ISSN : 0974 - 7435 Volume 10 Issue 10

2014

BioTechnology

An Indian Journal

FULL PAPER

BTAIJ, 10(10), 2014 [4596-4605]

Grey forecasting model in the application of world women's pentathlon performance prediction research

Su Jin[1*], Tiecheng Guo[2]

[1]Institute of Physical Education, Northeast Normal University, Changchun130024, Jilin, (CHINA)

[2]Military sports department, Changchun University of science and technology, Changchun130000, Jilin, (CHINA)

ABSTRACT

Grey system theory is a kind of ideal method of dealing with the dynamic development problem of small sample, and the problems of competitive sports has a "poor information", "small sample" and "dynamic", and made the grey system in their studies than traditional more advantages of probability statistics and fuzzy mathematics, grey mathematics application in performance prediction and analysis of competitive sports more and more widely. This paper USES the literature material law, the statistics between 2001 ~ 2013 year calendar pentathlon world woman the best results. Woman pentathlon world best results, this paper deals with the GM(1, 1) grey model and gray model GM(1, 6). Based on GM(1, 1) model and compare the accuracy of the GM(1, 6) model and GM(1, N) model is studied in the application of sports competition, illustrates the method of GM(1, N) model in the application of multiple projects sports competition, select the GM(1, 6) prediction model for the application of the grey model to predict performance in competitive sports. At the same time, the application of gray forecast model GM(1, 6) has been established for screening, and ultimately determine to sample data from 2005 to 2013 in modeling, determine the world women's pentathlon performance prediction model.

KEYWORDS

women’s pentathlon; GM(1, 1) model; grey relational analysis; GM(1, 6) model.

© Trade Science Inc.

INTRODUCTION

In competitive sports, athlete competitive result influences factors are quite various, lots of factors are unknown, inexact, and factors affect and interact each other, it is difficult to make clear analysis on them one by one. But as a entirety system, competitive sports performance change and development have some internal rules, each factor comprehensive effects let competitive sports performance has uncertainties.

Traditional uncertain factors researching mathematical methods mainly are regression analysis, variance analysis, principal component analysis and other mathematical statistical methods to make statistical analysis of system. Mathematical statistics method requires a great deal of samples and data, data changing should have certain rules, relations among factors to be static and so on, requirements on system is higher, data samples should have better distribution rules, and its analysis cannot surely get effective statistical rules, even it gets statistical rules, in most cases it also cannot make analysis and prediction on system. Grey prediction system modeling according to grey system behavior features data, utilizes accumulating sequence to overcome original sequence volatility, randomness, excavate information data's explicit information and hidden information so that arrive at relative precise short-term prediction model. For sports competitive performance prediction problem's " poor data information" and "dynamics" as well as other features, grey prediction model has already become one of major method in competitive sports performance predicting, lots of scholars have made research and application on it, and got lots of achievements, as well as put forward many opinions. Among them, Ma Xiang-Hai (2012) researched on Chinese excellent decathlon athlete performance, by carrying out grey relational analysis on it and grey GM(1, 1) prediction model's modeling, he analyzed Chinese decathlon athlete development trend[1]; Sun Qiang (2012) by establishing Olympic Games men's 400m competition performance same dimension gray recurrence GM(1, 1) model, he solved grey model middle and long term competitive sports performance prediction[2];Wang Dao-Lin, Fan Xin-Sheng (2005) took Chinese 23th to 28th Olympic Games achieved gold medals numbers as original data to establish grey model, they stated grey model application in competitive sports competition[3]; Liu Jia-Jin (2006) analyzed the 14th to the 28th five sessions' Olympic Winter Games women's shot competitive performance development trend, he proposed GM(2, 1) is more fit for sports competitive performance prediction grey model establishing that possessed slightly swing data sequence[4];Liu Jia-Jin, Wang Dong, Liu Shun-Min (1999) by carrying out GM(1, 1)model group modeling on men's hammer player Bi Zhong-Nian best performance, they researched on grey model group prediction methods in sports competitive performance[5].

