Improvement and application of GM(1,1) model based on multivariable dynamic optimization

WANG Yuhong and LU Jie*
School of Business, Jiangnan University, Wuxi 214122, China

Abstract: For the classical GM(1,1) model, the prediction accuracy is not high, and the optimization of the initial and background values is one-sided. In this paper, the Lagrange mean value theorem is used to construct the background value as a variable related to $k$. At the same time, the initial value is set as a variable, and the corresponding optimal parameter and the time response formula are determined according to the minimum value of mean relative error (MRE). Combined with the domestic natural gas annual consumption data, the classical model and the improved GM(1,1) model are applied to the calculation and error comparison respectively. It proves that the improved model is better than any other models.

Keywords: grey prediction, GM(1,1) model, background value, grey system theory.

DOI: 10.23919/JSEE.2020.000024

1. Introduction

Since the theory of the grey system was put forward, the GM(1,1) model, as the core and base of the grey prediction theory [1], has been widely used in the concrete practice of production and life. However, in the practical application, the prediction accuracy is unstable or even deviates from the classical GM(1,1) model [2]. Therefore, many scholars studied the improvement of the GM(1,1) model from different angles, and achieved some results in the initial value, the background value, model parameter estimation and so on.

In the optimization of the initial value, considering that new information should play a key role in modeling, Luo [3] directly used $x^{(1)}(n)$ as the initial condition of the grey model, while Xu et al. [4] used $x^{(1)}(n)$ as the initial condition of the grey model respectively. Wang et al. [5] optimized the initial value of the model when the whitening equation was a non-homogeneous exponential function, and then constructed a new background value expression to reduce the model error and improve the prediction accuracy of the model.

In respect of the background value, the classical model has no strict theoretical basis for the construction of the background value $z^{(1)}(k)$, and the use scope is narrow, thus the scholars improve the background value from different angles. For the geometric significance of the integral, in order to reduce the error caused by the direct calculation of the whole trapezoidal area in the interval $[k-1, 1]$, Jiang et al. [6] calculated the sum of the interval integral by using the piecewise low order interpolation combined the complex trapezoid formula. Reference [7] used the irregular trapezoid area construction method to replace the traditional area construction method with the Riemann integral to optimize the background value of the traditional GM(1,1) model. Starting from the numerical characteristics of sequences, Peng et al. [8] abstracted the data sequence into non-homogeneous exponential functions to construct background values and the GM(1,1) model. To extend the applicability and accuracy of the classical model, Luo et al. [9] improved the background value under the condition that the original sequence spacing is inconsistent. In addition, some scholars proposed mathematical expressions of $z^{(1)}(k)$ from different angles, and they also improved the model predicting accuracy to a certain extent in [10–12]. In the estimation of model parameters, Meng et al. [13] used the particle swarm optimization algorithm, Lee et al. [14] used the genetic algorithm to optimize the classical model, and these methods achieved good results.

It can be seen that in the aspect of background value optimization, the existing research mainly focuses on the improvement of the adjacent mean construction method [15]. In fact, when the whitening equation

\[
\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b
\]  

(1)
is integrated on both sides in interval \([k - 1, 1]\), we can see that the background value \(z^{(1)}(k)\) varies with \(k\), and the proximity of \(z^{(1)}(k)\) to the latter reflects the applicability and accuracy of the GM(1,1) model. Although some studies show the dynamic in the improvement of the background value, there is little relationship with \(k\).

At the same time, the traditional GM(1,1) is used as an exponential prediction model [16], and the essence is a static equation with \(x^{(1)}(1)\) as a fixed point [17]. According to the principle of the least squares, the fitting curve does not necessarily pass through the pre-artificially defined initial points. The oldest data are not closely related to the future [18]. Although some scholars put forward \(x^{(1)}(n)\) as a fixed point, the effect of these methods still remains to be tested when the model involves multiple variables [19].

In summary, in the study of the classical GM(1,1) model improvement, most scholars aim at optimizing one aspect of the classical models. In the process of modeling, the more variables there are, the higher the accuracy of the prediction will be [20]. If the initial value and the background value of the model are regarded as variables, the more variables are optimized, the higher the accuracy of the model is [21]. In this paper, we first improve the construction of the background value of the classical GM(1,1) model, and apply the Lagrange mean value theorem to construct the background value as a function of \(k\); at the same time, the initial value is no longer fixed. The background value and the initial value are set as variables, and the minimum value of the average relative error is taken as the principle. The optimal parameter and the time response formula are determined by the dynamic optimization; finally, the improved model is used to predict with the domestic natural gas annual consumption data.

