http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (15): 3335-3339, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.3335.3339

Application of Single Sequence First-Order Linear Dynamic Model in Performance Prediction of the Olympic Games Men's 400 m Freestyle

Wei Zhang
Department of Sports, Tianjin Polytechnic University, 300387, Tianjin, China

Abstract: The Olympics is the biggest event in the international sports meeting. The athletes getting medals in the Olympic Games can not only demonstrate personal style, but also win the honor for their motherland. Thus each country's cultivation of athletes has a strong force. In order to make their own country's athletes reach the world's top level, it is not enough just to achieve the predecessors' performance. Since, each game may have a new breakthrough, sports scores forecast becomes a demand of sports development strategy and athletes training methods. This study uses the champion performance data of men's 400m freestyle in recent 10 Olympic Games, establishes a single sequence first-order linear dynamic model as a predictive model and conducts single sequence residuals identify forecast using the differential dynamic modeling method. It predicts the champion score of men's 400 m freestyle in the 31st Olympic Games is 220.16 sec; the score shows that the achievements of the 31st champion will not break the record, but close to the last record in the Olympic Games. This study provides a reference for the selection and training of next Olympics athletes in various countries around the world and lays a foundation for the relevant scientific research.

Key words: First-order linear model, gray theory, accuracy testing point by point

INTRODUCTION

There are many uncertain factors in the athlete's performance prediction process. If it can be enumerated, the number of predicted independent variables would become bigger increasing forecast calculation complexity (Tan and Liu, 2003). Moreover we still cannot clearly figure out the variable relationship of all influencing factors with modern science. At this time, using the gray prediction model in gray theory is a much better research method (Liu, 2004). In the forecasting system there is a substance motion form commonly existing in the objective world, of exercise. Like sports, it is a fundamental property of the material existence. Its research object is the comprehensive role of some known factors and some unknown factors (Yu and Liu, 2001). Because the achievements of every athlete in the Olympic Games are determined by known factors and unknown factors, such as the best results that athletes usually can play, the assessed index value of athlete's professional and technical level, the mental quality of athletes peacetime training and ability level of each special sport, these factors having taken place and after assessment are known factors (Sun *et al.*, 2006); performance degree on the sports arena and the psychological fluctuations in big event and other undetermined factors are collectively regarded as the unknown factors (Yang, 2005). According to the theory of gray system model, we can only study the gray amount changing over time within a certain range and the race achievements affected by these factors, then build predictive models using these data (Xu, 2001).

The determination of men's 400 m freestyle cultivating decision-making is constrained by many factors, but the determination of the training objectives is important reference for the decision-making determination (Zhang and Feng, 2013). This study collects the champion performance data of men's 400 m freestyle in recent 10 Olympic Games on the basis of predecessors, establishes a single sequence first-order linear dynamic model, predicts the champion performance of the 31st Olympic champion in this event, provides target value for the national athletes training decision-making and has actual reference value.

RESEARCH OBJECTS AND METHODS

Research objects: This study takes the men's 400 m freestyle swimming champion achievement data of the recent 10 sessions Summer Olympic Games as research objects (Li and Xiao, 2004); the statistical data is from the official website of the Chinese Olympic Committee http://www.olympic.cn/games/list2.html. The data is reliable, as shown in Table 1.

Table 1: List of men's 400 m freestyle swimming champion achievements from the 21st to the 30th Summer Olympic Games

Olympic session No.	Venue	Champion nationality	Champion Name	Achievement	Score (sec)
21th, 1976	Montreal	United States	Goodell Brian Stuart	3:51.93	231.93
22th, 1980	Moscow	Soviet Union	Salnikov Vladimir	3:51.31	231.31
23th, 1984	Los Angeles	United States	Dicarlo George Thomas	3:50.91	230.91
24th, 1988	Seoul	Democratic Republic of Germany	Dassler Uwe	3:46.95	226.95
25th, 1992	Barcelona	CIS	SADOVYI Evgueni	3:45.00	225.00
26th, 1996	Atlanta	New Zealand	Loader Danyon Joseph	3:47.97	227.97
27th, 2000	Sydney	Australia	Ian Thorpe	3:40.59	220.59
28th, 2004	Athens	Australia	Ian Thorpe	3:43.10	223.10
29th, 2008	Beijing	Korea	Park Tea Hwan	3:41.86	221.86
30th, 2012	London	China	Sun Yang	3:40.14	220.14

Research methods:

- **Literature material method:** The downloading papers from CNKI full-text periodical database, tutorials and courseware of gray system theory and tutorial of matrix calculation do adequate preparation for this paper's theoretical correctness
- **Mathematical statistics method:** Use the Excel spreadsheet processing software and function calculator to compute the statistics data
- **Mathematical model method:** Establish a single sequence first-order linear dynamic model and take it as predictive model; use differential dynamic modeling method to conduct single sequence residuals identify prediction.

