Advanced Engineering Forum
ISSN: 2234-991X, Vols. 6-7, pp 97-103
doi:10.4028/www.scientific.net/AEF.6-7.97 Online: 2012-09-26

Multiple Variables Time Series Adaptive Prediction Model Based on Grey Theory

Li Xiaolei[1,a] ,Ma Jiezhong[1,b] ,Guo Yangming[1,c]

[1]Northwesten Polytechnical University. Xi'an, Shaanxi Province, 710072, China

[a]465628547@qq.com, [b]majz@nwpu.edu.cn, [c]yangming_g@nwpu.edu.cn

Keywords: Grey Prediction; Multi-Variable Time Series; Prediction Model

Abstract. Fault prediction is critical to ensure the safety and reliability of complex system. The reported fault prediction methods have achieved some success in practical applications. Generally, the information used in fault prediction is always mined from multi-variable time series and small simple data. Thus, based on grey prediction theory, an adaptive prediction model with multi-variable small simple time series data is proposed. In this method, after analyzing the disadvantages of $GM(1,1)$ model, we modify the initial values and background values of $GM(1,1)$ model, and then the interrelations and characteristics of the multiple variables time series are taken into account. The results of experiment with a certain complex system show that the model has good prediction precision, which will be useful in applications.

Introduction

Fault diagnosis and prediction of complex system has attached more and more attention, because it is very significant to ensure normal operation of complex system, especially under limited information. Therefore, fault prediction used to safeguard complex system's safety has been extensively studied. But, fault prediction of complex system often faces to solve the small sample problem. Moreover, one fault mode is often represented by several variable data, and one variable data may include several different fault modes. So it is necessary to consider several multiple relevant variables data used to fault prediction at the same time However, the known fault prediction methods have the following drawbacks: (1) These methods only considered one variable time series or separately considered the development and changes of several variable time series. And the methods always required large sample data, which is not suitable to solve small sample data. (2) Although some methods took the relations of relevant variable time series into account, they can not be applied to practical engineering applications for their complex.

In recent years, grey prediction[1,2] based on grey theory has become a hot research filed. It is demonstrated by theory and application that grey prediction model is appropriate to engineering application, because it only need little number of sample data and its expression and calculation are very simple.

$GM(1,1)$ model is the most commonly used prediction model. With an analysis of the disadvantage of $GM(1,1)$ model, an adaptive prediction model with multiple variable small sample time series is proposed. This model considers the relationship among the several time series and develops the initial values and background values of $GM(1,1)$. The experiments are conducted by a complex system.

A Brief Introduce of $GM(1,1)$ Model

Assume an original time series $X^{(0)}$,

$$X^{(0)}= (x^{(0)}(1),\ x^{(0)}(2),\dots,\ x^{(0)}(n)) \tag{1}$$

A new time series $X^{(1)}$ is generated by the accumulated generating operation.

$$X^{(1)}= (x^{(1)}(1),\ x^{(1)}(2),\dots,\ x^{(1)}(n)) \tag{2}$$

where $x^{(1)}(k)=\sum_{i=1}^{k} x^{(0)}(k)$, $k=1,2,\ldots,n$.

$GM(1,1)$ is based on the theory of gray number whitenization. So, let the background value be $z^{(1)}(k)=0.5(x^{(1)}(k)+x^{(1)}(k-1))$. Then we can get the basic form of $GM(1,1)$ model and its whitenization equations.

$$x^{(0)}(k)+az^{(1)}(k)=b,\quad k=2,3,\ldots,n \tag{3}$$

$$\frac{dx^{(1)}(t)}{dt}+ax^{(1)}(t)=b \tag{4}$$

The parameters a and b can be calculated by the least squares method.

$$\hat{a}=\begin{bmatrix} a & b \end{bmatrix}^T=(B^TB)^{-1}B^TY \tag{5}$$

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

The prediction equations are shown as follows.

$$\hat{x}^{(1)}(k+1)=(x^{(0)}(1)-b/a)e^{-ak}+b/a \tag{6}$$

$$\hat{x}^{(0)}(k+1)=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k) \tag{7}$$

where $k=1,2,\ldots,n$, and $\hat{x}^{(0)}(1)=x^{(0)}(1)$.

