

Grey GM (1,1) model based on combinatorial buffer operator

HE Peng-xiang^{1,2}, SUN Sheng-xiang¹

¹Department of Management Engineering and Equipment Economy, Naval University of Engineering, Wuhan, Hubei, China
²91306 PLA, Shanghai, China

Abstract : Considering that some emerging industries have not developed for a long time, the amount of data available for forecasting future economic problems is relatively limited, complex and changeable. Based on the principle of combinatorial prediction, a combinatorial buffer operator based on different order buffer operators is proposed, and the correlation area between the generated sequence and the original sequence after the buffer operator is used as weighting criterion. The grey GM (1,1) prediction model based on the combined buffer operator was established, which effectively overcame the influence of abnormal data and restored the change rule of data series. The average prediction error of the data in literature [7] by using the combined buffer operator established in this paper was 0.98%. Compared with 6.89%, 11.59% and 1.30% of the original method, the prediction accuracy is significantly improved.

1 the introduction

Grey system theory is a method to study uncertain data with small data and poor information, which was established by Chinese scholar Professor Deng Julong in 1982. Sequence generation is an important means to deal with uncertain information in the grey system theory. When the system is strongly disturbed by some external factors in operation, the buffering operator can effectively overcome the impact of impact disturbance on the real change law of the original data series, weaken the uncertainty of the original data series, so as to show the potential law.

GM (1, 1) model is the basic model of grey prediction theory, but in the use of grey GM (1, 1) model, because the system behavior data, and the effects of shock disturbance and distortion, and therefore cannot reflect the variety regulation system of the right, so when we use the original data modeling analysis, first of all we should deal with the raw data. Since the concept of buffering operator and the three axioms of buffering operator were proposed in literature [1-2], scholars at home and abroad have carried out a lot of research on the construction of new buffering operators and the action properties of buffering operators on sequences. Literature [3-4] constructed a series of basic strengthening and weakening buffer operators. When these buffer operators are applied to practical problems, a more appropriate buffer operator should be selected according to the actual data sequence. However, after the buffer operator is determined, its action intensity is still fixed and it is impossible to achieve effective adjustment of data sequence. For how to control the intensity of buffer

operator function so as to find the real law of data sequences, literature [5] first constructed the idea of variable weight which introduced into the buffer operator and constructed respectively variable weight weakening buffer operators and variable weight strengthening buffer operators, based on a certain error minimization criterion optimization model is set up and use genetic algorithm to solve. Literature [6] proposed a kind of weakening variable weight buffer operator based on average growth rate, and solved the variable weight coefficient of buffer operator by using grey correlation analysis and particle swarm optimization algorithm. Literature [7] considered the more important role of new information in the prediction and combined the principle of new information firstly to construct the full information variable weight buffer operator, then takes minimizing the average prediction error as the optimization goal, and uses genetic algorithm to optimize the weight regulating factors. Literature [8-13] constructed a series of variable weight buffer operators, which achieved good prediction effect when applied to practical problems, and improved the stability and accuracy of prediction results to a certain extent. Literature [14] proposed the combination of the same buffer operator for the first time, and analyzed and studied the combination properties of the buffer operator. However, the weight selection of the combined buffer operator only adopted human subjective judgment and lacked theoretical basis. Literature [15] analyzed and studied that the buffer operator could improve the prediction accuracy of the model because it could weaken the influence of abnormal data on the change rule of the data. At present, the theoretical research on buffer operators mainly focuses on the construction of new variable weight buffer operators for different problems.

Pengxiang He;1213605430@qq.com

By using intelligent algorithm to optimize the weight factors, the prediction accuracy is improved, while the study on the action strength of classical practical buffer operators of different orders is ignored. For the same prediction problem, The degree of data expansion or shrinkage will vary greatly after the action of buffer operators of different orders, so we can combine the same buffer operator of different orders to more accurately restore the original law of data.

