Improved GM(1,1) model based on Simpson formula and its applications

Xin Ma\textsuperscript{a,b}, Wenqing Wu \textsuperscript{a,c,*}, Yuanyuan Zhang\textsuperscript{a}

\textsuperscript{a}School of Science, Southwest University of Science and Technology, Mianyang 621010, P.R. China
\textsuperscript{b}State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Southwest Petroleum University, Chengdu 610500, P.R. China
\textsuperscript{c}V.C. \& V.R. Key Lab of Sichuan Province, Sichuan Normal University, Chengdu 610068, P.R. China

Abstract

The classical GM(1,1) model is an efficient tool to make accurate forecasts with limited samples. But the accuracy of the GM(1,1) model still needs to be improved. This paper proposes a novel discrete GM(1,1) model, named GM\textsubscript{SD}(1,1) model, of which the background value is reconstructed using Simpson formula. The expression of the specific time response function is deduced, and the relationship between our model and the continuous GM(1,1) model with Simpson formula called GM\textsubscript{SC}(1,1) model is systematically discussed. The proposed model is proved to be unbiased to simulate the homogeneous exponent sequence. Further, some numerical examples are given to validate the accuracy of the new GM\textsubscript{SD}(1,1) model. Finally, this model is used to predict the Gross Domestic Product and the freightage of Lanzhou, and the results illustrate the GM\textsubscript{SD}(1,1) model provides accurate prediction.

Keywords: Grey forecasting, GM\textsubscript{SD}(1,1) model, Simpson formula, Small sample, Background value

\textsuperscript{*}To cite this document: Ma Xin, Wu Wenqing, Zhang Yuanyuan, Improved GM(1,1) model based on Simpson formula and its applications, Journal of Grey System, 2019, 31(4): 33-46.
\textsuperscript{*}Corresponding author. Wenqing Wu
Email address: wwqing0704163.com (Wenqing Wu)
1. Introduction

Grey system theory is an effective method to analyze uncertain problems with small samples and poor information, which was founded by professor Deng Julong [1]. The principle of the grey theory is “grey box” of which some information is known and the rest is unknown. Comparing with other methods, such as knowledge-driven method (Xiao et al. [2]), fuzzy systems (Wu et al. [3]), hybrid forecasting system (Du et al. [4], Ma et al. [5]), coupling mathematical model (Wang et al. [6, 7]), the grey model only needs little origin data having simple calculation process and satisfactory forecasting accuracy. Due to this important feature, it has been successfully applied in various fields. Its applications include, but are not limited to, the inverted pendulum control (Huang and Huang [8]), the semiconductor manufacturing layout (Chang et al. [9]), the stock price forecasting (Chen et al. [10]), the energy production (Wang et al. [11, 12], Zeng et al. [13], Zhou and He [14]), the energy consumption (Ma and Liu [15], Wu et al. [16, 17]), the industrial pollutant emission (Ma et al. [18]), the health expenditure of China (Wu et al. [19]), and China’s electricity consumption (Zeng [20], Wu et al. [21]).

In 1982, Deng presented the classical continuous GM(1,1) model of which procedures start with a differential equation called whitening equation. By discretizing the whitening equation and employing the least squares method, system parameters are estimated. Then simulation values and prediction values are computed with the help of the whitening equation and system parameters. Over the past three decades, a great number of univariate grey forecasting models have been proposed based on Deng’s pioneer work. Some excellent models in this area are those NNGBM(1,1) (Chen et al. [10], Zhang et al. [22]), GGM(1,1) (Zhou and He [14]), DGM(1,1) (Xie and Liu [23], Zeng et al. [24]), NGBM(1,1) (Chen et al. [25], Kong and Ma [26], Pei et al. [27], Salehi and Dehnavi [28]), SAGM(1,1) (Truong and Ahn [29]), NGM(1,1,k) (Cui et al. [30], Zeng and Li [31]). Recently, He and Wang [32] studied the continuous GM(1,1) model, i.e., GM_{SC}(1,1) model where the background value was derived by utilizing the
Simpson numerical integration formula. But their model has been shown inaccurate in some applications and biased for the homogeneous exponent sequence. Thus the optimization of the grey model and the improvement of the grey system theory have acquired a lot of attention. For instance, an efficient way to improve the effectiveness of the grey models was the development of discrete grey models DGM(1,1) (Xie and Liu [23]). For more details, the readers are directed to Xia et al. [33], Ma [34] and Wang and Phan [35].

In this paper, we focus on a kind of discrete GM(1,1) model called GM_{SD}(1,1) model where the background value is computed employing the Simpson numerical integration formula. Its solutions of time response function and restored values, properties, and applications are derived. We also study the forecast stability problem of the discrete GM_{SD}(1,1) model and discuss its causes from continuous to discrete in detail. That our model is also unbiased to simulate the homogeneous exponent sequence is proved. Finally, we simulate and forecast the Gross Domestic Product and the freightage of Lanzhou by using four kinds of GM(1,1) models.

The remainder of this paper is organized as follows. Section 2 gives a brief overview of the continuous GM(1,1) model. Its solutions and properties of GM_{SD}(1,1) model are derived in Section 3. Section 4 discusses the validation of the GM_{SD}(1,1) model. Applications are provided in Section 5. Conclusions are drawn in Section 6.

2. The basis of GM(1,1) model

This section gives a brief overview of the classical continuous GM(1,1) model. Suppose an original non-negative series be \( X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)) \) and the \( x^{(0)}(k) \) represents the behavior of the data at the time index \( k \) for \( k = 1, 2, \ldots, n \). Deng [1] proposed the GM(1,1) model is the following linear differential equation

\[
\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b, \quad (1)
\]
where the $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i)$, $k = 1, 2, \ldots, n$ are the first-order accumulated generating operating (1-AGO) series of $X^{(0)}$, the $a$ and $b$ are system parameters. Eq. (1) is also called the whitening equation of the GM(1,1) model.

