Multi-variable weakening buffer operator and its application

Lifeng Wu*[a,c], Sifeng Liu[b], Yingjie Yang[b], Lihua Ma[c], Hongxia Liu[c]

[a] *College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China*

[b] *Centre for Computational Intelligence, De Montfort University, Leicester, LE1 9BH, UK*

[c] *College of Economics and Management, Hebei University of Engineering, Handan, 056038, China*

[c] *Art College, Hebei University of Engineering, Handan, 056038, China*

Abstract:

To weaken the disturbances of multi-variable and reveal the real situation, it is proved that the essence of the weakening buffer operator (abbreviated as WBO) can weaken the disturbance of one variable. According to this, the multi-variable weakening buffer operator is put forward. The multi-variable weakening buffer operator can satisfy the desire to use the freshest data and its buffer effect is obvious when the sample size is small. Four real cases show that the proposed multi-variable weakening buffer operator has higher forecasting performances.

Keywords: forecasting; grey system theory; weakening buffer operator; multiple linear regression; energy demand forecasting

1. Introduction

Forecasting the future values of time series data plays a very important role in our research. Superior forecasting ability is an important characteristic of successful managers in complex and uncertain environments. With the rapid developments of science and technology, managers can better understand the future situation and make right strategies and plans by getting more information in advance. Thus, some forecasting methods have been developed, for example, moving

[0]*Corresponding Author: E-mail address: wlf6666@126.com.

average, support vector clustering, neural networks and time series [7, 24]. However, these methods require a large amount of data and must meet some assumptions, and they are invalid when the observed data available is of a small sample size.

Moreover, in some cases, the time series may exhibit a jumping phenomenon due to the changing of policy. Therefore, it is difficult to fit a reasonable mathematic model. In such a case, it is most important to weaken the impact of policy. Considering the above reasons, based on the fact that the recent data could provide more information than the distant data (More recent data are typically more relevant, especially for short-term forecasts [9]), more emphasis has been placed on data which is more recent. Take the energy demand in China from 1985 to 2006 as an example. Actually, the growth rate of energy demand in China from 1985 to 2002 is tempered. From 2003 to 2006, Chinese total energy demand increased abruptly.(During this period, China has already entered into the heavy chemical industry stage, becoming one member of World Trade Organization. The rapid development of heavy and chemical industry accelerates Chinese economic growth. The growth rate of Chinese economy was above 10%. The rapid development of economy accelerates Chinese energy demand growth.) To obtain better forecasting results from 2004 to 2006, we must give more weight to the data of 2003.

Grey weakening buffer operator is proposed to cope with the disturbance. It was first introduced in early 1990s by Liu [23]. Since then, WBO has been widely and successfully applied to various systems [5,15,16,17,26]. Its improved form appeared simultaneously [1,3,4,6,8,13,14,19,22,25]. However, traditional WBOs can only deal with a single variable, and little research has been conducted on multiple variables. The multi-variable weakening buffer operator model is put forward by giving more weight to newer information in this paper.

The rest of this paper is organized as follows. In Section 2, the essence of a variable WBO is introduced. In Sections 3, the multi-variable weakening buffer operator model is proposed. In Section 4, the advantages of the new model over the traditional model is clarified by four real cases.

The conclusions of this work are discussed in Section 5.

2. The essence of a variable WBO

With the changes of economic development, the historical data tends to deviate from the current situations. Grey buffer operator can weaken these disturbances and reveal the real situation. Its definition is as follows.

2.1. GM(1,1) model with the WBO

Definition [13] Given a raw data sequence $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}$, $X^{(0)}D = \{x^{(0)}(1)d, x^{(0)}(2)d, \cdots, x^{(0)}(n)d\}$, where

$$x^{(0)}(k)d = \frac{x^{(0)}(k) + x^{(0)}(k+1) + \cdots + x^{(0)}(n)}{n-k+1}, \tag{1}$$

D is a first order WBO.

