

# A New Method for Quarterly-Data Predictions Based on the Extended Grey Model GM(1, 1, $\exp \times \sin$ , $\exp \times \cos$ ) and Its Application in China's Quarterly GDP Prediction

Maolin CHENG\*

*School of Mathematical Sciences, Suzhou University of Science and Technology, Suzhou 215009, China*  
*E-mail: chengml53@qq.com*

Bin LIU

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

**Abstract** Grey prediction is vital in statistical prediction with wide applications. However, most grey prediction methods focus on annual predictions of the monotonic time series instead of the seasonal time series. The paper uses the extended model of the grey GM(1, 1) model to predict the seasonal time series. Some improvements have been made in two aspects to improve the prediction accuracy of the model. 1) We introduce seasonal multiple factors to transform the original time series, which improves the adaptability of the seasonal data to the model. The transformed series conforms to the law presented by the model. 2) The seasonal data are in superimposed sine and cosine fluctuations with tendencies. Therefore, the paper extends the grey action quantity of the traditional GM(1, 1) model. The newly extended grey model is called the GM(1, 1,  $\exp \times \sin$ ,  $\exp \times \cos$ ) model, which is provided with the parameter optimization methods and time response equations. According to the proposed modeling method, we establish a GM(1, 1,  $\exp \times \sin$ ,  $\exp \times \cos$ ) model for China's quarterly gross domestic product (GDP) with high accuracy.

**Keywords** GM(1, 1,  $\exp \times \sin$ ,  $\exp \times \cos$ ) model; parameter estimation; time-response equation; prediction accuracy

## 1 Introduction

The grey prediction is an important method of statistical prediction. Currently, the grey prediction has been used widely in the fields of industry, agriculture, commerce and economy<sup>[1,2]</sup> and other fields, such as environment, energy, society and military<sup>[3–7]</sup>. The grey prediction models used widely currently include the GM(1, 1) model<sup>[8,9]</sup>, the GM(1, 1) power model<sup>[10]</sup>, the GM(1,  $N$ ) model<sup>[11]</sup>, and the GM( $N$ , 1) model<sup>[12]</sup>. In the grey prediction models, the GM(1, 1) model is an important type, but it has big prediction errors sometimes and thus has some limitations in applications.

Generally, the traditional GM(1, 1) model is applicable to the prediction of monotonous time sequence, and many scholars have made related studies to improve the prediction precision

---

Received December 1, 2022, accepted January 31, 2023

Supported by National Natural Science Foundation of China (11401418)

\*The corresponding author

of GM(1, 1) model. Wang and Lu<sup>[13]</sup> used the Lagrange's mean value theorem to construct the background value considered as a variable, meanwhile set the initial value as a variable, and determined the minimum of average relative error with the time response equation to build a grey GM(1, 1) model. The example showed that the improved model was superior to other models. Chen and Zhu<sup>[14]</sup> constructed the background value using the continuous piecewise rational linear interpolation spline and built a novel GM(1, 1) model to offer a more rational formula for the calculation of background value. The novel GM(1, 1) model had higher effectiveness and accuracy in terms of data processing compared with the classical GM(1, 1) model. Mi, et al.<sup>[15]</sup> considered China's cross-border e-commerce as the classical research object, introduced the new information priority principle in the grey model and proposed an improved grey GM(1, 1) model to explore the object's dynamic variation law as a whole and predict the object's future development trend. To improve the fitting and prediction performance of unequal-interval GM(1, 1) model, Tang and Lu<sup>[16]</sup> analyzed the main factors of source of error of the model and, on the basis, proposed an improved unequal-interval GM(1, 1) model based on the grey derivative and accumulated generating. The models and methods above all improved the modeling precision of monotonous time sequence.

To use the GM(1, 1) model for the prediction of fluctuating sequence, especially the seasonal data, some scholars have improved the modeling method of GM(1, 1) model. Wang and Sun<sup>[17]</sup> proposed a new method based on the X12-GM(1, 1) combined model. The method first used the X12 model to get the data's long-term trend factor, periodic variation factor and seasonal factor, and then fit and predicted the long-term trend factor and periodic variation factor using the GM(1, 1) model and finally multiplied the fitting value and prediction value of GM(1, 1) by the seasonal factor to get the original sequence's fitting value and prediction value. The traditional time sequence prediction methods had some shortcomings in the future prediction due to modern finance's volatility and instability, so to solve the problem, reference [18] proposed an unbiased GM(1, 1) hybrid method based on the Taylor's series approximation to deal with the incomplete, noisy and uncertain data in the system. China's wind power generation shows an exponential growth trend and a seasonal fluctuating pattern, in which case the traditional model can't predict it accurately. However, the grey model GM(1, 1) can capture the exponential growth trend and the Hodrick-Prescott filter can deal with the seasonal factors. In this case, Qian and Wang<sup>[19]</sup> proposed a novel seasonal prediction method integrating the HP filter into the grey model GM(1, 1). Finally, they verified the effectiveness and practicability of improved unequal-interval GM(1, 1) model through examples. To solve the traditional grey model's poor performance in predicting oscillatory sequences common in the reality, such as the seasonal data predicting problem, Zeng<sup>[20]</sup> proposed a GM(1, 1|sin) model and a GM(1, 1|sin) power model, gave the parameter calculation formula with the least squares criterion, constructed a nonlinear optimization model with the average simulation relative error minimization as the object, got the optimal parameter using the particle swarm optimization algorithm, and finally applied the novel model to the simulation and prediction of urban traffic and new high-tech product export volume. Results showed that the novel models had higher simulation precision and were more adaptable to the prediction and analysis on oscillatory sequence. To solve the problem that some grey models failed to predict the small-sample oscillatory sequences, Wang<sup>[21]</sup> proposed

a small-sample oscillatory sequence grey prediction method based on the Fourier series. The method first built a GM(1, 1) power model for the original sequence to describe the general trend of system and fit the residual sequence using the Fourier series, then constructed a new time response function using the sum of the former two, and finally built a nonlinear optimization model with the average error minimization as the object to solve and get the optimal parameter. The application example proved that the method could improve the grey model's prediction precision for small-sample oscillatory sequence effectively. The models and method above all improved the modeling precision of fluctuating time sequence to some extent.