MODEL BUILDING

Data processing theory: The frequently-used gray system generation methods contains: The accumulated generating, the regressive generating, the mean generating and the level than generating and so on. This study adopts the accumulated generating to process data. The accumulated generating means by data accumulation at each time between series to obtain new data series. The series before accumulation is known as the original series, the series after accumulation is known as the generated series. The generation method is a means to make the gray data become white data. Through accumulated generation we can see the development trend in data accumulation process, thus the regularity of the original data becomes stronger.

Use $x^{(0)}$ to present original series, then $x^{(0)} = [x^{(0)}(1), x^{(0)}(2), x^{(0)}(3),\ldots, x^{(0)}(n)]$. Use $x^{(1)}$ to present generating series, then $x^{(1)} = [x^{(1)}(1), x^{(1)}(2), x^{(1)}(3),\ldots, x^{(1)}(n)]$. Wherein, $x^{(1)}(k)$ satisfies Eq. 1:

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i);\ k = 1, 2, 3,\ldots, n \quad (1)$$

Equation 1 is called the one-time accumulation, namely the 1-AGO (Accumulating Generation Operator).

Table 2: List of data processing

Olympic session No.	Champion scores (sec)	Undifferentiated data	Differentiated original data	Accumulated generating data
21	231.93	220	11.93	11.93
22	231.31	220	11.31	23.24
23	230.91	220	10.91	34.15
24	226.95	220	6.95	41.10
25	225.00	220	5.00	46.10
26	227.97	220	7.97	54.07
27	220.59	220	0.59	54.66
28	223.10	220	3.10	57.76
29	221.86	220	1.86	59.62
30	220.14	220	0.14	59.76

But sometimes one-time accumulation is insufficient to reflect its regularity. The series needs times accumulation, its expression relationship is shown in Eq. 2:

$$\begin{aligned} x^{(r)}(k) &= \sum_{i=1}^{k-1} x^{(r-1)}(i) + x^{(r-1)}(k) = x^{(r-1)}(k-1) + x^{(r-1)}(k) \\ x^{(r)}(k) &= \sum_{i=1}^{k} x^{(r-1)}(i) = \sum_{i=1}^{k}\left(\sum_{j=1}^{i} x^{(r-2)}(k)\right) \end{aligned} \quad (2)$$

The accumulated generating series enables any non-negative series to turn into non-swing non-minus ascending series. In this study, using the one-time accumulation is enough. Use accumulated generating theoretical to compare the characteristics of the raw data and generating data in this article, as shown in Table 2.

Modeling theory: GREY MODEL (n, h) is abbreviated as GM (n, h) model, where n represents the differential order, h represents the number of variables. This study creates a single sequence first-order linear dynamic model, in which n and h is 1. It is called GM (1.1) model and its expression form is shown in Eq. 3 below:

$$\frac{dx}{dt} + ax = u \quad (3)$$

By the definition of differential coefficient we know that:

$$\frac{dx}{dt} = \lim_{\Delta t \to 0} \frac{x(t+\Delta t) - x(t)}{\Delta t}$$

when Dt is a small units like 1, we have the expression form as Eq. 4:

$$\frac{dx}{dt} \approx \frac{\Delta x}{\Delta t} = x(t+1) - x(t) = \Delta^{(1)}(x(x+1)) \tag{4}$$

Wherein:

$$\frac{\Delta x}{\Delta t}$$

represents the one-time regressive generation of x(k+1) and also the binary combination equivalent value of x(k+1) and x(k). This combination is called even couple denoted as [x(k+1), x(k)]. Define a mapping F as shown in Eq. 5:

$$F: [x(k+1),\ x(k)] \rightarrow \frac{dx}{dt} \tag{5}$$

Define z(t) as the background value of:

$$\frac{dx}{dt}$$

at the t moment; take even average value as background value, then:

$$z(t) = \frac{1}{2}[x(k) + x(k+1)]$$

According to Eq. 1-5, we can obtain differential equation in albino form as shown in Eq. 6 below:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u \tag{6}$$

