Journal of Systems Science and Information
Dec., 2014, Vol. 2, No. 6, pp. 543–552
DOI: 10.1515/JSSI-2014-0543

Optimizing the Grey GM(1,N) Model by Rebuilding All the Back Ground Values

Xin MA
School of Science, Southwest Petroleum University, Chengdu 610500*, China*
E-mail: *cauchy*7203*@gmail.com*

Zhibin LIU
School of Science, Southwest Petroleum University, Chengdu 610500*, China*
E-mail: *liuzhibinswpi@vip.sina.com*

Yishen CHEN
China Petroleum Engineering Construction Corporation, Beijing, China
E-mail: *tsinghua*616*@*163*.com*

Abstract The GM(1,N) model is a very important prediction model of the grey system. But the inherent defect of GM(1,N), which may cause very large error, is still there. This paper analyzes the source of the error of GM(1,N) and reveals that it's all the back ground values that effect the precision and applicability of GM(1,N). Three methods are employed to revise the GM(1,N) model. The simulation test shows the new models perform with higher precision and robustness. Even in some extreme cases, in which the original GM(1,N) is invalid, the new models are still valid and perform well.

Keywords grey system theory; GM(1,N) model; back ground value; Gauss-Legendre formula

1 Introduction

The prediction models[1] are important parts of the grey system models. Being easy to implement and simply structured, the grey prediction models have been applied to a range of areas, varying in economics[2], industry[3, 4], aerography[5] and so on. But there still exist some defects in the grey prediction models, which appeals a plenty of researches.

Tan's[6] research revealed that it was the structure of back ground value of the GM(1,1) model effects the precision of the model, and also for the generalized models[7]. Luo, Liu, et al.[8] pointed out that the back ground value was the integration of the 1-AGO sequence, and induced the formulation of the back ground value. Based on the research of Luo, researchers employed a plenty of different numeric integral formulas to structure the back ground value, such as Gauss-Legendre formula[9, 10], Simpson formula[11], Newtown-Cortes[12], and the revised models all performed well than the original GM(1,1). Other researchers such as Li[13], Li[14] and Wang[15], structured a lot of formulations to compute the back ground value and improved the precision of the GM(1,1) model and its generalized models. These researches proved that restructuring of the back ground value was capable to improve the precision of the GM(1,1) model.

Received February 10, 2014, accepted April 23, 2014

As an extension model of GM(1,1), the GM(1,N) is more complex and more useful in the multiple regression problems. It was also used in varied of areas, such as industry[16], economics[17], biology[18] and so on. But the GM(1,N) is still not available in a plenty of situations, in which the error is very large. Several works have been employed to try to solve this problem. As is sharing a lot of similarities to GM(1,1), researchers tried to restructure the back ground value of GM(1,N). Liu[19] structured a formulation of the background value for GM(1,N), which was proved to perform better in predicting the upside-down of the road. Shen, etc.[20] used the Newton-Cotes formula and Gauss-Legendre formula to structure the back ground value, which was proved to be better performing in prediction of the transportation noise. But as its complexity, restructuring the back ground value is still limited, which will be shown in this paper, in some cases, the revised back ground value is still not available to improve the precision of the GM(1,N).

In order to extend the validation of the GM(1,N), Zhai, etc.[21] introduced the MGM(1,N) model, which analysed the grey relationship value of the input data and the output data, and then built N GM(1,1) models. Some improving works have also been done on this model[22–24], which indeed presented higher precision and robustness. However, these methods still failed to revise the GM(1,N) itself, the stationed defects still exist, of which the essence has still not been revealed. And obviously, there has been no way to overcome the defects.

The rest of this paper is organized as follows, Sec.2 introduces the principles of the GM(1,N) model; Sec.3 analyses the essence of the defects of GM(1,N); Sec.4 gives the method to improve the precision and robustness of GM(1,N); Sec.5 presents some simulation tests and some conclusions are drawn in Sec.6.