Currently the studies on GM(1, 1) model mainly focus on the background value, the grey derivative, the parameter optimization and the model's extrapolation. Most methods use the GM(1, 1) model to study the prediction of monotonous time sequence and have good results. However, some methods fail to deal with seasonal time sequence prediction problems. The paper explores the problem and uses the grey GM(1, 1) model's extended model to predict the seasonal time sequence. In fact, the following three main factors influencing the grey model's prediction precision: First, the model's structure should meet the variation law of time sequence; second, the time sequence should meet the law presented by the model; third, the modeling method used should be scientific and rational. To use the GM(1, 1) model to predict seasonal data and improve modeling precision, the paper makes improvements in the three aspects. First, the paper introduces a seasonal multiple factor to make a new transformation of original time sequence to increase the seasonal data's adaptability to the model. Next, considering the seasonal time sequence shows a superimposed sine-cosine fluctuation state with a trend variation, so the paper extends the traditional GM(1, 1) model's grey action. The newly extended grey model is called the GM(1, 1, exp $\times$ sin, exp $\times$ cos) model. To avoid big average simulation relative error (MAPE<sub>1</sub>) or big average prediction relative error (MAPE<sub>2</sub>) of the model, we define the objective function as the minimum of max(MAPE<sub>1</sub>, MAPE<sub>2</sub>), and get the model's parameters using an optimization method. And then, we make simulations and predictions using the time response equation derived. With the model and method proposed, the paper builds GM(1, 1, exp $\times$ sin, exp $\times$ cos) models for China's quarterly GDP. Results show that the models have high precision.

## 2 Establishment Method of the Traditional Grey GM(1, 1) Model

Assuming original time series  $x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$ , its first-order cumulative-generation series (1-AGO) is  $x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$ , where  $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)$ ,  $k = 1, 2, \dots, n$ .

**Definition 1**<sup>[22]</sup>  $z^{(1)} = \{z^{(1)}(1), z^{(1)}(2), \dots, z^{(1)}(n)\}$  is called the series of background values, where  $z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k) + x^{(1)}(k-1))$ ,  $k = 2, 3, \dots, n$ .

**Definition 2**<sup>[22]</sup>  $x^{(0)}(k) + az^{(1)}(k) = b$  is called the grey differential equation of the grey GM(1, 1) model.

**Definition 3**<sup>[22]</sup>  $\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b$  is called the whitening equation of the grey GM(1, 1) model.

**Theorem 1**<sup>[22]</sup> The parameter estimate of the grey GM(1,1) model is

$$\begin{pmatrix} a \\ b \end{pmatrix} = (B'B)^{-1}B'Y, \quad (1)$$

where

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

**Theorem 2**<sup>[22]</sup> The time response function of the grey GM(1,1) model is

$$\hat{x}^{(1)}(t) = \left(x^{(1)}(1) - \frac{b}{a}\right)e^{-a(t-1)} + \frac{b}{a}. \quad (3)$$

The simulated and predicted values of the original series can be obtained according to time response function series  $\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$ .

### 3 Forms of Extended Grey Model GM(1,1, exp×sin, exp×cos)

From the perspective of structure, the traditional grey GM (1,1) model has a time response equation in the form of corrected exponential curve, so it can achieve good prediction results for the time sequence presenting slow exponential curve variations. The model is generally adapted to the prediction of monotonous time sequence but has poor prediction results for the seasonal data in the fluctuating state. Therefore, to improve seasonal data's modeling precision, we must improve the original data sequence, the model's structure and the modeling method.

$x^{(0)}(t)$  is transformed to make the seasonal data more adaptable to the model. Seasonal multiple factors are introduced and can be divided into constant and linear factors.

**Definition 4** Let original time series  $x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$ ,  $D_i$  is called a constant seasonal multiple factor, i.e.,  $D_i = \begin{cases} d_1, i = 1, 5, \dots, 4m-3, \dots \\ d_2, i = 2, 6, \dots, 4m-2, \dots \\ d_3, i = 3, 7, \dots, 4m-1, \dots \\ d_4, i = 4, 8, \dots, 4m, \dots \end{cases} \quad (m = 1, 2, 3, \dots).$

New time series  $x^{(D)} = \{x^{(0)}(1)D_1, x^{(0)}(2)D_2, \dots, x^{(0)}(n)D_n\}$ , and new cumulative-generation series  $x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$ , where  $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)D_i$ , and  $k = 1, 2, \dots, n$ .

**Definition 5** Let original time series  $x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$ ,  $D_i$  is called a linear seasonal multiple factor, i.e.,  $D_i = \begin{cases} d_1 + h_1i, i = 1, 5, \dots, 4m-3, \dots \\ d_2 + h_2i, i = 2, 6, \dots, 4m-2, \dots \\ d_3 + h_3i, i = 3, 7, \dots, 4m-1, \dots \\ d_4 + h_4i, i = 4, 8, \dots, 4m, \dots \end{cases} \quad (m = 1, 2, 3, \dots; i = 1, 2, 3, \dots).$  New time series  $x^{(D)} = \{x^{(0)}(1)D_1, x^{(0)}(2)D_2, \dots, x^{(0)}(n)D_n\}$ , and new cumulative-

generation series  $x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$ , where  $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)D_i$ , and  $k = 1, 2, \dots, n$ .

The seasonal time series is generally in superimposed sine and cosine fluctuations with variable trends. The traditional gray model structure is extended to meet the needs of such modeling.