Analysis of Grey Model $GM(1,1)$

From the section 1, we can conclude that $GM(1,1)$ has the following disadvantages.

(1) The Ref.[3] pointed out that $GM(1,1)$ prediction model is essentially an extrapolation method, because it uses the exponential curve to fit the accumulated time series and makes the first data of the historic data as the initial values. The least squares method is utilized to perform the fitting process. But, according to the principle of the least squares, the fitting curve does not necessarily pass through the first data point. In addition, the new information priority principle[1] shows that the new information has greater value for prediction. So it is not a good choice to select the first data value as the initial value in $GM(1,1)$ model.

(2) $GM(1,1)$ prediction model is only used a single time series and it can not represent the interaction and collaborative development among multiple variables time series. So some scholars proposed grey multi-variable model $MGM(1,n)$ [4] based on $GM(1,1)$ model.

(3) The parameters a and b of the grey model $GM(1,1)$ [5] have great influence on the prediction accuracy of the grey model $GM(1,1)$. The values of a and b depend on the original time series and the structure form of background values, that is, $z^{(1)}(k)$ in the original matrix B is one of the key factors affecting the prediction accuracy.

Under normal conditions, $z^{(1)}(k)$ is generated by the average generation method. Fig. 1 shows that the method is appropriate while the time interval is very short and the data changes smoothly.

From Fig. 1, the conformation of $z^{(1)}(k)$ can be regarded as the area of ABCD. In fact, if the original time series $x^{(0)}(k)$ is concave, $x^{(1)}(k)$ is always concave too. When the fitting curve of grey model is exponential curve, its corresponding area at the interval $[k-1,k]$ is always less than the area of trapezoid ABCD. This way, the value of $z^{(1)}(k)$ is always larger than its actual value, therefore, the model generates hysteresis error. And the obvious change of original time series will lead to larger hysteresis error, this means that $GM(1,1)$ could get high prediction accuracy only when the absolute value of the parameter a is very small.

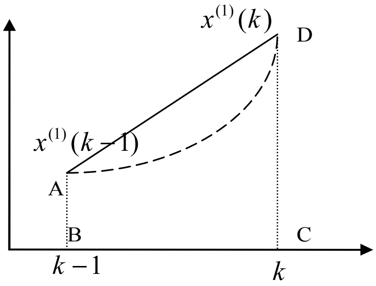

Fig. 1 Conformation figure of $z^{(1)}(k)$

Adaptive Multi-variable Time Series Prediction Model

Multi-variable Prediction Model. Several simulation experiments and applications show that $MGM(1,n)$ model is better than $GM(1,1)$ model in fitting and prediction accuracy. But, $MGM(1,n)$ model is not suitable for the changing sharply variables data. Especially, its fitting results are poor when variables change near a certain value, which leads to its poor applicability.

Using the idea of Ref. [4], we establish multi-variable time series prediction model based on $GM(1,1)$ model. Here, the original condition is modified with the i-*th* element ($x_i^{(1)}(l)$) and the background value is also modified.

The process of the proposed method is shown as follows.

$x_i^{(0)}(k)$, $i=1,2,...,n$, is the i-*th* variable time series. $x_i^{(1)}(k)$ is generated by accumulated generating operation.

$$x_i^{(1)}(k)=\sum_{j=1}^{k} x_i^{(0)}(j) \tag{8}$$

where $k=1,2,...,m$。

$$X^{(0)}(k)=(x_1^{(0)}(k),x_2^{(0)}(k),...,x_n^{(0)}(k))^T \quad X^{(1)}(k)=(x_1^{(1)}(k),x_2^{(1)}(k),...,x_n^{(1)}(k))^T$$

According to the Eq.(3) and Eq.(4), the n-elements-first-order-differential-equations are established.

$$\begin{cases} \dfrac{dx_1^{(1)}}{dt}=a_{11}x_1^{(1)}+a_{12}x_2^{(1)}+...+a_{1n}x_n^{(1)}+b_1 \\ \dfrac{dx_2^{(1)}}{dt}=a_{21}x_1^{(1)}+a_{22}x_2^{(1)}+...+a_{2n}x_n^{(1)}+b_2 \\ \quad ... \\ \dfrac{dx_n^{(1)}}{dt}=a_{n1}x_1^{(1)}+a_{n2}x_2^{(1)}+...+a_{nn}x_n^{(1)}+b_n \end{cases} \tag{9}$$