2. Error analysis of classical GM(1,1) model

According to the whitening equation of the basic form of the GM(1,1) model [22], when (1) is integrated on both sides in interval \([k - 1, 1]\),

\[
x^{(0)}(k) + a \int_{k-1}^{k} x^{(1)}(t)dt = b,
\]

then we can get

\[
z^{(1)}(k) = \int_{k-1}^{k} x^{(1)}(t)dt.
\]

The background value of the classical model is the trapezoidal area, however, its actual value should be the area of the curved trapezoid surrounded by the curve \(x^{(1)}(t)\) of \(\int_{k-1}^{k} x^{(1)}(t)dt\) in the interval \([k - 1, 1]\) and the \(t\)-axis. As shown in Fig. 1, the difference between the two areas is the source of the background value errors in the classical model.

![Fig. 1 Source of background value errors in GM(1,1) model](image)

The calculation formula of the background value of the classical model is expressed as follows:

\[
z^{(1)}(k) = 0.5(x^{(1)}(k) + x^{(1)}(k - 1)).
\]

When the change of the accumulated generating sequence is gentle and the time interval is small, it is appropriate to adopt the above method to calculate [23]. However when the fluctuation of the first-order accumulated generating sequence is large, the above method will cause greater errors [24]. From the view of sequence generation characteristics, the classical model artificially stipulates that old information and new information are equally important, which is not in line with reality [25].

In this paper, the background value is regarded as a variable, and the parameters of the background value and the specific form of time response are determined when the mean relative error (MRE) reaches the minimum value, which can significantly reduce the error caused by human factors and improve the prediction accuracy.

At the same time, on the basis of optimizing the background value, the initial value is also optimized. As an exponential prediction model, the GM(1,1) model is essentially a static equation with \(x^{(1)}(1)\) as a fixed point. Compared with the dynamic equation, the static equation does not have the characteristics of independency of benchmark selection, step size and inherent consistency [26]. The fitting effect is usually worse than that of the dynamic equation [27]. For the above considerations, the initial value is also set as a variable to improve the fitting accuracy.

3. Improvement of GM(1,1) model

3.1 Brief introduction to classic GM(1,1)

**Definition 1** [28] Set a non-negative original sequence

\[
X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)).
\]

\(X^{(1)}\) is a cumulative generation sequence of \(X^{(0)}\):

\[
X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n)),
\]

\(X^{(1)}\) is a cumulative generation sequence of \(X^{(0)}\):

\[
X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n)),
\]
\[ x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), \quad k = 1, 2, \ldots, n. \]  

The original form of the GM(1,1) model is
\[ x^{(0)}(k) + ax^{(1)}(k) = b. \]

**Definition 2** [28] As \( X^{(0)} \) and \( X^{(1)} \) are shown in Definition 1, set
\[ Z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \ldots, z^{(1)}(n)), \]

among which (4) is satisfied, then the basic form of the GM(1,1) model is
\[ x^{(0)}(k) + az^{(1)}(k) = b. \]

Its whitening equation is (1).

**Theorem 1** [28] Set the parameter column \( \hat{a} = (a, b)^T \), and
\[
B = \begin{bmatrix}
-z^{(1)}(2) & 1 \\
-z^{(1)}(3) & 1 \\
\vdots & \vdots \\
-z^{(1)}(n) & 1 \\
\end{bmatrix}, \quad Y = \begin{bmatrix}
x^{(0)}(2) \\
x^{(0)}(3) \\
\vdots \\
x^{(0)}(n) \\
\end{bmatrix}.
\]

Then the least square parameter estimation of the GM(1,1) model satisfies \( \hat{a} = (B^T B)^{-1} B^T Y \).

The time response sequence of the GM(1,1) model is
\[ \hat{x}^{(1)}(k) = (x^{(0)}(1) - b/a)e^{-a(k-1)} + b/a. \]

Its reduction value is
\[
\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) = (x^{(0)}(1) - b/a)(e^{-a(k-1)} - e^{-a(k-2)}),
\]

among which \( k = 2, 3, \ldots, n \).

It can be seen from (12), the prediction accuracy of the GM(1,1) model depends on the selection of the fixed points and the values of parameters \( a, b \), and the values of \( a, b \) depend on the construction of the background value. Combining the initial and background values can significantly improve the accuracy of the model.

