbation  $\Delta$  and record  $Q = \eta + \Delta$ . We consider  $Q$  as the parameter value of model to get the fitting residual of  $r$ , and consider corresponding sub-sample variance as  $m_r^2$ , freedom degree as  $n_r$  and population variance as  $\delta_r^2$  to construct the statistics

$$F = \frac{\delta_r^2}{\delta_s^2} \cdot \frac{m_s^2}{m_r^2}, \quad (47)$$

$F \sim F(n_s - 1, n_r - 1)$ . We suppose  $H_0 : \delta_r^2 = \delta_s^2$ ;  $H_1 : \delta_r^2 \neq \delta_s^2$ . If it is on the confidence level of  $\alpha$ ,

$$\frac{m_s^2}{m_r^2} \leq F_{\frac{\alpha}{2}}(n_s - 1, n_r - 1), \quad (48)$$

$H_0$  is accepted while  $H_1$  is rejected, i.e., the model is considered as a robust model on the confidence level of  $1 - \alpha$ .

For the grey GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model built in Example 1, the paper gives  $\eta$  a small perturbation  $\Delta = 1\%$ , and calculates and gets  $\frac{m_s^2}{m_r^2} = 1.0531$ . Because  $F_{\frac{\alpha}{2}}(n_s - 1, n_r - 1) = F_{0.01}(11, 11) = 4.50$ ,  $\frac{m_s^2}{m_r^2} \leq F_{\frac{\alpha}{2}}(n_s - 1, n_r - 1)$ . Therefore, the model is robust.

For the grey GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model built in Example 2, the paper gives  $\eta$  a small perturbation  $\Delta = 1\%$ , and calculates and gets  $\frac{m_s^2}{m_r^2} = 4.2651$ . Because  $F_{\frac{\alpha}{2}}(n_s - 1, n_r - 1) = F_{0.01}(11, 11) = 4.50$ ,  $\frac{m_s^2}{m_r^2} \leq F_{\frac{\alpha}{2}}(n_s - 1, n_r - 1)$ . Therefore, the model is robust.

## 6.2 Precision Evaluation

The precision evaluation is an evaluation on the fitting accuracy of model. It generally analyzes the practical fitting results and calculates some statistics, such as the model's coefficient of determination  $R^2 = 1 - \frac{\sum (Y_t - \hat{Y}_t)^2}{\sum (Y_t - \bar{Y})^2}$ , mean absolute relative error  $\text{MAPE} = \frac{1}{N} \sum_{t=1}^N |\frac{Y_t - \hat{Y}_t}{Y_t}|$ , Theil index of inequality  $\mu = \frac{\sqrt{\frac{1}{n} \sum (Y - \hat{Y})^2}}{\sqrt{\frac{1}{n} \sum Y^2 + \frac{1}{n} \sum \bar{Y}^2}}$ , and so on.

For the grey GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model built for Example 1, the paper calculates and gets the coefficient of determination  $R^2 = 0.9893$  close to 1, the mean absolute relative error  $\text{MAPE} = 2.26\%$  less than 3%, and the Theil index of inequality  $\mu = 0.0281$  close to 0, so the model has high precision.

For the grey Riccati model built for Example 2, the paper calculates and gets the coefficient of determination  $R^2 = 0.9952$  close to 1, the mean absolute relative error  $\text{MAPE} = 1.92\%$  less than 3%, and the Theil index of inequality  $\mu = 0.0177$  close to 0, so the model has high precision.

## 7 Results and Discussions

In Example 1, the paper builds four grey models for China's natural gas consumption. We can see that the GM(2, 1, 0.5,  $\sum \exp(ct)$ ) model proposed, the GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model proposed, the conventional GM(2, 1) model, and the extended grey power model proposed in [3] have the average simulation relative errors of 2.97%, 2.26%, 6.23% and 3.43%, respectively. In the models, the GM(2, 1, 0.5,  $\sum \exp(ct)$ ) model proposed and the GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model proposed both have the average simulation relative errors less than 3%, with high precision; the GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model has the highest simulation precision; the conventional GM(2, 1) model has the poorest simulation precision. It indicates the improved method proposed has the precision which is significantly improved compared with that of the conventional method and is superior to that of related models.