The approximation of $dx^{(1)}(t)/dt$ is taken as
\[
\frac{dx^{(1)}(t)}{dt} = \lim_{\Delta t \to 1} \frac{x^{(1)}(t) - x^{(1)}(t - \Delta t)}{\Delta t} \approx x^{(1)}(t) - x^{(1)}(t - 1) = x^{(0)}(t),
\]
and the background values of $x^{(1)}(t)$ are defined as
\[
z^{(1)}(t) = \frac{1}{2} \left( x^{(1)}(t) + x^{(1)}(t - 1) \right).
\]

Thus the differential Eq. (1) can be approximately rewritten as the following difference equation
\[
x^{(0)}(t) + ax^{(1)}(t) = b. \tag{2}
\]

Employing the least squares estimation method, from Eq. (2) by considering $t = 2, 3, \ldots, n$, the model parameters $a$ and $b$ can be given below
\[
\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \left( \Lambda^T \Lambda \right)^{-1} \Lambda^T \eta, \tag{3}
\]
where $\Lambda$ and $\eta$ are defined as follow
\[
\Lambda = \begin{pmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(\nu) & 1 \end{pmatrix}, \quad \eta = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(\nu) \end{pmatrix},
\]
where $\nu$ is the number of samples that are used to build the grey models, and the left $n - \nu$ samples are used to test.

Solving Eq. (1), the time response function can be expressed by
\[
\hat{x}^{(1)}(k + 1) = \left( x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a}, \quad k = 1, 2, \ldots, n - 1. \tag{4}
\]

Then the restored values of $\hat{x}^{(0)}(k + 1)$ can be estimated by inverse accumulated generating operation (IAGO) which is given by
\[
\hat{x}^{(0)}(k + 1) = \hat{x}^{(1)}(k + 1) - \hat{x}^{(1)}(k), \quad k = 1, 2, \ldots, n - 1, \tag{5}
\]
or

\[ x^{(0)}(k + 1) = \frac{e^a - 1}{a} \left( b - ax^{(0)}(1) \right) e^{-ak}, \quad k = 1, 2, \ldots, n - 1. \quad (6) \]

As presented above, once given the sample data, the system parameters in Eq. (1) obtained. The output series also predicted with system parameters and input series by the Eqs. (4)-(6).

3. The discrete GM$_{SD}(1,1)$ model

3.1. Representation of the discrete GM$_{SD}(1,1)$ model

This subsection derives the discrete GM$_{SD}(1,1)$ model with Eq. (1) and the Simpson numerical integration formula. Considering the integration of Eq. (1) in the interval \([k - 1, k + 1]\), it follows

\[ \int_{k-1}^{k+1} dx^{(1)}(t) + a \int_{k-1}^{k+1} x^{(1)}(t)dt = b \int_{k-1}^{k+1} dt, \quad k = 2, 3, \ldots, n - 1. \quad (7) \]

It follows from Eq. (7) that

\[ x^{(1)}(k + 1) - x^{(1)}(k - 1) + a \int_{k-1}^{k+1} x^{(1)}(t)dt = 2b, \quad k = 2, 3, \ldots, n - 1. \quad (8) \]

By utilizing the Simpson numerical integration formula, Eq. (8) can be expressed as

\[ x^{(1)}(k + 1) - x^{(1)}(k - 1) + a \frac{x^{(1)}(k - 1) + 4x^{(1)}(k) + x^{(1)}(k + 1)}{3} = 2b, \quad k = 2, 3, \ldots, n - 1. \quad (9) \]

Eq. (9) turns to be

\[ (a + 3) x^{(1)}(k + 1) + 4ax^{(1)}(k) + (a - 3) x^{(1)}(k - 1) - 6b = 0, \]

\[ k = 2, 3, \ldots, n - 1. \quad (10) \]

It follows from Eq. (10) that

\[ x^{(1)}(k + 1) - wx^{(1)}(k) = \frac{a - 3}{w(a + 3)} \left( x^{(1)}(k) - wx^{(1)}(k - 1) \right) + \frac{6b}{a + 3}, \quad k = 2, 3, \ldots, n - 1. \quad (11) \]
where \( w = \sqrt{\frac{a+3}{a+3} - 2n} \).

Iterating Eq. (11) by itself, we have that

\[
x^{(1)}(k+1) - wx^{(1)}(k) = a - 3 \left\{ \frac{a - 3}{w(a + 3)} \left[ x^{(1)}(k - 1) - wx^{(1)}(k - 2) \right] + \frac{6b}{w(a + 3)} \right\} + \frac{6b}{a + 3}
\]

\[
= \left( \frac{a - 3}{w(a + 3)} \right)^2 \left[ x^{(1)}(k - 1) - wx^{(1)}(k - 2) \right] + \frac{6b}{a + 3} \sum_{i=0}^{1} \left( \frac{a - 3}{w(a + 3)} \right)^i
\]

\[
= \left( \frac{a - 3}{w(a + 3)} \right)^{k-1} \left[ x^{(1)}(2) - wx^{(1)}(1) \right] + \frac{6b}{a + 3} \sum_{i=0}^{k-2} \left( \frac{a - 3}{w(a + 3)} \right)^i
\]

\[
= \lambda^{k-1} \left[ x^{(1)}(2) - wx^{(1)}(1) \right] + \mu \frac{1 - \lambda^{k-1}}{1 - \lambda}, \quad (12)
\]

where \( \lambda = \frac{a-3}{w(a+3)} \), \( \mu = \frac{6b}{a+3} \).