### 3.2 Improvement of grey GM(1,1)

We can get \( X^{(1)} \) by accumulation through (5), that is (7). The whitening equation of the GM(1,1) model is (1).

Since \( X^{(1)} \) is continuous in interval \([k-1, 1]\), and derivable in \((k-1, 1)\), according to Lagrange’s mean value theorem, we can know that \( \hat{x}^{(1)}_k \in (k-1, k) \) satisfies the following equation:
\[ x^{(1)}(k) - x^{(1)}(k-1) = \frac{dx^{(1)}(t)}{dt} \bigg|_{t=\xi_k}. \]

Let \( t = \xi_k \) in (1), due to \( x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1) \), we can get
\[ x^{(0)}(k) + ax^{(1)}(\xi_k) = b. \]

Since \( X^{(0)} \) is a non-negative sequence, \( X^{(1)} \) is a monotone increasing sequence.
\[ x^{(1)}(\xi_k) \in (x^{(1)}(k-1), x^{(1)}(k)) \]
\[ x^{(1)}(\xi_k) = a_kx^{(1)}(k-1) + (1 - a_k)x^{(1)}(k) \]
where \( 0 \leq a_k \leq 1, k = 1, 2, \ldots, n \). Then bring \( x^{(1)}(\xi_k) \) into (19),
\[ b = x^{(0)}(k) + a(a_kx^{(1)}(k-1) + (1 - a_k)x^{(1)}(k)). \]

Replace the theoretical value \( a_k \) in (17) with the actual value \( \tilde{a}_k \), then we can get
\[ b + \varepsilon_k = x^{(0)}(k) + a(\tilde{a}_kx^{(1)}(k-1) + (1 - \tilde{a}_k)x^{(1)}(k)) \]
where \( \varepsilon_k \) is the error caused by the replacement. Use the least square method to make \( \sum_{i=2}^{n} \varepsilon_k^2 \) get the minimum, then the values of \( a \) and \( b \) can be calculated.
\[
\begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y
\]
among which
\[
B = \begin{bmatrix}
-\tilde{a}_2x^{(1)}(1) - (1 - \tilde{a}_2)x^{(1)}(2) \\
-\tilde{a}_3x^{(1)}(2) - (1 - \tilde{a}_3)x^{(1)}(3) \\
\vdots \\
-\tilde{a}_nx^{(1)}(n-1) - (1 - \tilde{a}_n)x^{(1)}(n) \\
\end{bmatrix}, \quad Y = \begin{bmatrix}
x^{(0)}(2) \\
x^{(0)}(3) \\
\vdots \\
x^{(0)}(n) \\
\end{bmatrix}^T,
\]

where \( \tilde{a}_k \) is a function of \( k \), that is, \( \tilde{a}_k = f(k) \). Because the specific forms of \( \tilde{a}_k \) and \( f(k) \) are unknown, \( 0 \leq \tilde{a}_k \leq 1 \) and \( k = 1, 2, \ldots, n \), we can assume \( \tilde{a}_k = \alpha \beta^\gamma \) \( (0 \leq \alpha \leq 1, \beta = 1, 2, \ldots) \), so that \( \tilde{a}_k \) could be an arbitrary value of the interval \([0, 1]\).

**Theorem 2** If \( z^{(1)}(k) \) in GM(1,1) is constructed as follows:
\[ z^{(1)}(k) = \alpha \beta^\gamma x^{(1)}(k-1) + (1 - \alpha \beta^\gamma)x^{(1)}(k), \]
\[ 0 \leq \alpha \leq 1; \beta = 1, 2, \ldots \]
among which \( k = 2, 3, \ldots, n \), the error of the background value increases with the increasing of \( \beta \).