The paper takes world women's pentathlon performance prediction as an example, it makes statistics of 2001～2013 such 13 years' world women's pentathlon best performances, makes grey relational analysis of their total performance and other each event performance, establishes GM(1, 1)and GM(1, 6) grey prediction model, and researches grey model's application in competitive sports by predicting.

WORLD WOMEN'S PENTATHLON PREVIOUS CHAMPIONS' PERFORMANCES ANALYSIS

In order to predict world women's pentathlon competitive performance development trend, the paper makes statistics of year 2001～2013 such 13 years' world women's pentathlon champions'

performances, and based on them, it establishes grey relational analysis and grey prediction model. The statistical data is as TABLE 1:

TABLE 1 : Year 2001～2013 previous women's pentathlon highest performance table[6]

Year	Total score	60m hurdle (s)	High jump (m)	Shot (m)	Long jump (m)	800m (s)
2001	4850	8.25	1.8	16.31	6.69	143.32
2002	4622	8.56	1.78	13.66	6.42	143.59
2003	4933	8.19	1.89	14.48	6.61	135.58
2004	4759	8.48	1.88	15.08	6.45	141.69
2005	4948	8.19	1.93	13.29	6.65	133.47
2006	4713	8.31	1.82	13.21	6.44	132.94
2007	4944	8.2	1.88	14.43	6.59	133.04
2008	4867	8.54	1.99	13.85	6.41	136.42
2009	4784	8.3	1.86	13.77	6.54	136.63
2010	4937	8.04	1.9	14.01	6.44	132.55
2011	4723	8.11	1.8	14.81	6.34	138.99
2012	5013	8.38	1.84	16.51	6.57	131.15
2013	4851	8.5	1.89	13.4	6.57	130.53

Women's pentathlon competition is composed of 60m hurdle, high jump, shot, long jump and 800m five events, its final result is got by calculating the five events' performances with certain methods. In order to research on women's pentathlon total performance and establish prediction model on it, it can firstly analyze its total scores development trend.

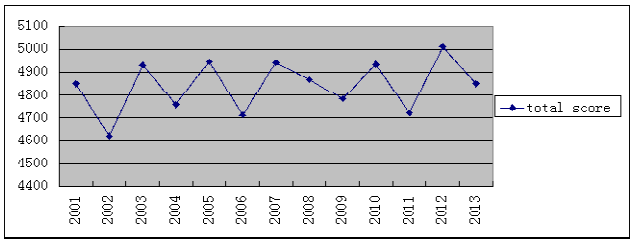

Figure 1: Year 2001～2013 previous women's pentathlon highest performance total scores change chart

By Figure 1, it can see that previous total scores are among 4600～5100, and their changes is on volatility, it goes up and down, fluctuates, so irregular. And data quantity is little, information quantity is little, and women's pentathlon competition performance development change is a dynamic process, traditional mathematical method is difficult to analyze it and judge its change trend. Grey theory uses accumulating sequence to do data mining on original data sequence, and applies generated sequence into modeling, so that mines its internal rules from less quantity, changing data, establishes model and analyzes world women's pentathlon annual best performance change and development trend.

WORLD WOMEN'S PENTATHLON ANNUAL BEST PERFORMANCE GM(1, 1) MODEL

GM(1, 1) model is grey prediction model's most widely applied model. In grey theory applications, 29.08% researches apply into grey prediction model, and among them 75.61% of used grey prediction model is GM(1, 1) grey prediction model, and its thesis quantities is increasing at annual

19.69%～21.08% speed. GM(1, 1) model method is simple, application range is wide, it can be applied into lots of traditional mathematical method unsolved problems fields, and grey prediction model takes the leading position. Due to world women's pentathlon total performance data features, the paper adopts GM(1, 1) model to carry out model analysis of it.

Data test

By data statistics, it carries out grey model original data sequence generation on year 2001～2013 world women's annual best performance total performance. Generated data sequence is as following:

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \cdots, x^{(0)}(13))$$

In order to ensure GM(1, 1) grey prediction model's accuracy, it needs to carry out testing on modeling obtained original sequence data. If its original sequence ultimate rate $\lambda(k)$ falls in the interval $(e^{-\frac{2}{n+1}}, e^{\frac{2}{n+2}})$, then it can use its data to carry out grey prediction modeling. Otherwise it should process with original data. Input $n=13$ and it can get interval （0.8669, 1.1426）.