Current research on variable weights buffer operator weight mainly concentrated in the minimum average relative error, the grey correlation maximum or better smoothness as constraint conditions, using intelligent optimization algorithms, definite, but the resulting weight coefficient is not depend on the time variable, in combination of the weight of time series changes over time is different, is a random variable, And if we only consider when modeling error and minimum, and ignore the weight of the uncertainty, it obviously is not consistent with what is actually, given the new information priority principle of grey system theory, we should put the time-varying factors into the study of weight coefficient, literature [16] [17] and literature will be introduced to the related area of the weight of combination forecast analysis. Demand for logistics freight volume and armored equipment combined forecasting model is established to study respectively, compared with using intelligent optimization algorithm has stronger time-varying characteristic, therefore this article build composite buffer operator introduces the ideas of variable weight combination forecast, when applying correlation area to different order to determine the weighing values of buffer operator. In practical application, according to the specific situation of different data sequences, the buffer operator which is more suitable for the development law of data sequences of practical problems can be selected as the basic buffer operator according to the prediction error of the first-order buffer sequence. Finally, the problems in literature [7] are studied for example.

2 Grey GM (1,1) model

Let the small-sample time data series be), where $k=1,2,\dots,N$. The steps of solving the problem using the grey GM (1,1) model are as follows:

Step 1: Solve the first-order accumulation generation sequence, $X^{(1)}$ is the first-order accumulation generation sequence of $X^{(0)}$.

$$X^{(1)} = (X^{(1)}(1), X^{(1)}(2), \dots, X^{(1)}(n)), \text{ and}$$

$$X^{(1)}(k) = \sum_{i=1}^k X^{(0)}(i), \quad k=1,2,\dots,n$$

Step 2: Build a first-order differential equation model

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b, \quad -a \text{ is the development coefficient}$$

and b is the grey variable.

Step 3: Estimate the parameter vector $\hat{a} = [a, b]^T$ by

$$\text{least square method: } \hat{a} = (B^T B)^{-1} B^T Y$$

$$Y = \begin{bmatrix} X^{(0)}(2) \\ X^{(0)}(3) \\ \vdots \\ X^{(0)}(n) \end{bmatrix}, \quad B = \begin{bmatrix} -X^{(1)}(2) & 1 \\ -X^{(1)}(3) & 1 \\ \vdots & \vdots \\ -X^{(1)}(n) & 1 \end{bmatrix}$$

Step 4: Solve the model. Solve the values of A and B and substitute them into Equation (1). Set initial value for the $X^{(0)}(1)$, Then, the time response of GM (1,1) model is as follows:

$$\hat{x}^{(1)}(k) = (x^{(0)}(1) - \frac{b}{a})x^{-a(k-1)} + \frac{b}{a}, \quad k=1,2,\dots,n$$

Then through accumulation and reduction to generate:

$$\hat{X}^{(0)}(k) = \hat{X}^{(1)}(k) - \hat{X}^{(1)}(k-1)$$

3 Combination buffer operator based on weight change

Theorem 1 Let the original data sequence $XD_1 = (x(1)d_1, x(2)d_1, \dots, x(n)d_1)$ be a first-order weakened buffer sequence, $XD_1D_2 = (x(1)d_1d_2, x(2)d_1d_2, \dots, x(n)d_1d_2)$ is a second-order weakening buffer sequence, And so on, $XD_1D_2 \dots D_m = (x(1)dd_2 \dots d_m, x(2)dd_2 \dots d_m, \dots, x(n)dd_2 \dots d_m)$ is m -order weakening buffer sequence, then:

$$x(k)d = \frac{w_1x(k)d_1 + w_2x(k)d_1d_2 + \dots + w_mx(k)d_1d_2 \dots d_m}{w_1 + w_2 + \dots + w_m}$$

Then, when X is a monotonic growth sequence, a monotonic attenuation sequence and an oscillating sequence, D is all a weakening buffer operator.

Proof: D_1 is taken as the weakening buffer operator to prove that D is the weakening buffer operator.

It is easy to find that the buffer operator constructed by the analytic expression of $x(k)d$ satisfies the three axioms of buffer operator.

