Sports Performance Prediction Based on Variable Weight Support Vector Machine

Kuifeng GAO

Department of Social Science and Public Education, Shangdong Medical College, Linyi Shangdong, 276005, China

***Abstract* — In this paper we consider the application of grey model to solve more effectively and accurately problems which are difficult to solve by traditional statistical methods, such as competitive sports performance prediction problems with 'poor information', 'small sample' and underlying 'dynamics'. However, the application of grey model here also has many problems, and traditional GM (1, 1) model cannot reach the corresponding precision in the prediction of many applications. We therefore further consider in this paper the application of GM (1, 1) grey model based on residual correction in competitive sports. We do this by taking the annual best performances of World Men's 200 Meters from 2003 to 2013 as samples, where the prediction precision reaches the first-class precision and prediction modeling for best performance of World Men's 200 Meters in 2014. Our results show the prediction performance obtained is 20.59 seconds, which shows good agreement with actual figures.**

Keywords - sports performance prediction; grey vector model; residual correction; data mining; sports event decision-making

I. INTRODUCTION

Since 1982 when Professor Deng Julong founded grey mathematics, application of grey mathematics and its theoretical research have had great development and the modeling fields have involved decision-making, control, planning and prediction and its application field has involved medical treatment, economy, education and sports and other relevant fields just in over 30 years. However, in the athletics field, application of grey model develops very rapidly. According to statistics of relevant papers, in current 5 years, the number of competitive sports papers with grey model as research methods will grow at a rate from 9.81% to 10.10% per year [1]. The research and application of grey model has become one of topics with most rapid development in application in the field of sports. In this paper, performance of World Men's 200 Meters is taken as example to research application of grey model in competitive sports performance prediction. Establishment of GM (1, 1) grey prediction model of residual correction is determined finally with research and analysis, which improves application and development of grey model in competitive sports field at the same time of providing reference for Men's 200 Meters sport in our country.

II. ANALYSIS ON PERFORMANCES OF CHAMPIONS OF WORLD MEN'S 200 METERS OVER THE YEARS

In order to predict the development trend of athletic performance of World Men's 200 Meters, the method of documentary is used in this paper for statistics of annual best performances of World Men's 200 Meters from 2003 to 2013 and based on this, the GM(1, 1) grey prediction model of residual correction is conducted. Its statistical data is shown in Table 1:

TABLE 1. BEST PERFORMANCES TABLE[7] OF WOMEN'S PENTATHLON FROM 2003 TO 2013

Year	2003	2004	2005	2006	2007	2008
Performance	20.13	20.58	20.69	20.91	20.58	20.91
Year	2009	2010	2011	2012	2013	
Performance	20.94	20.74	20.82	20.7	20.63	

For sports performance prediction, only the modeling of performance data of different development and change trend is conducted with different and corresponding gray models needed, the prediction results can reach corresponding prediction precision. In order to build a prediction model of high precision for annual best performance of World Men's 200 Meters, the change and development trend of annual best performances of World Men's 200 Meters from 2003 to 2013 needs to be analyzed first so that the corresponding grey prediction model can be determined.

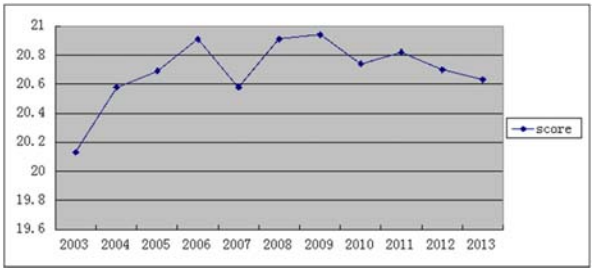

Figure 1. Change diagram of best performances of men's 200 meters from 2003 to 2013

