Grey model with time varying weighted generating operator

Lifeng Wu*[a], Sifeng Liu[a], Yingjie Yang[b]

[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*

Abstract:

A grey model with time varying weighted generating operator is put forward in order to fully extract information concealed in recent data. This model increases the weight of new data and relieve the influence of the the data fluctuation. The relationship between the sample size and the error from the inverse time varying weighted generating operator is discussed. Compared with traditional grey forecasting models, the results of practical numerical examples demonstrate that this grey model performs well in forecasting problems with limited data, and can provide reliable and acceptable accuracy for future prediction.

Keywords: grey forecasting; grey generating operator; failure frequency forecasting; electricity consumption; tuberculosis incidence

1. Introduction

Forecasting a future trend is an important issue in fields, including economics [1], energy [2,3], engineering [4] and so on. To obtain a reliable forecast, certain laws in terms of system development must be discovered on the basis of real observed data. Therefore, building a predictor that can describe system evolution is necessary. Traditional approaches, such as linear regression, autoregressive moving-average model, are built on the assumption concerning the statistical distribution of big data. However, in some cases, due to the limitation of information and knowledge, only part

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

of the system data could be fully realized, especially in the the early stages of a process.

To overcome this problem, Deng proposed grey forecasting model to catch the system development tendency [5]. Because the systems with completely known information are regarded as white, those systems with completely unknown information are regarded as black, and the systems with partially known information and partially unknown information are regarded as grey in the theory of control [6]. Grey forecasting model is a time series model that deals with the system with partially known information and partially unknown information. Unlike statistical methods, grey forecasting model extracts the unknown information by the accumulated generating operator and is suitable to the limited data problem [7].

As a simple model, grey first-order differential equation with one variable (GM(1,1) model) has already been widely applied in many problems. Such as, energy demand [8], software stage effort [9], control system [10,11], fashion color trend forecasting [12], etc. However, the fitting and forecasting precision of traditional GM(1,1) model is unacceptable when the data is fluctuant. This has motivated many researchers to increase the prediction and simulation accuracy, such as, extended grey models [5], optimized grey models [13], hybrid model [8]. However, these models do not take into account the principle of new information priority [14]. How to reflet the importance of new information is a crucial problem, because new information is more important to predict the tendency of development from scanty data.

Therefore, a grey model with a novel generating operator is proposed in Section 2. The advantages of the grey model over the traditional models are proved by three real cases in Section 3. Some conclusions are presented in the final Section.

2. The GM(1,1) model with time varying weighted generating operator

For the original data sequence $X = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}$, a new sequence $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\}$ can be generated by the accumulated generating operator as $x^{(1)}(k) =$

$\sum_{i=1}^{k} x^{(0)}(i)(k = 1, 2, \cdots, n)$. The first-order differential equation with a variable,

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b,$$

is the whitenization equation of the original form of the GM(1,1) model

$$x^{(0)}(k) + az^{(1)}(k) = b.$$

where $z^{(1)}(k) = \frac{x^{(1)}(k)+x^{(1)}(k-1)}{2} (k = 2, 3, \cdots, n.)$

The solution of $\frac{dx^{(0)}(t)}{dt} + ax^{(0)}(t) = b$ is $x^{(0)}(t) = (x^{(0)}(1) - \frac{b}{a})e^{-at} + \frac{b}{a}$. Hence $x^{(0)}(k) = (x^{(0)}(1) - \frac{b}{a})e^{-ak} + \frac{b}{a}$ is called the time response equation of GM(1,1). Then we can get the simulative value by $\hat{x}^{(0)}(k + 1) = \hat{x}^{(1)}(k + 1) - \hat{x}^{(1)}(k)$.

The existing grey models all utilize the accumulated generating operator, which leads to pay more attention to the old information, not consider the importance of new information. This operator do not consider the influence of time specially and the time factor should not be ignored. To conserve the time information and highlight the importance of new information, the GM(1,1) model with time varying weighted generating is introduced in this section. Each data is weighted by its time, thus we give the following definition:

Definition 1 Given an original sequence $X = \{x(1), x(2), \cdots, x(n)\}$, the time varying weighted operator is

$$y(k) = \frac{kx(k) + (k + 1)x(k + 1) + \cdots + nx(n)}{\frac{(n-k+1)(n+k)}{2}}.$$