2 The principles of GM(1,N)

Definition 1 [1]

Set the sequence

$$X_1^{(0)} = (x_1^{(0)}(1), x_1^{(0)}(2), \cdots, x_1^{(0)}(n))$$

as the feature sequence of the system, the sequences

$$\begin{aligned} X_2^{(0)} &= (x_2^{(0)}(1), x_2^{(0)}(2), \cdots, x_2^{(0)}(n)), \\ X_3^{(0)} &= (x_3^{(0)}(1), x_3^{(0)}(2), \cdots, x_3^{(0)}(n)), \\ &\vdots \\ X_N^{(0)} &= (x_N^{(0)}(1), x_N^{(0)}(2), \cdots, x_N^{(0)}(n)) \end{aligned}$$

as the reliance sequences. Generally, all these sequences are nonnegative sequences.

Definition 2 [1] *A sequence $X_i^{(1)} = (x_i^{(1)}(1), x_i^{(1)}(2), \cdots, x_i^{(1)}(n))$ is the 1-AGO (Accumulated Generating Operation) sequence of the $X_i^{(0)}$, which satisfies $X_i^{(1)}(k) = \sum_{m=1}^{k} X_i^{(0)}(m)$.*

Definition 3 [1]

$$x_1^{(0)}(k) + a z_1^{(1)}(k) = \sum_{i=1}^{N} b_i x_i^{(1)}(k) \tag{1}$$

is called the original GM(1,N) model, where $z_1^{(1)}(k) = 0.5(x_1^{(1)}(k) + x_1^{(1)}(k-1))$ is called the mean generation of consecutive neighbors sequence.

Theorem 1 [1] *The least squares estimation for $\hat{\boldsymbol{a}} = [a, b_2, \cdots, b_N]$ of the GM(1,N) model satisfies*

$$\hat{\boldsymbol{a}} = (\boldsymbol{B}^{\boldsymbol{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\boldsymbol{T}}\boldsymbol{Y} \tag{2}$$

where

$$\boldsymbol{B} = \begin{bmatrix} -z_1^{(1)}(2) & x_2^{(1)}(2) & \cdots & x_N^{(1)}(2) \\ -z_1^{(1)}(3) & x_2^{(1)}(3) & \cdots & x_N^{(1)}(3) \\ \vdots & \vdots & & \vdots \\ -z_1^{(1)}(n) & x_2^{(1)}(n) & \cdots & x_N^{(1)}(n) \end{bmatrix}, \boldsymbol{Y} = \begin{bmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \\ \vdots \\ x_1^{(0)}(n) \end{bmatrix}.$$

Definition 4 [1] *The*

$$\frac{\mathrm{d}x_1^{(1)}}{\mathrm{d}t} + ax_1^{(1)} = b_2x_2^{(1)} + b_3x_3^{(1)} + \cdots + b_Nx_N^{(1)} \tag{3}$$

is called the whiten equation of the GM(1,N) model, also called image equation.

Theorem 2 [1]

a) The solution of the whiten equation is

$$\begin{aligned} x_1^{(1)}(t) &= \mathrm{e}^{-at}\left[\sum_{i=2}^{N}\int b_ix_i^{(1)}(t)\mathrm{e}^{at}\mathrm{d}t + x_1^{(0)}(0) - \sum_{i=2}^{N}\int b_ix_i^{(0)}(t)\mathrm{e}^{at}\mathrm{d}t\right] \\ &= \mathrm{e}^{-at}\left[x_1^{(0)}(0) - t\sum_{i=2}^{N}b_ix_i^{(0)}(t) + \sum_{i=2}^{N}\int b_ix_i^{(1)}(t)\mathrm{e}^{at}\mathrm{d}t\right] \end{aligned} \tag{4}$$

b) When the $X_i^{(0)}(i = 2, 3, \cdots, N)$ does not fluctuate violently, the $\sum_{i=1}^{N} b_ix_i^{(1)}(k)$ could be regarded as grey constant, thus the approximation time responding formulation is

$$\hat{x}_1^{(1)}(k+1) = \mathrm{e}^{-ak}\left(x_1^{(1)}(0) - \frac{1}{a}\sum_{i=2}^{N}b_ix_i^{(1)}(k+1)\right) + \frac{1}{a}\sum_{i=2}^{N}b_ix_i^{(1)}(k+1) \tag{5}$$

where $x_1^{(1)}(0)$ values $x_1^{(1)}(1)$.