The sequence $\{x^{(0)}(1)d, x^{(0)}(2)d, \cdots, x^{(0)}(n)d\}$ is given by WBO of Eq.(1). A new series $X^{(1)}D = \{x^{(1)}(1)d, x^{(1)}(2)d, \cdots, x^{(1)}(n)d\}$ can be generated by the first-order accumulated generating operator as $x^{(1)}(k)d = \sum_{i=1}^{k} x^{(0)}(i)d, k = 1, 2, \cdots, n$. Since the original form of GM(1,1) model $x^{(0)}(k)d + az^{(1)}(k) = b$, where $z^{(1)}(k) = \frac{x^{(1)}(k)d + x^{(1)}(k-1)d}{2}, k = 2, 3, \cdots, n$. [11, 20]. The least squares estimate of a and b can be obtained by

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (B^T B)^{-1} B^T Y,$$

where

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

The prediction expression of GM(1,1) model can be obtained as $\hat{x}^{(1)}(t) = [x^{(0)}(1) - \frac{\hat{b}}{\hat{a}}]e^{-\hat{a}t} + \frac{\hat{b}}{\hat{a}}$ by solving the differential $\frac{dx^{(1)}}{dt} + ax^{(1)} = b$. The prediction at $k+1$ can be obtained as $\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$.

2.2. The essence of WBO

A lemma is given in order to discuss the essence of WBO.

Lemma 1 [2] Assume that $A \in C^{n\times n}, \delta A \in C^{n\times n}, b \in C^n, \delta b \in C^n$, vector norm $\|\cdot\|$ and matrix norm $\|\cdot\|$ are tolerant. If a matrix norm $\|\cdot\|$ followed $\|A^{-1}\|\|\delta A\| < 1$, then the solutions of linear system equations $AX = b$ and $(A+\delta A)(X+\delta x) = b + \delta b$ satisfy

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\|A\|\|A^{-1}\|}{1-\|A\|\|A^{-1}\|\frac{\|\delta A\|}{\|A\|}}(\frac{\|\delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|})$$

Theorem 1 For the GM(1,1) model of original data $\{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}$ based on WBO, if the rth data is disturbed, that is $\hat{x}^{(0)}(r) = x^{(0)}(r) + \epsilon_r, r = 1, 2, \cdots, n$. L_r is the $\|\delta B\|_2$ resulted from this disturbance, T_r is the $\|\delta Y\|_2$ resulted from this disturbance, the relative perturbation bound of the parameter estimation is $\frac{\|B\|\|B^{-1}\|}{1-\|B\|\|B^{-1}\|\frac{\|L_r\|}{\|B\|}}(\frac{\|L_r\|}{\|B\|}+\frac{\|T_r\|}{\|Y\|})$ by Lemma 1. Then the relative perturbation bound of the parameter estimation is larger while the more recent data is perturbed.

Proof. Since

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} = -\begin{bmatrix} 1 & \frac{1}{2} & 0 & \cdots & 0 \\ 1 & 1 & \frac{1}{2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x^{(0)}(1)d & -1 \\ x^{(0)}(2)d & 0 \\ \vdots & \vdots \\ x^{(0)}(n)d & 0 \end{bmatrix}$$

$$= -\begin{bmatrix} 1 & \frac{1}{2} & 0 & \cdots & 0 \\ 1 & 1 & \frac{1}{2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ 0 & \frac{1}{n-1} & \frac{1}{n-1} & \cdots & \frac{1}{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{2} \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x^{(0)}(1) & -n \\ x^{(0)}(2) & 0 \\ \vdots & \vdots \\ x^{(0)}(n) & 0 \end{bmatrix}$$

$$= -\begin{bmatrix} \frac{1}{n} & \frac{1}{n}+\frac{1}{2(n-1)} & \frac{1}{n}+\frac{1}{2(n-1)} & \cdots & \frac{1}{n}+\frac{1}{2(n-1)} \\ \frac{1}{n} & \frac{1}{n}+\frac{1}{n-1} & \frac{1}{n}+\frac{1}{n-1}+\frac{1}{2(n-2)} & \cdots & \frac{1}{n}+\frac{1}{n-1}+\frac{1}{2(n-2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{n} & \frac{1}{n}+\frac{1}{n-1} & \frac{1}{n}+\frac{1}{n-1}+\frac{1}{2(n-2)} & \cdots & \sum_{k=2}^{n}\frac{1}{k}+\frac{1}{2} \end{bmatrix} \begin{bmatrix} x^{(0)}(1) & -n \\ x^{(0)}(2) & 0 \\ \vdots & \vdots \\ x^{(0)}(n) & 0 \end{bmatrix}$$

and

$$Y = \begin{bmatrix} x^{(0)}(2)d \\ x^{(0)}(3)d \\ \vdots \\ x^{(0)}(n)d \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{n-1} & \frac{1}{n-1} & \cdots & \frac{1}{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{2} \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x^{(0)}(1) \\ x^{(0)}(2) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}.$$