From the figure, it can be seen that change range of best performance of World Men's 200 Meters is from 20.13 to 20. 94, and from 2003 to 2006, the performance change showed a downward trend in general and in 2007, the performances increased, which is higher compared with that of 2006 and 2008. Performance of 2009 has not great change compared with that of 2008, however, from 2009 to 2013, the performance change showed upward trend. Because the

change showed volatility with increase and reduction and it fluctuated unsteadily without rule to be followed, moreover, the data amount and information amount are less, at the same time, development and change of annual best performance of World Men's 200 Meters is a dynamic process, and the requirements of traditional mathematical methods such as regression analysis, interpolating fit and other traditional mathematical methods in sample data are higher with need of the sample containing large amount of data, certain change rule and static development, it has been difficult to analyze it and judge the change trend. Through the analysis on the trend of its development and change, it is decided to use GM (1, 1) grey prediction model to build model for it in this paper and to predict the best performance of World Men's 200 Meters in 2014.

III. GM(1,1) MODEL OF BEST PERFORMANCE OF WORLD MEN'S 500 METERS

A. Inspection of data

Original data sequence of grey model is generated for best performances of World Men's 200 Meters from 2003 to 2013 through data statistics. Generated data sequence is shown as follows:

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \cdots, x^{(0)}(11))$$

In order to ensure the accuracy and precision of GM (1, 1) grey prediction model, inspection processing needs to be conducted to original sequence data obtained in modeling. Step ratio $\lambda(k)$ of the original data sequence is required to be in the range of $(e^{-\frac{2}{n+1}}, e^{\frac{2}{n+2}})$, or the data can not be used for grey prediction modeling. If the data sequence of step ratio is not in the required range, original data sequence shall be processed. Range (0.8465，1.1663) can be obtained by substituting $n = 11$ into it.

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, k = 2,3,\cdots,11$$

From the formula of step ratio of data sequence

It can be obtained that data sequence of step ratio obtained is in the range (0.9781，1.0096), thus, original data sequence can be used for direct modeling of prediction system.

B. Establishment of GM(1,1) model

(1) Cumulative and mean processing for new data sequence

Cumulative processing is conducted to new data sequence by accumulating operator $AG0$ to weaken its randomness and cumulative data sequence can be obtained.

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

The generated mean data sequence $z^{(1)}(k)$ is obtained with formula $z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k)$.

(2) Construction of data matrix B and data vector Y

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

(3)Calculation of $\hat{u}$

$$\hat{u} = (a,b)^T = (B^TB)^{-1}B^TY = \begin{pmatrix} -0.0002 \\ 20.7291 \end{pmatrix}$$

(4)Establishment of model

$$\frac{dx^{(1)}}{dt} - 0.0002x^{(1)} = 20.7291$$

It can be obtained

$$x^{(1)}(k+1) = 122821e^{-0.000168803k} - 122801$$

（5）The obtained grey model prediction value is

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

Where $k = 2,3,\cdots,13$, then it can be obtained that the prediction vale of best performances of Men's 200 Meters from 2003 to 2013 is shown in following table:

TABLE 2. PREDICTION TABLE OF BEST PERFORMANCES OF MEN'S 200 METERS FROM 2003 TO 2013

Year	2003	2004	2005	2006
Performance	20.1300	20.7342	20.7388	20.7412
Year	2007	2008	2009	2010
Performance	20.7447	20.7482	20.7517	20.7553
Year	2011	2012	2013	
Performance	20.7588	20.7623	20.7658	

C. Inspection of GM(1,1) model

All inspection indexes of model are shown as follows:

From Table 1, it can be obtained that the maximum relative error of the prediction model is 0.9%, and the maximum deviation of step ratio is 0.0217.

TABLE 3. INSPECTION TABLE OF GM(1,1) MODEL

Year	Original value	Model value	Residual	Relative error	Deviation of step ratio

2003	20.13	21.3000	0	0	
2004	20.58	20.7342	-0.1542	0.0075	0.0217
2005	20.69	20.7377	-0.0477	0.0023	0.0051
2006	20.91	20.7412	0.1688	0.0081	0.0104
2007	20.58	20.7447	-0.1647	0.0080	-0.0162
2008	20.91	20.7482	0.1618	0.0077	0.0156
2009	20.94	20.7517	0.1883	0.0090	0.0013
2010	20.74	20.7553	-0.0153	0.0007	-0.0098
2011	20.82	20.7588	0.0612	0.0029	0.0037
2012	20.70	20.7623	-0.0623	0.0030	-0.0060
2013	20.63	20.7658	-0.1358	0.0066	-0.0036