**Definition 6** Supposing original time series  $x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$ , new cumulative-generation series  $x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$ , and new background value  $z^{(1)}(k) = \int_{k-1}^k x^{(1)}(t)dt$ . Then  $x^{(1)}(k) + a_1 z^{(1)}(k) = c_0 + \sum_{i=1}^p e^{b_i k} (c_i \sin(s_i k) + f_i \cos(s_i k))$  is called the grey differential equation of extended grey model GM(1,1, exp×sin, exp×cos).

**Definition 7**  $\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = c_0 + \sum_{i=1}^p e^{b_i t} (c_i \sin(s_i t) + f_i \cos(s_i t))$  is called the whitening equation of extended grey model GM(1,1, exp×sin, exp×cos).

In particular, it becomes the whitening equation of the traditional grey GM(1,1) model when  $c_i = 0, f_i = 0$ .

#### 4 Time Response Equation of Extended Grey Model GM(1,1, exp×sin, exp×cos)

**Theorem 3** When the series is recorded as  $x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$  and the whitening equation of extended grey model GM(1,1, exp×sin, exp×cos) is  $\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = c_0 + \sum_{i=1}^p e^{b_i t} (c_i \sin(s_i t) + f_i \cos(s_i t))$ , then its time response equation is

$$\begin{aligned} x^{(1)}(t) = e^{-a(t-1)} & \left\{ x^{(1)}(1) + \frac{c_0(e^{a(t-1)} - 1)}{a} \right. \\ & + \sum_{i=1}^p \frac{c_i e^{-a} e^{(a+b_i)t} [(a+b_i) \sin(s_i t) - s_i \cos(s_i t)]}{(a+b_i)^2 + s_i^2} \\ & - \sum_{i=1}^p \frac{c_i e^{-a} e^{(a+b_i)t} [(a+b_i) \sin(s_i) - s_i \cos(s_i)]}{(a+b_i)^2 + s_i^2} \\ & + \sum_{i=1}^p \frac{f_i e^{-a} e^{(a+b_i)t} [(a+b_i) \cos(s_i t) + s_i \sin(s_i t)]}{(a+b_i)^2 + s_i^2} \\ & \left. - \sum_{i=1}^p \frac{f_i e^{-a} e^{(a+b_i)t} [(a+b_i) \cos(s_i) + s_i \sin(s_i)]}{(a+b_i)^2 + s_i^2} \right\}. \end{aligned} \quad (4)$$

*Proof*

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = c_0 + \sum_{i=1}^p e^{b_i t} (c_i \sin(s_i t) + f_i \cos(s_i t)), \quad (5)$$

i.e.

$$\frac{dx^{(1)}(t)}{dt} = -ax^{(1)}(t) + c_0 + \sum_{i=1}^p e^{b_i t} (c_i \sin(s_i t) + f_i \cos(s_i t)), \quad (6)$$

its solution is

$$x^{(1)}(t) = e^{\int_1^t -ad\delta} \left[ x^{(1)}(1) + \int_1^t e^{\int_1^\theta ad\delta} \left[ c_0 + \sum_{i=1}^p e^{b_i \theta} (c_i \sin(s_i \theta) + f_i \cos(s_i \theta)) \right] d\theta \right]$$

$$\begin{aligned}
&= e^{-a(t-1)} \left[ x^{(1)}(1) + \int_1^t c_0 e^{a(\theta-1)} d\theta + e^{-a} \sum_{i=1}^p c_i \int_1^t e^{(b_i+a)\theta} \sin(s_i \theta) d\theta \right. \\
&\quad \left. + e^{-a} \sum_{i=1}^p f_i \int_1^t e^{(b_i+a)\theta} \cos(s_i \theta) d\theta \right] \\
&= e^{-a(t-1)} \left\{ x^{(1)}(1) + \frac{c_0(e^{a(t-1)} - 1)}{a} + \sum_{i=1}^p \frac{c_i e^{-a} e^{(a+b_i)t} [(a+b_i) \sin(s_i t) - s_i \cos(s_i t)]}{(a+b_i)^2 + s_i^2} \right. \\
&\quad - \sum_{i=1}^p \frac{c_i e^{-a} e^{(a+b_i)t} [(a+b_i) \sin(s_i) - s_i \cos(s_i)]}{(a+b_i)^2 + s_i^2} \\
&\quad + \sum_{i=1}^p \frac{f_i e^{-a} e^{(a+b_i)t} [(a+b_i) \cos(s_i t) + s_i \sin(s_i t)]}{(a+b_i)^2 + s_i^2} \\
&\quad \left. - \sum_{i=1}^p \frac{f_i e^{-a} e^{(a+b_i)t} [(a+b_i) \cos(s_i) + s_i \sin(s_i)]}{(a+b_i)^2 + s_i^2} \right\}. \tag{7}
\end{aligned}$$

## 5 Parameter Optimization Method of Extended Grey Model GM(1,1, exp×sin, exp×cos)

**Theorem 4** Let the grey differential equation of extended grey model GM(1, 1, exp×sin, exp×cos) be  $x^{(1)}(k) + a_1 z^{(1)}(k) = c_0 + \sum_{i=1}^p e^{b_i k} (c_i \sin(s_i k) + f_i \cos(s_i k))$ .