$Y = \{y(1), y(2), \cdots, y(n)\}$ is the time varying weighted generating sequence. It can be expressed as by matrix:

$$[y(1), y(2), \cdots, y(n)] = [x(1), x(2), \cdots, x(n)] \begin{bmatrix} \frac{1}{\frac{(n+1)n}{2}} & 0 & \cdots & 0 \\ \frac{2}{\frac{(n+1)n}{2}} & \frac{2}{\frac{(n+2)(n-1)}{2}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{n}{\frac{(n+1)n}{2}} & \frac{n}{\frac{(n+2)(n-1)}{2}} & \cdots & 1 \end{bmatrix}$$

The correspondent inverse operator of time varying weighted is

$$[x(1),x(2),\cdots,x(n)]=[y(1),y(2),\cdots,y(n)]\begin{bmatrix} \frac{(n+1)n}{2} & 0 & 0 & \cdots & 0 \\ -\frac{(n-1)(n+2)}{2} & \frac{(n-1)(n+2)}{4} & 0 & \cdots & 0 \\ 0 & -\frac{(n-2)(n+3)}{4} & \frac{(n-2)(n+3)}{6} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

Definition 2 For the original sequence $X=\{x(1),x(2),\cdots,x(n)\}$, the time varying weighted generating sequence is $Y=\{y(1),y(2),\cdots,y(n)\}$.

$$y(k+1)-y(k)+az(k)=kb$$

is the original form of GM(1,1) model with time varying weighted operator (WGM(1,1)), where $z(k)=\frac{y(k)+y(k+1)}{2}(k=1,2,\cdots,n-1)$. The least squares estimate of a and b can be obtained by

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}=(A^TA)^{-1}A^TB \tag{1}$$

where

$$B=\begin{bmatrix} y(2)-y(1) \\ y(3)-y(2) \\ \vdots \\ y(n)-y(n-1) \end{bmatrix}, A=\begin{bmatrix} -z(2) & 2 \\ -z(3) & 3 \\ \vdots & \vdots \\ -z(n) & n \end{bmatrix}$$

If we set $\hat{y}(1)=y(1)$, the solution of the whitenization equation $\frac{dy}{dt}+ay=tb$ can be expressed as

$$y(t+1)=[y(1)-\frac{\hat{b}}{\hat{a}}+\frac{\hat{b}}{\hat{a}^2}]e^{-\hat{a}t}+\frac{\hat{b}}{\hat{a}}t-\frac{\hat{b}}{\hat{a}^2}$$

The simulation and forecasting value can be derived by applying the inverse time varying weighted operator

$$[\hat{x}(1),\hat{x}(2),\cdots,\hat{x}(n)]=[\hat{y}(1),\hat{y}(2),\cdots,\hat{y}(n)]\begin{bmatrix} \frac{(n+1)n}{2} & 0 & 0 & \cdots & 0 \\ -\frac{(n-1)(n+2)}{2} & \frac{(n-1)(n+2)}{4} & 0 & \cdots & 0 \\ 0 & -\frac{(n-2)(n+3)}{4} & \frac{(n-2)(n+3)}{6} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

A Theorem is given in order to explain the fact that WGM(1,1) pays more attention to the new information.

Theorem 1 Assume that $X = \{x(1), x(2), \cdots, x(n)\}$ is the original sequence, A and B is the same as Eq.(1). According to the Lemma 1 from paper [7],

$$L[x(k)] = \frac{\kappa_\dagger}{\gamma_\dagger}(\frac{\|\delta A\|_2}{\|A\|}\|x\| + \frac{\|\delta B\|}{\|A\|} + \frac{\kappa_\dagger}{\gamma_\dagger}\frac{\|\delta A\|_2}{\|A\|}\frac{\|r_x\|}{\|A\|}), k = 1, 2, \cdots, n.$$

is the perturbation bound when ϵ is regarded as a disturbance of $x(k)$ ($\epsilon \neq 0$ and $\|A^\dagger\|_2\|\delta A\|_2 < 1$), then $L[x(k)]$ will increase when $k \to n$.