c) The minus formulation is

$$\hat{x}_1^{(0)}(k+1) = \hat{x}_1^{(1)}(k+1) - \hat{x}_1^{(1)}(k) \tag{6}$$

d) The difference reduction formulation is

$$\hat{x}_1^{(0)}(k) = -az_1^{(1)}(k) + \sum_{i=2}^{N}b_i\hat{x}_i^{(1)}(k) \tag{7}$$

3 The analysis on the source of error for GM(1,N)

Consider the integration of the equation (3), which is

$$x_1^{(1)}(k) - x_1^{(1)}(k-1) + a\int_{k-1}^{k}x_1^{(1)}(t)\mathrm{d}t = \sum_{i=2}^{N}b_i\int_{k-1}^{k}x_i^{(1)}(t)\mathrm{d}t.$$

As $x_1^{(0)}(k) = x_1^{(1)}(k) - x_1^{(1)}(k)$, we have

$$x_1^{(0)}(k) + a\int_{k-1}^{k}x_1^{(1)}(t)\mathrm{d}t = \sum_{i=2}^{N}b_i\int_{k-1}^{k}x_i^{(1)}(t)\mathrm{d}t \tag{8}$$

To the Definition 3, equation (8) is equivalent to the original GM(1,N) model, i.e. the equation (1).

It is obvious that the $\int_{k-1}^{k} x_1^{(1)}(t)\mathrm{d}t$ has been taken place by $z_1^{(1)}(k) = 0.5(x_1^{(1)}(k) + x_1^{(1)}(k-1))$ and the $\int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t$ has been taken place by $x_i^{(1)}(k)$. The $z_1^{(1)}(k)$ is also called the back ground value. In this paper, we call the $z_i^{(1)}(k)$ as the i^{th} back ground value, thus the $z_1^{(1)}(k)$ will be called as the first back ground value in the rest of this paper.

3.1 Errors from the first back ground value

According to the research of Tan[6], the $z_1^{(1)}(k) = 0.5(x_1^{(1)}(k) + x_1^{(1)}(k-1))$ is the 2-point trapezoid formula, which performs a very little algebra precision in integration. And researches proved that, a higher precision integration formula could overcome this defect. Fig.1 shows the reason why the low precision integration formula will cause a higher error.

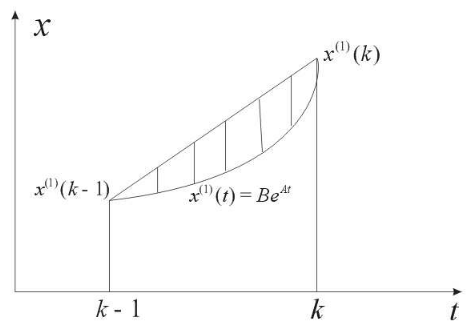

Figure 1 The indication of error which caused by the structure of the first back ground value

3.2 Errors from the second to the N^{th} back ground value

However, no research has paid attention to the right side of the equation (8). The original theorem of GM(1,N) made an assumption "the $x_i^{(1)}(k)$ does not fluctuate violently", and then the $\int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t (i = 2, 3, \cdots, N)$ was taken by $x_i^{(1)}(k)$ itself. But researches have never shown that in what kind of conditions the sequence could be regarded as "not fluctuating violently", and how close is the $\int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t (i = 2, 3, \cdots, N)$ to the $x_i^{(1)}(k)$. Thus, it is still difficult to explain why the GM(1,N) model could not perform well in a plenty of cases.