If the seasonal multiple factors and  $\alpha, b_i, s_i$  are known, then parameter estimate is

$$\hat{B} = \begin{pmatrix} a_1 \\ c_0 \\ c_1 \\ f_1 \\ \vdots \\ c_p \\ f_p \end{pmatrix} = (X'X)^{-1}X'Y, \tag{8}$$

where

$$X = \begin{pmatrix} -\alpha \sum_{i=1}^1 x^{(0)}(i)D_i - (1-\alpha) \sum_{i=1}^2 x^{(0)}(i)D_i & 1 & z_1^{(1)}(2) & z_1^{(2)}(2) & \cdots & z_p^{(1)}(2) & z_p^{(2)}(2) \\ -\alpha \sum_{i=1}^2 x^{(0)}(i)D_i - (1-\alpha) \sum_{i=1}^3 x^{(0)}(i)D_i & 1 & z_1^{(1)}(3) & z_1^{(2)}(2) & \cdots & z_p^{(1)}(3) & z_p^{(2)}(3) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -\alpha \sum_{i=1}^{n-1} x^{(0)}(i)D_i - (1-\alpha) \sum_{i=1}^n x^{(0)}(i)D_i & 1 & z_1^{(1)}(n) & z_1^{(2)}(2) & \cdots & z_p^{(1)}(n) & z_p^{(2)}(n) \end{pmatrix}, \tag{9}$$

and

$$Y = \begin{pmatrix} x^{(0)}(2)D_2 \\ x^{(0)}(3)D_3 \\ \vdots \\ x^{(0)}(n)D_n \end{pmatrix}, \tag{10}$$

$$z_i^{(1)}(k) = \frac{e^{b_i k}(b_i \sin(s_i k) - s_i \cos(s_i k))}{b_i^2 + s_i^2} - \frac{e^{b_i(k-1)}(b_i \sin(s_i(k-1)) - s_i \cos(s_i(k-1)))}{b_i^2 + s_i^2}, \quad (11)$$

and

$$z_i^{(2)}(k) = \frac{e^{b_i k}(b_i \cos(s_i k) + s_i \sin(s_i k))}{b_i^2 + s_i^2} - \frac{e^{b_i(k-1)}(b_i \cos(s_i(k-1)) + s_i \sin(s_i(k-1)))}{b_i^2 + s_i^2}. \quad (12)$$

*Proof* The whitening equation of extended grey model GM(1, 1, exp  $\times$  sin, exp  $\times$  cos) is

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = c_0 + \sum_{i=1}^p e^{b_i t}(c_i \sin(s_i t) + f_i \cos(s_i t)), \quad (13)$$

$$\frac{dx^{(1)}(t)}{dt} = -ax^{(1)}(t) + c_0 + \sum_{i=1}^p e^{b_i t}(c_i \sin(s_i t) + f_i \cos(s_i t)), \quad (14)$$

i.e.,

$$\int_{k-1}^k \frac{dx^{(1)}(t)}{dt} dt = -a \int_{k-1}^k x^{(1)}(t) dt + c_0 + \sum_{i=1}^p c_i \int_{k-1}^k e^{b_i t} \sin(s_i t) dt + \sum_{i=1}^p f_i \int_{k-1}^k e^{b_i t} \cos(s_i t) dt. \quad (15)$$

When the background value is recorded as  $z^{(1)}(k) = \int_{k-1}^k x^{(1)}(t) dt$ ,

$$\begin{aligned} z_i^{(1)}(k) &= \int_{k-1}^k e^{b_i t} \sin(s_i t) dt \\ &= \frac{e^{b_i k}(b_i \sin(s_i k) - s_i \cos(s_i k))}{b_i^2 + s_i^2} - \frac{e^{b_i(k-1)}(b_i \sin(s_i(k-1)) - s_i \cos(s_i(k-1)))}{b_i^2 + s_i^2}, \end{aligned} \quad (16)$$

and

$$\begin{aligned} z_i^{(2)}(k) &= \int_{k-1}^k e^{b_i t} \cos(s_i t) dt \\ &= \frac{e^{b_i k}(b_i \cos(s_i k) + s_i \sin(s_i k))}{b_i^2 + s_i^2} - \frac{e^{b_i(k-1)}(b_i \cos(s_i(k-1)) + s_i \sin(s_i(k-1)))}{b_i^2 + s_i^2}. \end{aligned} \quad (17)$$

Then

$$x^{(1)}(k) - x^{(1)}(k-1) = -az^{(1)}(k) + c_0 + c_1 z_1^{(1)}(k) + d_1 z_2^{(1)}(k) + \cdots + c_p z_p^{(1)}(k) + d_p z_p^{(1)}(k), \quad (18)$$

i.e.

$$x^{(0)}(k) D_k = -az^{(1)}(k) + c_0 + c_1 z_1^{(1)}(k) + d_1 z_2^{(1)}(k) + \cdots + c_p z_p^{(1)}(k) + d_p z_p^{(1)}(k). \quad (19)$$

Since  $z^{(1)}(k) \approx \alpha_k x^{(1)}(k-1) + (1-\alpha_k)x^{(1)}(k) = \alpha_k \sum_{i=1}^{k-1} x^{(0)}(i) D_i + (1-\alpha_k) \sum_{i=1}^k x^{(0)}(i) D_i$ , Theorem 4 is derived from the law of least squares.

We can calculate the simulated and predicted values of  $x^{(0)}$  by Theorem 5 below after estimating parameters.

**Theorem 5** *The simulated value of  $x^{(0)}$  is*

$$\hat{x}^{(0)}(k) = (\hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1))/D_k, \quad k = 2, 3, \dots, m. \quad (20)$$

The predicted value is

$$\hat{x}^{(0)}(k) = (\hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1))/D_k, \quad k = m+1, 3, \dots, n. \quad (21)$$

*Proof* When the original time series is recorded as  $x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$ , new time series  $x^{(D)} = \{x^{(0)}(1)D_1, x^{(0)}(2)D_2, \dots, x^{(0)}(n)D_n\}$ , and new cumulative-generation series  $x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$ .

