ISSN : 0974 - 7435 Volume 10 Issue 4

BioTechnology
An Indian Journal

FULL PAPER

BTAIJ, 10(4), 2014 [749-755]

Application study of Grey GM (1, 1) model on the prediction of world elite athletes' long jump performance

Jinzhu Li

Institute of Physical Education, Gannan Normal University, Ganzhou 341000, Jiangxi, (CHINA)

ABSTRACT

It uses document literature method and mathematical statistics method, analyzes the annual best performance in the world long jump from 2000 to 2013. By using GM (1, 1) model, GM (2, 1) model and GM (1, 1) model group, it conducts comparative analysis on the results of the three gray modeling, and in particular carries through a detailed study on the application of the three in athletic performance prediction. The results show that: for the forecasting problem of sports performance whose time series do not swing strongly, the GM (2, 1) prediction model is not applicable. GM (1, 1) model is more suitable for the prediction problem application that the athletic performance's time series have stronger exponent law. By comparison study, for the prediction issues with a relatively large number of statistical data, GM (1, 1) model groups are more conducive to improving the prediction accuracy of the athletic performance in this paper, so it makes the gray model more flexible in practical application.

KEYWORDS

Men's long jump;
Gray theory;
GM (1, 1) model;
GM (2, 1) model;
Performance prediction.

INTRODUCTION

For results prediction of athletic sports, the methods used in the past in addition to data statistics also include data fitting function, curvilinear regression and time series analysis method. This method does function fitting through statistical performance of a lot of competition and team members, obtains the approximate model that can reflect the variation law of sports results, and predicts the sports results through the fitting function; and these methods are usually difficult to achieve high prediction precision are due to the small amount of data. Meanwhile, the Sports Performance is affected by many factors, its results have more ups and downs, fluctuate a lot, there are no rules to follow, and thus the traditional forecasting methods are difficult to predict them effectively. The building of wrong sports prediction model would bring serious negative impact on the development of competitive sports, so we must be very careful for the prediction modeling of athletic sports. Grey prediction model processes the raw data to model and predict through the generated new series, which can predict the "small sample", "poor information" and other issues number in the prediction problem, and has a wide range of applications in the forecast.

Since the gray theory was created, gray model has been widely used in control, decision making, optimi-

FULL PAPER

zation and forecasting and other fields. It has made great progress especially in the prediction field of the "small sample" and "poor information", a large number of domestic and foreign scholars study and explore its applications. With the continuous development of the world sport culture and sport industry, the athletic performance also receives concern from people. Yuan Jian-guo (1992) carried through gray modeling study on our national women's 500m short track speed skating records, analyzed the GM gray modeling method for the long-term results forecast[1]. Liu Jia-jin, through modeling of the 14th to the 28th women's shot put racing results, analyzed the GM ((2,1)) gray modeling method for the sporting events prediction issues at the turning point[2]. Through analysis and research on the basic idea of the gray modeling, the advantages of gray modeling and other aspects, Liu Di study its application in sports[3]. Through modeling the 400m achievements of the 20th to the 29th Olympic Games, Sun Qiang conducts gray prediction study on its achievements[4]. Through the analysis of the achievement of women's long-distance running and the establishment of gray prediction model, Yang Bing-long and Du Xin-jing study the growth law of the performance of women's long-distance running[5].

Based on previous studies, this paper carries through statistics on the best annual performance of the world men's long jump, builds and analyzes the three gray models on the data, studies the application of grey model in sports performance prediction, and points out the characteristics of the three kinds of models. The study of the gray forecasting model also provides reference for the application of gray model in the prediction of athletic performance.

DATA STATISTICS AND ANALYSIS OF DATA DEVELOPMENT TRENDS

This paper collects the best annual performance of world men's long jump from 2000 to 2013, and carries through statistical analysis of its development trends. By studying its development trends, we use three different gray prediction models to conduct comparative analysis on the results.