If it happens that disturbance $\hat{x}^{(0)}(2) = x^{(0)}(2) + \epsilon_2$, matrix B becomes

$$\hat{B} = -\begin{bmatrix} \frac{1}{n} & \frac{1}{n}+\frac{1}{2(n-1)} & \frac{1}{n}+\frac{1}{2(n-1)} & \cdots & \frac{1}{n}+\frac{1}{2(n-1)} \\ \frac{1}{n} & \frac{1}{n}+\frac{1}{n-1} & \frac{1}{n}+\frac{1}{n-1}+\frac{1}{2(n-2)} & \cdots & \frac{1}{n}+\frac{1}{n-1}+\frac{1}{2(n-2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{n} & \frac{1}{n}+\frac{1}{n-1} & \frac{1}{n}+\frac{1}{n-1}+\frac{1}{2(n-2)} & \cdots & \sum\limits_{k=2}^{n}\frac{1}{k}+\frac{1}{2} \end{bmatrix} \begin{bmatrix} x^{(0)}(1) & -n \\ x^{(0)}(2)+\epsilon_2 & 0 \\ \vdots & \vdots \\ x^{(0)}(n) & 0 \end{bmatrix}$$

$$= B - \begin{bmatrix} 0 & (\frac{1}{n}+\frac{1}{2(n-1)})\epsilon_2 \\ 0 & (\frac{1}{n}+\frac{1}{n-1})\epsilon_2 \\ \vdots & \vdots \\ 0 & (\frac{1}{n}+\frac{1}{n-1})\epsilon_2 \end{bmatrix},$$

then $\delta B = -\begin{bmatrix} 0 & (\frac{1}{n}+\frac{1}{2(n-1)})\epsilon_2 \\ 0 & (\frac{1}{n}+\frac{1}{n-1})\epsilon_2 \\ \vdots & \vdots \\ 0 & (\frac{1}{n}+\frac{1}{n-1})\epsilon_2 \end{bmatrix}, L_2 = \|\delta B\|_2 = \sqrt{(\frac{1}{2(n-1)}+\frac{1}{n})^2 + (n-2)(\frac{1}{n-1}+\frac{1}{n})^2}|\epsilon_2|;$

vector Y becomes

$$\hat{Y} = \begin{bmatrix} 0 & \frac{1}{n-1} & \frac{1}{n-1} & \cdots & \frac{1}{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{2} \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x^{(0)}(1) \\ x^{(0)}(2)+\epsilon_2 \\ \vdots \\ x^{(0)}(n) \end{bmatrix} = Y + \begin{bmatrix} 0 & \frac{1}{n-1} & \frac{1}{n-1} & \cdots & \frac{1}{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{2} \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 0 \\ \epsilon_2 \\ \vdots \\ 0 \end{bmatrix}$$

then $\delta Y = -\begin{bmatrix} \frac{\epsilon_2}{n-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, T_2 = \|\delta Y\|_2 = \frac{|\epsilon_2|}{n-1}.$

Similarly, if it happens that disturbance $\hat{x}^{(0)}(r) = x^{(0)}(r)+\epsilon_r, r = 3, 4, \cdots, n$, we obtain L_r and T_r. It is easy to find that L_r and T_r from $\frac{\|B\|\|B^{-1}\|}{1-\|B\|\|B^{-1}\|\frac{\|L_r\|}{\|B\|}}(\frac{\|L_r\|}{\|B\|}+\frac{\|T_r\|}{\|Y\|})$ are increasing functions of r. So the relative perturbation bound of the parameter estimation $(\frac{\|B\|\|B^{-1}\|}{1-\|B\|\|B^{-1}\|\frac{\|L_r\|}{\|B\|}}(\frac{\|L_r\|}{\|B\|}+\frac{\|T_r\|}{\|Y\|}))$ is an increasing function of r, that is to say, the relative perturbation bound of parameter estimation is larger, while the more recent data is perturbed and the perturbation of each data is equable.