Note that

\[
x^{(1)}(k+1) - w^{k-1} x^{(1)}(2)
\]

\[
= \sum_{j=0}^{k-2} w^j \left[ x^{(1)}(k - j + 1) - wx^{(1)}(k - j) \right]
\]

\[
= \sum_{j=0}^{k-2} w^j \left\{ \lambda^{k-j-1} \left[ x^{(1)}(2) - wx^{(1)}(1) \right] + \mu \frac{1 - \lambda^{k-j-1}}{1 - \lambda} \right\}
\]

\[
= \sum_{j=0}^{k-2} w^j \lambda^{k-j-1} \left[ x^{(1)}(2) - wx^{(1)}(1) \right] + \mu \frac{1 - \lambda^{k-j-1}}{1 - \lambda} \sum_{j=0}^{k-2} w^j (1 - \lambda^{k-j-1}). \quad (13)
\]

We have the 1-AGO series \( \hat{X}^{(1)} \) of discrete GM\(_{SD}(1,1) \) is

\[
\hat{x}^{(1)}(k+1) = w^{k-1} x^{(1)}(2) + \left[ x^{(1)}(2) - wx^{(1)}(1) \right] \sum_{j=0}^{k-2} w^j \lambda^{k-j-1}
\]

\[
+ \mu \frac{1 - \lambda^{k-j-1}}{1 - \lambda} \sum_{j=0}^{k-2} w^j (1 - \lambda^{k-j-1}), \quad k = 2, 3, \ldots, n - 1. \quad (14)
\]
Apply first-order inverse accumulation operation to obtain the simulation and forecasting value

\[ \hat{x}^{(0)}(k + 1) = \hat{x}^{(1)}(k + 1) - \hat{x}^{(1)}(k) \]
\[ = w^{k-2}(w - 1)x^{(1)}(2) \]
\[ + \left[ x^{(1)}(2) - wx^{(1)}(1) \right] \left( \lambda - 1 \right) \sum_{j=0}^{k-2} w^j \lambda^{k-2-j} + w^{k-2} \]
\[ + \mu \sum_{j=0}^{k-2} w^j \lambda^{k-2-j}, \quad k = 1, 2, \ldots, n - 1. \quad (15) \]

Now the discrete GMSD(1,1) model has been constructed, and the whole modeling procedure is analyzed.

3.2. Parameters estimation of the discrete GMSD(1,1) model

From the definition of 1-AGO, we have that

\[ x^{(1)}(k + 1) - x^{(1)}(k - 1) = x^{(0)}(k + 1) + x^{(0)}(k). \]

By the Simpson numerical integration formula, the background value of \( X^{(1)} \) is defined as

\[ z^{(1)}(k) = \frac{x^{(1)}(k - 1) + 4x^{(1)}(k) + x^{(1)}(k + 1)}{6}. \]

Thus the Eq. (9) can be rewritten as below

\[ x^{(0)}(k + 1) + x^{(0)}(k) + az^{(1)}(k) = 2b. \quad (16) \]

Employing the least squares estimation method, from Eq. (16) by considering \( k = 2, 3, \ldots, n - 1 \), the model parameters \( a \) and \( b \) can be given as

\[ \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \left( B^T B \right)^{-1} B^T Y, \quad (17) \]

where \( B \) and \( Y \) are defined as follows

\[ B = \begin{pmatrix} \frac{x^{(1)}(2) + 4x^{(1)}(3)}{6} \\ \frac{x^{(1)}(3) + 4x^{(1)}(4)}{6} \\ \vdots \\ \frac{x^{(1)}(n-2) + 4x^{(1)}(n-1) + x^{(1)}(n)}{6} \end{pmatrix}, \quad Y = \begin{pmatrix} \frac{x^{(0)}(2) + x^{(0)}(3)}{2} \\ \frac{x^{(0)}(3) + x^{(0)}(4)}{2} \\ \vdots \\ \frac{x^{(0)}(n-1) + x^{(0)}(n)}{2} \end{pmatrix}. \]
3.3. Difference between GM\textsubscript{SC}(1,1) and GM\textsubscript{SD}(1,1) models

This subsection discusses the difference between the continuous GM\textsubscript{SC}(1,1) model and the discrete GM\textsubscript{SD}(1,1) model. In the paper of He and Wang \cite{32}, the time response function is expressed by

\[ \hat{x}^{(1)}(k + 1) = \left( x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a}, \quad k = 1, 2, \ldots, n - 1, \]  

(18)

and the restored values of \( \hat{x}^{(0)}(k + 1) \) is given by

\[ \hat{x}^{(0)}(k + 1) = \hat{x}^{(1)}(k + 1) - \hat{x}^{(1)}(k) = \frac{e^a - 1}{a} \left( b - ax^{(0)}(1) \right) e^{-ak}, \]

\[ k = 1, 2, \ldots, n - 1. \]  

(19)

They are the same as the ones of the classical continuous GM(1,1) model provided in Section 2. The system parameters \( a \) and \( b \) in Eqs. 18 and 19 are derived from the least squares estimation solution of the Eq. 16. Obviously, the function 18 must coincidence with the difference Eq. 16, otherwise the continuous GM\textsubscript{SC}(1,1) model will not be accurate. Substituting the expression 18 into the Eq. 16, the left side of Eq. 16 turns to be

\[ L(t) = x^{(0)}(k + 1) + x^{(0)}(k) + ax^{(1)}(k) \]

\[ = x^{(1)}(k + 1) - x^{(1)}(k) + a \left( \frac{x^{(1)}(k - 1) + 4x^{(1)}(k) + x^{(1)}(k + 1)}{3} \right) \]

\[ = \frac{1}{3} \left( (a + 3) x^{(1)}(k + 1) + 4ax^{(1)}(k) + (a - 3) x^{(1)}(k - 1) \right) \]

\[ = \frac{1}{3} \left( x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} \left[ (a + 3) + 4ae^a + (a - 3)e^{2a} \right] \]

\[ + \frac{b}{3a} \left[ (a + 3) + 4a + (a - 3) \right] \]

\[ = \frac{1}{3} \left( x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} \left[ (a + 3) + 4ae^a + (a - 3)e^{2a} \right] + 2b. \]  

(20)

The right side of Eq. 16 is

\[ R(t) = 2b. \]  

(21)