Let $A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$, $B2=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$

Eq.(9) can be written in matrix form.

$$\frac{dX^{(1)}}{dt} = AX^{(1)} + B2 \tag{10}$$

In order to indentify the parameters A and $B2$, the Eq.(9) and Eq.(10) need to be discretized. According to the definition of differential, the forward and backward difference forms of Eq.(9) and Eq.(10) are expressed as follows.

Forward Difference Form

$$X^{(1)}(t+\Delta t) - X^{(1)}(t) = AX^{(1)}(t) + B2 \tag{11}$$

Backward Difference Form

$$X^{(1)}(t+\Delta t) - X^{(1)}(t) = AX^{(1)}(t+\Delta t) + B2 \tag{12}$$

where, Δt is unit time interval.

Following the grey theory, $X^{(1)}$ is used to yield the background values. We propose that the general differential form with（11）×ω＋（12）×（1-ω）, and then

$$X^{(1)}(t+\Delta t) - X^{(1)}(t) = A(\omega X^{(1)}(t) + (1-\omega)X^{(1)}(t+\Delta t)) + B2 \tag{13}$$

where, ω in the Eq.(13) is weight factor. $0 \le \omega \le 1$. Let $a_i = \left[a_{i1}, a_{i2}, ..., a_{in}, b_i\right]^T$, $i = 1,2,...,n$. The identification value of a_i can be obtained by least squares method.

$$\hat{a}_i = \begin{bmatrix} \hat{a}_{i1} & \hat{a}_{i2} & \cdots & \hat{a}_{in} & \hat{b}_i \end{bmatrix}^T = (L^T L)^{-1} L^T Y_i \tag{14}$$

where $$L = \begin{bmatrix} \omega x_1^{(1)}(1) + (1-\omega)x_1^{(1)}(2) & \omega x_2^{(1)}(1) + (1-\omega)x_2^{(1)}(2) & \cdots & \omega x_n^{(1)}(1) + (1-\omega)x_n^{(1)}(2) & 1 \\ \omega x_1^{(1)}(2) + (1-\omega)x_1^{(1)}(3) & \omega x_2^{(1)}(2) + (1-\omega)x_2^{(1)}(3) & \cdots & \omega x_n^{(1)}(2) + (1-\omega)x_n^{(1)}(3) & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \omega x_1^{(1)}(m-1) + (1-\omega)x_1^{(1)}(m) & \omega x_2^{(1)}(m-1) + (1-\omega)x_2^{(1)}(m) & \cdots & \omega x_n^{(1)}(m-1) + (1-\omega)x_n^{(1)}(m) & 1 \end{bmatrix} \tag{15}$$

$Y_i = \begin{bmatrix} x_i^{(0)}(2) & x_i^{(0)}(3) & \cdots & x_i^{(0)}(m) \end{bmatrix}^T$

The identification values of A and $B2$ are:

$$\hat{A} = \begin{bmatrix} \hat{a}_{11} & \hat{a}_{12} & \cdots & \hat{a}_{1n} \\ \hat{a}_{21} & \hat{a}_{22} & \cdots & \hat{a}_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{a}_{n1} & \hat{a}_{n2} & \cdots & \hat{a}_{nn} \end{bmatrix}, \quad \hat{B}2 = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \vdots \\ \hat{b}_n \end{bmatrix} \tag{16}$$

Let the l-th element ($x_i^{(1)}(l)$) in $x_i^{(1)}$ as the original condition. Then continuous-time response of Eq.(10) is presented as follows.

$$X^{(1)}(t) = e^{At}X^{(1)}(l) + A^{-1}(e^{At} - I)B2 \tag{17}$$

where $e^{At} = I + At + \frac{A^2}{2!}t^2 + ... = I + \sum_{k-1}^{\infty} \frac{A^k}{k!}t^k$.