Analysis of the annual best performance of men's long jump

The statistics data in TABLE 1 shows that the best annual performance of the world men's long jump is in upward trend for the three years from 2000 to 2002; And it is in decline trend for the six years from 2002 to 2007; it is in decline trend for the three years from 2007 to 2009; and it is into a downward trend for the three years from 2009 to 2011; it turns upward trend from 2011 to 2013.

TABLE 1 : The annual best performance of the world men's long jump[6]

Year	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**
Results (m)	8.4	8.43	8.59	8.43	8.41	8.37	8.36
Year	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**
Results (m)	8.31	8.42	8.71	8.27	8.21	8.28	8.31

Best annual results of World Men's Long Jump are in oscillating trend, its causes are various; various disturbances in the game is difficult to analyze, and the undulation and oscillation of its performance data also brings difficulties and uncertainties for its functional fitting. Using traditional mathematical analysis methods to conduct mathematical modeling on its performance prediction, it is difficult to obtain satisfying mathematical models.

Analysis on the model selection of the annual best performance for the men's long jump

For the modeling of uncertain systems, the three most common methods are the probability statistics, fuzzy mathematics and gray math. The common feature of these three modeling methods is the uncertainty of the study systems, but there are obvious differences between the three. The research object of fuzzy mathematic is mainly "not clear cognition", and the study object of probability statistics has phenomenon of "random uncertainty", which requires a large amount of statistical data, gray system mainly studies "less data modeling", the study object focuses on " clear extension and not clear connotation." Relatively speaking, the information of long jump is less; the extension is clear comparatively speaking, gray forecasting model is more applicable to their modeling and prediction.

Comparative analysis of GM (1, 1) model and GM (2, 1) model

In the gray GM (M, N) model, GM (1, 1) model and GM (2, 1) model are the most commonly used two modeling models. In GM (1, 1) model, the corresponding differential equations to the gray differential equations are first order differential equations; and in GM (2, 1) model, the corresponding differential equations to the gray differential equations are second order differential equations. In the gray model, GM (1, 1) model is more suitable for the sequence which has strong exponent laws after an accumulation, that is can only model for monotonic data. For sequence of non-monotonic swing development, GM (2, 1) model will be more applicable[7].

Developments and changes of the athletic performance have strong stage characteristics. The improvement of athletes' competitive level has a period, but over time, its competitive level will have downturns. Development of sports is also a process; even there will be low ebb in the course of development. Application of gray prediction model starts from the time-series of athletic performance; the athletic performance data in different development stages will inevitably affect the different judgments by predictors on the development trends of the data, and it will certainly create a different gray prediction model. Here, our country can use the GM (1, 1) model to predict their group, but its prediction accuracy is unsatisfactory because of the complexity of the development trend of the results.

When the development of athletic performance is in the transition period, the time sequence of the results will be non-monotonic swing, effective GM (1,1) prediction model cannot be established; and as the GM (2,1) prediction model itself has its characteristics, which is an effective model for its sports performance modeling.

Analysis on the gray model selection of annual best scores for men's long jump

The selection of the gray model depends on the swing case of the modeled data columns, GM (1,1) model and GM (2,1) model are both applicable when modeling for a relatively smooth data columns, but the GM (1, 1) prediction model has certain advantages. When modeling for the problem whose data has swing trend, the GM (2, 1) prediction model has a good prediction accuracy; the greater the swings is, the more applicable GM (2,1) model is[7].

As can be seen from Figure 1, the swing range of the best annual performance of 2001 to 2003 and 2007 to 2010 is very large. And the performance trends from 2011 to 2013 have changed. Best achievement data sequence of men's Long Jump swings slightly large; GM (1, 1) model and GM (2, 1) are both applicable.

