

## Study on Optimization for Grey Forecasting Model

Ying Li<sup>1</sup>, Min-an Tang<sup>1</sup>, Tao Liu<sup>1</sup>, Min-an Tang<sup>2</sup>

<sup>1</sup>School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou, Gansu, China

<sup>2</sup>School of Mechanical and Electronical Engineering, Lanzhou University of Technology, Lanzhou, Gansu, China

**Abstract**—In order to improve the prediction accuracy of GM(1,1) model, function transformation  $f(x) = c \ln x + d$  was applied to improve the smoothness of the original sequence. Considering affecting of the initial value and the background value selection to forecasting precision of the model, this paper puts forward the idea of optimization value from the three aspects: smoothness, the background value and the initial data sequence at the same time, and obtained the improved GM model. The model was applied to the bearing sleeve wear prediction, and compared with the condition of single models. The simulation results show that the improved model has smaller error and higher accuracy, the new model prediction accuracy is above 99.8%, and the validity and practicability of the method is illustrated, which enriches the optimization theory of gray model and broadens the application scope of grey model.

**Keywords**—grey theory; GM(1,1) model; data transformation; initial condition; background value; prediction accuracy.

### I. INTRODUCTION

As one of the important part of grey system theory, grey model has been widely used in a number of fields. GM(1, 1) model with its small sample size required for modeling, simple and convenient calculation, and therefore has more advantages than other traditional forecasting methods. In order to improve the accuracy of the model fitting and forecasting, many scholars have done a lot of researches on increasing the smoothness of original data series, improving method of parameter estimation of the model and optimizing the background values and initial conditions.

For the research of improving grey model prediction, in order to improve the smoothness of the original sequence, many function transformation were presented in the literatures [1-9], which include logarithmic function transformation, sine transform, cosine transform, power function transformation, tangent function transformation, tangent-exponential function transformation, power function-exponential function transformation, negative exponential function transformation, and achieved satisfactory results in practical applications. It can be seen from the time response equation that factors which influence simulation results and prediction accuracy of the model, depends on the selection of the initial value and parameters  $a$  and  $b$ . Yet the parameter values are also dependent on background values. Therefore, the initial value select and whether the structure of background is reasonable directly affect simulation value and prediction

accuracy of the grey model. The traditional GM(1, 1) model regards the first data in historical as the initial condition, which is equal to fitting curve must be after the first data point in the historical data, but that does not necessarily fit the facts. Literatures [10-17] studied the method of improving the prediction accuracy of the model in the aspect of improving background values and initial conditions. However, the contribution to unilateral improve model accuracy is limited, and in many ways the optimization model accuracy can be greatly improved. Based on the previous researches, this paper combined function transformation technology which literature [9] proposed with the method of optimizing initial conditions and original background values which the literature [14] proposed to improve GM(1, 1) model, thereby, the model accuracy is further improved.

### II. TRADITIONAL GREY FORECASTING MODEL ESTABLISHMENT

Assume that there is a set of non-negative original data sequence  $x^{(0)}$ :

$$x^{(0)}_{(k)} = (x^{(0)}_{(1)}, x^{(0)}_{(2)}, x^{(0)}_{(3)}, \dots, x^{(0)}_{(n)}) \quad (1)$$

Where,  $x^{(0)}_{(k)} \geq 0, k = 1, 2, 3, \dots, n$ . Then,  $x^{(1)}$  is  $x^{(0)}$  once accumulated generation sequence:

$$x^{(1)}_{(k)} = (x^{(1)}_{(1)}, x^{(1)}_{(2)}, x^{(1)}_{(3)}, \dots, x^{(1)}_{(n)}), \quad (2)$$

Where,  $x^{(1)}_{(k)} = \sum_{i=1}^k x^{(0)}_{(i)}, k = 1, 2, 3, \dots, n$ .  $z^{(1)}$  is mean generation of consecutive neighbors sequence of series  $x^{(1)}$ :

$$z^{(1)}_{(k)} = (z^{(1)}_{(2)}, z^{(1)}_{(3)}, z^{(1)}_{(4)}, \dots, z^{(1)}_{(n)}) \quad (3)$$

Where,

$$z^{(1)}_{(k)} = \frac{1}{2}(x^{(1)}_{(k)} + x^{(1)}_{(k-1)}), \quad k = 2, 3, 4, \dots, n$$

To establish the first order whitenization differential equation

of series  $x^{(1)}$ :  $\frac{dx^{(1)}}{dt} + ax^{(1)} = b$ , and solving parameters

$a$  and  $b$  by the least square method.

$$\hat{\alpha} = (B^T B)^{-1} B^T Y_N = [a, b]^T \quad (4)$$

Where,  $a$  and  $b$  are the parameters to be identified.