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)D_i \quad (22)$$

and

$$\hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) = \sum_{i=1}^k x^{(0)}(i)D_i - \sum_{i=1}^{k-1} x^{(0)}(i)D_i = x^{(0)}(k)D_k, \quad (23)$$

so

$$\hat{x}^{(0)}(k) = \frac{(\hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1))}{D_k}, \quad k = 2, 3, \dots, n. \quad (24)$$

We can determine seasonal multiple factors and  $\alpha, b_i, s_i$  in Theorem 4 by optimization. Supposing there are  $n$  years of observations, the observations from the 1<sup>st</sup> to  $m^{\text{th}}$  Years are used for modeling, and those from  $m+1^{\text{th}}$  to  $n^{\text{th}}$  years are used for predictions. Simulated and predicted values are recorded as  $\hat{x}^{(0)}(k)$  ( $k = 2, 3, \dots, m$ ) and  $\hat{x}^{(0)}(k)$  ( $k = m+1, 3, \dots, n$ ), respectively; its average-simulation relative error  $\text{MAPE}_1 = \frac{1}{m-1} \sum_{k=2}^m \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%$ ; average-prediction relative error  $\text{MAPE}_2 = \frac{1}{n-m} \sum_{k=m+1}^n \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%$ . The objective function is defined as the minimum value of  $\max(\text{MAPE}_1, \text{MAPE}_2)$  to reduce simulation and prediction errors.

The optimization problem is as follows:

If the seasonal factor is constant, the following objective function exists.

$$\min_{d_1, d_2, d_3, d_4, \alpha, b_i, s_i} \text{MAPE} = \max(\text{MAPE}_1, \text{MAPE}_2), \quad (25)$$

If the seasonal factor is linear, the following objective function exists.

$$\min_{d_1, d_2, d_3, d_4, h_1, h_2, h_3, h_4, \alpha, b_i, s_i} \text{MAPE} = \max(\text{MAPE}_1, \text{MAPE}_2), \quad (26)$$



Constraints are as follows:

$$\text{s.t.} \left\{ \begin{array}{l} \text{MAPE}_1 = \frac{1}{m-1} \sum_{k=2}^m \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%, \\ \text{MAPE}_2 = \frac{1}{n-m} \sum_{k=m+1}^n \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%, \\ \hat{x}^{(0)}(t) = (\hat{x}^{(1)}(t) - \hat{x}^{(1)}(t-1))/D_t, \\ \hat{x}^{(1)}(t) = e^{-a(t-1)} \left\{ x^{(1)}(1) + \frac{c_0(e^{a(t-1)} - 1)}{a} \right. \\ \quad + \sum_{i=1}^p \frac{c_i e^{-a} e^{(a+b_i)t} [(a+b_i) \sin(s_i t) - s_i \cos(s_i t)]}{(a+b_i)^2 + s_i^2} \\ \quad - \sum_{i=1}^p \frac{c_i e^{-a} e^{(a+b_i)t} [(a+b_i) \sin(s_i) - s_i \cos(s_i)]}{(a+b_i)^2 + s_i^2} \\ \quad + \sum_{i=1}^p \frac{f_i e^{-a} e^{(a+b_i)t} [(a+b_i) \cos(s_i t) + s_i \sin(s_i t)]}{(a+b_i)^2 + s_i^2} \\ \quad \left. - \sum_{i=1}^p \frac{f_i e^{-a} e^{(a+b_i)t} [(a+b_i) \cos(s_i) + s_i \sin(s_i)]}{(a+b_i)^2 + s_i^2} \right\}, \\ (a, c_0, c_1, d_1, \dots, c_p, d_p)^T = (X'X)^{-1} X'Y, \\ 0 \leq b_i \leq 1, d_1 > 0, d_2 > 0, d_3 > 0, d_4 > 0. \end{array} \right. \quad (27)$$

It is a nonlinear optimization problem, which can be solved by MATLAB software.

## 6 Grey Modeling of China's Quarterly GDP

GDP is the final result of the production activities of all resident units in a country (or region) within a certain period. It is the core index of national economic accounting and an important index to measure the economic status and development level of a country or region.

The National Bureau of Statistics announced that China's GDP in 2019 was 99.0865 trillion yuan on January 17, 2020, ranking second in the world; per capita GDP reached a new level of 10,000 US dollars for the first time. According to the released data on January 17, 2022, China's GDP in 2021 was 114.367 trillion yuan. It was up 8.1% from the previous year and a two-year average of 5.1% in constant prices. However, how does GDP grow in quarters? The changes in quarterly GDP show a volatile growth and are difficult to be predicted accurately; therefore, it is of great significance to study the method of quarterly GDP prediction. The paper uses the proposed method to build a grey prediction model for China's quarterly GDP, which is recorded as  $x^{(0)}(t)$  with the unit of yuan. The data are from the *China Statistical Yearbook* (see Table 1 for the specific values).

If the seasonal factor is constant, the GM(1,1,exp×sin, exp×cos) model is established below.

$$\begin{aligned} & (d_1, d_2, d_3, d_4, \alpha, b_1, s_1, b_2, s_2) \\ & = (1.0609, 1.0202, 1.0004, 0.9577, 0.9974, -3.1213, 0.4113, -0.0821, -0.2763), \end{aligned} \quad (28)$$

and

$$(a_1, c_0, c_1, f_1, c_2, f_2) \\ = (-0.024359305, 187024.56, 13109133.0, 0.014891823, -27408.592, 0.014891778). \quad (29)$$

Then the time response equation is

$$\hat{x}^{(1)}(t) = 7855799.0e^{0.0243593t} - 86378.76e^{-0.0821037t} \cos(0.276282t) \\ - 33285.29e^{-0.0821037t} \sin(0.276282t) - 535696.1e^{-3.1213407t} \cos(0.411283t) \\ - 4097274.0e^{-3.1213407t} \sin(0.411283t) - 7677755.0. \quad (30)$$

Simulated and predicted values of the original series are calculated by  $\hat{x}^{(0)}(k) = \frac{\hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)}{D_k}$ ; relative errors of simulated and predicted values of the original series in each period are calculated by  $RE(t) = \left| \frac{x^{(0)}(t) - \hat{x}^{(0)}(t)}{x^{(0)}(t)} \right| \times 100\%$ ; the average-simulation relative error is calculated by  $MAPE_1 = \frac{1}{m-1} \sum_{k=2}^m \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%$ ; the average-prediction relative error is calculated by  $MAPE_2 = \frac{1}{n-m} \sum_{k=m+1}^n \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%$ . Table 1 shows the related results.