**Proof** Let the whitening equation (1) be integrated on both sides in interval \([k-1, 1]\),
\[ \int_{k-1}^{k} \frac{dx^{(1)}(t)}{dt} dt + a \int_{k-1}^{k} x^{(1)}(t) dt = b. \]
That is
\[ b = x^{(1)}(k) - x^{(1)}(k - 1) + a \int_{k-1}^{k} x^{(1)}(t) \, dt. \quad (21) \]

Comparing (4) and (19), the actual value of the background value of the GM(1,1) model should be equal to the area of the curved trapezium formed by the curve \( x^{(1)}(t) \) in the interval \([k - 1, 1]\) and the t axis [28]. When the background value is constructed in (19), the difference between them is the source of the model error. Let
\[ c = \int_{k-1}^{k} x^{(1)}(t) \, dt - x^{(1)}(k) \]
measure the error, that is
\[ c = \int_{k-1}^{k} x^{(1)}(t) \, dt - \alpha \sqrt[\beta]{x^{(1)}(k)} - \alpha \sqrt[\beta]{x^{(1)}(k)}(x^{(1)}(k) - x^{(1)}(k - 1)) + \int_{k-1}^{k} x^{(1)}(t) \, dt - 1. \quad (22) \]

Assume \( \beta_2 \geq \beta_1 \geq 1 \), then
\[ c_{\beta_2} - c_{\beta_1} = (\alpha^{\beta_2} \sqrt[\beta_2]{x^{(1)}(k)} - \alpha^{\beta_1} \sqrt[\beta_1]{x^{(1)}(k)})(x^{(1)}(k) - x^{(1)}(k - 1)). \quad (23) \]

\( X^{(1)} \) is a monotone increasing sequence, thus \( x^{(1)}(k) - x^{(1)}(k - 1) \geq 0 \), \( k = 2, 3, \ldots, n \), when \( \beta_2 \geq \beta_1 \), \( \alpha^{\beta_2} \sqrt[\beta_2]{x^{(1)}(k)} \geq \alpha^{\beta_1} \sqrt[\beta_1]{x^{(1)}(k)} \), as \( 0 \leq \alpha \leq 1 \), thus \( 0 \leq \alpha^{\beta_2} \leq \alpha^{\beta_1} \). In summary,
\[ c_{\beta_2} - c_{\beta_1} = (\alpha^{\beta_2} \sqrt[\beta_2]{x^{(1)}(k)} - \alpha^{\beta_1} \sqrt[\beta_1]{x^{(1)}(k)})(x^{(1)}(k) - x^{(1)}(k - 1)) \geq 0. \quad (24) \]

\[ \text{MRE} = \frac{1}{n-1} \sum_{k=2}^{n} \Delta(k) = \frac{1}{n-1} \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| = \frac{1}{n-1} \left| \frac{x^{(0)}(k) - \left( x^{(1)}(m) - \frac{b}{a} \right) \left( e^{-a(k-m)} - e^{-a(k-m-1)} \right)}{x^{(0)}(k)} \right| \quad (28) \]

among which \( m = 1, 2, \ldots, n \).

The MRE can be calculated according to different \( \alpha \) values, and find out \( \hat{\alpha} \) corresponding to the least mean relative error between the fitting sequence \( \hat{X}^{(0)} \) and the real time sequence, then the values of parameters \( a, b \) could be determined and brought into (12) to get the best fitting equation to predict.

In order to compare the difference between the proposed method and other adaptive parameter grey models, on the basis of referring to previous literature, the adaptive multi-parameter grey model (AMGM) [29] and the feature adaptive GM(1,1) model (FAGM) [4] are selected, and the calculation results are compared according to the actual examples.

The AMGM is an adaptive multi-parameter prediction model with equal-dimensional grey number compensation. This method enriches one-step prediction data into the original sequence and removes the oldest data to form a new equal-dimension sequence. A new multi-parameter prediction model is established for this series and the next step is forecasted. This model is called the adaptive multi-parameter prediction model. Because there is no strict theoretical foundation for the determination of initial values in classical models, in the FAGM, the selection of initial
values is not the first in the sequence, but based on the minimum value of the MRE, which greatly reduces the error caused by human intervention. Thus, we compare the above two selected models with the proposed model to illustrate the rationality and effectiveness of our proposed model in the empirical analyses section.

4. Empirical analyses

Natural gas is a high quality and efficient clean energy. Vigorously promoting the natural gas industry is the only way for China to speed up the construction of a modern clean and efficient energy system [30]. With the changes in the energy consumption structure of residents in the process of urbanization in China, as well as the promotion of the related environmental policies and the “coal to gas” project, the demand for the natural gas is exuberant, and the price of non-resident gas decreases, the enthusiasm of the enterprise is increased, and the consumption of the natural gas is increasing rapidly.