Calculated values of the multi-variable time series prediction model are shown as follows.

$$\hat{X}^{(1)}(k) = e^{\hat{A}(k-l)}X^{(1)}(l) + \hat{A}^{-1}(e^{\hat{A}(k-l)} - I)\hat{B}2 \tag{18}$$

where $k = 1,2,...$, $1 \le l \le m$. So,

$$\hat{X}^{(0)}(k) = \hat{X}^{(1)}(k) - \hat{X}^{(1)}(k-1), \quad k = 2,3,... \tag{19}$$

From the above derivation, it can be seen that multi-variable time series prediction model based on grey prediction modeling theory, takes the interrelations and characteristics of the multiple variables time series into account. The model overcomes the drawbacks in $GM(1,1)$ and it is the promotion of $GM(1,1)$ prediction model with n variables. The new model is not simple combination of several separate $GM(1,1)$ models and it is also different from $GM(1,n)$ model.

Adaptive Multi-variable Time Series Prediction Model. With the growth of prediction period, the disturbances and random factors in the future will have great impact on the prediction performance. According to the new information priority principle, we propose adaptive multi-variable time series prediction model with recursive compensation. In the proposed model, the new prediction data are added into the prediction time series and the oldest ones are dropped. This way, the new time series are used to execute the next prediction.

For the proposed model, we need to get the optimal value of ω to improve the prediction accuracy. Now, the main methods include gray relational grade-fuzzy nearness, genetic algorithm, particle swarm optimization(PSO), etc[6,7]. POS[7] is an evolutionary computation technique. PSO algorithm has a simple concept and it is easy to realize. Thus, PSO receive more attention in recent years. In this study, we use POS to obtain the optimal ω.

The prediction steps of adaptive multi-variable time series prediction model are described as follows.

(1) Perform accumulated generating operation with the original multi-variable time series.

(2) Select the initial values and let $l=1$, 2,…, n, calculate the parameters A, $B2$ and ω depending on the fitting accuracy of the model.

(3) Determine the values of A, $B2$, ω and l. Establish the best prediction model and make one step prediction.

(4) Add one step prediction values into the time series and drop the oldest data. Repeat the steps (1), (2), (3) and (4).

Case Analysis

In order to estimate the performance of the proposed model, we select 10 sets of original data from a complex system with three variables. Then, the predicting results of $MGM(1,n)$ [4] and the proposed model are compared.

Table 1 Original data

No.	$x_1^{(0)}(k)$	$x_2^{(0)}(k)$	$x_3^{(0)}(k)$
1	9.60	10.23	9.96
2	8.99	10.55	9.76
3	8.38	10.86	9.39
4	7.78	11.15	8.87
5	7.18	11.43	8.23
6	6.60	11.69	7.50
7	6.03	11.93	6.72
8	5.48	12.16	5.92
9	4.95	12.37	5.14
10	4.44	12.56	4.28

We use the first 8 sets data in Table 1 to establish multi-variable time series prediction model and analyze the model's fitting accuracy. The last 2 sets data are used to test the prediction performance.

The average relative error (ARE) of fitting values of the model, shown in Eq.(21), are measured as the criterion to test simulation accuracy.

$$\sigma = \frac{\sum_{i=1}^{n}\sum_{k=1}^{m} w_i(k)}{nm} \tag{21}$$

where $w_i(k) = |\hat{x}_i^{(0)}(k) - x_i^{(0)}(k)| / x_i^{(0)}(k)$, $i = 1,2,...,n$, $k = 1,2,...,m$.

The prediction results of $x_1^{(0)}(k)$, $x_2^{(0)}(k)$ and $x_3^{(0)}(k)$ using $MGM(1,1)$ described in Ref.[4] are shown in Table 2.