Wherein a, u are undetermined parameters, the discretization of Eq. 6 is shown in Eq. 7:

$$\Delta^{(1)}(x^{(1)}(k+1) + az^{(1)}(x(k+1))) = u \tag{7}$$

In Eq. 7 $\Delta^{(1)}(x^{(1)}(k+1))$ is the regressive generating sequence of $x^{(1)}$ at moment (k+1); $z^{(1)}(x(k+1))$ is the background value of:

$$\frac{dx^{(1)}}{dt}$$

at moment (k+1).

Due to:

$$\begin{aligned} &\Delta^{(1)}(x^{(1)}(k+1)) = x^{(1)}(k+1) - x^{(1)}(k) = x^{(0)}(k+1) \\ &z^{(1)}(k+1) = \frac{1}{2}(x^{(1)}(k+1) + x^{(1)}(k)) \end{aligned} \tag{8}$$

Equation 8 is brought into Eq. 7 and we can obtain Eq. 9:

$$x^{(0)}(k+1) = a\left[-\frac{1}{2}(x^{(1)}(k) + x^{(1)}(k+1))\right] + u \tag{9}$$

Expand Eq. 9 and we can obtain Eq. 10:

$$Y = \begin{bmatrix} x^0(2) \\ x^0(3) \\ x^0(4) \\ \vdots \\ x^0(n) \end{bmatrix} \quad B = \begin{bmatrix} -\frac{1}{2}(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -\frac{1}{2}(x^{(1)}(2) + x^{(1)}(3)) & 1 \\ -\frac{1}{2}(x^{(1)}(3) + x^{(1)}(4)) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}(x^{(1)}(n-1) + x^{(1)}(n)) & 1 \end{bmatrix} \tag{10}$$

Make $F = [a\ u]^T$ as the undetermined identified parameter vector and then Eq. 10 can be written as Eq. 11:

$$Y = B\Phi \tag{11}$$

Parameter vector Φ: we can use exponential smoothing forecast values to estimate the expression of vector Φ and it satisfies Eq. 12:

$$\hat{\Phi} = [\hat{a}\ \hat{u}] = (B^T B)^{-1} B^T Y \tag{12}$$

Substitute the solved parameters into Eq. 6, we can obtain the discrete solution as shown in Eq. 13:

$$\hat{x}^{(1)}(k+1) = \left[x^{(1)}(1) - \frac{\hat{u}}{\hat{a}}\right]e^{-\hat{a}k} + \frac{\hat{u}}{\hat{a}} \tag{13}$$

Restored to the original data, the result is shown in Eq. 14:

$$\hat{x}^{(0)}(k+1) = (1 - e^{-\hat{a}})\left[x^{(1)}(1) - \frac{\hat{u}}{\hat{a}}\right]e^{-\hat{a}k} \tag{14}$$

Relative error test model: After expression $\hat{x}^{(0)} = [\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \hat{x}^{(0)}(3),..., \hat{x}^{(0)}(n)]$ is obtained using G (1.1) model, the expression of residual series E is shown in Eq. 15 below:

$$E = [e(1), e(2), e(3),..., e(n)] = x^{(0)} - \hat{x}^{(0)} \tag{15}$$

In Eq. 15 $e(k) = x^{(0)}(k)-\hat{x}^{(0)}(k)$, k = 1, 2, 3,..., n.

The relative error:

$$rel(k) = \frac{e(k)}{x^{(0)}(k)} \times 100\%,\ k = 1, 2, 3, ..., n$$

the expression of average error rel is shown in Eq. 16:

$$rel = \frac{1}{n}\sum_{k=1}^{n}|rel(k)| \tag{16}$$

Posterior difference test model: The variance of original sequence $x^{(0)}$ and residual sequence E, respectively expressed as S^2_1 and S^2_2, as shown in Eq. 17:

$$\begin{cases} S_1^2 = \frac{1}{n}\sum_{k=1}^{n}\left[x^{(0)}(k) - \bar{x}\right]^2,\ \bar{x} = \frac{1}{n}\sum_{k=1}^{n} x^{(0)}(k) \\ S_2^2 = \frac{1}{n}\sum_{k=1}^{n}\left[e(k) - \bar{e}\right]^2,\ \bar{e} = \frac{1}{n}\sum_{k=1}^{n} e(k) \end{cases} \tag{17}$$