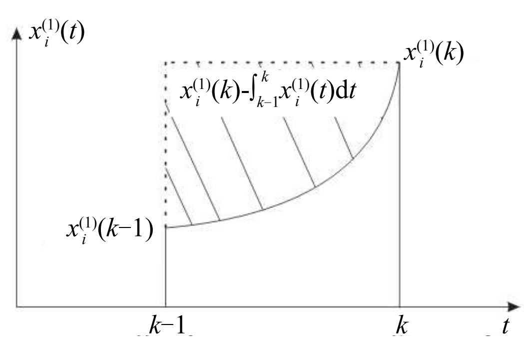

Figure 2 The indication of error which caused by the structure of the $i^{\text{th}}(>1)$ back ground value

Firstly, we consider the geometrical meaning of the i^{th} back ground value $z_i^{(1)}(k) = \int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t$ $(i = 2, 3, \cdots, N)$. As is shown in Fig.2, the $\int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t$ is the area of the trapezoid with curve side, of which the vertices are k, $k-1$, $x_i^{(1)}(k)$ and $x_i^{(1)}(k-1)$, and $x_i^{(1)}(k)$ is the area of the rectangle, of which the length is $x_i^{(1)}(k)$ and width is 1. The shadow area is the redundancy when the $\int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t$ is taken place by $x_i^{(1)}(k)$. This is obviously another reason which cause some significant error to GM(1,N) model.

According to the geometrical analysis, we could draw some results, which reflect the relationship between the error and the original sequence. To prove these results, we need to overview the definition of the convex and concave function and the Hadamard theorem.

Definition 5 *If* $\frac{f(x_1)+f(x_2)}{2} \geq f(\frac{x_1+x_2}{2})$ *or* $\frac{f(x_1)+f(x_2)}{2} \leq f(\frac{x_1+x_2}{2})$ *, then* $f(x)$ *is a convex or concave function, respectively.*

Theorem 3 (Hadmard Theorem) *If* $f(x)$ *is a convex or concave function, there must be*

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b f(x)\mathrm{d}x \leq \frac{f(a)+f(b)}{2}$$

or

$$f\left(\frac{a+b}{2}\right) \geq \frac{1}{b-a}\int_a^b f(x)\mathrm{d}x \geq \frac{f(a)+f(b)}{2}$$

respectively.

Remark 4 If the 1-AGO sequence $X_i^{(1)} = (x_i^{(1)}(1), x_i^{(1)}(2), \cdots, x_i^{(1)}(n))$ is a linear sequence, i.e. $x_i^{(1)}(k) = \alpha k + \beta (i = 1, 2, \cdots, N, k = 1, 2, \cdots, n)$, then $x_i^{(1)}(k) - \int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t = 0.5x_i^{(0)}(k)$.

Proof As $x_i^{(1)}(k) = \alpha k + \beta$, thus

$$x_i^{(1)}(k) - \int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t = (\alpha k + \beta) - \left\{\frac{\alpha[k^2 - (k-1)^2]}{2} + \beta\right\} = 0.5\alpha = 0.5x_i^{(0)}(k).$$

Remark 5 If the 1-AGO sequence $X_i^{(1)} = (x_i^{(1)}(1), x_i^{(1)}(2), \cdots, x_i^{(1)}(n))$ is a concave sequence, then $x_i^{(1)}(k) - \int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t < 0.5x_i^{(0)}(k)$.

Proof As $X_i^{(1)} = (x_i^{(1)}(1), x_i^{(1)}(2), \cdots, x_i^{(1)}(n))$ is a strict concave sequence, $x_i^{(1)}(t)$ approaches to a concave function. To the Hadamard theorem and differential mean value theorem, we have

$$\int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t = x_i^{(1)}(\xi_k) > \frac{x_i^{(1)}(k) + x_i^{(1)}(k-1)}{2}, \text{ where } k-1 < \xi_k < k.$$

Then

$$x_i^{(1)}(k) - \int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t < x_i^{(1)}(k) - \frac{x_i^{(1)}(k) + x_i^{(1)}(k-1)}{2} = 0.5x_i^{(0)}(k).$$

Remark 6 If the 1-AGO sequence $X_i^{(1)} = (x_i^{(1)}(1), x_i^{(1)}(2), \cdots, x_i^{(1)}(n))$ is a convex sequence, then $x_i^{(1)}(k) - \int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t > 0.5x_i^{(0)}(k)$.