Due to the limited domestic gas resources, the growth of the natural gas demand is reflected in the sharp rise in external dependence. Data from the National Development and Reform Commission (NDRC) show that in 2017, China’s natural gas consumption increased by 15.3% year on year, and the annual import volume increased by nearly 30%, which led to China’s natural gas dependence in 2017 by up to 39%. When the consumption of the natural gas increases, the construction of natural gas storage in China is still in the initial stage, the rapid growth of the natural gas demand can only be made up by increasing import.

It can be seen that the natural gas, as an important energy source in China, not only provides strong support for life and production, but also relates to social security and stability. Forecasting the annual consumption of the natural gas in China not only helps maintain the balance between supply and demand of the natural gas, but also provides a basis for the formulation of major national policies. From the website of NDRC, we obtain the annual consumption data of the natural gas in China (2005—2016: one billion m$^3$).

$$X(0) = (48.2, 59.3, 73.0, 84.1, 92.6, 111.2, 137.1, 150.9, 171.9, 188.4, 197.3, 205.8)$$

First, we use the classic GM(1,1) model to calculate

$$a = -0.1148, \quad b = 61.9253.$$  

Take the results into the time response equation and we can get

$$\hat{x}^{(1)}(k) = 587.4536e^{0.1148(k-1)} - 539.254.$$  

According to (27), we can calculate the fitting sequence and MRE.

Bring the original sequence into the AMGM model, with the change of $\alpha$ in the interval [0, 1], Fig. 2 can be obtained.

Calculation with Matlab shows that when $\alpha = 0.83$, the minimum value of the MRE is 0.056.

$$a = -0.1367, \quad b = 52.4176,$$

$$\hat{x}^{(1)}(k) = 431.5546e^{0.1367(k-1)} - 383.355.$$  

According to (27), we can calculate the fitting sequence and MRE of the AMGM model.

Bring the original sequence into the FAGM model with the change of $m$, Fig. 3 can be obtained.

Calculation with Matlab shows that when $m = 2$, the minimum value of the MRE is 0.071, bring $m = 2$ and $x^{(1)}(2) = 107.5$ to the time response equation, we can get

$$x^{(1)}(k) = 646.7536e^{0.1148(k-2)} - 539.254.$$  

Fig. 2 Change of MRE with $\alpha$

Fig. 3 Change of MRE with $m$
According to (10) and (27), we can calculate the fitting sequence and MRE of the FAGM model.

Then, the original sequence is introduced into the proposed prediction model. Let $\alpha$ change in the interval $[0, 1]$, assign $\alpha$ repeatedly, and add a small amount greater than zero $\Delta \alpha = 0.01$ once at a time until $\alpha = 1$. At the same time, $m = 1, 2, \ldots, 12$, calculate the MRE values when $m$ and $\alpha$ change. According to the calculation of Matlab, the minimum value of the MRE is 0.05 when $\alpha = 0.7, m = 5$. The MRE changes with $m$ and $\alpha$, which is shown in Fig. 4.

Put $\alpha = 0.7, m = 5$ into the time response equation

$$x^{(1)}(k) = 731.1e^{0.134 \cdot 8(k-5)} - 373.9034.$$  

The fitting error (FE) and relative error (RE) of the AMGM, FAGM, and the proposed model are shown in Table 1.

| Year | Raw data | GM(1.1) | FE | RE | AMGM | FE | RE | FAGM | FE | RE | The proposed model | FE | RE |
|------|----------|---------|----|----|------|----|----|------|----|----|---------------------|----|----|
| 2005 | 48.2     | –       | –  | –  | –    | –  | –  | –    | –  | –  | –                   | –  | –  |
| 2006 | 59.3     | 71.49   | 0.21| 0.06| 70.16| 0.18| 63.94| 0.08 |
| 2007 | 73.0     | 80.19   | 0.10| 0.01| 78.70| 0.07| 73.16| 0.00 |
| 2008 | 84.1     | 89.94   | 0.07| 0.01| 88.28| 0.05| 83.73| 0.00 |
| 2009 | 92.6     | 100.89  | 0.09| 0.03| 99.02| 0.07| 95.81| 0.03 |
| 2010 | 111.2    | 113.17  | 0.02| 0.02| 111.07| 0.00| 109.63| 0.01 |
| 2011 | 137.1    | 126.94  | 0.07| 0.08| 124.59| 0.09| 125.45| 0.08 |
| 2012 | 150.9    | 142.38  | 0.06| 0.05| 143.63| 0.05| 139.75| 0.07 |
| 2013 | 171.9    | 159.71  | 0.07| 0.04| 156.76| 0.09| 143.55| 0.05 |
| 2014 | 188.4    | 179.14  | 0.05| 0.01| 175.83| 0.07| 187.96| 0.00 |
| 2015 | 197.3    | 200.94  | 0.02| 0.09| 197.23| 0.00| 215.09| 0.09 |
| 2016 | 205.8    | 225.40  | 0.10| 0.21| 221.23| 0.07| 246.12| 0.19 |
| MRE  | –        | – 0.080 | –  | 0.056| –    | 0.071| –   | 0.050 |

From Table 1, we can see that the error of the proposed model is smaller than that of other models, and the proposed model achieves good prediction results.