**Table 1** Relevant calculation results of grey modeling of China's quarterly GDP

Time	$t$	$x^{(0)}(t)$	Model 1 proposed in the paper (seasonal factors are constant)		Model 2 proposed in the paper (seasonal factors are linear)	
			Simulated value	Relative error %	Simulated value	Relative error %
Q1 2017	1	181867.7	-	-	-	-
Q2 2017	2	201950.3	204319.08	0.83	204480.45	0.442
Q3 2017	3	212789.3	216000.69	1.47	217104.7	1.52
Q4 2017	4	235428.7	230777.6	2.35	228872.73	0.0246
Q1 2018	5	202035.7	214572.12	0.106	211781.52	1.79
Q2 2018	6	223962.2	228514.5	0.0122	227404.47	0.0362
Q3 2018	7	234474.3	237363.41	1.19	235821.96	0.00309
Q4 2018	8	258808.9	251497.44	1.46	251358.68	0.428
Q1 2019	9	217168.3	229709.84	0.299	231970.31	0.242
Q2 2019	10	241502.6	241399.22	2.02	245448.41	0.186
Q3 2019	11	251046.3	248797.77	0.932	247599.66	2.02
Q4 2019	12	276798.0	262923.56	0.822	259278.2	2.6
Q1 2020	13	205244.8	240605.07	10.5	235891.09	8.03
Q2 2020	14	248347.7	254256.14	0.352	251683.57	0.0259
Q3 2020	15	264355.7	264194.38	0.0981	259713.31	2.48
Q4 2020	16	295618.8	281905.85	0.432	282478.28	0.093
Q1 2021	17	247985.0	260600.93	0.947	263151.09	0.0908
			Predicted value	Relative error %	Predicted value	Relative error %
Q2 2021	18	281528.0	278060.33	3.19	285064.63	0.0669
Q3 2021	19	289919.3	291390.28	0.47	293832.16	0.528
Q4 2021	20	324237.4	313052.52	0.81	318124.89	3.16
Average-simulation relative error (Q1 2017-Q1 2021)			-	1.49	-	1.25
Average-prediction relative error (Q2 2021-Q4 2021)			-	1.48	-	1.24
Average overall relative error (Q1 2017-Q4 2021)			-	1.49	-	1.25

If the seasonal factor is linear, the GM(1, 1, exp×sin, exp×cos) model is established below.

$$\begin{aligned} & (d_1, d_2, d_3, d_4, h_1, h_2, h_3, h_4, \alpha, b_1, s_1, b_2, s_2) \\ & = (1.0695, 1.0177, 1.0044, 1.0044, -0.0004, -0.0003, 0.0002, \\ & \quad -0.0013, 0.9587, -0.1969, 0.5745, -0.1740, -0.5496), \end{aligned} \quad (31)$$

and

$$\begin{aligned} & (a_1, c_0, c_1, f_1, c_2, f_2) \\ & = (-0.02062895, 198407.2, 267758.6, 0.012507122, 220082.32, 0.012506868). \end{aligned} \quad (32)$$

The time response equation is

$$\begin{aligned} \hat{x}^{(1)}(t) = & 9650765.3e^{0.02062895t} + 355825.08e^{-0.17401504t} \cos(0.54957526t) \\ & + 126023.19e^{-0.17401504t} \sin(0.54957526t) \\ & - 407591.03e^{-0.19693091t} \cos(0.57454776t) \\ & - 154339.54e^{-0.19693091t} \sin(0.57454776t) - 9617900.7. \end{aligned} \quad (33)$$

$\hat{x}^{(0)}(k) = \frac{(\hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1))}{D_k}$  is used to calculate simulated and predicted values of the original series, relative errors of simulated and predicted values in each period, and average-simulation and average-prediction relative errors (see Table 1 for relative results).

We compare the modeling accuracy with other models by calculations. The ARIMA model is established to obtain:

$$(1 + 0.311952B)(1 - B)^2(1 - B^4)x_t^{(0)} = 5815.32 + (1 + B)\varepsilon_t. \quad (34)$$

AIC = 270.2888 is minimum. Table 2 shows the simulated and predicted values of the original series calculated from the model as well as relative errors and average relative errors of each period.

The superposed trigonometric function model is established as follows:

$$\begin{aligned} x_t^{(0)} = & 199019.43 + 4026.4962t - 19708.122 \sin\left(2\pi \times \frac{4}{17}t\right) - 13868.303 \sin\left(2\pi \times \frac{8}{17}t\right) \\ & - 4864.0584 \cos\left(2\pi \times \frac{4}{17}t\right) - 5924.239 \sin\left(2\pi \times \frac{8}{17}t\right) + \varepsilon_t. \end{aligned} \quad (35)$$

The test quantities of the model are as follows:  $R^2 = 0.8996371$ ,  $F = 19.72046$ ,  $P = 0.00003719183$ .

To compare the model proposed with the grey models of other references, we make calculations. For the seasonal data, we build a grey GM(1, 1, sin) model with the improved method proposed by reference [20], and then get the following time response equation:

$$\begin{aligned} \hat{x}^{(1)}(k) = & 12587335.7e^{0.0167k} + 1004.8227 \\ & \times [7.9501 \cos(-7.9501k) - 0.0167 \sin(-7.9501k)] - 12413456.5. \end{aligned} \quad (36)$$

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 3. Table 3 shows the relative errors and average relative errors in the periods.