Proof In a similar way to the Remark 5, we have

$$x_i^{(1)}(k) - \int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t > x_i^{(1)}(k) - \frac{x_i^{(1)}(k) + x_i^{(1)}(k-1)}{2} = 0.5x_i^{(0)}(k).$$

It could be easily seen based on the above three results that only when the 1-AGO sequence $X_i^{(1)}$ is a strict concave sequence and the original sequence $X_i^{(0)}$ is very small, the $\int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t$ could be taken place by $x_i^{(1)}(k)$, and then the least squares estimation of the Eq.(1) is close enough to its real value. Thus, the assumption in Theorem 2(b), which says "the $x_i^{(1)}(k)$ does not fluctuate violently" could be translated to " the $x_i^{(1)}(k)$ is a concave sequence and not too large". But if the $X_i^{(1)}$ is linear or convex, the difference between $x_i^{(1)}(k)$ and $\int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t$ could not be ignored anymore, and the original operations of GM(1,N) may cause very large error, and this is exactly the inherent defect of the original GM(1,N) theory. Hence, the assumption in Theorem 2(b) is a too strong condition, and could not be satisfied in a lot of situations. But we want the GM(1,N) to be available in more cases, such as when the $x_i^{(1)}(k)$ is concave or linear sequence. To extend the GM(1,N) and overcome its inherent defect original GM(1,N), we need to consider the following theorem.

Theorem 7 [1] *If the original sequence $X_i^{(0)}$ is a non-negative quasi-smooth sequence, its 1-AGO sequence $X_i^{(1)}$ satisfies the approximation exponential law.*

A so-called "non-negative quasi-smooth sequence" is used most frequently in the grey modeling. Within the above discussion and Theorem 7, we could see that for any non-negative quasi-smooth sequence $X_i^{(0)}$, the original theory of GM(1,N) is not available at all. However, as the 1-AGO sequence $X_i^{(1)}$ satisfies the approximation exponential law, it is reasonable to set $x_i^{(1)}(t) = B_i \mathrm{e}^{A_i t}$, and then the i^{th} background value is $\int_{k-1}^{k} B_i \mathrm{e}^{A_i t}\mathrm{d}t$. Now, the i^{th} background value is an integration of the exponential function. If an appropriate numeric formula is employed to compute its real value, the precision and applicability of GM(1,N) could be improved. The following section will present the details of rebuilding all the background values using different numeric formula.

4 Methods to improve the GM(1,N)

According to the analysis, it is the back ground values that effect the error of the GM(1,N) model. Thus, to improve the GM(1,N) model, we need to choose better ways to compute all the back ground values. Being similar to the first back ground value, the second to N^{th} back ground values are essentially integrations. Thus, using appropriate numeric integral formula could reduce the error of GM(1,N).

4.1 Trapezoid formula

In the original GM(1,N), the first back ground value is set as

$$z_1^{(1)}(k) = 0.5(x_1^{(1)}(k) + x_1^{(1)}(k-1)),$$

which is essentially the trapezoid formula of the $z_1^{(1)}(k) = \int_{k-1}^{k} x_1^{(1)}(t)\mathrm{d}t$. Thus, the second to N^{th} back ground values could also be computed by the trapezoid formula, i.e. for all the i, we have

$$z_i^{(1)}(k) = \int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t = 0.5(x_i^{(1)}(k) + x_i^{(1)}(k-1)) \tag{9}$$

4.2 Logarithmic form

As is shown in Theorem 7, the 1-AGO $x_i^{(1)}(k)$ satisfies the approximation exponential law, thus to the knowledge of grey system, it could be fitted by an exponential function. Set

$x_i^{(1)}(t) = B_i \mathrm{e}^{A_i t}$, thus

$$z_i^{(1)}(k) = \int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t = \frac{1}{A_i}(B_i \mathrm{e}^{A_i k} - B_i \mathrm{e}^{A_i (k-1)}) = \frac{1}{A_i}(x_i^{(1)}(k) - x_i^{(1)}(k-1)) \tag{10}$$

and also

$$\frac{x_i^{(1)}(k)}{x_i^{(1)}(k-1)} = \frac{B_i \mathrm{e}^{A_i k}}{B_i \mathrm{e}^{A_i (k-1)}} \tag{11}$$

thus

$$A_i = \ln x_i^{(1)}(k) - \ln x_i^{(1)}(k-1) \tag{12}$$

Combining the (11) and (12), we have

$$z_i^{(1)}(k) = \frac{x_i^{(1)}(k) - x_i^{(1)}(k-1)}{\ln x_i^{(1)}(k) - \ln x_i^{(1)}(k-1)} \tag{13}$$