In Example 1, the GM(2, 1, 0.5,  $\sum \exp(ct)$ ) model proposed, the GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model proposed, the conventional GM(2, 1) model and the extended grey power model proposed in [3] have the average prediction relative errors of 2.96%, 2.25%, 6.23% and 3.42%, respectively.



The GM(2, 1, 0.5,  $\sum \exp(ct)$ ) model and the GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model both have the average prediction relative error less than 3% with high prediction precision. The GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model have the smallest prediction error.

From the perspective of overall relative error, the GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model proposed has the highest precision and the average overall relative error of only 2.26%, followed by the GM(2, 1, 0.5,  $\sum \exp(ct)$ ) model (2.97%), the extended grey power model proposed in [3] (3.43%) and the conventional GM(2, 1) model (6.23%). It indicates the improved method proposed can improve the model's precision compared with the conventional method. The extended grey power model proposed in [3] also has high precision but its precision is lower than that of the GM(2, 1, 0.5,  $\sum \exp(ct)$ ) model proposed and the GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model proposed.

In Example 2, the paper builds three grey models for the natural gas consumption of Chongqing City, China. From the perspective of average simulation relative error, we can see the GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model has the highest simulation precision. In terms of average simulation relative error, the results ranked in ascending order are as follows: The GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model (1.92%), the grey GM(2, 1) model proposed in [8] (2.25%) and the grey GM(2, 1) model proposed in [22] (3.06%). From the perspective of average prediction relative error, we can see the GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model also has the highest precision. The results ranked in ascending order are as follows: The GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model (1.92%), the grey GM(2, 1) model proposed in [8] (3.87%), and the grey GM(2, 1) model proposed in [22] (8.98%). In this case, from the perspective of average overall relative error, we can see that the GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model also has the highest precision with the average overall relative error of only 1.92%. For the other models, the results ranked in ascending order are as follows: The grey GM(2, 1) model proposed in [8] (2.68%) and the grey GM(2, 1) model proposed in [22] (4.64%). The results indicate the models built with the improved method proposed have high precision and are superior to related models.

Example 1 and Example 2 show that the model and method proposed have high reliability and effectiveness. The main reason is the GM(2, 1,  $\alpha$ ,  $\sum \exp(ct)$ ) model has high adaptability to data and can adapt to the variation of various data, especially the fast growing data. Next, the paper makes a data transformation for the accumulated generating sequence of original sequence to make the data better adapt to the model's structure. In addition, the paper uses an optimization method for parameter estimation, thus greatly improving the model's precision.

## 8 Conclusions

The conventional grey GM(2, 1) model built for the fast growing sequence has big errors, so the paper proposes an improved method to build the grey GM(2, 1) model. The paper makes improvements from the following two aspects: First, the paper improves the accumulated generating sequence of original time sequence, i.e., making a new transformation; next, the paper extends the structure of conventional grey GM(2, 1) model, i.e., building the GM(2, 1,  $\sum \exp(ct)$ ) model. Theoretically, the new transformation for accumulated generating sequence is the extension of conventional accumulated generating sequence. For the same model, the improved method enhances data's adaptability to the model and thus improves modeling precision significantly. In addition, the newly built GM(2, 1,  $\sum \exp(ct)$ ) model is the extension of conventional

grey GM(2, 1) model. It extends the grey action into a superposed exponential function to get the better adaptability to data. In this way, the model built for fast growing sequence may have higher precision which is significantly improved compared with that of the conventional model.

As for the modeling method, the paper optimizes the parameters in the new transformed of accumulated generating sequence and the parameters in the model simultaneously and the optimization objective is taking the minimum of  $\max(\text{MAPE}_1, \text{MAPE}_2)$ , in which case the simulation and prediction errors are both small.

As for the practical application, the paper builds a GM(2, 1,  $\sum \exp(ct)$ ) model for China's natural gas consumption in Example 1. The model has high precision superior to that of conventional GM(2, 1) model and other models in the comparisons. To further prove the effectiveness of the method proposed, the paper builds a GM(2, 1,  $\sum \exp(ct)$ ) model for the natural gas consumption of Chongqing city, China, in Example 2. The model built also has high precision superior to that of two other improved GM(2, 1) models.