(1) Qualified residual model

Relative error sequence is

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

then mean relative error can be obtained as

$$\overline{\Delta} = \frac{1}{\mathrm{n}} \sum \Delta_k = 0.0051$$

(2) Qualified model of association degree

Absolute association degree g is the absolute association degree of original sequence

$$g = 0.9947$$

$x^{(0)}$ and corresponding grey prediction sequence $\hat{x}^{(0)}$, and it is obtained that

(3) Qualified model of mean square deviation ratio

S_1^2 and S_2^2 are variances of the original sequence $x^{(0)}$ and the residual sequence $\varepsilon(k)$, and the mean square deviation ratio can be obtained as

$$C = S_2 / S_1 = 0.5710$$

Inspection data of model precision can be obtained as shown in following Table 2 from data of Table 1:

TABLE 4. PRECISION TABLE OF GM (1,1) PREDICTION MODEL

Precision class	Relative error	Absolute association degree	Mean square deviation ratio
Third class	0.0051	0.9947	0.5710

From Table 4, it can be known that for the performance prediction of World Men's 200 Meters by using traditional GM (1, 1) grey model directly, the relative error is 0.0051 with first-class precision and the absolute association degree is 0.9947 with first-class precision and mean square deviation ratio is 0.5710 with third-class precision, and in the general, the model prediction precision is better. In order to guarantee the precision of established prediction model, in this paper, the precision of minimum precision index among three indexes is taken as that of prediction model. Then from Table 4, it can be known that prediction precision of GM (1, 1) grey prediction model is at the third class.

IV. ESTABLISHMENT OF GM(1,1) MODEL OF RESIDUAL SEQUENCE

A very important problem in application of grey model is the precision of prediction model. In practical applications, the precision of prediction model of grey model can not meet corresponding requirements due to the change and development characteristics of data itself, model application and knowledge quantity of problem solver himself and other limitation, and there is exception for competitive sports performance prediction model. How to solve the prediction precision problem of gray model and how to improve and process grey model effectively is a significant problem in development of grey model. From computing results in the last chapter, it can be known that for grey model modeling for annual best performance of World Men's 200 Meters, its inspection index of mean square deviation ratio is 0.5289 with third-class precision, and prediction precision of prediction model is lower. There have been many methods and theories of performance aiming at improvement ways of prediction precision problem of grey model, and most of them are to process sample data and select albinism differential equation of grey model. In order to improve prediction precision of annual best performance of World Men's 200 Meters, GM (1, 1) grey prediction model is amended by conducting grey prediction of residual sequence for the model in this paper.

A. Establishment of GM(1,1) prediction model of residual sequence

TABLE 5. PREDICTION MODEL RESIDUALS OF BEST PERFORMANCES OF MEN'S 200 METERS FROM 2003 TO 2013

Year	2003	2004	2005	2006	2007	2008
Performance	0	-0.1542	-0.0477	0.1688	-0.1647	0.1618
Year	2009	2010	2011	2012	2013	
Performance	0.1883	-0.0153	0.0612	-0.0623	-0.1358	

Residual refers to difference between the prediction value and actual value of grey prediction model, and grey prediction model of residual sequence refers to that residual sequence is taken as the original sequence of grey prediction model for modeling of GM(1,1) model to obtain the prediction model, and then residual prediction model is used to amend GM(1,1) grey prediction model established originally to improve the precision of prediction model. From above calculation, the residual of annual best performance of World Men's 200 Meters is shown in Table 5。

Residuals from 2007 to 2013 are taken to generate residual sequence of best performance of World Men's 200 Meters:

$$\varepsilon_0 = (\varepsilon_0(1), \varepsilon_0(2), \cdots, \varepsilon_0(7))$$

From Table 5, it can be known that parts of residual values are negative and can not be used for modeling of GM (1, 1) grey model directly. At the same time, if grey prediction modeling is conducted for them, step ratio of the sequence $\lambda(k)$ will be required to be in range $(e^{-\frac{2}{n+1}}, e^{\frac{2}{n+2}})$, or the data can not be used for grey prediction modeling. Where $n = 7$, range (0.7788，1.2488) can be obtained by substituting it.