4.3 Gauss-Legendre formula

The two points Gauss-Legendre formula is

$$\int_{-1}^{1} f(x)\mathrm{d}x \approx f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right) \tag{14}$$

Take $x_i^{(1)}(t) = B_i \mathrm{e}^{A_i t}$ and $A_i = \ln x_i^{(1)}(k) - \ln x_i^{(1)}(k-1)$ into formula (14), we have

$$z_i^{(1)}(k) = \int_{k-1}^{k} x_i^{(1)}(t)\mathrm{d}t = 0.5 x_i^{(1)}(k-0.5) \left[\left(\frac{x_i^{(1)}(k-1)}{x_i^{(1)}(k)}\right)^{\frac{1}{2\sqrt{3}}} + \left(\frac{x_i^{(1)}(k)}{x_i^{(1)}(k-1)}\right)^{\frac{1}{2\sqrt{3}}}\right] \tag{15}$$

Set $x_i^{(1)}(k-0.5) = 0.5(x_i^{(1)}(k) + x_i^{(1)}(k-1))$, thus

$$z_i^{(1)}(k) = 0.25[x_i^{(1)}(k) + x_i^{(1)}(k-1)] \left[\left(\frac{x_i^{(1)}(k-1)}{x_i^{(1)}(k)}\right)^{\frac{1}{2\sqrt{3}}} + \left(\frac{x_i^{(1)}(k)}{x_i^{(1)}(k-1)}\right)^{\frac{1}{2\sqrt{3}}}\right] \tag{16}$$

5 Simulation test

The raw data NO.1 to NO.7 is taken from reference [25], and raw data NO.8 is taken from reference [1]. All the raw data is shown in Table 1 to Table 8 below.

Table 1 Raw data NO.1

t	1	2	3	4	5	6	7	8
X_1	9941	10608	10389	10160	10427.8	10495.8	10563.8	10679.3
X_2	4306.9	4526.3	4570.6	4694.7	4840.2	4980.4	5016	5287.1

Table 2 Raw data NO.2

t	1	2	3	4	5	6	7	8
X_1	4526.3	4570.6	4306.9	4694.7	4840.2	4980.4	5016	5287.1
X_2	14462.8	14931.5	14870.1	18138.4	19613.4	21522.3	24658.1	28044

Table 3 Raw data NO.3

t	1	2	3	4	5	6	7	8	9	10
X_1	405	366	356	357	334	362	371	380	381	399
X_2	1766	1860	1969	2062	2103	2301	2560	2847	3265	3756
X_3	2210	2253	2366	2475	2622	2936	3255	3587	4140	4760

Table 4 Raw data NO.4

t	1	2	3	4	5	6	7	8
X_1	4582	4940	5431	6096	6660	7335	8265	9258
X_2	3621	4012	4504	4918	5476	6301	7077	7665

Table 5 Raw data NO.5

t	1	2	3	4	5	6	7	8
X_1	146.9	204	181.8	277	348	426.8	511	539.7
X_2	7.069	9.276	8.725	13	17.16	21.327	27.082	30.24

Table 6 Raw data NO.6

t	1	2	3	4	5	6	7	8	9	10
X_1	7.19	7.44	7.84	8.78	8.7	11.02	12.12	13.94	16.1	17.12
X_2	2831.9	3175.5	3522.4	3878.4	3442.3	4710.7	5285.9	6229.7	7770.6	8749.3

Table 7 Raw data NO.7

t	1	2	3	4	5	6
X_1	0.9166	1.09	1.203	1.249	1.318	1.3
X_2	174.06	257.4	292.9	339.5	419.2	408.4