In the calculation results, when $\alpha = 0.7$ and $m = 5$, the average relative error is minimized, which is not consistent with $\alpha = 0.5$ and $m = 1$ of the classical GM(1,1) model. It also proves that the prediction accuracy will be higher if the initial and background values of the model are considered as variables to optimize.

When $\beta = 1, \hat{\alpha}_k = \alpha^k (0 \leq \alpha \leq 1)$, the MRE achieves the minimum value. In fact, the larger the $\beta$ value, the smaller the bias on new information. In this paper, the optimal solution is obtained when $\beta = 1$. At the same time, when the fitting accuracy gets close to that of the classical model, the prediction accuracy is much higher than that of the classical model. It also proves that the emphasis on new information is helpful to improve the prediction accuracy. It can also be found by numerical calculation that the minimum value obtained by the MRE generally increases with the increase of $\beta$ when $\alpha$ and $m$ change in the definition domain. Fig. 5 and Fig. 6 can be obtained by comparing three improved methods and classical models.

Overall, the error of the proposed model is significantly
smaller than that of the classical model and other two improved methods. With the deepening of the improvement, the model error shows a decreasing trend, which verifies the effectiveness of the model optimization.

\[
\alpha_0 = \alpha^k \quad (0 \leq \alpha \leq 1).
\]

Therefore, although we do not know the specific form of \( f(k) \), we also artificially put forward a specific form of \( f(k) \) in the article, but as long as it can meet the requirement that values can be freely chosen in the range of \([0,1]\), it is feasible.

![Fig. 6 Overall average error of four methods](image)

### 5. Conclusions

This paper improves the classical model. In this paper, the whitening equation of the classical GM(1,1) model is transformed and deformed by the Lagrange mean value theorem and the interpolation coefficient method, which reduces the error of the model and improves the prediction accuracy of the model. At the same time, the validity of the model is verified by a practical example.

The Lagrange mean value theorem reflects the relationship between the global average rate of change of the derivative function in a closed interval and the local rate of change of a point. In [30], the differential part of the whitening equation of the GM(1,1) model was replaced by the Lagrange mean value theorem to eliminate the error caused by approximate approximation. At the same time, the background value is no longer generated by the adjacent mean, but is replaced by the interpolation coefficient method.

This paper improves the classical model and improves the prediction accuracy significantly, but it still needs to be discussed in some aspects. On the construction of background values, in the process of model equivalent substitution \( \hat{\alpha}_k = \hat{\alpha}_k \quad (0 \leq \alpha \leq 1) \), \( \alpha \) changes in the interval \([0,1]\), \( \hat{\alpha}_k \) is a function of \( k \), that is, \( \hat{\alpha}_k = f(k) \). Because the specific forms of \( \hat{\alpha}_k \) and \( f(k) \) are unknown, \( 0 \leq \hat{\alpha}_k \leq 1 \) and \( k = 1, 2, \ldots, n \), let \( \hat{\alpha}_k = \alpha^{\beta} \quad (0 \leq \alpha \leq 1, \beta = 1, 2, \ldots) \), so that \( \hat{\alpha}_k \) could be an arbitrary value of the interval \([0,1]\). Because the specific form of \( f(k) \) is not clear, we only know that \( f(k) \) needs to be guaranteed a free value in the range of \([0,1]\). Thus, we set up a specific form to meet the requirement, and through the mathematical proof, we finally get \( \hat{\alpha}_k = \alpha^k \quad (0 \leq \alpha \leq 1) \). Therefore, although we do not know the specific form of \( f(k) \), we also always draw two conclusions. On the one hand, the specific form of \( f(k) \) is not unique, as long as it meets the requirements; on the other hand, the change of the specific form of \( f(k) \) will not affect the accuracy of the model. It can be seen from two aspects. Firstly, the only requirement for \( f(k) \) in the process of model construction is that it can be freely chosen in the range of \([0,1]\). Secondly, through the method put forward, the final transformation of the model into the background value is a function related to \( k \), and the initial value is related to \( m \). By adjusting different \( k \) and \( m \) values, the minimum MRE can be obtained. At this time, the value of unknown parameters in the model can be determined, and the final form of the model can be determined, and the optimal fitting equation can be obtained for prediction. Therefore, the change of the specific form of \( f(k) \) will not affect the accuracy of the model.