$a$  is development coefficient and  $b$  is grey action quantity, which are reflect the relationship between the growth rate and the data of original data sequence.

Where,

$$B = \begin{bmatrix} -z^{(1)}_{(2)} & 1 \\ -z^{(1)}_{(3)} & 1 \\ \vdots & \vdots \\ -z^{(1)}_{(n)} & 1 \end{bmatrix}, Y_N = \begin{bmatrix} x^{(0)}_{(2)} \\ x^{(0)}_{(3)} \\ \vdots \\ x^{(0)}_{(n)} \end{bmatrix}. \quad (5)$$

Determining parameters  $a, b$  and solving the differential equations, then we can get model GM(1, 1):

$$\hat{x}^{(1)}_{(k+1)} = \left( x^{(0)}_{(1)} - \frac{b}{a} \right) e^{-ak} + \frac{b}{a}, k = 1, 2, 3, \dots, n. \quad (6)$$

The reduction value is:

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

where,

$$k = 1, 2, 3, \dots, n. \quad (7)$$

### III. OPTIMIZED GM(1,1) MODEL

#### A. Original Sequence Data Transformation

The original data sequence by some function transforming to reduce its smooth ratio, however inadequate date transformation can satisfy smoothness requirement but not necessarily get higher accuracy. The non-negative function transformation  $f(x) = c \ln x + d$ ,  $c \geq \max(x_{(k)}, k = 1, 2, \dots, n)$ ,  $x_{(k)} > e$ , which meet the requirement of decreasing its smooth ratio, adjusting compression ratio, keeping character of concave and convex unchanged, reduction error does not increase was put forwards in literature [9]. It is verified that the function transformation to deal with the original data sequence which can improve the smoothness of the initial sequence, under the circumstance of remaining the basic rules of the original data sequence unchanged.

#### B. Background Value and Initial Value Optimized

For optimizing background value, the literature [10]

used  $z^{(1)}_{(k)} = \frac{x^{(1)}_{(k)} - x^{(1)}_{(k-1)}}{\ln x^{(1)}_{(k)} - \ln x^{(1)}_{(k-1)}}$  instead of

$z^{(1)}_{(k)} = \frac{1}{2}(x^{(1)}_{(k)} + x^{(1)}_{(k-1)})$ , to minimize the error

caused by the background value, and enhance precision of modeling sequences to a certain extent. While initial value is another important factor of effecting simulation and prediction accuracy of GM(1, 1) model. The literature [13] based on new information priority principle and minimum information theory of grey system theory, the initial

condition of the original GM(1, 1) model  $x^{(0)}_{(1)}$  by instead with  $x^{(1)}_{(n)}$  which is the  $n$ th component of series  $x^{(1)}$  so that original sequence information can be fully reflected, and thus overcomes the disadvantages of GM(1, 1) modeling has nothing to do with  $x^{(0)}_{(1)}$ . Since the background value and the initial value optimization can improve the accuracy of the model, and the two are independent with each other. Therefore, optimize the background value firstly and then replace the original date  $x^{(0)}_{(1)}$  in model

$\hat{x}^{(0)}_{(k+1)} = (1 - e^a) \left( x^{(0)}_{(1)} - \frac{b}{a} \right) e^{-ak}$  with  $x^{(1)}_{(n)}$  and gets

the new time response sequence

$$\hat{x}^{(0)}_{(k+1)} = (1 - e^a) \left( x^{(1)}_{(n)} - \frac{b}{a} \right) e^{-ak}. \quad \text{The literature [14]}$$

was studied the method of optimizing initial conditions and background value, which get a new model was suitable for low growth index sequence modeling, as well as applies to high growth index sequence modeling, and the fitting precision was very high.

#### C. The Modeling Process of the Optimized GM(1,1) Model

The new modeling process of GM(1, 1) model is as follows: Suppose a non-negative original data sequence  $y^{(0)} : y^{(0)}_{(k)} = (y^{(0)}_{(1)}, y^{(0)}_{(2)}, y^{(0)}_{(3)}, \dots, y^{(0)}_{(n)})$ ,  $k = 1, 2, 3, \dots, n$ . Do once an accumulation of data sequence  $y^{(0)}$  and gets the data sequence  $y^{(1)} : y^{(1)}_{(k)} = (y^{(1)}_{(1)}, y^{(1)}_{(2)}, y^{(1)}_{(3)}, \dots, y^{(1)}_{(n)})$ ,  $k = 1, 2, 3, \dots, n$ . Inspect smoothness and quasi-exponent of the original series and analyzes the feasibility of modeling.