**Table 2** Calculation results of grey modeling of China's quarterly GDP by other models

Time	$t$	$x^{(0)}(t)$	ARIMA model		superimposed trigono-metric model	
			Simulated value	Relative error %	Simulated value	Relative error %
Q1 2017	1	181867.7	200401.68	10.2	186248.15	2.41
Q2 2017	2	201950.3	209985.41	3.98	207717.93	2.86
Q3 2017	3	212789.3	209535.92	1.53	229121.98	7.68
Q4 2017	4	235428.7	211969.99	9.96	222674.16	5.42
Q1 2018	5	202035.7	210722.9	4.3	191844.84	5.04
Q2 2018	6	223962.2	226955.23	1.34	226712.66	1.23
Q3 2018	7	234474.3	225930.53	3.64	234145.93	0.14
Q4 2018	8	258808.9	228466.58	11.7	254176.62	1.79
Q1 2019	9	217168.3	226690.67	4.38	204030.28	6.05
Q2 2019	10	241502.6	245495.86	1.65	241448.36	0.0225
Q3 2019	11	251046.3	243720.04	2.92	242766.33	3.3
Q4 2019	12	276798.0	246558.18	10.9	277448.57	0.235
Q1 2020	13	205244.8	244340.21	19.0	225987.81	10.1
Q2 2020	14	248347.7	272476.69	9.72	250103.33	0.707
Q3 2020	15	264355.7	264845.96	0.185	257285.48	2.67
Q4 2020	16	295618.8	265667.55	10.1	290990.29	1.57
Q1 2021	17	247985.0	261730.28	5.54	256681.57	3.51
			Predicted value	Relative error %	Predicted value	Relative error %
Q2 2021	18	281528.0	282405.06	0.312	254698.59	9.53
Q3 2021	19	289919.3	277482.97	4.29	276168.36	4.74
Q4 2021	20	324237.4	284833.74	12.2	297572.42	8.22
Average-simulation relative error (Q1 2017-Q1 2021)			-	6.54	-	3.22
Average-prediction relative error (Q2 2021-Q4 2021)			-	5.58	-	7.50
Average overall relative error (Q1 2017-Q4 2021)			-	6.40-	-	3.86

Then, we build a grey GM(1,1) power model using the improved method proposed by reference [21] and get the following time response equation:

$$\begin{cases} \tilde{x}^{(1)}(k) = [5078682.5 + 31563.77e^{-0.005357336(k-1)}]^{1.1685}, \\ \varepsilon(k) = 118920.4 - 99282.22 \cos(0.0663k) - 74884.79 \sin(0.0663k), \\ \hat{x}^{(0)}(k) = \tilde{x}^{(1)}(k) - \tilde{x}^{(1)}(k-1) + \varepsilon(k). \end{cases} \quad (37)$$

With  $\hat{x}^{(0)}(k)$ , we calculate and get the original sequence's simulation and prediction values which are shown in Table 3. Table 3 gives the relative errors and average relative errors in the periods.

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

**Table 3** Calculation results of grey modeling of China's quarterly GDP by the methods of other references

Time	$t$	$x^{(0)}(t)$	Model proposed by Reference [20]		Model proposed by Reference [21]	
			Simulated value	Relative error %	Simulated value	Relative error %
Q1 2017	1	181867.7	-	-	-	-
Q2 2017	2	201950.3	203384.9	0.71	201950.3	$4.43 \times 10^{-8}$
Q3 2017	3	212789.3	208604.18	1.97	212961.92	0.0811
Q4 2017	4	235428.7	229434.26	2.55	219824.0	6.63
Q1 2018	5	202035.7	228184.23	12.9	224712.31	11.2
Q2 2018	6	223962.2	215697.42	3.69	228547.07	2.05
Q3 2018	7	234474.3	227584.91	2.94	231814.23	1.13
Q4 2018	8	258808.9	246167.79	4.88	234804.86	9.27
Q1 2019	9	217168.3	239231.95	10.2	237707.82	9.46
Q2 2019	10	241502.6	230644.3	4.5	240652.31	0.352
Q3 2019	11	251046.3	248038.53	1.2	243729.76	2.91
Q4 2019	12	276798.0	262244.11	5.26	247006.14	10.8
Q1 2020	13	205244.8	251231.03	22.4	250529.23	22.1
Q2 2020	14	248347.7	248402.49	0.0221	254333.33	2.41
Q3 2020	15	264355.7	269377.14	1.9	258442.26	2.24
Q4 2020	16	295618.8	277755.39	6.04	262871.48	11.1
Q1 2021	17	247985.0	264910.54	6.83	267629.59	7.92
			Predicted value	Relative error %	Predicted value	Relative error %
Q2 2021	18	281528.0	268901.38	4.49	272719.36	3.13
Q3 2021	19	289919.3	291049.48	0.39	278138.65	4.06
Q4 2021	20	324237.4	293045.11	9.62	283880.97	12.4
Average simulation relative error (Q1 2017-Q1 2021)			-	5.50	-	6.22
Average prediction relative error (Q2 2021-Q4 2021)			-	4.83	-	6.55
Average overall relative error (Q1 2017-Q4 2021)			-	5.39	-	6.27

Table 1, Table 2, Table 3 and Figure 1 show that the extended grey models built with the method proposed for seasonal data have high simulation and prediction precision which is much higher than that of the ARIMA model, the superimposed trig function model, the grey model proposed by reference [20] and the grey model proposed by reference [21]. It indicates the model and method proposed have high reliability and effectiveness.