(1) When X is a monotonic growing sequence, since D is a weakening operator, $x(k) \leq x(k)d_1 \leq x(k)d_1d_2 \leq \dots \leq x(k)d_1d_2 \dots d_m$ can be written according to the weakening buffer operator properties, then

$$x(k)d \geq \frac{w_1x(k)d_1 + w_2x(k)d_1 + \dots + w_mx(k)d_1}{w_1 + w_2 + \dots + w_m} = x(k)d_1 \geq x(k)$$

Therefore, it can be proved that when X is a monotonic growing sequence, D is a weakening buffer operator.

(2) Similarly, it can be proved that, when X is a monotonic attenuation sequence, $x(k) \geq x(k)d$, D is a weakening buffer operator.

(3) When X is an oscillating sequence, let $x(a) = \max\{x(k) | k=1,2,\dots,n\}$ This is where $x(k)$ gets the maximized;

$x(b) = \min\{x(k) | k=1,2,\dots,n\}$ This is where $x(k)$ gets the minimum;

Then

$$x(a)d = \frac{w_1x(a)d_1 + w_2x(a)d_1d_2 + \dots + w_mx(a)d_1d_2 \dots d_m}{w_1 + w_2 + \dots + w_m} \leq$$

$$\frac{w_1x(a)d_1 + w_2x(a)d_1 + \dots + w_mx(a)d_1}{w_1 + w_2 + \dots + w_m} = x(a)$$

By the same token, $x(b)d \geq x(b)$. Therefore, for an oscillating sequence, the buffering operator D can obviously reduce its amplitude, so when X is an oscillating sequence, D is a weakening buffering operator.

In summary, it can be proved that when the basic buffer operator is a weakening buffer operator, the combined buffer operator D is also a weakening buffer operator.

Theorem 2: Let the original data sequence $XD_1 = (x(1)d_1, x(2)d_1, \dots, x(n)d_1)$ be a first-order weakened buffer sequence, $XD_1D_2 = (x(1)d_1d_2, x(2)d_1d_2, \dots, x(n)d_1d_2)$ is a second-order weakening buffer sequence, And so on, $XD_1D_2 \dots D_m = (x(1)dd_2 \dots d_m, x(2)dd_2 \dots d_m, \dots, x(n)dd_2 \dots d_m)$ is m-order weakening buffer sequence, then:

$$x(k)d = (x(k)d_1)^{p_1} (x(k)d_1d_2)^{p_2} \dots (x(k)d_1d_2 \dots d_m)^{p_m}$$

Among them, $p_1 + p_2 + \dots + p_m = 1$

When X is a monotonic growth sequence, monotonic attenuation sequence and oscillating sequence, D is also strengthening (weakening) buffer operator. (The proof is similar to Theorem 1)

4 Weight analysis

4.1 Correlation Area

Correlation area refers to the area between the two polylines corresponding to the sequence after the original data sequence and the buffer operator. The correlation area reflects the fitting effect of time t and time t+1. The smaller the correlation area is, the smaller the overall fitting error is and the better the prediction accuracy is. Therefore, we can use the correlation area as the weighting criterion of the combined buffer operator.

(1) For monotonic growth and monotonic decay sequences, $(x(t) - x(t)d)(x(t+1) - x(t+1)d) > 0$, The area composed of the four points $(t, x(t)), (t, x(t)d), (t+1, x(t+1))$ and $(t+1, x(t+1)d)$ is just a trapezoid. The area of the trapezoid is the correlation area between the original sequence and the buffer operator, which is denuded as $S(t+1)$, then $S(t+1) = 0.5 \times [|x(t) - x(t)d| + |x(t+1) - x(t+1)d|]$

(2) For the oscillating sequence, when $(x(t) - x(t)d)(x(t+1) - x(t+1)d) < 0$

$$S(t+1) = 0.5 * [|x(t) - x(t)d| + |x(t+1) - x(t+1)d|]$$

When $(x(t) - x(t)d)(x(t+1) - x(t+1)d) < 0$

$(t, x(t)), (t, x(t)d), (t+1, x(t+1))$ and $(t+1, x(t+1)d)$ of four points is the area of the two triangles, Let the intersection

Table 1 FDI and growth rate in Zhejiang Province from 2000 to 2007

Year	1	2	3	4	5	6	7	8
Investment / \$10,000	161266	221162	316002	544936	668128	772271	888935	1036576
The growth rate / %	—	37.14	42.88	72.45	22.61	15.59	15.11	16.61