Proof Because all the matrix norms are equivalent, let the tolerance norm be the matrix norm for the sake of convenience. If ϵ is regarded as a disturbance of $x(1)$, then

$$B + \delta B = \begin{bmatrix} (\frac{2}{(n+1)n} - \frac{2}{(n+2)(n-1)})\sum\limits_{k=2}^{n} kx(k) - \frac{2}{(n+1)n}(x(1)+\epsilon) \\ y(3) - y(2) \\ \vdots \\ y(n) - y(n-1) \end{bmatrix},$$

$$\delta B = \begin{bmatrix} -\frac{2}{n(n+1)}\epsilon \\ 0 \\ \vdots \\ 0 \end{bmatrix};$$

Because

$$A = \begin{bmatrix} -z(2) & 2 \\ -z(3) & 3 \\ \vdots & \vdots \\ -z(n) & n \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & 0 & \cdots & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{1}{2} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} y(1) & -\frac{3}{2} \\ y(2) & -\frac{5}{2} \\ \vdots & \vdots \\ y(n) & -\frac{2n+1}{2} \end{bmatrix}$$

we have

$$A + \delta A = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & 0 & \cdots & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{1}{2} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{2}{(n+1)n}\sum\limits_{k=2}^{n} kx(k) + \frac{2}{(n+1)n}(x(1)+\epsilon) & -\frac{3}{2} \\ y(2) & -\frac{5}{2} \\ \vdots & \vdots \\ y(n) & -\frac{2n+1}{2} \end{bmatrix}$$

$$= A + \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & 0 & \cdots & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{1}{2} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{2}{(n+1)n}\epsilon & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}$$

$$= A + \begin{bmatrix} -\frac{\epsilon}{(n+1)n} & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}$$

thus $\|\delta B\|_2 = \frac{2|\epsilon|}{(n+1)n}$, $\|\delta A\|_2 = \sqrt{\lambda_{max}(\delta A^T \delta A)} = \frac{|\epsilon|}{(n+1)n}$.

If ϵ is regarded as a disturbance of $x(2)$, thus $\|\delta B\|_2 = \sqrt{(\frac{4}{(n+1)n})^2 + (\frac{4}{(n-1)(n+2)})^2}|\epsilon|$, $\|\delta A\|_2 = \sqrt{\frac{1}{n^2(n+1)^2} + (\frac{1}{(n-1)(n+2)})^2} 4|\epsilon|$.

Similarly, if ϵ is regarded as a disturbance of $x(k)$, we also obtain $\|\delta A\|_2$ and $\|\delta B\|_2$. Obviously, $\|\delta A\|_2$ and $\|\delta B\|_2$ will increase when $k \to n$. So $\frac{\kappa_\dagger}{\gamma_\dagger}$ and $\|\delta A\|_2$ are all increasing function of k, and other variables do not change. We obtain that the perturbation bound $L[x(k)]$ is the increasing function of k, i.e. $L[x(k)]$ will increase when $k \to n$.

$L[x(k)]$ will increase when $k \to n$, which indicates that $x(k)$ has more influence on the parameters (a and b) than $x(k-1)$ and $x(n)$ has most influence on the parameters (a and b). So we conclude that WGM(1,1) can pay more attention to the new information.

A Theorem is given in order to discuss the error from the inverse time varying weighted operator.

Theorem 2 $Y = \{y(1), y(2), \cdots, y(n)\}$ is viewed as time varying weighted generating sequence for original sequence X. $\hat{Y} = \{\hat{y}(1), \hat{y}(2), \cdots, \hat{y}(n)\}$ is the fitted value of Y. $\hat{X} = \{\hat{x}(1), \hat{x}(2), \cdots, \hat{x}(n)\}$ is the fitted value of X. If $|y(k) - \hat{y}(k)| < \epsilon$ for $\forall k = 1, 2, \cdots, n$, then $|x(k) - \hat{x}(k)| < (2n - 2k + 1)\epsilon$

Proof For $\forall k = 1, 2, \cdots, n, |y(k) - \hat{y}(k)| < \epsilon$, then $|x(k) - \hat{x}(k)| = |\frac{(n-k+1)(n+k)}{2k} y(k) - \frac{(n-k)(n+k+1)}{2k} y(k+1) - \frac{(n-k+1)(n+k)}{2k} \hat{y}(k) + \frac{(n-k)(n+k+1)}{2k} \hat{y}(k+1)| = |\frac{(n-k+1)(n+k)}{2k} y(k) - \frac{(n-k+1)(n+k)}{2k} \hat{y}(k) + \frac{(n-k)(n+k+1)}{2k} \hat{y}(k+1) - \frac{(n-k)(n+k+1)}{2k} y(k+1)| < \frac{(n-k+1)(n+k)}{2k} |y(k) - \hat{y}(k)| + \frac{(n-k)(n+k+1)}{2k} |\hat{y}(k+1) - y(k+1)| = \frac{(n-k+1)(n+k)}{2k} \epsilon + \frac{(n-k)(n+k+1)}{2k} \epsilon = \frac{n^2-k^2+n}{k} \epsilon$.