The theoretical and practical applications prove that the grey modeling method proposed has high effectiveness and reliability. The two improved methods can be generalized to the modeling of other grey models. The key in future research is finding a proper transformation form of accumulated generating sequence and the model forms, especially the extended form of grey action, for different time sequences.

## References

- [1] Tong Q. Weighted non-equal interval gray GM(1, 1) model based on function  $\cot(x^\alpha)$  transformation and its application. *Mathematics in Practice and Theory*, 2021, 51(13): 209–215.
- [2] Wang Z X, Zhao Y F. GM(1, 1) model with seasonal dummy variables and its application. *Systems Engineering — Theory & Practice*, 2020, 40(11): 2981–2990.
- [3] Cheng M L, Liu B. An extended grey GM(1, 1) power model and its application. *Journal of Statistics and Information*, 2021, 36(10): 3–11.
- [4] Luo Y H, Chen Y J. Linear time-varying parameter non-equidistant GM(1, 1) power model and its application. *Systems Engineering*, 2021, 39(5): 152–158.
- [5] Xie M, Wu L F. Short-term traffic flow prediction based on GM(1,  $N$ ) power model optimized by rough set algorithm. *Mathematics in Practice and Theory*, 2021, 51(9): 241–249.
- [6] Wang J F, Tian C M. Time delayed GM(1,  $N$ ) model for mixed-frequency data. *Operations Research and Management Science*, 2021, 30(12): 123–127.
- [7] Zhao X Q, Liu Y Z, Zhang L W. The forecasts of temperature and humidity of the barn based on the improved GM(2, 1). *Journal of the Chinese Cereals and Oils Association*, 2012, 27(8): 85–87.
- [8] Zeng L, Luo S G. A discrete GM(2, 1) model with fractional-order accumulation and its application. *Journal of Chongqing Normal University (Natural Science)*, 2021, 38(5): 73–80.
- [9] Tang L W, Lu Y Y. The optimization of GM(2, 1) model based on parameter estimation of grade difference format. *Systems Engineering — Theory & Practice*, 2018, 38(2): 502–508.
- [10] Su H J, Shao Y. Optimization method of time response coefficient of GM(2, 1) model. *Statistics and Decision*, 2018, 34(8): 24–27.
- [11] Zhao X Q, Chen H L. Prediction formula's defect of GM(2, 1) and its improvement. *Journal of Wuhan University of Technology*, 2006, 28(10): 125–127.
- [12] Su H, Wei Y, Shao Y. On optimizing time response sequence of grey model GM(2, 1). *Journal of Grey System*, 2011, 23(2): 119–126.
- [13] An Q W. Two GM(2, 1) improved model to predict air quality index. 5th International Conference on Advanced Materials and Computer Science (ICAMCS 2016), 2016: 422–428.

- [14] Shen J H, Zhao X R. Improvement of GM(2, 1) model by minimum squares. Journal of Harbin Engineering University, 2001, 22(4): 64–66.
- [15] Zeng X Y, Xiao X P. Research on morbidity problem of accumulating method GM(2, 1) model. Systems Engineering and Electronics, 2006, 17(4): 542–544.
- [16] Xiao X P, Guo J. The morbidity problem of GM(2, 1) model based on vector transformation. Journal of Grey System, 2014, 26(3): 1–11.
- [17] Liu C L, Chen J, Qian J J, Sun Y L, Han X M. Optimum grey action quantity for GM(2, 1) model. Mathematics in Practice and Theory, 2017, 47(18): 177–184.
- [18] Shui N X, Dong T H, Sha Z. On the model of grey system GM(2, 1). Systems Engineering, 1990, 8(6): 31–35.
- [19] Yong H, Wei Y. The optimization of the non-equigap DGM(2, 1) model. Journal of Grey System, 2011, 14(1): 41–46.
- [20] Shao Y, Su H J. On approximating grey model DGM(2, 1). Journal of Grey System, 2012, 1(4): 8–13.
- [21] Xu N, Dang Y G. An optimized grey GM(2, 1) model and forecasting of highway subgrade settlement. Mathematical Problems in Engineering, 2015, 2015: 606707.
- [22] Cheng M L, Shi G J. Modeling and application of grey model GM(2, 1) based on linear difference equation. Journal of Grey System, 2019, 31(2): 37–50.
- [23] Liu S F. Grey system theory and its applications. 9th ed. Beijing: Science Press, 2021.