By sequence ultimate rate formula:

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, k = 2,3,\cdots,13$$

It can get that obtained ultimate rate formula falls in the interval （0.9370, 1.0499）, it can directly use original sequence to model on prediction system.

GM(1, 1) model establishment

Accumulating and mean handling with new data column

By accumulating operator $AG0$ carrying out accumulating with new data sequence, weaken its randomness, and can get accumulating sequence:

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(13))$$

And by formula $z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k)$, it solves its average value generation sequence $z^{(1)}(k)$.

Construct data matrix B and data vector Y:

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(11) & 1 \end{bmatrix}, Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(11) \end{bmatrix}$$

Calculate $\hat{u}$:

$$\hat{u} = (a,b)^T = (B^T B)^{-1} B^T Y = \begin{pmatrix} 0.0022 \\ 19.2322 \end{pmatrix}$$

Model establishment:

$$\frac{dx^{(1)}}{dt} + 0.0022x^{(1)} = 19.2322$$

It can solve:

$$x^{(1)}(k+1) = 2163990e^{-0.002208k} - 2159140$$

Then it can solve year 2001～2013 women's pentathlon best performance predicted value as following TABLE 2:

TABLE 2: Year 2001～2013 women's pentathlon best performance prediction table

Year	2001	2002	2003	2004	2005	2006	2007
Total score	4850.0	4782.6	4793.2	4803.8	4814.4	4825.0	4835.7
Year	2008	2009	2010	2011	2012	2013	
Total score	4846.4	4857.1	4867.8	4878.6	4889.4	4900.2	

GM(1, 1) model testing

Model's each test indicator is as following TABLE 3 show.

TABLE 3: GM(1, 1) model testing table

Year	Original value	Model value	Residual	Relative error	Ultimate rate deviation
2001	4850	4850.0	0	0	
2002	4622	4728.6	-160.6	0.0347	-0.0516
2003	4933	4793.2	139.8	0.0283	0.0610
2004	4759	4803.8	-44.8	0.0094	-0.389
2005	4948	4814.4	133.6	0.0270	0.0361
2006	4713	4825.0	-112.0	0.0238	-0.0522
2007	4944	4835.7	108.3	0.0219	0.0446
2008	4867	4846.4	20.6	0.0042	-0.0181
2009	4784	4857.1	-73.1	0.0153	-0.0196
2010	4937	4867.8	69.2	0.0140	0.0288
2011	4723	4878.6	-155.6	0.0329	-0.0476
2012	5013	4889.4	123.6	0.0247	0.0558
2013	4851	4900.2	-49.2	0.0101	-0.0357

By TABLE, it can get prediction model relative error is not above 3%, ultimate deviation is not above0.7.

Residual qualified model

Relative error sequence is:

$\Delta = (\Delta_1, \Delta_1, \cdots, \Delta_1)$

Then it can solve average relative error is:

$$\bar{\Delta} = \frac{1}{n}\sum \Delta_k = 0.0190$$

Correlation degree qualified model

Absolute correlation degree g is original sequence $x^{(0)}$ and corresponding grey prediction sequence $\hat{x}^{(0)}$ absolute correlation degree, it solves:

$$g = 0.9066$$

Mean square error ratio qualified model

S_1^2 and S_2^2 are respectively original sequence $x^{(0)}$ and residual sequence $\varepsilon(k)$ variances, it can solve mean square error ratio value as:

$$C = S_2 / S_1 = 0.9466$$

By TABLE 3 data, it can solve model precise test data as following TABLE 4.

TABLE 4 : GM(1, 1) prediction model precise table

Precise grade	Relative error	Absolute correlation degree	Mean square error ratio value
Grade Four	0.0190	0.9066	0.9466

By TABLE 4, it is clear that directly uses traditional GM(1, 1) grey model to predict world women's pentathlon annual best performance, its relative error is 0.0190 that is second grade precise, absolute correlation degree is 0.9066 that is first grade precise, mean square error ratio is 0.9466 that is four grade precise. To ensure established prediction model precise, it takes each indicator lowest precise as prediction precise. Therefore, it is clear that GM(1, 1) grey prediction model predicted precise is grade four, it is difficult to precisely make effective prediction on world women's pentathlon annual best performance.