The largest error from the inverse time varying weighted generating operator is $\frac{n^2-k^2+n}{k}\epsilon$, which indicates that the error will be reduced when $k \to n$. From the restoring error viewpoint, WGM(1,1) model can indeed pay more attention to the new information. The largest error $\frac{n^2-k^2+n}{k}\epsilon$ can also explore the influence of sample size, indicating that the bigger n is, the larger the error is.

Remark 1

Each data $y(k)$ would change when $x(1)$ is disturbed by ϵ. So it is certaint that the simulative value ($\hat{x}(k)$) will change under WGM(1,1). We can say that WGM(1,1) can make better use of the available data. However, the simulative value of non-homogenous discrete grey model will not change when $x(1)$ is disturbed by ϵ, i.e., the non-homogenous discrete grey model can not use the first data $x(1)$ [15].

Remark 2 It is obvious the traditional GM(1,1) is an exponential model. It is decreased or increased. However, the monotonicity of the simulative sequence $\hat{X}$ is uncertain when we construct WGM(1,1) model. Its monotonicity is determined by the actual data, as stated in Case 2 below. Because

$$\hat{x}(k) = \frac{(n-k+1)(n+k)}{2k}\hat{y}(k) - \frac{(n-k)(n+k+1)}{2k}\hat{y}(k+1) = \frac{(n-k+1)(n+k)}{2k}\{[y(1) - \frac{\hat{b}}{\hat{a}} + \frac{\hat{b}}{\hat{a}^2}]$$
$$\times e^{-\hat{a}(k-1)} + \frac{\hat{b}}{\hat{a}}(k-1) - \frac{\hat{b}}{\hat{a}^2}\} - \frac{(n-k)(n+k+1)}{2k}\{[y(1) - \frac{\hat{b}}{\hat{a}} + \frac{\hat{b}}{\hat{a}^2}]e^{-\hat{a}k} + \frac{\hat{b}}{\hat{a}}k - \frac{\hat{b}}{\hat{a}^2}\}$$

From the above equation we come to the conclusion that the simulative sequence $\hat{x}(k)(k = 1, 2, \cdots, n)$ is not always an exponential-growth model. So the monotonicity of the simulative value X is uncertain.

Actually, to remove the influence of initial condition on modelling and to further enhance the accuracy of WGM(1,1), it is assumed that $y(k) = \beta_1 e^{-\hat{a}(k-1)} + \beta_2 k + \beta_0$, where $\beta_1, \beta_2, \beta_0$ are the unknown parameters. These parameters can be obtained by the least squares method:

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_0 \end{bmatrix} = (D^T D)^{-1} D^T F$$

where

$$B = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(n) \end{bmatrix}, F = \begin{bmatrix} e^{-\hat{a}} & 1 & 1 \\ e^{-2\hat{a}} & 2 & 1 \\ \vdots & \vdots & \\ e^{-n\hat{a}} & n & 1 \end{bmatrix}$$

Thus, $\hat{y}(k) = \hat{\beta}_1 e^{-\hat{a}(k-1)} + \hat{\beta}_2 k + \hat{\beta}_0$, which is a 4-parametric exponential model.

3. Validation of the WGM(1,1) model

In this section, three cases are used to verify the applicability of the proposed model with criterion: mean absolute percentage error (MAPE $= 100\% \frac{1}{n} \sum_{k=1}^{n} |\frac{x(k)-\hat{x}(k)}{x(k)}|$).

Case 1. The failure rate forecasting example [16]

Shao et al. applied GM(1,1) to predict the failure rate of weapon spare parts. We adopted the same sample as reference [16] to construct WGM(1,1). The unit of data is the frequency per hour. As Table 1 shown, MAPE of WGM(1,1) is smaller than that of GM(1,1). Hence the former has better forecasting accuracy.

Table 1 The fitting values and MAPE of different grey models

Year	Actual value	GM(1,1) value	WGM(1,1) value
4	0.15	0.15	0.168
5	0.3342	0.3868	0.399
6	0.655	0.5893	0.627
7	1.1158	0.8976	1.014
MAPE		15.11	10.3
8	1.5320	1.3671	1.510
9	2.201	2.0823	2.089
MAPE		8.08	3.28

Case 2. The Russia electricity consumption forecasting example [13]

Li et al. applied adaptive GM(1,1) to predict the electricity consumption of Russia. The WGM(1,1) is built by the same sample as Reference [13]. Its results is given in Table 2 and Table 3.