Let \( \phi(a) = (a + 3) + 4ae^a + (a - 3)e^{2a} \), we obtain the following numerical result displayed in Table I and Fig. I.
Table 1: Computation results of function \( \phi(a) \) under different values of \( a \)

| \(a\) | \(\phi(a)\) | \(a\) | \(\phi(a)\) | \(a\) | \(\phi(a)\) |
|------|-----------|------|-----------|------|-----------|
| 0.00 | 0         | 0.35 | \(2.4989\times10^{-4}\) | 0.70 | 0.0115    |
| 0.05 | \(1.0952\times10^{-8}\) | 0.40 | \(5.1310\times10^{-4}\) | 0.75 | 0.0172    |
| 0.10 | \(3.6857\times10^{-7}\) | 0.45 | \(9.7400\times10^{-4}\) | 0.80 | 0.0251    |
| 0.15 | \(2.9440\times10^{-6}\) | 0.50 | 0.0017     | 0.85 | 0.0358    |
| 0.20 | \(1.3053\times10^{-5}\) | 0.55 | 0.0029     | 0.90 | 0.0503    |
| 0.25 | \(4.1922\times10^{-5}\) | 0.60 | 0.0048     | 0.95 | 0.0696    |
| 0.30 | \(1.0981\times10^{-4}\) | 0.65 | 0.0076     | 1.00 | 0.0950    |

Figure 1: Function \( \phi(a) \) for different values of \( a \)

One checks easily that when \(|a|\) approximately to zero, the first term of Eq. (20) is approximately to zero. In this situation, we can say \( L(t) = R(t) \). However, when \(|a|\) is large (\(a \geq 1.5\)), the errors \( L(t) - R(t) \) will be quite large. That implies the function (18) will not coincidence with the difference Eq. (16), and the continuous GM\(_{SC}(1,1)\) model may not be accurate. On the other hand, the discrete function (14) is exactly the solution of the difference Eq. (16). This means the performance of the discrete GM\(_{SD}(1,1)\) model is not limited to the value of system parameters.

In the above analysis, the difference between the continuous GM\(_{SC}(1,1)\) model and the discrete GM\(_{SD}(1,1)\) model is that the modelling accuracy of
the former depends on system parameters’ value, while the later does not. This is the advantage of the discrete model compared to the continuous one.

3.4. Unbiased property of the discrete GM\(_{(1,1)}\) model

This subsection proves the discrete GM\(_{(1,1)}\) model is unbiased to simulate the homogeneous exponent sequence. Set the sequence is \(\{rq^k, k = 1, 2, \ldots, n\}\), then the following original sequence as \(X^{(0)} = (rq, rq^2, \ldots, rq^n)\). One checks easily that

\[
x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i) = \frac{rq(1-q^k)}{1-q}, k = 1, 2, \ldots, n.
\]

The 1-AGO of \(X^{(0)}\) is given by

\[
X^{(1)} = \left\{ \frac{rq(1-q^2)}{1-q}, \frac{rq(1-q^3)}{1-q}, \ldots, \frac{rq(1-q^n)}{1-q} \right\}.
\]

Substituting these values into the matrix \(B\) and \(Y\), it follows that

\[
B = \begin{bmatrix}
-\frac{6rq-rq^2(1+4q+q^2)}{6(1-q)} & 1 \\
\frac{6rq-rq^2(1+4q+q^2)}{6(1-q)} & 1 \\
\vdots & \vdots \\
\frac{6rq-rq^{n-1}(1+4q+q^2)}{6(1-q)} & 1 \\
\end{bmatrix}, \quad Y = \begin{bmatrix}
\frac{rq^2(1+q)}{2} \\
\frac{rq^3(1+q)}{2} \\
\vdots \\
\frac{rq^{n-1}(1+q)}{2} \\
\end{bmatrix}.
\]

After some calculations, we known that

\[
\begin{bmatrix}
a \\
b \\
\end{bmatrix} = (B^T B)^{-1} B^T Y = \begin{bmatrix}
\frac{3(1-q^2)}{1+4q+q^2} \\
\frac{3rq(1+q)}{1+4q+q^2} \\
\end{bmatrix}.
\]

(22)

From Eq. (22), we can easily obtain

\[
w = q, \quad \lambda = -\frac{q + 2}{2q + 1}, \quad \mu = \frac{3rq(1 + q)}{2q + 1}.
\]
Substituting the three values into the Eq. (15), it yields that

\[
\hat{x}^{(0)}(k+1) = q^{k-2}r q \frac{(1-q^2)}{1-q} (q-1) \\
+ \left( r q \frac{(1-q^2)}{1-q} - qr q \right) \left( -\frac{3q + 3}{2q+1} \sum_{j=0}^{k-2} w^j \lambda^{k-2-j} + q^{k-2} \right) \\
+ \frac{3qr(q+1)}{2q+1} \sum_{j=0}^{k-2} w^j \lambda^{k-2-j} \\
= -rq^{k-1} + rq^{k+1} + qr \left( -\frac{3(q+1)}{2q+1} \sum_{j=0}^{k-2} w^j \lambda^{k-2-j} + q^{k-2} \right) \\
+ \frac{3qr(q+1)}{2q+1} \sum_{j=0}^{k-2} w^j \lambda^{k-2-j} \\
= rq^{k+1} = x^{(0)}(k+1).
\]

Eq. (23) indicates that the homogeneous exponent simulative unbiased property is met.

3.5. Modelling evaluation criteria

To examine the prediction accuracy of the GM(1,1) model, the absolute percentage error (APE) and the mean absolute percentage error (MAPE) are adopted in this paper. They are defined as follows

\[
\text{APE}(k) = \frac{\left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%}{}, \quad k = 2, 3, \ldots, n, \quad (24)
\]

\[
\text{MAPE} = \frac{1}{m - \ell + 1} \sum_{k=\ell}^{m} \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%{}, \quad m \leq n. \quad (25)
\]

From Eq. (24), APE\((k), k = 2, 3, \ldots, \nu\) is referred to as the absolute simulation percentage error at time \(k\), while APE\((k), k = \nu + 1, \nu + 2, \ldots, n\) is referred to as the absolute prediction percentage error at time \(k\). Further, when \(\ell = 2, m = \nu\), the MAPE is the mean absolute simulation percentage error termed MAPE\(_{\text{simu}}\), when \(\ell = \nu + 1, m = n\), the MAPE is the mean absolute prediction percentage error termed MAPE\(_{\text{pred}}\), and when \(\ell = 2, m = n\), the MAPE is the overall mean absolute percentage error termed MAPE\(_{\text{over}}\).
4. Validation of the GM_{SD}(1,1) model

This section provides some numerical examples to validate the accuracy of the GM_{SD}(1,1) model compared to the classical GM(1,1) model, the DGM(1,1) model and the GM_{SC}(1,1) model.