The essence of WBO can attach more importance to the more recent data. Attaching more importance to the more recent data is likely to have better forecasting performance, which is consistent with the priority theory of new information in grey system theory. According to the essence of WBO, the multi-variable weakening buffer operator is put forward.

3. Multi-variable weakening buffer operator

The sequence $\{x_1(k)d, x_2(k)d, \cdots, x_m(k)d, y(k)d\}$ is given by the weakening buffer operator of the observations $\{x_1(k), x_2(k), \cdots, x_m(k), y(k)\}$. Let $\{y(1)d, y(2)d, \cdots, y(k)d\}$ be the sequence from the regression model given by

$$y(k)d = \beta_0 + \beta_1 x_1(k)d + \beta_2 x_2(k)d + \cdots + \beta_m x_m(k)d \tag{2}$$

where $\beta_0, \beta_1, \beta_2, \cdots, \beta_m$ are the regression parameters, Eq.(2) is called as multi-variable weakening buffer operator. The ordinary least squares estimators of the parameters are

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_n \end{bmatrix} = (X^T X)^{-1} X^T Y,$$

where

$$X = \begin{bmatrix} 1 & x_1(1)d & \cdots & x_m(1)d \\ 1 & x_1(2)d & \cdots & x_m(2)d \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_1(n)d & \cdots & x_m(n)d \end{bmatrix}, Y = \begin{bmatrix} y(1)d \\ y(2)d \\ \vdots \\ y(n)d \end{bmatrix}$$

Then the model $y(k) = \hat{\beta}_0 + \hat{\beta}_1 x_1(k) + \hat{\beta}_2 x_2(k) + \cdots + \hat{\beta}_m x_m(k)$ is used to predict the variable $y(k)$. The flow chart of the multi-variable weakening buffer operator model is shown in Fig. 1.

Fig.1. The flow chart of the multi-variable weakening buffer operator model

Because all matrix norms are equivalent, they are substantially consistent. The matrix m_1 norm and its tolerant vector norm l_1 are used in order to analyze clearly.

Theorem 2 For the multi-variable regression model with WBO $y(k)d = \beta_0 + \beta_1 x_1(k)d + \beta_2 x_2(k)d + \cdots + \beta_m x_m(k)d$, if the rth data is disturbed, that is $\hat{x}_i(r) = x_i(r) + \epsilon_i (i = 1, 2, \cdots, m), \hat{y}(r) = y(r) + \epsilon_0 (r = 1, 2, \cdots, n)$. L_r is the $\|\delta X\|_{m_1}$ resulted from the disturbance $\hat{x}_i(r) = x_i(r) + \epsilon_i (i = 1, 2, \cdots, m)$, T_r is the $\|\delta Y\|_1$ resulted from the disturbance $\hat{y}(r) = y(r) + \epsilon_0 (r = 1, 2, \cdots, n)$, the relative perturbation bound of the parameter estimation is $\frac{\|X\|\|X^{-1}\|}{1 - \|X\|\|X^{-1}\| \frac{\|L_r\|}{\|X\|}} (\frac{\|L_r\|}{\|X\|} + \frac{\|T_r\|}{\|Y\|})$ by Lemma 1. Then the relative perturbation bound of the parameter estimation is larger while the more recent data is perturbed.

Proof. Since

$$X=\begin{bmatrix}1 & x_1(1)d & \cdots & x_m(1)d\\ 1 & x_1(2)d & \cdots & x_m(2)d\\ \vdots & \vdots & \vdots & \vdots\\ 1 & x_1(n)d & \cdots & x_m(n)d\end{bmatrix}=\begin{bmatrix}\frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n}\\ 0 & \frac{1}{n-1} & \cdots & \frac{1}{n-1}\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}\begin{bmatrix}1 & x_1(1) & \cdots & x_m(1)\\ 1 & x_1(2) & \cdots & x_m(2)\\ \vdots & \vdots & \vdots & \vdots\\ 1 & x_1(n) & \cdots & x_m(n)\end{bmatrix},$$

$$Y=\begin{bmatrix}y(1)d\\ y(2)d\\ \vdots\\ y(n)d\end{bmatrix}=\begin{bmatrix}\frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n}\\ 0 & \frac{1}{n-1} & \cdots & \frac{1}{n-1}\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}\begin{bmatrix}y(1)\\ y(2)\\ \vdots\\ y(n)\end{bmatrix}.$$