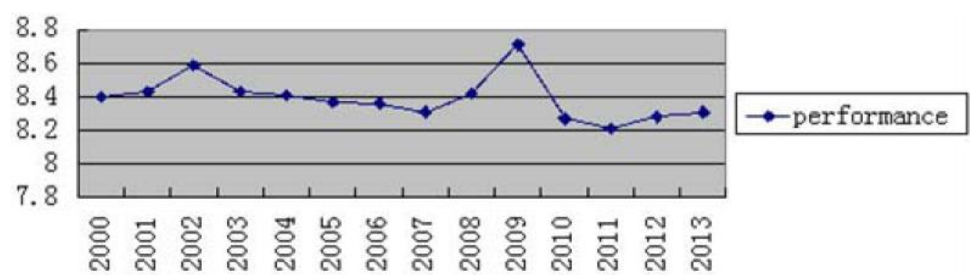

Figure 1: The change trends of the annual best performance for the world's men's long jump

ESTABLISH THE GM (1, 1) MODEL OF WORLD MEN'S LONG JUMP PERFORMANCE

Class ratio test

Establish the annual best performance data time series of the world's men's long jump as follows:

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(14))$$

$$= (8.4, 8.43, 8.59, 8.43, 8.41, 8.37, 8.36, 8.31, 8.42, 8.71, 8.27, 8.21, 8.28, 8.31)$$

(1)Calculate the class ratio

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

$$\lambda = (\lambda(2), \lambda(3), \cdots, \lambda(14))$$

$$= (0.9964, 0.9814, 1.0190, 1.0024, 1.0048, 1.0012, 1.0060, 0.9869, ($$

FULL PAPER

0.9667 ,1.0532 ,1.0073 ,0.9915 ,0.9964)

The class ratio judgment: when $\lambda(k) \in (e^{-\frac{2}{n+1}}, e^{\frac{2}{n+2}})$, the series $x^{(0)}$ can be applied to predictions of GM (1, 1) model. It can be obtained all $\lambda(k) \in (0.8676, 1.1426)$ through the calculation, among them $k = 2, 3, \cdots, 14$.

GM (1, 1) modeling

(1) Do $1-AG0$ on the original data and obtain: =(8.4000, 16.8300, 25.4200, 33.8500, 42.2600, 50.6300, 58.9900, 67.3000, 75.7200, 84.4300, 92.7000, 100.9100, 109.1900, 117.5000)

(2) Construct data matrix Band data vector Y:

$$B = \begin{bmatrix} -\frac{1}{2}(x^{(1)}(1)+x^{(1)}(2)) & 1 \\ -\frac{1}{2}(x^{(1)}(2)+x^{(1)}(3)) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}(x^{(1)}(13)+x^{(1)}(14)) & 1 \end{bmatrix}, Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(14) \end{bmatrix}$$

(3) Calculate $\hat{u}$: $\hat{u} = (a,b)^T = (B^T B)^{-1} B^T Y = \begin{pmatrix} 0.0019 \\ 8.5092 \end{pmatrix}$

(4) Modeling: $\frac{dx^{(1)}}{dt} + 0.0019x^{(1)} = 8.5092$

We can obtain: $x^{(1)}(k+1) = -4588.9e^{-0.00185} + 4597.3$

(5)By the generated series value and the model, restore $\hat{x}^{(0)}(k+1)$, and obtain:

$x^{(1)}$= (8.4000, 8.4858, 8.4701, 8.4545, 8.4388, 8.4232, 8.4077, 8.3921, 8.3766, 8.3611, 8.3456, 8.3302, 8.3148, 8.2994).

Model test

The test index values of the model are as follows in TABLE 2:

Can be drawn from TABLE 2, the maximum error of the model is 4.1%, its prediction accuracy is relatively high.