4.1. Validation of GM_{SD}(1,1) and GM_{SC}(1,1) models

This subsection provides an example to verify the accuracy of the GM_{SD}(1,1) model and the GM_{SC}(1,1) model to simulate and predict the homogeneous exponent sequence. Let \( x^{(0)}(k) = rq^k, \ k = 1, 2, \ldots, 12, \ r > 0, \) where parameter \( r \) is randomly generated in \([1, 15]\) by the discrete uniform distribution, and parameter \( q \) is given in the intervals \([0.1, 5.0]\) by the step 0.01. We define the following notation in the sequel

\[
\varepsilon = |\hat{a} - a| + |\hat{b} - b|, \tag{26}
\]

where \( \hat{a} \) and \( \hat{b} \) are the estimated parameters of GM_{SD}(1,1) and GM_{SC}(1,1) models, and parameters \( a \) and \( b \) are the provided determined of Eq.(22).

Employing the above specific parameters, the graphs are depicted in Fig. 2.

It can be seen in Fig. 2 that the maximum \( \varepsilon \) is only \( 1.6172 \times 10^{-11} \), which is obvious a truncation error by computer.

![Figure 2: The values of \( \varepsilon \) for different values of \( q \) and \( r \)](image-url)
Further, fixed parameters $r = 0.05$ and $q = 2.25$, results of the $\text{GM}_{\text{SC}}(1,1)$ and the $\text{GM}_{\text{SD}}(1,1)$ models are listed in Table 2. We observe from Table 2 that the maximum absolute simulation percentage error of $\text{GM}_{\text{SC}}(1,1)$ and $\text{GM}_{\text{SD}}(1,1)$ models are, respectively, 1.0397% and $0.1559 \times 10^{-12}\%$, and the maximum absolute prediction percentage error of $\text{GM}_{\text{SC}}(1,1)$ and $\text{GM}_{\text{SD}}(1,1)$ models are, respectively, 2.1037% and $0.0945 \times 10^{-12}\%$. Obviously, the APE of the $\text{GM}_{\text{SD}}(1,1)$ model is caused by the round-off error of computer, while the APE of the $\text{GM}_{\text{SC}}(1,1)$ model is caused by its inconsistency.

| $k$ | actual values | $\text{GM}_{\text{SC}}(1,1)$ model | $\text{GM}_{\text{SD}}(1,1)$ model |
|-----|---------------|-----------------------------------|-----------------------------------|
|     | values        | APE($k$)$\%$                      | values                            | APE($k$)$\%$                      |
| 1   | 0.1125        | 0.1125                            | 0.0000                            | 0.1125                            |
| 2   | 0.2531        | 0.2531                            | 0.0000                            | 0.2531                            |
| 3   | 0.5695        | 0.5667                            | 0.5034                            | 0.5695                            |
|     |               |                                   |                                   | $0.1559 \times 10^{-12}$          |
| 4   | 1.2814        | 1.2727                            | 0.6825                            | 1.2814                            |
|     |               |                                   |                                   | $0.0520 \times 10^{-12}$          |
| 5   | 2.8833        | 2.8584                            | 0.8613                            | 2.8833                            |
|     |               |                                   |                                   | $0.0462 \times 10^{-12}$          |
| 6   | 6.4873        | 6.4199                            | 1.0397                            | 6.4873                            |
|     |               |                                   |                                   | $0.0274 \times 10^{-12}$          |
| 7   | 14.5965       | 14.4187                           | 1.2178                            | 14.5965                           |
|     |               |                                   |                                   | $0.0122 \times 10^{-12}$          |
| 8   | 32.8420       | 32.3837                           | 1.3956                            | 32.8420                           |
|     |               |                                   |                                   | $0.0216 \times 10^{-12}$          |
| 9   | 73.8946       | 72.7321                           | 1.5731                            | 73.8946                           |
|     |               |                                   |                                   | $0.0192 \times 10^{-12}$          |
| 10  | 166.2628      | 163.3528                          | 1.7503                            | 166.2628                          |
|     |               |                                   |                                   | $0.0342 \times 10^{-12}$          |
| 11  | 374.0914      | 366.8821                          | 1.9271                            | 374.0914                          |
|     |               |                                   |                                   | $0.0456 \times 10^{-12}$          |
| 12  | 841.7056      | 823.9990                          | 2.1037                            | 841.7056                          |
|     |               |                                   |                                   | $0.0945 \times 10^{-12}$          |
|     | MAPE$_{\text{simu}}$ | 0.7717                            |                                   | $1.5565 \times 10^{-13}$          |
|     | MAPE$_{\text{pred}}$ | 1.6613                            |                                   | $3.7893 \times 10^{-14}$          |
|     | MAPE$_{\text{over}}$ | 1.3055                            |                                   | $5.0888 \times 10^{-14}$          |

4.2. Validation of $\text{GM}_{\text{SD}}(1,1)$ and other grey models

This subsection further illustrates the advantage of the $\text{GM}_{\text{SD}}(1,1)$ model by using some real cases. We consider the numerical example from the paper [36] to predict total electricity consumption in China during 2005-2014. Data from 2005 to 2011 are applied to develop different grey models, while data from
2012 to 2014 are applied to test. Results are presented in Table 3 showing that the GM_{SP}(1,1) model outperforms the other grey models in this example.