If it happens that disturbance $\hat{x}_i(n)=x_i(n)+\epsilon_i(i=1,2,\cdots,m),\hat{y}(n)=y(n)+\epsilon_0$, matrix X becomes

$$\hat{X}=\begin{bmatrix}\frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n}\\ 0 & \frac{1}{n-1} & \cdots & \frac{1}{n-1}\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}\begin{bmatrix}1 & x_1(1) & \cdots & x_m(1)\\ 1 & x_1(2) & \cdots & x_m(2)\\ \vdots & \vdots & \vdots & \vdots\\ 1 & x_1(n)+\epsilon_1 & \cdots & x_m(n)+\epsilon_m\end{bmatrix}$$

$$=X+\begin{bmatrix}\frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n}\\ 0 & \frac{1}{n-1} & \cdots & \frac{1}{n-1}\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}\begin{bmatrix}0 & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & \epsilon_1 & \cdots & \epsilon_m\end{bmatrix},$$

then

$$\delta X=\begin{bmatrix}\frac{\epsilon_1}{n} & \frac{1}{n} & \cdots & \frac{\epsilon_m}{n}\\ 0 & \frac{\epsilon_1}{n-1} & \cdots & \frac{\epsilon_m}{n-1}\\ \vdots & \vdots & \vdots & \vdots\\ 0 & \epsilon_1 & \cdots & \epsilon_m\end{bmatrix},$$

$\|\delta X\|_{m_1}=(1+\frac{1}{n-1}+\frac{1}{n})|\epsilon_1|+(1+\frac{1}{n-2}+\frac{1}{n-1}+\frac{1}{n})|\epsilon_2|+\cdots+\sum\limits_{k=1}^{n}\frac{1}{k}|\epsilon_m|$; vector Y becomes

$$\hat{Y}=\begin{bmatrix}\frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n}\\ 0 & \frac{1}{n-1} & \cdots & \frac{1}{n-1}\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}\begin{bmatrix}y(1)\\ y(2)\\ \vdots\\ y(n)+\epsilon_0\end{bmatrix}=Y+\begin{bmatrix}\frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n}\\ 0 & \frac{1}{n-1} & \cdots & \frac{1}{n-1}\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}\begin{bmatrix}0\\ 0\\ \vdots\\ \epsilon_0\end{bmatrix}=Y+\begin{bmatrix}\frac{\epsilon_0}{n}\\ \frac{\epsilon_0}{n-1}\\ \vdots\\ \epsilon_0\end{bmatrix}$$

then

$$\delta Y=\begin{bmatrix}\frac{\epsilon_0}{n}\\ \frac{\epsilon_0}{n-1}\\ \vdots\\ \epsilon_0\end{bmatrix},$$

$\|\delta Y\|_1=\sum\limits_{k=1}^{n}\frac{|\epsilon_0|}{k}$.

Similarly, if it happens that disturbance $\hat{x}_i(n-1)=x_i(n-1)+\epsilon_i(i=1,2,\cdots,m),\hat{y}(n-1)=y(n-1)+\epsilon_0$, we obtain $L_{n-1}=\|\delta X\|_{m_1}=(\frac{1}{n-1}+\frac{1}{n})|\epsilon_1|+(\frac{1}{n-2}+\frac{1}{n-1}+\frac{1}{n})|\epsilon_2|+\cdots+\sum\limits_{k=2}^{n}\frac{1}{k}|\epsilon_m|$, $T_{n-1}=\|\delta Y\|_1=\sum\limits_{k=2}^{n}\frac{|\epsilon_0|}{k}$.

If it happens that disturbance $\hat{x}_i(r) = x_i(r) + \epsilon_i (i = 1, 2, \cdots, m), \hat{y}(r) = y(r) + \epsilon_0, r = 1, 2, \cdots, n-2$, we obtain $L_r = \|\delta X\|_{m_1} = (\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{r})|\epsilon_1| + (\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{r-1})|\epsilon_2| + \cdots + \sum\limits_{k=n-r+1}^{n} \frac{|\epsilon_m|}{k}$, $T_r = \|\delta Y\|_1 = \sum\limits_{k=n-r+1}^{n} \frac{|\epsilon_0|}{k}$.