Posteriori differential ratio:

$$C = \frac{S_2}{S_1}$$

the expression to calculate the small error probability is shown in Eq. 18:

$$p = P\{|e(k)-\bar{e}|<0.6745\ S_1\} \tag{18}$$

In Eq. 17 and 18 C and p are two important indicators of posteriori difference test. The smaller indicator C is the better. Because when the value of C is smaller, it indicates that although the original number is very discrete, the difference between the calculated value and the actual value of the model are not too discrete. The bigger indicator p is the better. Because when p is larger, it indicates that difference between residuals and the average value of residuals is less than the given value $0.6745S_1$ showing the evenly distributed fitting values. Generally in accordance with C and p two evaluation indicators, the model accuracy can be divided into four levels, as shown in Table 3.

Table 3: Reference table of the accuracy testing level

Accuracy of the model class	C	p
First (well) C≤0.35	0.95≤p	
Second (qualified)	0.35<C£0.5	0.08≤p<0.95
Third (reluctantly)	0.5<C£0.65	0.70≤p<0.80
Fourth (unqualified)	0.65<C	p<0.70

RESULTS ANALYSIS

Solving parameter vector: Substituting the data in, result is shown in Eq. 19 below:

$$\hat{\Phi} = [\hat{a}\ \hat{u}] = \left(\begin{bmatrix} 17.585 & 1.00 \\ 28.965 & 1.00 \\ 37.625 & 1.00 \\ 43.600 & 1.00 \\ 50.085 & 1.00 \\ 54.365 & 1.00 \\ 56.210 & 1.00 \\ 58.960 & 1.00 \\ 59.690 & 1.00 \end{bmatrix}\begin{bmatrix} 17.585 & 1.00 \\ 28.965 & 1.00 \\ 37.625 & 1.00 \\ 43.600 & 1.00 \\ 50.085 & 1.00 \\ 54.365 & 1.00 \\ 56.210 & 1.00 \\ 58.960 & 1.00 \\ 59.690 & 1.00 \end{bmatrix}^T\right)^{-1}\begin{bmatrix} 17.585 & 1.00 \\ 28.965 & 1.00 \\ 37.625 & 1.00 \\ 43.600 & 1.00 \\ 50.085 & 1.00 \\ 54.365 & 1.00 \\ 56.210 & 1.00 \\ 58.960 & 1.00 \\ 59.690 & 1.00 \end{bmatrix}\begin{bmatrix} 11.31 \\ 10.91 \\ 6.95 \\ 5.00 \\ 7.97 \\ 0.59 \\ 3.1 \\ 1.86 \\ 0.14 \end{bmatrix} \tag{19}$$

$$\hat{\Phi} = [\hat{a}\ \hat{u}] = \begin{bmatrix} 3328.2885 & & -2041.1969 & \\ 9301.1419 & & -7045.3318 & \\ 23426.2747 & & -20119.3965 & \\ 34881.4575 & & -30339.3232 & \\ 64985.598 & & -59691.503 & \\ 175928.882 & 6 & -169839.96 & 7 \\ 158092.200 & 7 & -151548.109 & 3 \\ 386447.811 & 8 & -379585.752 & 7 \\ 386447.811 & 8 & -379585.752 & 7 \end{bmatrix}$$

Predict the cumulative series and the original series: The predicted accumulated series $\hat{x}^{(1)}$ and the predicted original series $\hat{x}^{(0)}$ can be obtained in accordance with Eq. 13 and 14, as shown in Table 4.