ESTABLISH THE GM (2, 1) MODEL OF THE WORLD MEN'S LONG JUMP AND COMPARE IT WITH THE GM (1, 1) MODEL

Establish the GM (2, 1) model of the World Men's Long Jump

(1) Establish GM (2, 1) model as follows:

Its original sequence x $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(14))$, its gray differential equations can be established as follows:

Where, $\alpha^{(1)}x^{(0)} = (\alpha^{(1)}x^{(0)}(1), \alpha^{(1)}x^{(0)}(2), \cdots \alpha^{(1)}x^{(0)}(14))$

TABLE 2 : The test table of GM (1, 1) model

Number	Year	Original value	Model value	Residual error	Relative error	Class ratio deviation
1	2000	8.40	8.4000	0	0	
2	2001	8.43	8.4858	-0.0558	0.0066	0.0054
3	2002	8.59	8.4701	0.1199	0.0140	0.0204
4	2003	8.43	8.4545	-0.0245	0.0029	-0.0171
5	2004	8.41	8.4388	-0.0288	0.0034	-0.0005
6	2005	8.37	8.4232	-0.0532	0.0064	-0.0029
7	2006	8.36	8.4077	-0.0477	0.0057	0.0007
8	2007	8.31	8.3921	-0.0821	0.0099	-0.0042
9	2008	8.42	8.3766	0.0434	0.0052	0.0149
10	2009	8.71	8.3611	0.3489	0.0401	0.0351
11	2010	8.27	8.3456	-0.0756	0.0091	-0.0513
12	2011	8.21	8.3302	-0.1202	0.0146	-0.0054
13	2012	8.28	8.3148	-0.0348	0.0042	0.0103
14	2013	8.31	8.2994	0.0106	0.0013	0.0055

$$\alpha^{(1)}x^{(0)} + a_1 x^{(0)} + a_2 z^{(0)} = b$$

$$z^{(1)} = (z^{(1)}(1), z^{(1)}(2), \cdots, z^{(1)}(14))$$

$$z^{(1)}(k) = x^{(1)}(k) + x^{(2)}(k-1), k = 1,2,\cdots,14$$

$$\alpha^{(1)}x^{(0)}(k) = x^{(0)}(k) - x^{(0)}(k-1), k = 1,2,\cdots,14$$

(2) Substitute the best annual performance of the world men's long jump into GM (2,1) model, the test indicators are as follows in TABLE 3:

As can be obtained through above TABLE 3, for the forecasting of the best annual performance for the world men's long jump, the error of GM (2, 1) prediction model is large, and it is completely not applicable.

Compare the GM (2, 1) model and GM (2, 1) model of the world men's long jump performance

As can be obtained from TABLE 3, the predicted results by GM (2, 1) model and the actual performance differ greatly; the prediction on best annual performance of world men's long jump is completely not applicable. The maximum error of the prediction by GM (1, 1) model is only 4.1%, and its prediction accuracy is first level, which is fully applicable to predict its future performance.

THE APPLICATION OF GM (1, 1) MODEL GROUP IN PREDICTION

In practical forecast problems, GM (1, 1) model groups have merits of less data needs, short-term and high prediction accuracy, and thus it receives widespread attention in the field of sports[8]. In practical forecast problems, even if the data after GM (1, 1) model accumulation changes exponentially, the smoothness of the sequence is better, defects like large randomness of its data are still possible, the effect is not very satisfactory in the forecast of athletic performance. The applications of GM (1, 1) model group will effectively solve this problem, and will make the application of gray forecast in athletic performance more mature.

Establish the GM (1, 1) model group

Through the settlement of the best annual performance of the world men's long jump, it takes the performance of year 2013 as base point and takes data of different dimensions to build GM (1, 1) model, and the prediction equation group is as follows:

$$\begin{cases} x^{(1)}(k+1) = -4588.9e^{-0.00185t} + 4597.3 & (14\ d\ model) \\ x^{(1)}(k+1) = -3943.41e^{-0.00215t} + 3951.84 & (13\ d\ model) \\ x^{(1)}(k+1) = -5446.87e^{-0.00149t} + 5655.46 & (12\ d\ model) \\ x^{(1)}(k+1) = -5533.07e^{-0.00152t} + 5541.5 & (11\ d\ model)0 \\ x^{(1)}(k+1) = -5112.17e^{-0.00165t} + 5120.58 & (10\ d\ model)0 \\ x^{(1)}(k+1) = -3787.95e^{-0.00223t} + 3796.32 & (9\ d\ model) \\ x^{(1)}(k+1) = -2545.43e^{-0.00332t} + 2553.79 & (8\ d\ model)0 \\ x^{(1)}(k+1) = -1310.49e^{-0.00651t} + 1318.80 & (7\ d\ model) \\ x^{(1)}(k+1) = -893.624e^{-0.00958t} + 902.04 & (6\ d\ model)0 \\ x^{(1)}(k+1) = 3573.12e^{-0.00230t} - 3564.41 & (5\ d\ model) \\ x^{(1)}(k+1) = -1355.50e^{-0.00604t} - 1347.23 & (4\ d\ model) \end{cases}$$