According to the data in Table 1, it is easy to find that

of two broken lines be $(t_x, x(t_x))$, then

$$S(t+1) = 0.5 * [|x(t) - x(t)d| * (t_x - t) + |x(t+1) - x(t+1)d| * (t+1 - t_x)]$$

$$t_x = t + \frac{|x(t) - x(t)d|}{|x(t) - x(t)d| + |x(t+1) - x(t+1)d|}$$

Therefore, we can calculate the variable weight coefficient of each generated sequence from time 2 to time n. However, since the data changes at two adjacent moments are considered in the calculation of the associated area, we cannot calculate the associated area for the initial data at time 1. For the data at the initial moment, the average value method is adopted to calculate.

4.2 Determination of variable weight coefficient

The larger the correlation area is, the weaker the correlation between the original sequence and the generated sequence, and the larger the fitting error is. Therefore, a smaller weight should be given. Therefore, we take the reciprocal of the correlation area at time t as the eigenvalue at that time, and let

$$d_i(t+1) = \frac{1}{s_i(t+1)}, \quad i = 1, 2, \dots, n$$

Then the weight coefficient of the generated sequence acted by the ith order buffer operator at the moment t+1 is

$$w_i(t+1) = \frac{d_i(t+1)}{d_1(t+1) + d_2(t+1) + \dots + d_m(t+1)}, \quad i = 1, 2, \dots, m$$

Therefore, we can calculate the variable weight coefficient of each generated sequence from time 2 to time n. However, since the data changes at two adjacent moments are considered in the calculation of the associated area, we cannot calculate the associated area for the initial data at time 1. For the data at the initial moment, the average value method is adopted to calculate.

5 Analysis of calculation examples

The data of foreign investment in Zhejiang Province in literature [7] are used to illustrate the practical application value of the combined buffer operator proposed in this paper. And compared with the prediction accuracy of the first-order buffer operator D1, the second-order buffer operator D2 and the variable weight buffer operator D4 proposed in the original paper, to verify the practicability and superiority of the proposed combined buffer operator.

the growth rate of investment from 2001 to 2003 is

obviously higher than that from 2004 to 2007. Therefore, we need to use the weakening buffer operator to preprocess the original data to weaken the growth trend of the series. The geometric average weakening buffer operator commonly used in reference 1 was selected to preprocess the original data from 2000 to 2004, and the data from 2005 to 2007 were taken as the test data, and the average prediction error from 2005 to 2007 was used

as the basis to judge the prediction accuracy.

Meanwhile, the weighting criterion of the combined buffer operator based on the associated area is adopted. According to Equation (1), the data sequence after the action of the first-order and second-order combined geometric average weakening buffer operator is obtained as follows:

Table 2 Data sequences after the action of the three weakening buffering operators

Year	First-order buffering operator D1	Second order buffering operator D1	The weight of D1	The weight of D2	Combinatorial buffer operator D3
2000	333142	482272	0.5	0.5	407707
2001	399394	529002	0.64	0.36	445746
2002	486368	580955	0.62	0.38	522154
2003	603396	634938	0.61	0.39	615760
2004	668128	668128	0.61	0.39	668128

Based on the above three groups of data from 2000 to 2004, the GM (1,1) model is adopted to model and the

investment data from 2005 to 2007 is predicted as follows:

Table 3 Comparison of prediction accuracy of the four models

Year	First-order buffering operator D1		Second order buffering operator D1		Combinatorial buffer operator D3		Variable weight buffer operator D4	
	Predictive value	Relative error /%	Predictive value	Relative error /%	Predictive value	Relative error /%	Predictive value	Relative error /%
2005	806037	4.37	729211	5.58	776118	0.50	752294	2.59
2006	954168	7.34	788014	11.35	886937	0.22	888935	0.00
2007	1129521	8.97	851560	17.85	1013579	2.22	1050395	1.33
Average relative error /%		6.89		11.59		0.98		1.30