On improving the accuracy of the model, in the process of building and calculating the model, \( \alpha \) and \( m \) are chosen to make the model achieve the optimal effect. In fact, there are countless possibilities for \( \alpha \) and \( m \). The value of \( \alpha \) and \( m \) is an interval rather than just a certain number which can make the fitting error of the modified model less than that of the classical model. In the interval, the fitting error of the improved model is less than that of the classical model, which proves two problems: on the one hand, the values of the parameters of the improved model are not unique. In this interval, no matter how to choose the value, the improved model outperforms the classical model. Thus the parameters only need to be selected in this interval, and we can definitely find a point to minimize the model error. On the other hand, the effect of the improved model is universal and will not be inappropriate for different sequences. Because the parameters of the improved model exist in an interval, and different sequences may only result in different interval ranges. Even in extreme cases, there is at least one point that makes the error of the improved model smaller than that of the classical model. In practice, the range of this interval is generally large, which can achieve good prediction results.

At the same time, the minimum value is obtained when \( m = 5 \) in the example. The difference between the classical model and the modified model is that, the classical model directly takes the first number of the cumulative se-
quence as a fixed point, while this article takes the minimum value to determine the initial value, which avoids the error caused by setting parameters artificially. With the change of $m$, the error of the model also changes. There is no strict theoretical basis for setting the value of $m$ in advance. This problem can be solved by choosing the $m$ value when the MRE reaches the minimum value.

In this paper, $\alpha$ and $m$ are set as variables, which greatly improves the accuracy of the model. Although the model error can be reduced by setting two variables, and the improved model is better than the classical model, it is still worth discussing whether the error can be reduced by adding more variables. Although it is believed that the accuracy can be improved by increasing variables as many as possible [31], there is no strict theoretical basis. Whether the accuracy of the model with three variables is better than that of the model with two variables, and other problems caused by increasing variables need to be further discussed.