| Year | Data | GM(1,1) | DGM(1,1) | GM_{SC}(1,1) | GM_{SD}(1,1) |
|------|------|---------|----------|--------------|--------------|
| 2005 | 24940.3 | 24940.3000 | 24940.3000 | 24940.3000 | 24940.3000 |
| 2006 | 28588.0 | 28678.0326 | 28701.0256 | 28588.0000 | 28588.0000 |
| 2007 | 32711.8 | 31558.7319 | 31586.3461 | 32127.3503 | 32080.3602 |
| 2008 | 34541.4 | 34728.7965 | 34761.7284 | 34825.5851 | 34472.6239 |
| 2009 | 37032.2 | 38217.2931 | 38256.3326 | 38190.8687 | 38490.1274 |
| 2010 | 41932.5 | 42056.2081 | 42102.2502 | 41881.3482 | 41543.9697 |
| 2011 | 47000.9 | 46280.7409 | 46334.7987 | 45928.4479 | 46203.8856 |
| 2012 | 49762.6 | 50929.6267 | 50992.8462 | 50366.6289 | 50042.7998 |
| 2013 | 54203.4 | 56045.4916 | 56119.1683 | 55233.6827 | 55485.5079 |
| 2014 | 56383.7 | 61675.2435 | 61760.8406 | 60571.0519 | 60258.6361 |
| MAPE_{simu} | 1.5675% | 1.5994% | 1.6293% | 1.7387% |
| MAPE_{pred} | 5.0428% | 5.1811% | 3.5137% | 3.2669% |
| MAPE_{over} | 2.7260% | 2.7933% | 2.3360% | 2.3118% |

5. Applications

In this section, the discrete GM_{SP}(1,1) model is used to predict the Gross Domestic Product (GDP) and the freightage of Lanzhou.

5.1. Forecasting the Gross Domestic Product of Lanzhou

Raw data of Lanzhou was collected from the website of the National Bureau of Statistics of China. The total Gross Domestic Product is measured in hundred million RMB. These real data from 2004 to 2009 are applied to build the prediction models, and the ones from 2010 to 2015 are applied for validation. The simulation and prediction results are listed in Table 4 while the errors are listed in Table 5 and Fig. 3.

From Tables 4 and 5 and Fig. 3 that four grey models have successfully caught the trend of the GDP. The GM_{SP}(1,1) model for the mean absolute prediction percentage error and the overall mean absolute percentage error are...
7.6118% and 5.0454%, respectively, which have the smallest errors among four grey models. Fig. 3 also indicates that the accuracy of GMSD(1,1) model is the best, and the GM(1,1) model is the worst.

### Table 4: Simulation and prediction results of GDP of Lanzhou

| Year | Data   | GM(1,1)   | DGM(1,1)  | GMSC(1,1) | GMSP(1,1) |
|------|--------|-----------|-----------|-----------|-----------|
| 2004 | 504.65 | 504.6500  | 504.6500  | 504.6500  | 504.6500  |
| 2005 | 567.04 | 568.6831  | 569.4678  | 567.0400  | 567.0400  |
| 2006 | 638.47 | 644.1549  | 645.1480  | 643.2803  | 642.0348  |
| 2007 | 732.76 | 729.6429  | 730.8858  | 731.7331  | 733.1665  |
| 2008 | 846.28 | 826.4761  | 828.0179  | 832.3484  | 831.2488  |
| 2009 | 926.00 | 936.1605  | 938.0585  | 946.7986  | 948.1637  |
| 2010 | 1100.40| 1060.4014 | 1062.7230 | 1076.9861 | 1076.0331 |
| 2011 | 1360.03| 1201.1307 | 1203.9551 | 1225.0746 | 1226.3916 |
| 2012 | 1563.80| 1360.5367 | 1363.9563 | 1393.5258 | 1392.7242 |
| 2013 | 1776.28| 1541.0980 | 1545.2212 | 1585.1395 | 1586.4313 |
| 2014 | 2000.94| 1745.6222 | 1750.5754 | 1803.1005 | 1802.4595 |
| 2015 | 2095.99| 1977.2894 | 1983.2206 | 2051.0319 | 2052.3253 |

### Table 5: Relative error values of GDP of Lanzhou by grey models (%)

| Year | GM(1,1) | DGM(1,1) | GMSC(1,1) | GMSP(1,1) |
|------|---------|----------|-----------|-----------|
| 2004 | 0       | 0        | 0         | 0         |
| 2005 | 0.2898  | 0.4281   | 0         | 0         |
| 2006 | 0.8904  | 1.0459   | 0.7534    | 0.5583    |
| 2007 | 0.4254  | 0.2558   | 0.1401    | 0.0555    |
| 2008 | 2.3401  | 2.1579   | 1.6462    | 1.7762    |
| 2009 | 1.0972  | 1.3022   | 2.2461    | 2.3935    |
| 2010 | 3.6349  | 3.4239   | 2.1278    | 2.2144    |
| 2011 | 11.6835 | 11.4758  | 9.9230    | 9.8261    |
| 2012 | 12.9980 | 12.7794  | 10.8885   | 10.9397   |
| 2013 | 13.2401 | 13.0080  | 10.7607   | 10.6880   |
| 2014 | 12.7999 | 12.5123  | 9.8873    | 9.9194    |
| 2015 | 5.6632  | 5.3802   | 2.1450    | 2.0832    |

**MAPE_{simu}** | **1.0086** | **1.0379** | **1.1869** | **1.1959** |
**MAPE_{pred}** | **9.9966** | **9.7633** | **7.6220** | **7.6118** |
**MAPE_{over}** | **5.9111** | **5.7972** | **5.0518** | **5.0454** |
5.2. Forecasting the freightage of Lanzhou

The raw data of the freightage of Lanzhou was also collected from the website of the National Bureau of Statistics of China. The total freightage is measured in ten thousand tons. Similarly, we divided these data into two groups, in which the first 6 samples are applied to build the prediction models, and the left samples are used to check and compare the forecasting results. The simulation and prediction results are listed in Table 6, while the errors are listed in Table 7 and Fig. 4.

We observe from Table 7 and Fig. 4 that the MAPE_{simu}, MAPE_{pred} and MAPE_{over} of GM_{SD}(1,1) model are 1.8007%, 6.3579% and 4.3325%, respectively. MAPE_{simu}, MAPE_{pred} and MAPE_{over} of GM(1,1) are 1.5632%, 8.4588% and 5.0110%, those of DGM(1,1) are 1.5687%, 8.4276% and 4.9981%, and those of GM_{SC}(1,1) model are 1.8499%, 6.3810% and 4.3383%, respectively.