Where, Smooth ratio test formula:  $\rho_{(k)} = \frac{y^{(0)}_{(k)}}{y^{(1)}_{(k-1)}}$ ,

$\rho_{(k)} \in (0, 0.5)$ . Quasiexponential test formula:

$\delta^{(1)}_{(k)} = \frac{y^{(1)}_{(k)}}{y^{(1)}_{(k-1)}}$ ,  $\delta^{(1)}_{(k)} \in (1, 1.5)$ . Do function

transportation for the original sequence  $f(x) = c \ln x + d$  to get the new

sequence  $x^{(0)} : x^{(0)}_{(k)} = (x^{(0)}_{(1)}, x^{(0)}_{(2)}, x^{(0)}_{(3)}, \dots, x^{(0)}_{(n)})$ ,  $k = 1, 2, 3, \dots, n$ . Do once a cumulative to the series  $x^{(0)}_{(k)}$

and get its generating sequence:  $x^{(1)}_{(k)} = (x^{(1)}_{(1)}, x^{(1)}_{(2)}, x^{(1)}_{(3)}, \dots, x^{(1)}_{(n)})$ , where,

$x^{(1)}_{(k)} = \sum_{i=1}^k x^{(0)}_{(i)}$ ,  $k = 1, 2, 3, \dots, n$ . To the series  $x^{(1)}$  do

mean generation of consecutive neighbors sequence  $z^{(1)}$  :

$$z^{(1)}_{(k)} = (z^{(1)}_{(2)}, z^{(1)}_{(3)}, z^{(1)}_{(4)}, \dots, z^{(1)}_{(n)}),$$

$$\text{where, } z^{(1)}_{(k)} = \frac{x^{(1)}_{(k)} - x^{(1)}_{(k-1)}}{\ln x^{(1)}_{(k)} - \ln x^{(1)}_{(k-1)}},$$

$k = 2, 3, 4, \dots, n$ . Establish the whitenization differential

equation  $\frac{dx^{(1)}}{dt} + ax^{(1)} = b$  which GM(1,1) model corresponded by using the first order accumulation generating sequence  $x^{(1)}$ , and using the method of least squares solution parameters  $a$  and  $b$  :

$$\hat{\alpha} = (B^T B)^{-1} B^T Y_N = [a, b]^T, \text{ where,}$$

$$B = \begin{bmatrix} -z^{(1)}_{(2)} & 1 \\ -z^{(1)}_{(3)} & 1 \\ \vdots & \vdots \\ -z^{(1)}_{(n)} & 1 \end{bmatrix}, Y_N = \begin{bmatrix} x^{(0)}_{(2)} \\ x^{(0)}_{(3)} \\ \vdots \\ x^{(0)}_{(n)} \end{bmatrix}.$$

Make an initial value of  $x^{(1)}_{(n)}$ , get the time response sequence solution of the differential equation:

$$\hat{x}^{(0)}_{(k+1)} = (1 - e^a) \left( x^{(0)}_{(1)} - \frac{b}{a} \right) e^{-ak}, \quad k = 0, 1, 2, \dots, n.$$

The reduction value is:

$$\hat{x}^{(0)}_{(k+1)} = (1 - e^a) \left( x^{(1)}_{(n)} - \frac{b}{a} \right) e^{-ak}, \quad k = 0, 1, 2, \dots, n.$$

Do inverse function transformation:  $\hat{y}^{(0)}_{(k)} = e^{\frac{\hat{x}^{(0)}_{(k)} - d}{c}}$ ,

and get the date series  $\hat{y}^{(0)}$  :

$$\hat{y}^{(0)}_{(k)} = (\hat{y}^{(0)}_{(1)}, \hat{y}^{(0)}_{(2)}, \hat{y}^{(0)}_{(3)}, \dots, \hat{y}^{(0)}_{(n)}),$$

$k = 1, 2, 3, \dots, n$ . Use the mean square error ratio  $C$  and small error probability  $p$  for model testing:  $C = \frac{S_1}{S_2}$ ,

$$\text{where, } S_1^2 = \frac{1}{n} \sum_{k=1}^n (x^{(0)}_{(k)} - \bar{x})^2,$$

$$S_2^2 = \frac{1}{n} \sum_{k=1}^n (\varepsilon_{(k)} - \bar{\varepsilon})^2,$$

$$\varepsilon_{(k)} = x^{(0)}_{(k)} - \hat{x}^{(0)}_{(k)}; p = P(|\varepsilon_{(k)} - \bar{\varepsilon}| < 0.6745S_1).$$

Mean square error ratio is as small as possible, and small error probability is bigger, the better. Model precision grade test table is as shown in table 1.