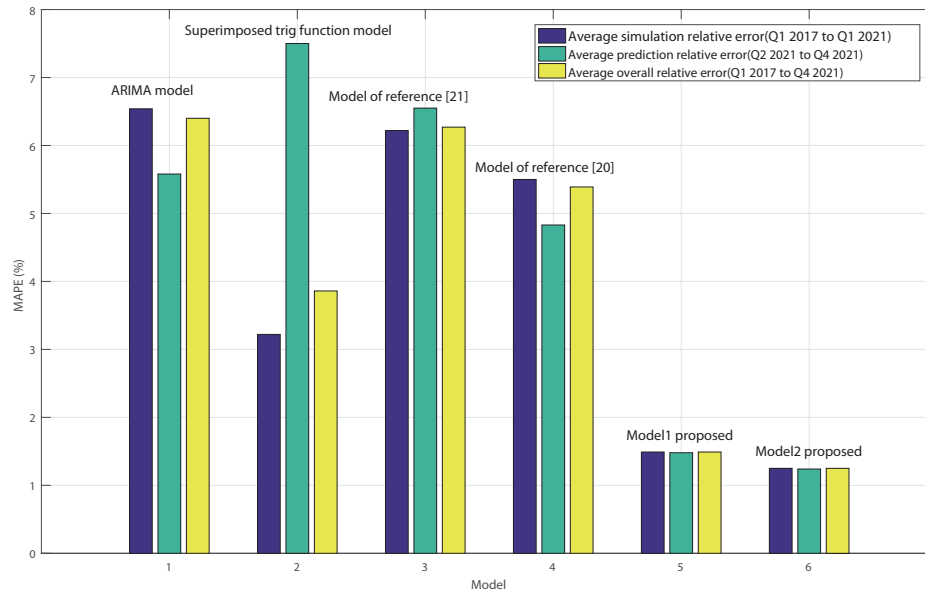


Figure 1 Histogram of average relative errors of six models

## 7 Conclusions

1) The original time series is transformed by introducing seasonal multiple factors. They are divided into constant and linear factors, which enhance the adaptation of the model for seasonal data.

2) The seasonal time series is in superimposed sine and cosine fluctuations with variable trends; therefore, the paper extends the grey action quantity of the traditional GM(1, 1) model. The newly extended grey model is called the GM(1, 1,  $\exp \times \sin$ ,  $\exp \times \cos$ ) model, which conforms to the regularity of seasonal time series and has high accuracy.

3) The paper presents the parameter optimization method of the GM(1, 1,  $\exp \times \sin$ ,  $\exp \times \cos$ ) model. Seasonal multiple factors and parameters contained in the model are optimized simultaneously, and minimum  $\max(\text{MAPE}_1, \text{MAPE}_2)$  is the objective function, which reduces the average-simulation and average-prediction relative errors of the model.

4) The paper builds the extended GM(1, 1,  $\exp \times \sin$ ,  $\exp \times \cos$ ) models with the method proposed. Results show that extended model 1 has an average simulation relative error of 1.49% and an average prediction relative error of 1.48%; extended model 2 has an average simulation relative error of 1.25% and an average prediction relative error of 1.24%. The errors are all small and extended model 2 has the highest precision. Moreover, extended model 1 and extended model 2 have the average simulation and prediction relative errors significantly smaller than those of the common ARIMA model, the superimposed trig function model, the grey model proposed by reference [20] and the grey model proposed by reference [21], indicating that the modeling method proposed has high reliability.

## References

- [1] Liu Z, Xie Y M, Dang Y G. A progress analysis of Chinese poverty alleviation based on improved GM(1, 1). *Systems Engineering — Theory & Practice*, 2019, 39(1): 2476–2486.

- [2] Cao Y, Yin, K D, Li X M. Prediction of direct economic loss caused by marine disasters based on the improved GM(1, 1) model. *Journal of Grey System*, 2020, 32(1): 133–145.
- [3] Xu N, Dang Y G. Characteristic adaptive GM(1, 1) model and forecasting of Chinese traffic pollution emission. *Systems Engineering — Theory & Practice*, 2018, 38(1): 187–196.
- [4] Qi Q, Li C. Application of GM(1, 1) model based on background value optimization in China civil aviation freight volume forecast. *Logistics Engineering and Management*, 2021, 43(10): 56–59.
- [5] Cai L L, Wu F, Lei D G. Pavement condition index prediction using fractional order GM(1, 1) model. *IEEE Transactions on Electrical and Electronic Engineering*, 2021, 16(8): 1099–1103.
- [6] Hu Y C. Energy demand forecasting using a novel remnant GM(1, 1) model. *Soft Computing*, 2020, 24(18): 13903–13912.
- [7] Huang C J, Cao Y Z, Zhou L. Application of optimized GM(1, 1) model based on EMD in landslide deformation prediction. *Computational and Applied Mathematics*, 2021, 40(8): 261.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] Xu N, Dang Y G. An optimized grey GM(2, 1) model and forecasting of highway subgrade settlement. *Mathematical Problems in Engineering*, 2015, 2015(1): 1–6.
- [13] Wang Y H, Lu J. Improvement and application of GM(1, 1) model based on multivariable dynamic optimization. *Journal of Systems Engineering and Electronics*, 2020, 31(3): 593–601.
- [14] Chen F Y, Zhu Y P. A new GM(1, 1) based on piecewise rational linear. *Engineering Letters*, 2021, 29(3): 849–855.
- [15] Mi C M, Wang Y J, Xiao L. Prediction on transaction amounts of China's CBEC with improved GM(1, 1) models based on the principle of new information priority. *Electronic Commerce Research*, 2021, 21(1): 125–146.
- [16] Tang L W, Lu Y Y. An improved non-equal interval GM(1, 1) model based on grey derivative and accumulation. *Journal of Grey System*, 2020, 32(2): 77–88.
- [17] Wang C S, Sun Z H. Monthly pork price forecasting method based on Census X12-GM(1, 1). *PLoS ONE*, 2021, 16(5): 1–13.
- [18] Kapila Tharanga Rathnayaka R M, Seneviratna D M K N. Taylor series approximation based unbiased GM(1, 1) hybrid statistical approach for forecasting stock market. *Journal of Grey System*, 2020, 32(3): 124–133.
- [19] Qian W Y, Wang J. An improved seasonal GM(1, 1) model based on the HP filter for forecasting wind power generation in China. *Energy*, 2020, 209: 118499.
- [20] Zeng L. Grey GM(1, 1|sin) power model based on oscillation sequences and its application. *Journal of Zhejiang University (Science Edition)*, 2019, 46(6): 697–704.
- [21] Wang Z X. Grey forecasting method for small sample oscillating sequences based on Fourier series. *Control and Decision*, 2014, 29(2): 270–274.
- [22] Liu S F. Grey system theory and its application. 9th Ed. Beijing: Science Press, 2021.