If it happens that disturbance $\hat{x}_i(r) = x_i(r) + \epsilon_i (i = 1, 2, \cdots, m), \hat{y}(r) = y(r) + \epsilon_0, r = 1, 2, \cdots, n$, L_r and T_r from $\frac{\|X\|\|X^{-1}\|}{1-\|X\|\|X^{-1}\|\frac{\|L_r\|}{\|X\|}}(\frac{\|L_r\|}{\|X\|} + \frac{\|T_r\|}{\|Y\|})$ are increasing functions of r. So the relative perturbation bound, $\frac{\|X\|\|X^{-1}\|}{1-\|X\|\|X^{-1}\|\frac{\|L_r\|}{\|X\|}}(\frac{\|L_r\|}{\|X\|} + \frac{\|T_r\|}{\|Y\|})$, is an increasing function of r, that is to say, the relative perturbation bound of the parameter estimation is larger, while the more recent data is perturbed and the perturbation of per period data is equable.

If the perturbation of per period data is equable, the relative perturbation bound of the parameter estimation is larger while the more recent data is perturbed. It means that the role of new information on cognition is better than the old information, that is to say the impact of data far from the prediction point on future prediction is gradually weakened or even nonexistent, while the data close to the prediction point will have a significant impact on prediction. It also means that this method gives new information more weight. As can be seen, $L_{r+1} - L_r$ and $T_{r+1} - T_r$ mean the change of relative perturbation bound, when $L_{r+1} - L_r$ and $T_{r+1} - T_r$ are larger, the relative perturbation bound changed greatly, the difference between the weight of $r+1$th data and the weight of rth data is larger. The size of $L_{r+1} - L_r$ and $T_{r+1} - T_r$ is determined by the size of sample n, if n is smaller, then $L_{r+1} - L_r$ and $T_{r+1} - T_r$ are larger; if n is larger, then $L_{r+1} - L_r$ and $T_{r+1} - T_r$ are smaller. That is to say, when n is small, the difference between the weight of old data and the weight of new data is more obvious; when n is large, the difference is not obvious, it can not highlight the role of WBO. So Eq.(2) is suitable for the small-data-set problems.

4. Experimentation results

In this section, to testify the proposed model, three criteria are used to evaluate the forecasting precision. They are mean absolute percentage error (MAPE $= 100\% \frac{1}{n} \sum\limits_{k=1}^{n} |\frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)}|$), absolute

mean error (AME $= \frac{1}{n}\sum_{k=1}^{n}|x^{(0)}(k) - \hat{x}^{(0)}(k)|$) and mean square error (MSE $= \frac{1}{n}\sum_{k=1}^{n}(x^{(0)}(k) - \hat{x}^{(0)}(k))^2$) .

Case 1: Energy demand forecasting in China [10]

We consider an example from paper [10] which provides the sample data. As the discussion in the introduction, data from 1985 to 2003 are used to construct the multi-variable weakening buffer operator model. Then data from 2004 to 2006 are predicted. Actual values and the forecasting values of four compared models are presented in Table 1. The fitting and prediction results of four models are plotted in Fig.2.

Fig.2. The fitting and prediction values by different models

Table 1 Forecasting results from different models (million ton)

Year	Actual value	GM(1,1)	BP network	support vector regression	multi-variable WBO
2004	203227	203730	197690	193940	194721
2005	224682	213420	212530	220250	220409
2006	246270	223560	237830	243560	245775
MAPE		21.1	3.9	2.5	1.9
AME		48061.5	8710	5476	4425
MSE		2459091257	83187691	37745031	30286051

As can be seen from Table 1, the multi-variable weakening buffer operator provides the lowest errors. This implies that the multi-variable weakening buffer operator can improve the prediction accuracy of the other models.

Case 2: Construction land demand forecasting example in Kunming[18]

We consider an example from paper [18]. The same sample is applied here to compare the precision. Actual values and fitting values of three compared models are presented in Table 2. The fitting and prediction results of three models are plotted in Fig.3.

As can be seen from Table 2, from a short-term forecasting viewpoint, multi-variable WBO has

the lowest errors from 2008 to 2009 compared with the other models. Therefore, we can conclude that the multi-variable WBO significantly enhances the precision in out-of-sample. Because the construction land demand in Kunming exhibited a jumping phenomenon when Qiu He took up the mayor post in 2007. Qiu supplied lots of land for the municipal construction. To predict the trend of construction land demand, it is most necessary to weaken the impact of policy. The multi-variable WBO made full use of the data from 2007, thus the impact of Qiu is eliminated.