Table 2 The fitting values and MAPE of different grey models (unit: KTOE)

Year	electricity consumption	GM(1,1) value	WGM(1,1) value
2000	52333	52333	53694
2001	53151	52952	52933
2002	53168	53561	53237
2003	54372	54177	54089
MAPE		0.5	0.9

As shown in Table 2 and Table 3, the MAPE of GM(1,1) is smaller than that of WGM(1,1) for the in-sample data, but the MAPE of GM(1,1) is larger than that of WGM(1,1) for the out-of-sample data. This demonstrates that WGM(1,1) can mitigate the drawbacks of GM(1,1), and can bring better forecasting results by capturing the characteristics of recent data. That is to say, simply reducing the simulation errors does not necessarily increase the forecasting accuracy. Meanwhile, AGM(1,1) [13] also cannot produce the desired forecasting results.

Table 3 The forecasting values and MAPE of different grey models (unit: KTOE)

Year	electricity consumption	GM(1,1)	AGM(1,1)[13]	WGM(1,1)
2004	55516	54800	54621	55382
2005	55898	55431	55195	57080
2006	58600	56069	55776	59166
2007	60281	56714	56362	61629
MAPE		3.1	3.55	1.4

Case 3. The tuberculosis incidence forecasting example [17]

Wu and Chen used the tuberculosis incidence during 2009-2012 as the in-sample data. The tuberculosis incidence of 2013 is regarded as out-of-sample. Actual value and the errors of two models are listed in Table 4. The results show that the forecasting accuracy of the WGM(1,1) is better than traditional GM(1,1), although the fitting accuracy of these two models are approximate.

Table 4 The forecasting values and MAPE of different grey models

Year	Actual value	GM(1,1)	WGM(1,1)
2009	184.4	184.4	191.4
2010	116.9	115.2	119.5
2011	96.3	99.6	95.2
2012	88.3	86.2	85.9
MAPE		2.4	2.3
2013	81.59	74.6	84.5
MAPE		8.6	3.6

Case 4. The tuberculosis incidence forecasting example [18]

Wu and Chen used the tuberculosis incidence during 2009-2012 as the in-sample data. The tuberculosis incidence of 2013 is regarded as out-of-sample. Actual value and the errors of two models are listed in Table 4. The results show that the forecasting accuracy of the WGM(1,1) is better than traditional GM(1,1), although the fitting accuracy of these two models are approximate.

Table 4 The forecasting values and MAPE of different grey models

Time	Actual value	GM(1,1)	WGM(1,1)	Model (1)	AR(1) value
1	0.2305	0.2305	0.2489	0.2761	
2	0.3197	0.3633	0.3485	0.3136	0.2974
3	0.3308	0.3869	0.3534	0.3512	0.3769
4	0.3714	0.4120	0.4316	0.3887	0.3868
5	0.4443	0.4388	0.4837	0.4262	0.4230
6	0.4868	0.4673	0.5229	0.4638	0.4881
7	0.5383	0.4977	0.5550	0.5013	0.5260
8	0.5310	0.5300	0.5826	0.5388	0.5719
9	0.6025	0.5644	0.6075	0.5764	0.5654
10	0.6528	0.6011	0.6305	0.6139	0.6292
11	0.6828	0.6401	0.6523	0.6514	0.6740
12	0.7055	0.6817	0.6732	0.6890	0.7008
13	0.6704	0.7260	0.6936	0.7265	0.7211
14	0.7140	0.7732	0.7136	0.7640	0.6897
MAPE		6.9	6.1	5.9	4.8
15	0.7343	0.8234	0.7334	0.8016	0.7070
16	0.7634	0.8769	0.7531	0.8391	0.7224
17	0.7885	0.9339	0.7728	0.8766	0.7361
MAPE		15.1	1.2	9.2	5.2

The properties of the time series is conserved

4. Conclusion

Grey generating operators play a significant role in grey forecasting models. However, the existing accumulated generating operator increases the weight of old data. It is well known that the objective of mining the past is to determine the future. Taking full advantage of every single piece of information is important, especially the recent information. The time varying weighted generating operator not only increases the weights of new data but also relieves the influence of numerical fluctuations. WGM(1,1) is developed to fully extract information concealed in new data, and it has a varying weight k that can realistically reflect the non-linearity of the data. This research further explores the performance of WGM(1,1) with three cases. The results show that WGM(1,1) is practical and reliable forecasting tool.