Here the GM_{SD}(1,1) model for the mean absolute prediction percentage error and the overall mean absolute percentage error are the smallest errors among four grey models. This also indicates that the accuracy of GM_{SD}(1,1) model is the best, the accuracy of GM_{SC}(1,1) model are inferior to GM(1,1) model and DGM(1,1) model and the GM(1,1) model is the worst to predict the freightage of Lanzhou.
Table 6: Simulation and prediction results of freightage of Lanzhou

| Year | Data | GM(1,1)  | DGM(1,1) | GMSC(1,1) | GMSD(1,1) |
|------|------|----------|----------|-----------|-----------|
| 2004 | 5786 | 5786.0000| 5786.0000| 5786.0000 | 5786.0000 |
| 2005 | 5973 | 6015.9317| 6017.4333| 5973.0000 | 5973.0000 |
| 2006 | 6262 | 6349.6357| 6351.3039| 6346.0123 | 6361.0051 |
| 2007 | 6840 | 6701.8503| 6703.6990| 6724.0011 | 6708.2685 |
| 2008 | 7207 | 7073.6022| 7075.6463| 7124.5042 | 7138.8492 |
| 2009 | 7332 | 7465.9753| 7468.2307| 7548.8625 | 7533.6395 |
| 2010 | 8032 | 7880.1133| 7882.5972| 7998.4969 | 8012.2088 |
| 2011 | 8882 | 8317.2236| 8319.9544| 8474.9130 | 8460.1700 |
| 2012 | 9728 | 8778.5804| 8781.5778| 8979.7061 | 8992.7976 |
| 2013 | 10531| 9265.5286| 9268.140| 9514.5662 | 9500.2731 |
| 2014 | 11147| 9779.4880| 9783.0839| 10081.2842| 10093.7664|

Table 7: Relative error values of freightage of Lanzhou by grey models (%)

| Year | GM(1,1) | DGM(1,1) | GMSC(1,1) | GMSD(1,1) |
|------|---------|----------|-----------|-----------|
| 2004 | 0       | 0        | 0         | 0         |
| 2005 | 0.7188  | 0.7439   | 0         | 0         |
| 2006 | 1.3995  | 1.4261   | 1.3416    | 1.5810    |
| 2007 | 2.0197  | 1.9927   | 1.6959    | 1.9259    |
| 2008 | 1.8509  | 1.8226   | 1.1447    | 0.9456    |
| 2009 | 1.8273  | 1.8580   | 2.9578    | 2.7501    |
| 2010 | 1.8910  | 1.8601   | 0.4171    | 0.2464    |
| 2011 | 6.3587  | 6.3279   | 4.5833    | 4.7493    |
| 2012 | 9.7597  | 9.7288   | 7.6922    | 7.5576    |
| 2013 | 12.0166 | 11.9854  | 9.6518    | 9.7875    |
| 2014 | 12.2680 | 12.2357  | 9.5606    | 9.4486    |

| | MAPE_{simu} | MAPE_{pred} | MAPE_{over} |
|---|-------------|-------------|-------------|
| 2004 | 1.5632      | 8.4588      | 5.0110      |
| 2005 | 1.5687      | 8.4276      | 4.9981      |
| 2006 | 1.8499      | 6.3810      | 4.3383      |
| 2007 | 1.8007      | 6.3579      | 4.3325      |

17
6. Conclusions

The current study studied the discrete GM(1,1) model with Simpson formula called GM_{SD}(1,1) model. Mathematical analysis is presented to indicate the difference between the GM_{SD}(1,1) model and the GM_{SC}(1,1) model. We also proved our model is unbiased to simulate the homogeneous exponent sequence. Applications are carried out to verify the performance of our model with the other three models. Computation results indicate that GM_{SD}(1,1) model provides accurate prediction, outperforming GM(1,1), DGM(1,1) and GM_{SC}(1,1) model.

It may be remarked here that the idea for GM_{SD}(1,1) model used in our paper can be used to analyze other grey forecasting model such as GM(1,n) or GMC(1,n) model. These are possible extensions and suggested directions for our future research.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (No.71901184, 71771033, 71571157, 11601357), the Humanities and Social Science Project of Ministry of Education of China (No.19YJCZH119), the Longshan academic talent research supporting program of SWUST (No.17LZXY20),
the Open Fund (PLN 201710) of State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation (Southwest Petroleum University), Applied Basic Research Program of Science and Technology Commission Foundation of Sichuan province (2017JY0159), the funding of V.C. & V.R. Key Lab of Sichuan Province (SCVCVR2018.08VS, SCVCVR2018.10VS), the Doctoral Research Foundation of Southwest University of Science and Technology (No.15zx7141, 16zx7140), and the National Statistical Scientific Research Project (No.2018LY42).

References

[1] J. Deng, Control problems of grey systems, Systems & Control Letters 1 (5) (1982) 288–294.

[2] Q. Xiao, C. Li, Y. Tang, L. Li, L. Li, A knowledge-deriven method of adaptively optimizing process parameters for energy efficient turning, Energy 166 (2019) 142–156.

[3] X. Wu, X. Ding, T. Lu, J. Wang, Topological dynamics of Zadeh’s extension on upper semi-continuous fuzzy sets, International Journal of Bifurcation and Chaos 27 (10) (2017) 1–13.

[4] P. Du, J. Wang, W. Yang, T. Niu, Multi-step ahead forecasting in electrical power system using a hybrid forecasting system, Renewable Energy 122 (7) (2018) 533–550.

[5] M. Ma, W. Cai, Y. Wu, China act on the energy efficiency of civil buildings (2008): a decade review, Science of The Total Environment 651 (15) (2019) 42–60.

[6] Y. Wang, X. Yi, Transient pressure behavior of a fractured vertical well with a finite-conductivity fracture in triple media carbonate reservoir, Journal of Porous Media 20 (8) (2017) 707–722.
[7] Y. Wang, X. Yi, Flow modeling of well test analysis for a multiple-fractured horizontal well in triple media carbonate reservoir, International Journal of Nonlinear Sciences and Numerical Simulation 19 (5) (2018) 439–457.