Table 8 Raw data NO.8

t	1	2	3	4	5
X_1	2.874	3.278	3.307	3.39	3.679
X_2	7.04	7.645	8.075	8.53	8.774

Table 9 Average growing speed and growing rate of reliance sequence

	NO.1	NO.2	NO.3	NO.4	NO.5	NO.6	NO.7	NO.8
Growing Speed	140.0	1940.0	187.4, 241.3	577.7	3.3	617.3	61.3	0.4
Growing Rate	2.99	10.14	8.54, 8.69	11.33	24.17	12.69	19.70	5.68

Table 10 The simulation results

	NO.1	NO.2	NO.3	NO.4	NO.5	NO.6	NO.7	NO.8
Original Model	4.56	101430.76	43.25	4.43	8.78	27.74	34106352.54	7.42
1st BGV-EF	4.67	160644.00	39.13	4.23	17.24	35.03	200743360.68	6.49
1st BGV-GL	4.78	31463.06	35.78	4.80	7.34	20.45	44566.62	9.19
A-BGV-TF	4.65	18.95	16.69	3.89	10.71	11.33	18.16	2.96
A-BGV-EF	4.65	19.00	16.69	3.90	10.79	11.47	18.20	2.91
A-BGV-GL	4.65	18.85	16.81	3.88	10.55	11.05	18.09	3.15

In Table 5, the growing speed is computed as $x^{(0)}(k) - x^{(0)}(k-1)$, and growing rate is $\frac{x^{(0)}(k)-x^{(0)}(k-1)}{x^{(0)}(k)} \times 100\%$, respectively. The raw data of which both the growing speed is larger than 50 and growing rate is larger than 10% is NO.2, 3, 4, 6, 7. According to analysis, the error may be very large using the GM(1,N) model.

Table 10 shows the simulation results, which are the mean percentage error (MPE) of each model. In Table 10, 1st BGV-EF indicates the GM(1,N) model of which the 1st back ground value is computed by the exponential form(Eq.(13)), and the 1st BGV-GL indicates the GM(1,N) model of which the 1st back ground value is computed by the Gauss-Legendre formula (Eq.(16)), respectively. The results shows that the models with revised 1st back ground value perform better than the original model in most cases. However, the original model and the models with revised 1st back ground value all shows very large errors for raw data NO.2, NO.3, NO.6 and NO.7, even the revised models show smaller errors. For raw data NO.2 and NO.7, especially, the models are totally invalid, of which the errors are extremely large. This indicates, the models with revised 1st back ground value doesn't overcome the inherent defect of GM(1,N) model.

Also in Table 10, the A-BGV-TF, A-BGV-EF and A-BGV-GL indicate the GM(1,N) model with all revised back ground value, which are computed by trapezoid formula (Eq.(9)), exponential form (Eq.(13)) and Gauss-Legendre formula (Eq.(16)), respectively. Compared with the original model and the models with revised 1st back ground value, the models with all revised back ground value perform much better in almost all the raw data. Especially for the raw data that the original model and models with revised 1st back ground value are invalid, the models with all revised back ground values are still valid and perform a very high precision. This indicates that the models with all revised back ground values have a higher precision and robustness than the original model and models with revised 1st back ground value.

However, the simulation results present another attribute of the models with all revised back ground values. For these models, the differences of errors are not significant, which means the new models are not sensitive to the form of the integration formulas.

6 Conclusions

The GM(1,N) model has some inherent defect which may cause significant error or even make the GM(1,N) model invalid. In this paper, we analyze the source of the error of GM(1,N) model, and indicate that it is the form of all the back ground values that effect the error of GM(1,N) model, and also point out it is "dangerous" to use the original GM(1,N)when the reliance sequences are non-negative quasi-smooth sequences.

To overcome this defect, we employ three types of numeric integral formula to compute all the back ground values. The simulation test indicates that the methods of this paper are valid, even in some extreme cases, in which the GM(1,N) is invalid, the revised models still perform well. Thus the methods of this paper enhance the precision and robustness of the GM(1,N) model.

References

[1] Liu S F, Lin Y, Forrest J, et al. Grey systems: Theory and applications. Springer, 2010.

[2] Zhou Z Y, Liu S F. Research on valuation of venture capital with grey prediction. Journal of Nanjing University of Aeronautics & Astronautics, 2004, 36(5): 644–648.