TABLE 3 : The test table of GM (2, 1) model

Number	Year	Original value	Model value	Residual error	Relative error	Class ratio deviation
1	2000	8.40	8.4	0	0	
2	2001	8.43	9	-0.67	-0.079	0
3	2002	8.59	9	-0.41	-0.047	0.1
4	2003	8.43	10	-1.57	-0.179	0.2
5	2004	8.41	13	-4.59	-0.54	0.5
6	2005	8.37	20	-11.63	-12	1.4
7	2006	8.36	40	-31.64	-32	3.8
8	2007	8.31	93	-84.69	-84	10.2
9	2008	8.42	234	-225.58	-226	26.8
10	2009	8.71	613	-604.29	-604	69.4
11	2010	8.27	1626	-1617.73	-1618	195.6
12	2011	8.21	4338	-4329.79	-4330	527.4
13	2012	8.28	11595	-11586.72	-11587	1399.4
14	2013	8.31	31018	31009.69	-31010	3731.7

Posterior error test and filtering accuracy analysis of GM (1, 1) model group

Posterior error test

Calculate the residual formula of the model $e(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$, after seek their residuals, and then obtain the variance S_1^2 and S_2^2 of $x^{(0)}(k)$ and its residual error $e(k)$, in order to obtain the variance ratio and the probability of small error:

$C = S_2 / S_1$ $P = P\left\{\left|e(k) - \bar{e}\right| < 0.6745 S_1\right\}$

Filtering accuracy analysis

Filtering relative error: $\Delta_n = \left|e(n) / x^{(0)}(n)\right|$

Filtering precision $= (1-\Delta_{(n)})*100$

Analysis of the best dimension GM (1, 1) model

To select the GM (1, 1) model with the most accurate prediction accuracy from GM (1, 1) model group, you need to do posterior error test and filtering accuracy values calculation on all dimensional GM (1, 1) models of the model group:

As can be seen from TABLE 4, only the 5-dimensional, 12-dimensional and 13-dimensional model accuracy are relatively high; and the filtering accuracy of five-dimensional model is the lowest, the 5-dimensional model can be ruled out. Through the 12-dimensional and 13-dimensional posterior error test, the 12-dimensional model is the first level, and the 13-dimensional model is the second level; so the 12-dimensional model is the best model for the performance prediction model of the world men's long jump.

Through the 12-dimensional modeling of gray prediction model, get its test indicators as follows in TABLE 5:

By using gray prediction model group, and the gray prediction model for the best annual performance of the world men's long jump changes from 14-dimensional prediction model to 12 dimensional prediction model, thus it greatly improves the accuracy of the prediction model. The application of gray prediction model group provides an effective method for the data screening in the setup process of the gray model, but also makes it more efficient and accurate in prediction of athletic performance, and provides reference for the future development trend of competitive sports and the decisions of related competitive sports.