[8] S.-J. Huang, C.-L. Huang, Control of an inverted pendulum using grey prediction model, IEEE Transaction on Industry Application 36 (2) (2000) 452–458.

[9] S.-C. Chang, H.-C. Lai, H.-C. Yu, A variable P value rolling grey forecasting model for Taiwan semiconductor industry production, Technological Forecasting & Social Change 72 (5) (2005) 623–640.

[10] C.-I. Chen, P.-H. Hsin, C.-S. Wu, Forecasting Taiwan’s major stock indices by the Nash nonlinear grey Bernoulli model, Expert Systems with Applications 37 (12) (2010) 7557–7562.

[11] Z. Wang, Q. Li, L. Pei, Grey forecasting method of quarterly hydropower production in China based on a data grouping approach, Applied Mathematical Modelling 51 (11) (2017) 302–316.

[12] Z. Wang, H. Zheng, L. Pei, T. Jin, Decomposition of the factors influencing export fluctuation in China’s new energy industry based on a constant market share model, Energy Policy 109 (10) (2017) 22–35.

[13] B. Zeng, H. Duan, Y. Bai, W. Meng, Forecasting the output of shale gas in china using an unbiased grey model and weakening buffer operator, Energy 151 (15) (2018) 238–249.

[14] W. Zhou, J. He, Generalized GM(1,1) model and its application in forecasting of fuel production, Applied mathematical Modelling 37 (9) (2013) 6234–6243.

[15] X. Ma, Z. Liu, The GMC(1, n) model with optimized parameters and its applications, Journal of Grey System 29 (4) (2017) 121–137.
[16] W. Wu, X. Ma, B. Zeng, Y. Wang, W. Cai, Application of the novel fractional grey model FAGMO(1,1,k) to predict China’s nuclear energy consumption, Energy 165 (2018) 223–234.

[17] W. Wu, X. Ma, B. Zeng, Y. Wang, W. Cai, Forecasting short-term renewable energy consumption of China using a novel fractional nonlinear grey Bernoulli model, Renewable Energy 140 (2019) 70–87.

[18] X. Ma, M. Xie, W. Wu, X. Wu, B. Zeng, A novel fractional time delayed grey model with grey wolf optimizer and its applications in forecasting the natural gas and coal consumption in Chongqing China, Energy 178 (2019) 487–507.

[19] W. Wu, X. Ma, Y. Zhang, Y. Wang, X. Wu, Analysis of novel FAGM(1,1,tα) model to forecast health expenditure of China, Grey Systems: Theory and Application 9 (2) (2019) 232–250.

[20] B. Zeng, Y. Tan, H. Xu, J. Quan, L. Wang, X. Zhou, Forecasting the electricity consumption of commercial sector in Hong Kong using a novel grey dynamic prediction model, Journal of Grey System 30 (1) (2018) 157–172.

[21] L. Wu, Y. Gao, Y. Yang, X. Chen, Using a novel multi-variable grey model to forecast the electricity consumption of Shandong province in China, Energy 157 (15) (2018) 327–335.

[22] L. Zhang, Y. Zheng, K. Wang, X. Zhang, Y. Zheng, An optimized Nash nonlinear grey Bernoulli model based on particle swarm optimization and its application in prediction for the incidence of Hepatitis B in Xinjiang, China, Computers in Biology & Medicine 49 (1) (2014) 67–73.

[23] N. Xie, S. Liu, Discrete grey forecasting model and its optimization, Applied Mathematical modelling 33 (2) (2009) 1173–1186.
[24] B. Zeng, H. Duan, Y. Zhou, A new multivariable grey prediction model with structure compatibility, Applied Mathematical Modelling 75 (2019) 385–397.

[25] C.-I. Chen, H. Chen, S. Chen, 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 13 (6) (2008) 1194–1204.

[26] L. Kong, X. Ma, Comparison study on the nonlinear parameter optimization of nonlinear grey Bernoulli model (NGBM(1,1)), Grey Systems: Theory and Application 8 (2) (2018) 210–226.

[27] L. Pei, Q. Li, Z. Wang, The NLS-based nonlinear grey Bernoulli model with an application to employee demand prediction of high-tech enterprises in China, Grey Systems: Theory and Application 8 (2) (2018) 133–143.

[28] M. Salehi, N. Dehnavi, Audit report forecast: an application of nonlinear grey Bernoulli model, Grey Systems: Theory and Application 8 (3) (2018) 295–311.

[29] D. Q. Truong, K. K. Ahn, An accurate signal estimator using a novel smart adaptive grey model SAGM(1,1), Expert Systems with Applications 39 (9) (2012) 7611–7620.

[30] J. Cui, S. Liu, B. Zeng, N. Xie, A novel grey forecasting model and its optimization, Applied Mathematical Modelling 37 (6) (2013) 4399–4406.

[31] B. Zeng, C. Li, Improved multi-variable grey forecasting model with a dynamic background-value coefficient and its application, Computers & Industrial Engineering 118 (4) (2018) 278–290.

[32] M. He, Q. Wang, Constructing the background value for GM(1,1) model based on Simpson formula (in chinese), Journal of Quantitative Economics 28 (4) (2011) 101–104.
[33] J. Xia, X. Ma, W. Wu, B. Huang, W. Li, Application of a new information priority accumulated grey model with time power to predict short-term wind turbine capacity, Journal of Cleaner Production 244 (2020) 118573.

[34] X. Ma, A brief introduction to the grey machine learning, Journal of Grey System 31 (1) (2019) 1–12.

[35] C.-N. Wang, V.-T. Phan, An improved nonlinear grey Bernoulli model combined with Fourier series, Mathematical Problems in Engineering 2015 (2015) 1–7.

[36] S. Ding, K. W. Hipel, Y. Dang, Forecasting China’s electricity consumption using a new grey prediction model, Energy 149 (15) (2018) 314–328.