Fig. 3. The fitting and prediction values by different models

Table 2 The fitted values and errors of different models

Year	Actual value	GM(1,1)	multivariate regression [18]	multi-variable WBO
2008	115086.2	112205.1	112637.7	112896.4
2009	133600.0	115267.0	125211.7	128600.7
MAPE		8.1	4.2	2.8
AME		10607	5418.4	3594.5
MSE		172199813	38179189	14894038

Case 3: The commercial housing average price forecasting example in Guangdong province [21]

We consider an example from paper [21] which provides the sample data. The commercial housing average price in Guangdong province exhibited a jumping phenomenon due to the recession in 2008. To predict the trend of the commercial housing average price, the multi-variable WBO can weaken the impact by making full use of the data from 2008. Thus, the data from 2001 to 2008 (in-sample data) are used to construct the different models. Then the value in 2009 (out-of-sample) is predicted. Actual values and the forecasting values of three compared models are presented in Table 3. The fitting and prediction results of three models are plotted in Fig.4.

Fig. 4. The fitting and prediction values by different models

Table 3 The fitted values and errors of different models

Time	Actual value	GM(1,1)	multivariate regression [21]	multi-variable WBO
2009	6513	6999.9	6418.5	6548.3
MAPE		4.66	4.09	1.91
AME		486.9	94.5	35.3
MSE		237147	8934	1249

As can be seen from Table 3, from a short-term forecasting viewpoint, the multi-variable WBO has better forecasting performance than the other models, it indicates that the multi-variable WBO can obtain accurate forecasting.

Case 4: The freight volume forecasting example in Handan city [12]

We consider an example from paper [12] which provides the sample data. In 2009, the government carried out a scheme for the development of logistics in Handan, thus the freight volume exhibited a jumping phenomenon. To predict the trend of freight volume, the multi-variable WBO will weaken the impact by making full use of the data from 2009. Therefore, the data from 1999 to 2009 (in-sample data) are used to construct the different models. Then the value from 2010 to 2012 (out-of-sample) are predicted. Actual values and the forecasting values of three compared models are presented in Table 4. The fitting and prediction results of three models are plotted in Fig.5.

Fig. 5. The fitting and prediction values by different models

Table 4 The fitted values and errors of different models

Time	Actual value	GM(1,1)	neural network [12]	multi-variable WBO
2010	25507	16717.0	26399.8	24893.4
2011	29276.6	17900.1	30242.7	29005.2
2012	34453.8	19166.9	36004.2	33916.7
MAPE		39.2	3.77	1.63
AME		11818	1136	474
MSE		1.47×10^8	1378070	246213

As can be seen from Table 4 and Fig.5, from a short-term forecasting viewpoint, the multi-variable WBO can effectively reduce the forecasting errors, which means that the multi-variable WBO reaches the objective of accurate forecasting.

5. Conclusion

As to the prediction, the most recent data is likely to carry more information than the older data. Therefore, the multi-variable WBO is constructed with preference for recent data, the prediction results must be very accurate. From the real cases shown above, it is found that the multi-variable WBO always has a better forecasting performance than the conventional GM(1,1) model and multivariate regression model. In this sense, it is suggested that it would be better to use the multi-variable WBO for practical small sample problems.

Acknowledgements

The authors are grateful to anonymous referees for their helpful and constructive comments on this paper. This work was supported by a Marie Curie International Incoming Fellowship within the 7th European Community Framework Programme (No.FP7-PIIF-GA-2013-629051), National Natural Science Foundation of China (No.71401051), the youth foundation of Humanities and Social Science of the Education Department Henan province (2015-QN-020), the Humanistic and Social Science Foundation of Ministry of Education (15YJA630017) and Natural Science Founda-

tion of Jiangsu Province (No.BK20130785, BK20130786).

References

[1] W.Z. Dai, Y. Su, New strengthening buffer operators and their applications based on prior use of new information, Acta Automatica Sinica 38 (2012) 1329-1334.

[2] H. Dai, The Matrix Theory, Beijing: Sciences Press, 2001.

[3] W.Z. Dai, Y. Su, Relationship between regulation degree and smoothness of buffer operators, Control and Decision 29 (2014) 158-162.