TABLE I THE MODEL PRECISION TEST TABLE

Model accuracy level	Mean square error ratio $C$	Small error probability $p$
1 (Good)	$C \leq 0.35$	$p > 0.95$
2 (Qualified)	$0.35 < C \leq 0.5$	$0.8 \leq p < 0.95$
3 (Basic eligibility)	$0.5 < C \leq 0.65$	$0.7 \leq p < 0.8$
4 (Unqualified)	$C < 0.65$	$p < 0.7$

#### IV. GM (1, 1) MODEL IN PREDICTION OF BEARING SLEEVE WEAR

In this section, we present wear the bearing sleeve with time prediction as an example. The bearing sleeve has been running in period, measuring and recording the amount of wear every half a month, the data is as shown in table 2.

TABLE II BEARING SLEEVE WEAR VOLUME [ $\mu m$ ]

No.	1	2	3	4	5	6	7	8
Wear volume	19.6	19.9	20.2	20.6	21.0	21.5	22.0	22.7

Firstly, test the smooth and level of original data sequence. When  $k > 3$ , smooth ratio are 0.3451, 0.2615, 0.2122, 0.1792, 0.1568, both within the range of (0, 0.5), and smooth ratio are decreasing, which satisfy with quasi smooth condition. When  $k > 3$ , the quasi exponent respectively are 1.3451, 1.2675, 1.2122, 1.1792, 1.1568, both within the range of (1, 1.5). So an accumulated generating sequence of the original sequence satisfied with quasi exponent, and GM(1, 1) model can be established.

##### A. Traditional GM (1, 1) Model

Established Traditional GM (1, 1) model:  $\hat{x}^{(1)}_{(k+1)} = 892.6776e^{0.0219k} - 873.0776$ , where,  $k = 0, 1, \dots, 7$ .

##### B. Function Transformation GM (1,1) Model

To establish GM (1, 1) model by function transforming  $f(x) = c \ln x + d$ , get transformed model:  $\hat{x}^{(1)}_{(k+1)} = 142.0473e^{0.0535k} - 134.5028$ , where,  $k = 0, 1, \dots, 7$ ,  $c = \max(x^{(0)}_{(k)}) = 22.7, d = -60$ .

##### C. Optimized GM (1, 1) Model Establishment

Using function transformation technique to transform the original data sequence, and then combined with the optimization method of background value and initial value, obtain the optimized GM(1, 1) model:  $\hat{x}^{(1)}_{(k+1)} = 206.6667e^{0.0536(k-8)} - 134.5299$ , where,  $k = 1, 2, \dots, 8$ .

A. Three Forecasting Models Precision Comparison

Compare and analysis the above three methods' results

TABLE III. COMPARISON OF THREE MODELS' SIMULATION VALUE AND RELATIVE ERROR

No.	Measured values	Original model simulated values	Relative errors %	Literature [9] model simulation values	Relative errors %	New method simulated values	Relative errors %
1	19.6						
2	19.9	19.7653	0.67699	19.8275	0.36430	19.8389	0.30704
3	20.2	20.2029	0.01436	20.2058	0.02871	20.2185	0.09158
4	20.6	20.7202	0.58368	20.6128	0.06214	20.6270	0.13107
5	21.0	21.0374	0.17841	21.0500	0.23811	21.0668	0.02063
6	21.5	21.6448	0.67357	21.5227	0.10562	21.5419	0.19488
7	22.0	21.9825	0.07946	22.0341	0.15500	22.0541	0.24591
8	22.7	22.5407	0.70134	22.5837	0.51413	22.6079	0.40573
The average relative error %			0.4154			0.2097	0.1995

B. Results of Three Methods' Forecasting Comparison

Figure 1 shows the comparison of three methods' forecast performance.