GM(1, 1) model group application

In order to more effective apply GM(1, 1) grey model to predict world women's pentathlon annual best performance, it adopts GM(1, 1) model group to model on its data, in the hope of more effective utilizing GM(1, 1) grey prediction model, more accurate researching on world women's pentathlon performance development change trend. GM(1, 1) grey model group is on the basis of GM(1, 1) grey model, it respectively carries out different dimensions' modeling on original statistical data. The paper takes year 2013 data as base point; it gradually cuts data from far and near, establishes four dimensions to thirteen dimensions' totally ten GM(1, 1) grey prediction models. By comparing each dimension relative error, absolute correlation degree and mean square error rate, it researches on GM(1, 1) prediction model's prediction precise on world women's pentathlon annual best performance. It gets precise table as TABLE 5.

TABLE 5 : GM(1, 1) model group precise test table

Dimension	Precise grade	Relative error	Absolute correlation degree	Mean square error ratio value

4	Grade three	0.0154	0.9146	0.7650
5	Grade three	0.0157	0.9755	0.9271
6	Grade four	0.0155	0.7128	0.9582
7	Grade three	0.0145	0.9786	0.9314
8	Grade three	0.0139	0.9958	0.8504
9	Grade three	0.0164	0.9842	0.9181
10	Grade three	0.0162	0.9937	0.9471
11	Grade three	0.0166	0.9881	0.9503
12	Grade three	0.0165	0.9980	0.8163
13	Grade three	0.0190	0.9066	0.9466

By TABLE 5, it is clear that GM(1, 1) grey prediction model's prediction modeling precise on world women's pentathlon annual best performance is quite low. In ten models, it has one model precise as grade four and all the rest are grade three. GM(1, 1) prediction model cannot effectively carry out prediction modeling on world women's pentathlon annual best performance.

WOMEN'S PENTATHLON PERFORMANCE CORRELATION DEGREE ANALYSIs

Women's pentathlon is composed of five events, every event performance will have certain effects on total performance, and its effects are both big and small, every sport event different development trend also surely affects pentathlon total performance development trend. Grey relational analysis is used to analyze total performance and each event sport performance dynamic relations with time changing and their features, which provides references for women's pentathlon total performance prediction.

Record world women's pentathlon annual best performance total score and each event performance sequence as:

$$x_i^{(0)} = (x_i^{(0)}(1), x_i^{(0)}(2), x_i^{(0)}(3), \cdots, x_i^{(0)}(13)) \quad i = 0,1,\cdots,5$$

Take world women's pentathlon annual best performance total score sequence $x_0^{(0)}$ as reference sequence, make mean transformation on each sequence, carry out standard processing with it, and solves its mean sequence.

Calculate its correlation coefficient: $\xi(k) = \dfrac{\min\limits_i \min\limits_k |x_0(k) - x_i(k)| + \rho \max\limits_i \max\limits_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \max\limits_i \max\limits_k |x_0(k) - x_i(k)|}$

Among them, ρ is resolution coefficient, the paper takes it as 0.5.

Calculate its correlation degree: $r_i = \dfrac{1}{n}\sum_{i=1}^{n} \xi_i(k)$

Input TABLE 1 statistical data into above steps, it can solve each event performance and total performance grey correlation degree values as TABLE 6.

TABLE 6 : Each event performance total performance grey correlation degree value

	Hurdle	High jump	Shot	Long jump	800m
Correlation degree	0.8516	0.8141	0.6174	0.8738	0.7752

By TABLE 6, it is clear that it is long jump performance that has largest influence on women's pentathlon total performance, its grey correlation degree is 0.8738, the secondary is hurdle, high jump,

800m and shot their grey correlation degree are respectively 0.8516, 0.8141, 0.7752, 0.6174. Grey correlation calculation can define each sport performance impact on women's all-around total performance. Though take different resolution coefficient will get different correlation degrees, their size order will not change, correlation sequence is essence of correlation analysis. Though each sport performance and total performance correlation degree are both big and small, each correlation degree has no big differences, when predict on women's pentathlon performance, each sport performance is important to total performance prediction.