TABLE 4 : The posterior error test and filtering accuracy values of the model group for the world men's long jump

Dimension	14	13	12	11	10	9	8	7	6	5	4
C	0.529	0.380	0.339	0.705	0.740	0.517	0.984	0.670	0.539	0.380	0.990
P	0.945	0.901	1.00	0.890	1.00	1.000	0.695	0.790	0.900	0.910	0.670
Level	3	2	1	4	4	3	4	4	3	2	4
Filtering accuracy	98.31%	99.46%	99.41%	98.52%	98.57%	98.46%	98.90%	98.48%	98.15%	99.34%	99.21%

TABLE 5 : The test table of 12 dimensional GM (1, 1) model

Number	Year	Original value	Model value	Residual error	Relative error	Class ratio deviation
1	2002	8.59	8.5900	0	0	
2	2003	8.43	8.4336	-0.0036	0.0004	-0.0175
3	2004	8.41	8.4210	-0.0110	0.0013	-0.0009
4	2005	8.37	8.4048	-0.0384	0.0046	-0.0033
5	2006	8.36	8.3959	-0.0359	0.0043	0.0003
6	2007	8.31	8.3833	-0.0733	0.0088	-0.0045
7	2008	8.42	8.3708	0.0492	0.0058	0.0145
8	2009	8.71	8.3583	0.3517	0.0404	0.0347
9	2010	8.27	8.3458	-0.0758	0.0092	-0.0516
10	2011	8.21	8.3334	-0.1234	0.0150	-0.0058
11	2012	8.28	8.3209	-0.0409	0.0049	0.0099
12	2013	8.31	8.3085	0.0015	0.0002	0.0051

CONCLUSIONS

The application of gray forecast model in sports must be established based on proper data analysis; through analysis of its data development trends, it starts from the model analytical factors, establishes the corresponding gray prediction model, keeps the scientifically of the gray prediction method in the applications of competitive sports as much as possible; for the forecast of sports scores, their information is less; due to the interference from other strong factors, the prediction accuracy of conventional analytical prediction method is not high; the characteristics of the gray model will enable it predict sports results with high accuracy, and provide a favorable reference for decision-making related to sports, etc.; when the statistical data is relatively large, the data can be filtered through the establishment of gray prediction model group; it chooses a gray forecasting model with higher precision accuracy to improve the prediction accuracy of the gray system.

REFERENCES

[1] Tu Chun-Jing, Du He-Ping, Wang-Wei; Gray prediction on chinese and foreign women's swimming accomploshments in the olympic games based on GM (1,1) model of multidimensional gray[J]. Journal of Jinggangshan University, **32(2)**, **(2011)**.

[2] Wu Xin; Study on the Interactive Development of Internet Media and Olympic in China [J]. Sports Culture Guide, **5**, **(2011)**.

[3] Zhang Zheng-Min; Prediction of the result of men's 100m run in the Olympic Games based on the grey GM (1, 1) model [J]. Journal of Physical Education, **18(4)**, 111-114 **(2011)**.

[4] Lin Dehua; The Result Analysis and Prediction of the Part Dominant Events of Track and Field in China on London Olympic Games[J]. Journal of Capital College of Physical Education, **23(6)**, 540-542, 549 **(2011)**.

[5] He Jiang-Hai; Beijing Olympic swimming event performance prediction based on Visual C++[J]. Journal of Wuhan Institute of Physical Education, **41(6)**, 63-65 **(2007)**.

[6] Jiang Yi-Peng; Prediction of Medals of China in the 28th Olympic Games [J]. Journal of Tianjin Institute of Physical Education, **19(2)**, 86-87 **(2004)**.

[7] Wu Xi-Lin, Li Jian-Chen; Prediction and Analysis of Chinese Track and Field Performance at 2008 Olympic Games[J]. Journal of Wuhan Institute of Physical Education, **39(6)**, 47-51 **(2005)**.

[8] Wang Guo-Fan, Zhao Wu, Liu Xu-Jun, Feng Shu-Hui, Xue Er-Jian, Chen Lin, Wang Bo; Olympic Performance Prediction based on GA and Regression Analysis [J]. China Sport Science and Technology, **47(1)**, **(2011)**.

[9] Fan Wen Jie, Zhou Ai Jie, Liu Jian Min; Grey Prediction of Track & Field Results at the 28th Olympic Games [J]. Journal of Beijing Sport University, **25(6)**, 847-849 **(2002)**.