With regard to directions for future research, one suggestion is to model grey multi-variable models and grey model with rolling mechanism by the time varying weighted generating operator. It is also suggested that the WGM(1,1) be applied to other practical fields, to further confirm its effectiveness.

Acknowledgements

The relevant researches done in this paper are supported by a Marie Curie International Incoming Fellowship within the 7th European Community Framework Programme (No. FP7-PIIF-GA-2013-629051), the National Natural Science Foundation of China (No.71171113 and 71171116), funding for talent introduction in NUAA (1009-YAH14003), Project of Social Science Foundation of the China (No.12AZD102).

References

[1] Falk Br?uning, Siem Jan Koopman. Forecasting macroeconomic variables using collapsed dynamic factor analysis. International Journal of Forecasting 2014,30(3): 572-584.

[2] Handa Yang, Ben Kurtz, Dung Nguyen, et al. Solar irradiance forecasting using a ground-based sky imager developed at UC San Diego. Solar Energy 2014, 103: 502-524.

[3] Weigang Zhao, Jianzhou Wang, Haiyan Lu. Combining forecasts of electricity consumption in China with time-varying weights updated by a high-order Markov chain model. Omega

2014, 45: 80-91.

[4] Giacomo Sbrana, Andrea Silvestrini. Random switching exponential smoothing and inventory forecasting. International Journal of Production Economics 2014,156: 283-294.

[5] Che-Jung Chang, Der-Chiang Li, Chien-Chih Chen, et al. A forecasting model for small non-equigap data sets considering data weights and occurrence possibilities. Computers & Industrial Engineering 2014,67:139-145.

[6] Sifeng Liu, Jeffrey Forrest, Yingjie Yang. Advances in grey systems research. The Journal of Grey System 2013(2):1-18.

[7] Lifeng Wu, Sifeng Liu, Ligen Yao, et al. The effect of sample size on the grey system model, Applied Mathematical Modelling 37(2013): 6577-6583.

[8] Guo-Dong Li, Shiro Masuda, Masatake Nagai. The prediction model for electrical power system using an improved hybrid optimization model. Electrical Power and Energy Systems (2013)44:981-987.

[9] Yong Wang, Qinbao Song, Stephen MacDonell, et al. Integrate the GM(1,1) and Verhulst models to predict software stage effort. IEEE Transactions on Systems, Man, and Cyybernetics part C: Applications and Reviews, 2009,39(6):647-658.

[10] Dinh Quang Truong, Kyoung Kwan Ahn, Nguyen Thanh Trung. Design of an advanced time delay measurement and a smart adaptive unequal interval grey predictor for real-time nonlinear control systems. IEEE Transactions on Industrial Electronics 2013,60(10):4574-4589.

[11] Whei-Min Lin, Chih-Ming Hong, Cong-Hui Huang, et al. Hybrid control of a wind induction generator based on grey-Elman neural network. IEEE Transactions on Control Systems Technology 2013,21(6):2367-2373.

[12] Tsan-Ming Choi, Chi-Leung Hui, Sau-Fun Ng, et al. Color trend forecasting of fashionable products with very few historical data. IEEE Transactions on Systems, Man and Cybernetics-Part C: Applications and Reviews 2012,42(6): 1003-1010.

[13] Li D.C., Chang C. J., Chen C. C., et al. 2012. Forecasting short-term electricity consumption using the adaptive grey-based approach-An Asian case. Omega 40(6):767-773.

[14] Lifeng Wu,Sifeng Liu,Ligen Yao,et al.,Grey system model with the fractional order accumulation. Communications in Nonlinear Science and Numerical Simulation 2013,18(7):1775-1785.

[15] Li-Feng Wu, Si-Feng Liu, Wei Cui, et al. Non-homogenous discrete grey model with fractional-order accumulation. Neural Comput & Applic, DOI 10.1007/s00521-014-1605-1.

[16] Shao Yanjun, Ma Chun-mao, Pan Hongxia, et al. Failure rate prediction research of weapon sparts based on discrete GM(1,1). Journal of Gun Launch & Control 2013,6: 96-99.

[17] Wu Yingmei, Chen Xiaoning. Application of grey model in prediction of tuberculosis incidence. China Tropical Medicine 2014,14(4):423-426.

[18] Forecast of power generation for grid-cnnected photovoltaic system based on grey theory and verification model. Energy and Power Engineering 2013, 5:177-181.