WOMEN'S PENTATHLON PERFORMANCE GM(1, 6)MODEL

By above analysis, it is clear that GM(1, 1) prediction model effects in world women's pentathlon annual best performance predicting application is not ideal, it is difficult to make effective prediction on its performance change development trend. By grey correlation analysis, it is clear that women's pentathlon performance is closely connected with its each event performance that belongs to multiple factors problems. In order to more precise predict its performance change trend, the paper establishes multiple variables' grey prediction model that takes women's pentathlon performance total performance and other five events' performance as variables, establishes GM(1, 6)grey prediction model for system, so that more precise and effective research on women's pentathlon performance development change trend.

To data sequence $x_i^{(0)},(i=0,1,\cdots,5)$, it carries out $1-AGO$ processing, weakens its randomness and gets world women's pentathlon annual best sport performance and each event performance accumulating sequence $x_i^{(1)},(i=0,1,\cdots,5)$.

To $x_0^{(1)}$ adjoining mean, it generates sequence $z_0^{(1)}=(z_0^{(1)}(2),z_0^{(1)}(3),\cdots,z_0^{(1)}(13))$, then it has:

$$B=\begin{bmatrix} -z_0^{(1)}(2) & x_1^{(1)}(2) & \cdots & x_5^{(1)}(2) \\ -z_0^{(1)}(3) & x_1^{(1)}(3) & \cdots & x_5^{(1)}(3) \\ \vdots & \vdots & \ddots & \vdots \\ -z_0^{(1)}(13) & x_1^{(1)}(13) & \cdots & x_5^{(1)}(13) \end{bmatrix}, Y=\begin{bmatrix} x_0^{(1)}(2) \\ x_0^{(1)}(3) \\ \vdots \\ x_0^{(1)}(13) \end{bmatrix}$$

Set $u=[a,b_1,b_2,\cdots,b_5]^T$, then GM(1, 6) model winterization equation is:

$$x_1^{(0)}(k)+az_1^{(1)}(k)=b_2x_2^{(1)}(k)+b_3x_3^{(1)}(k)+\cdots+b_5x_5^{(1)}(k)$$

It can solve by formula:

$$u=\begin{bmatrix} a \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}=(B^TB)^{-1}B^TY=\begin{bmatrix} 1.9 \\ -333.2 \\ 2621.5 \\ 162.3 \\ 1090.0 \\ -17.6 \end{bmatrix}$$

By winterization equation, it solves year 2001～2013 world women's pentathlon performance predicted value, and makes comparative analysis and test on solved predicted value and original value, it gets its model error values table as following TABLE 7.

TABLE 7 : GM(1, 6) model error value

Year	Original value	Model value	Residual	Relative error
2001	4850	4850.0	0	0
2002	4622	3995.5	626.5	0.1355
2003	4933	5456.6	-523.6	0.1061
2004	4759	4975.3	-216.3	0.0455
2005	4948	5003.1	-55.1	0.0111
2006	4713	4675.6	37.4	0.0079
2007	4944	4961.3	-17.3	0.0035
2008	4867	4868.1	-1.1	0.0002
2009	4784	4796.0	-11.9	0.0025
2010	4937	4897.8	39.2	0.0079
2011	4723	4698.3	24.2	0.0052
2012	5013	5057.1	-44.1	0.0088
2013	4851	44844.5	6.5	0.0013

By Table 7, it is clear that established GM(1, 6) grey prediction model by predicting and analyzing year 2001～2013world women's pentathlon annual highest performance, its maximum relative error is 13.55%, only year 2001 and 2002 such two years relative errors surpass 10%, other years' relative errors are relative ideal. In order to more precise test model accuracy, make use of solved prediction value to precede with grey prediction model each test indicator calculation; it solves each test indicator as TABLE 8:

TABLE 8 : GM(1, 6) prediction model precise table

Precise grade	Relative error	Absolute correlation degree	Mean square error ratio value
Grade four	0.0258	0.5018	0.7591

By Table 8, it is clear that model relative error is 0.0246 that precise grade is grade two, absolute correlation degree is 0.5018 that precise grade is grade four, mean square error ratio is 0.7591 that precise grade is grade four. Model precise is lower, it is difficult to make effective prediction, in order to establish precise prediction model, the paper applies GM(1, 6) grey model group to screen the model, each dimension model each test indicator and model precise grade is as following TABLE 9.