Relative error analysis: Substitute the dada from Table 3 into Eq. 15, we can obtain the residual series E as Eq. 19 shown below:

$$\left.\begin{array}{l} E = [e(1), e(2), e(3), ..., e(n)] = x^{(0)} - \hat{x}^{(0)} \\ e(1) = 0 \\ e(2) = 11.31 - 11.18 = 0.13 \\ e(3) = 10.91 - 10.71 = 0.200 \\ e(4) = 6.95 - 7.04 = -0.090 \\ e(5) = 5.00 - 4.73 = 0.270 \\ e(6) = 7.97 - 8.07 = -0.100 \\ e(7) = 0.59 - 0.31 = 0.280 \\ e(8) = 3.10 - 2.94 = 0.160 \\ e(9) = 1.86 - 1.85 = 0.010 \\ e(10) = 0.14 - 0.10 = 0.40 \end{array}\right\} \Rightarrow E^T = \begin{bmatrix} 0 \\ 0.13 \\ 0.200 \\ -0.090 \\ 0.270 \\ -0.100 \\ 0.280 \\ 0.160 \\ 0.010 \\ 0.040 \end{bmatrix} \tag{20}$$

Substitute data from Eq. 20 into 16, we can be obtain average error:

Table 4: Calculate the predicted series

Session No.	Original data	Predict the original data	Generating data	predicted generating data
22	220+11.31	220+11.18	23.24	23.11
23	220+10.91	220+10.71	34.15	33.95
24	220+6.95	220+7.04	41.10	41.19
25	220+5.00	220+4.73	46.10	45.83
26	220+7.97	220+8.07	54.07	54.17
27	220+0.59	220+0.31	54.66	54.38
28	220+3.10	220+2.94	57.76	57.60
29	220+1.86	220+1.85	59.62	59.61
30	220+0.14	220+0.10	59.76	59.72
31	\	220+0.16	\	59.92

$$\text{rel} = \frac{1}{n}\sum_{k=1}^{n}|\text{rel}(k)| = 0.044513375\%$$

Seen by the average error results, the error is very small, the model prediction is relatively good.

Posterior difference test: Substitute the dada from Table 3 into Eq. 17, then use WPS spreadsheet to calculate, we can be obtain:

$$\left.\begin{aligned} S_1^2 &= 15.891 \\ S_2^2 &= 0.01818 \end{aligned}\right\} \Rightarrow C = \frac{S_2}{S_1} = \frac{\sqrt{0.01818}}{\sqrt{15.891}} = 0.03382211\ \ 9582$$

Values of $|e(k)-\bar{e}|$ are all less than 0.6745 S_1 in the sample, so p»99.9%.

Put Posterior difference parameter C, p into accuracy test level in Table 2; it belongs to the First (Well) level. To sum up: the predicted results of men's 400m freestyle champion scores 220 +0.16 = 220.16 in the 31st Summer Olympics is relatively accurate.

CONCLUSION

Athletes training objectives occupies a very important position in the training cultivating decision-making and has a high reference value for accurate prediction on next top event's championship performance; The prediction model established in this study has high precision and well estimates the champion achievements of the next session's men's 400m freestyle; As for uncertain factors we can adopt gray prediction; use the existing law in data to research can achieve multiplier effect; With the development of science and technology, the gray variables gradually turn white. Gray prediction needs continuous improvement with the development of natural sciences which can greatly improve the forecast accuracy.

REFERENCES

Zhang, B. and Y. Feng, 2013. The special quality evaluation of the triple jump and the differential equation model of long jump mechanics based on gray correlation analysis. Int. J. Applied Math. Stat., 40: 136-143.

Yu, C.G. and Z.M. Liu, 2001. A research on evaluation indexes of sustainable development and human resources. J. Xi'an Inst. Phys. Educ., 1: 5-7.

Sun, G.L., M.L. Deng and C.C. Jia, 2006. Gray metabolism GM (1, 1) model and its forecast of china's import and export trends of construction ceramics. China Ceramics, 42: 24-26.

Liu, M.H., 2004. Countermeasures of developing the reserve forces of china's competitive swimmers for 2008 olympics. J. Xi'an Inst. Phys. Educ., 7: 6-10.

Xu, Q., 2001. The situation of chinese athetic swimming and countermeasures of their development. J. Xi'an Inst. Phys. Educ., 1: 78-80.

Yang, Y., 2005. Situation and development countermeasure analysis of competitive swimming from the 28th Athens Olympic games. J. Chengdu Phys. Educ. Inst., 1: 73-75.

Tan, Z.D. and P. Liu, 2003. The successful and deficient experience of China swimming from the Olympic games. J. Guangzhou Phys. Educ. Inst., 1: 81-87.

Li, Z.H. and L.P. Xiao, 2004. Chinese competition sport development strategy in 2010. J. Tianjin Inst. Phys. Educ., 3: 6-9.