References

[1] DENG J L. Grey system: social economy. Beijing: Defense Industry Press, 1985. (in Chinese)
[2] DANG Y G, FENG Y, DING S, et al. Grey clustering model of interval grey number based on kernel and gray level. Control and Decision, 2017, 32(10): 1844 – 1848. (in Chinese)
[3] LUO Y X. Stepwise optimization model and application of non-equidistant innovation GM(1,1). Systems Engineering — Theory and Practice, 2010, 30(12): 2254 – 2258. (in Chinese)
[4] XU N, DANG Y G. Feature adaptive GM(1,1) model and prediction modeling of traffic pollution emissions in China. Systems Engineering — Theory and Practice, 2018, 38(1): 187 – 196. (in Chinese)
[5] WANG Y H, LIU Q, TANG J R, et al. Optimization approach of background value and initial item for improving prediction precision of GM(1,1) model. Journal of Systems Engineering and Electronics, 2014, 25(1): 77 – 82.
[6] JIANG S Q, LIU S F, ZHOU X C. Optimization of background value of GM(1,1) model based on the complex trapezoidal formula. Control and Decision, 2014, 29(12): 2221 – 2225. (in Chinese)
[7] JIANG Y X, ZHANG Q S. Optimization of background value of GM(1,1) model. Chinese Management Science, 2015, 23(9): 146 – 152. (in Chinese)
[8] PENG Z B, ZHANG C, PENG W X, et al. GM(1,1) model background value construction methods and applications. Journal of Northeastern University (Natural Science Edition), 2017, 38(6): 869 – 873. (in Chinese)
[9] LUO Y X, CHE X Y. Improvement and application of initial value of non-equidistant GM(1,1) model. International Journal of Computer Science Issues, 2013, 47(9): 19 – 50.
[10] AMIRENEKOEEI K, ARDEHALIM M, SADRI A. Integrated resource planning for Iran: development of reference energy system, forecast, and long-term energy-environment plan. Energy, 2012, 46(1): 374 – 385.
[11] WANG Y H, DANG Y G, LI Y Q, et al. An approach to increase prediction precision of GM(1,1) model based on optimization of the initial condition. Expert Systems with Applications, 2010, 37(8): 5640 – 5644.
[12] LIU Q Y, YU D J. Non-equidistance and nonhomogeneous grey model NNFGM (1,1) with the fractional order accumulation and its application. Journal of Interdisciplinary Mathematics, 2017, 20(6/7): 1423 – 1426.
[13] MENG W, LIU S F, FANG Z G, et al. GM(1,1) order optimization model based on reciprocal fractional order operator. Control and Decision, 2016, 31(4): 661 – 666. (in Chinese)
[14] LEE Y S, TONG L. Forecasting energy consumption using a grey model improved by incorporating genetic programming. Energy Conversion and Management, 2011, 52(1): 147 – 152.
[15] MA X. A brief introduction to the grey machine learning. Journal of Grey System, 2019, 31(1): 1 – 12.
[16] MA X, LIU Z B. The kernel-based nonlinear multivariate grey model. Applied Mathematical Modelling, 2018, 56(1): 217 – 238.
[17] ZENG B, TAN Y T, XU H M, et al. Forecasting the electricity consumption of commercial sector in Hong Kong using a novel grey dynamic prediction model. Journal of Grey System, 2018, 30(1): 159 – 174.
[18] MA X, XIE M, WU W, et al. The novel fractional discrete multivariate grey system model and its applications. Applied Mathematical Modelling, 2019, 70(1): 402 – 424.
[19] MA X, LIU Z. The GMC(l,n) model with optimized parameters and its application. Journal of Grey System, 2017, 29(4): 122 – 138.
[20] MA X, LIU Z B. Research on the novel recursive discrete multivariate grey prediction model and its applications. Applied Mathematical Modelling, 2016, 40(7/8): 4876 – 4890.
[21] XIE N M, LIU S F. Discrete grey forecasting model and its optimization. Applied Mathematical Modelling, 2009, 33(2): 1173 – 1186.
[22] CHEN C I, CHEN H L, CHEN S P. Forecasting of foreign exchange rates of Taiwan’s major trading partners by novel nonlinear grey Bernoulli model NGBM(1,1). Communications in Nonlinear Science and Numerical Simulation, 2008, 13(6): 1194 – 1204.
[23] WU L F, LIU S F, CUI W, et al. Non-homogenous discrete grey model with fractional-order accumulation. Neural Computing and Applications, 2014, 25(5): 1215 – 1221.
[24] XIONG P P, DANG Y G, YAO T X, et al. Optimal modeling and forecasting of the energy consumption and production in China. Energy, 2014, 77(4): 623 – 634.
[25] PAO H T, FU H C, TSENG C L. Forecasting of CO$_2$ emissions, energy consumption and economic growth in China using an improved grey model. Energy, 2012, 40(1): 400 – 409.
[26] YANG B H, ZHAO J S. Fractional order discrete grey GM(1,1) power model and its application. Control and Decision, 2015, 30(3): 218 – 220.
[27] WU L F, GAO X H, FU B, et al. Review of grey GM(1,1) model. Practice and Understanding of Mathematics, 2017, 47(15): 227 – 233.
[28] LIU S F, YANG Y J, WU L F. Grey system theory and its applications. Beijing: Science Press, 2014. (in Chinese)
[29] ZHANG S J, CHEN S Y. GM(1,1) power model optimization method and its application. Systems Engineering, 2016, 34(8): 158 – 162. (in Chinese)
[30] XIE N M, WANG R Z. A historic review of grey forecasting models. Journal of Grey System, 2017, 29(4): 1 – 29.
[31] GUO Y M, JIANG H M, ZHAI Z J. Adaptive multi-parameter prediction model based on grey theory. Acta Aeronautica Et Astronautica Sinica, 2009, 30(5): 925 – 931. (in Chinese)
Biographies

WANG Yuhong was born in 1979. He received his Ph.D. degree from Nanjing University of Aeronautics and Astronautics in 2010. Now he is a professor of School of Business, Jiangnan University. His research interests include uncertainty system prediction and decision method, and grey system theory. E-mail: yuhongwang@jiangnan.edu.cn

LU Jie was born in 1993. He received his M.S. degree from Jiangnan University in 2019. His research interests include uncertainty system prediction and decision method, and grey system theory. E-mail: JIE0327@outlook.com