[4] Y. Gao, D.Q. Zhou, C.C. Liu, L. Zhang, Constructing methods of new buffer operators with variable weights and their inner link, System Engineering Theory & Practice 33 (2013) 489-497.

[5] C.L. Guo, X.X. Xu, Z.W. Gong, Co-integration analysis between GDP and meteorological catastrophic factors of Nanjing city bansed on the buffer operator, Natural Hazards 71 (2014) 1091-1105.

[6] R. Han, Z.P. Wu, Study on new weakening buffer operators, Journal of Communication University of China 22 (2015) 33-39.

[7] P.H.K. Ho, Forecasting tender price index under incomplete information, Journal of the Operational Research Society 64 (2013) 1248-1257.

[8] H.M. Jiang, Research and application on buffer operators with weighted logrithmic type, Sichuan University of Arts and Science Journal 25 (2015) 14-17.

[9] C. Kesten, J. Green, S. Armstrong, Simple versus complex forecasting: The evidence, Journal of Business Research 68 (2015) 1678-1685.

[10] H. Kong, P.R. Yang, J.H. Cheng, Forecasting model of energy demand based on Matlab support vector regression, Systems Engineering-Theory & Practice 31 (2011) 2001-2007.

[11] G.D. Li, S. Masuda, D. Yamaguchi, M. Nagai, A new reliability prediction model in manufacturing systems, IEEE Transactions on Reliability 59 (2010) 170-177.

[12] R.Y. Li, Using BP neural network to predict the freight volume of Handan City, Hebei University of Engineering, Handan, 2013.

[13] D.M. Li, X. Li, Research on a new weakening buffer operator of investment forecasting for constructing digital libraries, Journal of Nanjing Institute of Technology 10 (2012) 48-51.

[14] W. Li, H. Xie, Geometrical variable weights buffer GM(1,1) model and its application in forecasting of Chinas energy consumption, Journal of Applied Mathematics (2014)ID 131432.

[15] W. Li, Y.N. Yuan, D.X. Niu, long and medium term load forecasting based on grey model optimized by operator and time response function, Power System Protection and Control 39 (2011) 59-63.

[16] R.J. Liao, J.P. Yang, S. Grzybowski, Y.Y. Wang, J. Li, Forecasting dissolved gases content in power transformer oil based on weakening buffer operator and least square support vector machine-Markov, IET Generation, Transmission & Distribution 6 (2012) 142-151.

[17] D.P. Wang, B.W. Wang, Medium and long term load forecasting based on variable weights buffer grey model, Power System Technology 37 (2013) 167-171.

[18] K.X. Wang, Using grey-multiple regression model to prediction the construction land demand. Yunnan University of Finance and Economics, Kunming, 2012.

[19] Y. Wei, X.H. Kong, Constructing methods of several kinds of strengthening and weakening buffer operators and their inner link, Control and Decision 25 (2010) 196-203.

[20] L.F. Wu, S.F. Liu, L.G. Yao, S.L. Yan, The effect of sample size on the grey system model, Applied Mathematical Modelling 37 (2013) 6577-6583.

[21] M. Wu, Using multivariate regression to analyze the influencing factors of real estate prices, Guide to Business 7 (2012) 19-20.

[22] N. Xu, Y.G. Dang, Construction of buffer operator with smooth variable weight and its property, Control and Decision 29 (2014) 1262-1266.

[23] D. Yamaguchi, G.D. Li, M. Naga, A grey-based rough approximation model for interval data processing, Information Sciences 177 (2007) 4727-4744.

[24] Y.Y. Zhang, R.J. Liao, L.J. Yang, X.P. Deng, H.C. Cheng, C. Lv, A cost-effectiveness assessment model using grey correlation analysis for power transformer selection based on life cycle cost, Kybernetes 43 (2014) 5-23.

[25] H. Zhu, Z.R. Zhao, H.X. Qiu, J.D. Wei, Application of improved grey model with buffer operator in deformation forecasting, Journal of Water Resources and Architectural Engineering 11 (2013) 137-139.

[26] J.M. Zhu, D.T. Zhai, Z.W. Huang, Fracture healing stress prediction based on weakening buffer operator and GM(1,1) equal-dimension-new-information model, Chinese Journal of Biomedical Engineering 31 (2012) 268-275.