TABLE 9 : GM(1, 6) model group precise test table

Dimension	Precise grade	Relative error	Absolute correlation degree	Mean square error ratio value
4	Grade four	0.0716	0.5010	0.7109
5	Grade four	0.0516	0.5821	0.7215
6	Grade four	0.0428	0.5503	0.4796
7	Grade four	0.0413	0.7838	0.7085
8	Grade three	0.0353	0.9961	0.5282
9	Grade two	0.0319	0.9931	0.4087
10	Grade four	0.0296	0.9577	0.8124
11	Grade three	0.0271	0.9203	0.5393
12	Grade four	0.0271	0.9926	1.0007
13	Grade four	0.0258	0.5018	0.7591

By TABLE 9, it is clear that world women's pentathlon annual best performance grey prediction model only the nine dimensions precise arrives at grade two, then the paper finally defines to use its prediction model to carry out GM(1, 6) model modeling on data during year 2005 to 2013.

CONCLUSIONS

The paper analyzes competitive sports performance prediction, discusses grey prediction model's feasibility in competitive sports performance prediction modeling, it provides guidance for applied mathematics in competitive sports performance prediction and decision making as well as other aspects, drives applied mathematics application in competitive sports decision making, and lets competitive sports decision making more scientific; use statistical data modeling analysis, and apply grey model group model to model for competitive sports performance prediction, use grey model group to screen proper data samples dimensions to model, it more flexible applies grey prediction model, and improves grey prediction precise; researches on competitive sports comprehensive sports event. To multiple factor affected sports performances, it proposed to make grey relation analysis of them, make GM(1, N) grey prediction model, fully utilizes known data and conditions, and establishes precise prediction model.

REFERENCES

[1] Zhu Hong-bing, Liu Jian-tong, Wang Gang, etc,; Gray Forecast of GM(1, 1) Model and Its Application in Forecast of Sports Performance[J], Journal of Capital College of Physical Education, **15(1),** 118-121 **(2003)**.

[2] Wang Wei; Application Research about System Theory and Ma erkefu Chain Used in Predicting Results of Track and Field[J], Journal of Nanjing Institute of Physical Education, **18(6),** 85-87 **(2004)**.

[3] Zhao Yun hong, Zhou Yao, Equal Dimensions Gray Number Pattern and Its Forecast of World and National Records in Women 100 m Hurdles[J], China Sport Science and Technology, **38(2),** 39-40 **(2002)**.

[4] Chen Liang, Sai Qing-bin; An Analysis on the Developing Trend of the Performance of Track and Field in National Sports Games and Its Study with Grey Prediction[J], Journal of Capital College of Physical Education, **18(5)**, 85-88 **(2006)**.

[5] Song Ai-Ling, Chren Kai; Gray Model and Forecast of Hurdles the Best Results in the World and Chinese Liuxiang[J], Journal of Capital College of Physical Education, **15(4),** 68-69, 116 **(2003)**.

[6] Chen Liang, Tian De-bao, The Application Research of the Reconstructed GM(1, 1) Model of Background Value in Sports[J], Journal of Capital College of Physical Education, **18(6),** 127-128 **(2006)**.

[7] Tong Liping, Yuan Jianguo; GM(1, 1) Model of International and National Records in Women's 400 m Hurdling and Calculation[J], Journal of Shanghai Physical Education Institute, **25(2),** 44-47 **(2001)**.

[8] Liu Jia-jin; Research on the New Methods of Gray Envelope in Sports Field[J], Journal of Guangzhou Physical Education Institute, **26(2),** 54-56 **(2006)**.

[9] Liu Jia-jin et al.; Applied Study of GM(1, 1) Models in Sports Forecast[J], Zhejiang Sport Science, **21(1),** 60-64 **(1999)**.