[3] Zeng B. Research on electricity demand forecasting based on improved grey prediction model. Journal of Chongqing Normal University (Natural Science), 2012, 29(6): 99–104.

[4] Li H W, Mao W J. An optimized GM(1, 1) model based on bi-directional difference method and its

application in long-term power demand forecasting. Power System Protection and Control, 2011, 39(13): 53–58.

[5] Pai T Y, Lin S H, Yang P Y, et al. Predicting hourly ozone concentration time series in Dali area of Taichung city based on seven types of GM (1,1) model. Time Series Analysis, Modeling and Applications, Springer, 2013: 369–383.

[6] Tan G J. The structure method and application of background value in grey system GM(1,1) model (I). Systems Engineering — Theory & Practice, 2000, 20(4): 98–103.

[7] Tan G J. The structure method and application of background value in grey system GM(1,1) model (II). Systems Engineering — Theory & Practice, 2000, 20(5): 125–127.

[8] Luo D, Liu S F, Dang Y G. The optimization of grey model GM (1,1). Engineering Science, 2003, 8: 50–53.

[9] Jiang N, Liu X Y. A new method of background value-building and application in GM (1,1) model based on Gauss formula. Systems Engineering — Theory & Practice, 2004, 24(10): 123–126.

[10] Chen Y G, Yang D Y, Dai W Z. Modeling research of model GM (1,1) background value-building based on Gauss-Legendre quadrature and its application. Journal of Zhejiang Sci-Tech University, 2007, 4: 444–447.

[11] He M X, Wang Q. Constructing the background value for GM (1,1) model based on Simpson formula. Journal of Quantitative Economics, 2011, 4: 101–104.

[12] Li J F, Dai W Z. A new approach of background value-building and its application based on data interpolation and Newton-Cores formula. Systems Engineering — Theory & Practice, 2004, 24(10): 122–126.

[13] Li X Y, Li K, Shi H J, et al. Optimization of GM (1,1) prediction model based on background value and its application. Journal of University of Electronic Science and Technology of China, 2011, 6: 911–914.

[14] Li C F, Dai W Z. Determinator of the background level in the non-equidistant GM (1,1) model. Journal of Tsinghua University (Science and Technology), 2007, 47(2): 1729–1732.

[15] Wang Z S, Deng K Z. Study on grey Markov prediction model for old goaf residual subsidence. Journal of Geodesy and Geodynamics, 2010(6): 126–130.

[16] Fan A W, Pan Z Q, Wang W. Application of grey GM(1,N) model in Henan coal demand forecast. Coal Technology, 2011(10): 7–9.

[17] Li C M, Ding L Y. The overall assessment model and its application of coordinated development in small cities and towns based on GM (1,N). Resources Science, 2009, 7: 1181–1187.

[18] Zhang L T, Luo Y X. Multi-factored grey GM(1,N) model of experimentation data processing and its application. Journal of Machine Design, 2003, 20(3): 23–25.

[19] Liu H B, Xiang Y M, Nguyen H H. A multivariable grey model based on optimized background value and its application to subgrade settlement prediction. Applied Mechanics and Materials, 2013, 256: 1721–1725.

[20] Shen Y, Sun H Y, Li L P. Application of optimized GM(1,N) model on the forecast of traffic noise and precision analysis. Journal of Safety Science and Technology, 2012, 11: 27–32.

[21] Zhai J, Sheng J M, Feng Y J. The grey model MGM(1,n) and its application. Systems Engineering — Theory & Practice, 1997, 17(5): 109–113.

[22] Xiao Y C, Chen X H. Improvement of multivariable grey prediction formula. Systems Engineering — Theory & Practice, 2009, 29(6): 98–101.

[23] Zhou W, Fang Z G. Nonlinear optimization method of gray GM(1,N) model and application. Systems Engineering and Electronics, 2010, 32(2): 317–320.

[24] Li B, Jiang Y, Yang X W. MGM(1,n) model of optimized background. Journal of Southwest University (Natural Science Edition), 2013(1): 89–94.

[25] He M X, Wang Q. New algorithm for GM(1,N) modeling based on Simpson formula. Systems Engineering — Theory & Practice, 2013, 33(1): 199–202.