Table 2 Prediction results and relative errors of the $MGM(1,n)$ in Ref.[4]

No.	$x_1^{(0)}(k)$	ARE/%	$x_2^{(0)}(k)$	ARE/%	$x_3^{(0)}(k)$	ARE/%
9	4.9448	0.1051	12.3715	0.0121	5.0878	1.0156
10	4.4310	0.2027	12.5669	0.0549	4.2627	2.6781

In order to establish the proposed adaptive mutli-variable time series prediction model, the different initial values are chosen and the value of ω is optimized by POS. The parameters of POS are set as follows: population size M = 200, the acceleration coefficient $c_1 = c_2 = 2$, the inertia weight λ = 0.25, maximum allowable iterative times G = 20. According to the minimum fitting relative errors of the model, we set $x_i^{(1)}(1)$ as the initial value and ω=0.503. In this case, the fitting relative errors of the three variables time series are x_1 =0.0182%, x_2 =0.0084%, x_3 =0.0656%. The average relative error of the three variables time series is 0.0307%. The one step prediction results with the proposed model are shown in Table 3.

Table 3 One step prediction results and relative errors of the proposed model

No.	$x_1^{(0)}(k)$	ARE/%	$x_2^{(0)}(k)$	ARE/%	$x_3^{(0)}(k)$	ARE/%
9	4.9459	0.0828	12.3712	0.0097	5.0903	0.9669

One step prediction results are added into the time series and the oldest data is removed. The new initial values are chosen and the new value of ω is optimized using the POS. Here, $x_i^{(1)}(1)$ is initial condition and ω=0.503, the fitting relative errors of x_1, x_2 and x_3 are 0.0177%, 0.0086% and 0.0555% respectively. The average relative error is 0.0272%. Using this model, the one step prediction results and relative errors are shown in Table 4.

Table 4 One step prediction results and relative errors of the proposed model

No.	$x_1^{(0)}(k)$	ARE/%	$x_2^{(0)}(k)$	ARE/%	$x_3^{(0)}(k)$	ARE/%
10	4.4325	0.1689	12.5654	0.0430	4.2658	2.6073

From Table 2 to Table 4, we can see that the proposed adaptive multi-variable time series prediction model has higher fitting and prediction accuracy than $MGM(1,n)$ model.

Conclusions

In this work, we analyze the disadvantage of $GM(1,1)$ prediction model. An adaptive prediction model with several variables time series is established. The model applies fitting mean relative error as evaluation criterion and determines the optimal initial values and background values by PSO algorithm. The model has the following characteristics.

(1) The model uses several relative time series to enrich the prediction information, i.e., each time series not only provides relevant information for their own forecasting, but also provide the necessary information for the other series prediction. This characteristic is more valuable to small sample data.

(2) The model considers characteristics of time series and uses fitting mean relative error as evaluation criterion. The applicability of the model is improved using the optimal initial values and background values.

(3) According to the new information priority principle, the new data has great impact on prediction results. Thus, one step prediction data are added into the original time series and the oldest data are removed, which make the model have adaptive capability.

We conduct comparison experiments of a complex system using the proposed method. The results show that the proposed model has higher fitting and prediction accuracy than the prediction model proposed in Ref.[4]. It will be useful in actual applications.

References

[1] Deng Julong. The Primary Methods of Grey System Theory[M]. Wuhan: Huazhong University of Science and Technology Press, 2004. (in Chinese)

[2] Liu Sifeng, Dang Yaoguo, Fang Zhigeng. Application of Grey System Theory[M]. Beijing: Press of Science, 2004:24-27. (in Chinese)

[3] Zhang Dahai, Jiang Shifang, Shi Kaiquan. Theoretical Defect of Grey Prediction Formula and Its Improvement[J]. Systems Engineering-Theory & Practice, 2002, 22(8):140-142. (in Chinese)

[4] Zhai Jun, Sheng Jianming. Grey Model $MGM(1,n)$ and Its Application[J]. Systems Engineering-Theory & Practice, 1997，5：109-113. (in Chinese)

[5] Zhang Yi, Wei Yong, Xiong Changwei.One New Optimized Method of $GM(1,1)$ Model[J]. Systems Engineering-Theory & Practice, 2007, 4：141-146. (in Chinese)

[6] Liu Hong, Zhang Qishan. Life-span Prediction on Mechanical Products of Model Based on Particle Swarm Algorithm[J]. Journal of Machine Design, 2007, 24(10)：4-5,61. (in Chinese)

[7] Chen Guochu, Yu Jinshou. Particle Swarm Optimization Algorithm[J]. Information and Control, 2005, 3:318-324. (in Chinese)