It can be found from Table 3 that the average relative error of D1 and D2 is 6.89% and 11.59%, while the average error of buffer operator D3 is 0.98% after the weighted combination of D1 and D2, and the prediction accuracy reaches a higher level. At the same time, its prediction accuracy is lower than the average relative error of 1.30% of the full-information variable weight weakening buffer operator proposed in literature [7]. This is mainly because the first-order weakening buffer operator is not strong enough, so the acted data series is higher than the original data series as a whole, while the second-order weakening buffer operator is too strong. As a result, the acted data sequence is lower than the original data sequence as a whole. By combining the two sets of generated data with weights, the action intensity of the buffer operator can be better controlled, so that the newly generated data sequence can more accurately restore the real rules in the original data sequence, so as to obtain higher prediction accuracy. At the same time, aiming at the forecasting problem of a company's new product sales in literature [14], this paper selects the special strengthening buffer operator in literature [2] to process the original data, and uses the weighting method in Theorem 2 to obtain the combined special strengthening

buffer operator, and predicts the sales in June and July, the average prediction error is only 0.1%. The prediction accuracy is higher than the 0.65% prediction accuracy of the combined homosomal buffer operator used in literature [14]. Due to the space limitation, it will not be repeated here.

To sum up, we can find that higher prediction accuracy can be obtained by using either the weakened buffer operator or the strengthened buffer operator, the first weighting method or the second weighting method, which indicates that the combined buffer operator proposed in this paper has strong practicability. Through the combination of different order strengthen (weaken) the buffer operator can more effectively reduce the impact of the system disturbance, the more accurate the reduction of original data sequence rule, to avoid the first-order buffer operator strength is too weak to change the data sequence development trend and the intensity of higher-order buffer operator function is too strong lead to serious deviation from the original data sequence, The prediction performance of GM (1,1) model is improved.

6 Conclusion

This paper aims at the problem that it is difficult to collect and master the complete data in the current economic forecasting field, a combined buffer operator based on different order buffer operator is proposed. Considering the time-varying characteristics of the weight factor, a method based on the associated area is proposed to calculate the combined weight, which provides a new idea for solving the weight of the variable weight buffer operator. Taking the data in literature [7] as an example, the weakening and strengthening combined buffer operators constructed in this paper are used for modeling. An example shows that the variable weight combined buffer operator established in this paper can achieve a better prediction effect.

References

1. S.F.Liu. The Journal of Grey System. **3**,39-48,1 (1991) .
2. S.F.Liu. Grey Sys theory and its App. **8**(2017)
3. Y G.Guo,S F.Liu,B Liu. et al CHN.J.Managment.Sci. **12**:108-111(2004).
4. Y.G.Dang,B.Liu,Y.Q.Guan. Control and Decision. **20**,1332-1336,12(2005).
5. Z.X.Wang,Y.G.Dang,S.F.Liu. Control and Decision. **24**,1218-1222,8(2009).
6. W.Y.Qian,Y.G.Dang. Syst. Eng. **29**,105-110,1(2011).
7. Z.X.Wang. Zhejiang Univ.(Eng). **47**,1120-1128,6(2013).
8. X.M.Li,Y.G.Dang,Z.X.Wang. Sys Eng Theory&Pratice. **32**,2486-2492,11(2012).
9. N.Xu,Y.G.Dang. Control and Decision. **29**,1262-1266,7(2014).
10. J.F.Liu,S.F.Liu,Z.G.Fang. CHN.J.Managment.Sci. **24**,172-176,8(2016).
11. D.Luo,X.Wang. Grey Sys:Theory&App. **1**,148-157,2(2011).
12. Y.Wei,X.H.Kong,D.H.Hu. Grey Sys:Theory&App. **1**,178-185,1(2011).
13. W.Z,Dai,Z.H.Wu,A.P.Yang. Grey Sys:Theory&App. **1**,190-195,2(2011)
14. S.Liu,P.Li. Control and Decision. **31**,172-176,10(2016).
15. X.Y.Tian,G.Hong. ,Sys Eng Theory&Pratice. **8**,1(2018).
16. [C.Zhou,P.L.Zhang. Computer App. **32**,2628-2630,31(2016).
17. H.Li,T.L.Wang., Sys Eng&Elec Tech. **40**,2276-2281,10(2018).