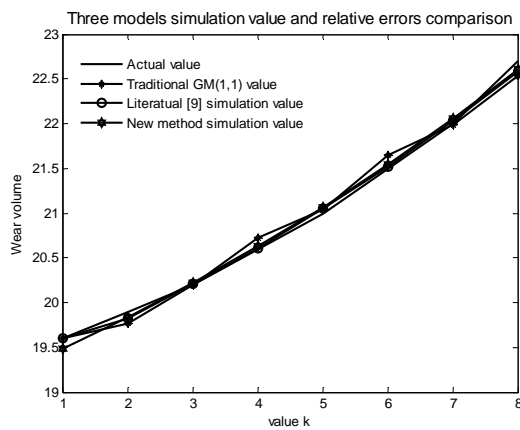


Figure 1. Predicted value and actual value comparison

From table 3 and figure 1 we can see the mean relative error obtained by the method proposed in this paper is smaller than the original model and literature [9] presented, and the new model prediction accuracy is above 99.8%. The new model mean square error ratio  $C = 0.126871642 < 0.35$ , small error probability  $p = 1 > 0.95$ , the accuracy of the model belongs to grade one. Therefore, we can conclude that the method proposed in this paper has higher simulation and prediction accuracy and certain practical value.

VI. CONCLUSIONS

This paper we have presented a new method of optimizing the traditional GM(1, 1) model by improving the smooth ratio of data series, the background value and initial

value at the same time, and get a new optimum grey model. Application examples compared the new improved models' results with improved single condition models' proves that the optimum model is better than improved single condition models, indicates that the new method this paper reported is practical and effective, which enrich the optimization theory of the gray model and broadens the application scope of the grey model.

VII. ACKNOWLEDGEMENTS

This work was financially supported by the National Natural Science Research Foundation (61263004), Innovation Planning Program of Gansu Province Science and Technology Research and Development Program (090GKCA009, 1304GKCA023), Science and Technology Research Projects of Lanzhou city (2013-4-18).

REFERENCES

- [1] Chen, T.J. A New Development of Grey Forecasting Model, J. China. *Systems Engineering*. 8 (1990) 50-52.
- [2] Cao, C., Fang, C.J. and Hu, Z.L. Grey Forecasting Model and its Application Based on The Sine Function Transformation, J. China. *Journal of Mathematics*. 33 (2013), 697-701.
- [3] Zheng, F. and Wei, Y. A new method to improve the smoothness of gray modeling data, J. China. *Statistics and Decision*. 9 (2007), 37-38.
- [4] Li, Q. The Further Generalization for Grey Forecasting Model, J. China. *System Engineering—Theory & Practice*. 13 (1993), 64-66.
- [5] Li, C.F. and Dai, W.Z. An Approach of the Grey Modeling Based on Transformation, J. China. *Systems Engineering*. 23 (2005), 110-111.
- [6] Guan, Y.Q. and Liu, S.F. An Approach to Grey Modeling Based on  $COI \times$  Transformation, J. China. *Systems Engineering*. 26 (2008), 89-93.
- [7] Chen, J. and Xu, C.X. An improvement of grey forecasting model, J. China. *Journal of Liaoning Normal University (Natural Science Edition)*. 28 (2005), 262-264.
- [8] He, B. and Meng, Q. Study on generalization for grey forecasting model, J. *System Engineering—Theory & Practice*. 22 (2002) 138-141.

- [9] Cui, L.Z. and Liu, S.F. Grey forecasting model based on data transformation technology, J. China. *Systems Engineering*. 28 (2010), 104-107.
- [10] Luo, D., Liu, S.F. and Dang, Y.G. The optimization of grey model GM(1,1), J. China. *Engineering Science*. (2003), 50-53.
- [11] Liao, F. Optimization integrated background value for new GM(1,1) model, J. China. *Mathematics In Practice and Theory*. 39 (2009), 107-113.
- [12] Luo, G.Z., Cui, Z.J. and Xie, N.M. An new improvement of grey GM(1,1) forecasting model, J. China. *Statistics and Decision*. 22 (2008), 11-13.
- [13] Dang, Y.G., Liu, S.F. and Liu, B. The GM models that  $x_{(n)}^{(1)}$  be taken as initial value, J. China. *Chinese Journal of Management Science*. (2005), 132-134.
- [14] Zhang, Y., Wei, Y. and Xiong, C.W. One new optimized method of GM(1,1) model, J. China. *System Engineering—Theory & Practice*. 27 (2007), 141-146.
- [15] Li, J.F. and Dai, W.Z. Research on the ameliorating GM(1,1) model and its application in the power quantity modeling of shanghai city, J. China. *System Engineering—Theory & Practice*. 3 (2005), 140-144.
- [16] Yin, F.P. The optimized GM models that  $x_{(n)}^{(1)}$  be taken as initial value, J. China. *Statistics and Decision*. (2010), 164-165.
- [17] Zhang, B. and Xi, G.Q. GM(1,1) model optimization based on the background and boundary value correction, J. China. *System Engineering—Theory & Practice*. 33 (2013), 682-688.