In order to conduct modeling of GM(1,1)grey prediction model for it, the residual sequence needs to be processed to obtain new data sequence meeting establishment condition of grey model. In this paper, translation transformation is conducted for residual sequence ε_0, namely a suitable positive value c is taken to meet $\varepsilon_1^{(0)}(k) = \varepsilon_0(k) + c$, and in this paper, make $c = 15$ so that the transformation sequence can be obtained as shown in Table 6:

TABLE 6. NEW RESIDUAL SEQUENCE VALUE

Year	2007	2008	2009	2010
Performance	-0.1647	0.1618	0.1883	-0.0153
Year	2011	2012	2013	
Performance	0.0612	-0.0623	-0.1358	

Through inspection of step ratio, it can be obtained that step ratio range of new residual sequence is [0.9785，1.0136], which meets establishment condition of GM(1,1)grey prediction model. Namely the modeling can be conducted for it according to above modeling steps of GM (1, 1) grey prediction model.

B. Result and inspection of GM(1,1) model of residual sequence

GM(1,1) modeling is conducted for the new residual sequence and assumed that the cumulative data sequence is $\varepsilon_1^{(1)}$, solution of the albinism differential equation can be obtained as:

$$\hat{\varepsilon}_1^{(1)}(k+1) = 3718.11 - 3703.27\exp(-.00410962k)$$

Comparative analysis and inspection are conducted for obtained prediction value and original value to get error value table of the model as follows:

Error value and residual value of obtained model in Table 7 are used to calculate residual GM(1,1) model to get all inspection indexes for calculation and the results obtained are shown in Table 8:

TABLE 7. ERROR VALUE OF GM(1,1) MODEL OF NEW RESIDUAL SEQUENCE

Year	Original value	Model value	Residual	Relative error	Deviation of step ratio
2007	14.8353	14.8353	0	0.0035	
2008	15.1618	15.1878	-0.0260	0.0002	0.0255
2009	15.1883	15.1255	0.0628	0.0025	0.0058
2010	14.9847	15.0635	-0.0788	0.0079	-0.0094
2011	15.0612	15.0017	0.0595	0.0052	0.0092
2012	14.9377	14.9402	-0.0025	0.0088	-0.0041
2013	14.8642	14.8789	-0.0147	0.0013	-0.0008

TABLE 8. PRECISION TABLE OF RESIDUAL GM(1,1) PREDICTION MODEL

Precision class	Relative error	Absolute association degree	Mean square deviation ratio
First class	0.0023	0.9978	0.3272

From Table 8, it can be known that relative error of model is 0.0023 with first-class precision and absolute association degree is 0.9978 with first-class precision and mean square deviation ratio is 0.3272 with first-class precision. Precision of residual GM(1, 1) grey system in conducting color prediction model is very high and it can be used for residual prediction.

Restore the residual sequence and assume the residual prediction values from 2003 to 2006 are zero and predictive residual value can be obtained as shown in Table 9:

TABLE 9. RESIDUAL PREDICTION VALUE OF BEST PERFORMANCE OF MEN'S 200 METERS FROM 2003 TO 2013

Year	2003	2004	2005	2006	2007	2008
Performance	0	0	0	0	-0.1647	0.1878
Year	2009	2010	2011	2012	2013	
Performance	0.1255	0.0635	0.0017	-0.0598	-0.1211	

V. ESTABLISHMENT OF GM(1,1) GREY PREDICTION MODEL BASED ON RESIDUAL CORRECTION

From calculation of last chapter, it can be known that the solution of albinism differential equation of new residual sequence is：

$$\hat{\varepsilon}_1^{(1)}(k) = 3718.11 - 3703.27\exp(-.00410962k)$$

However, solution of this albinism differential equation is obtained by conducting GM(1,1) grey prediction model for residual sequence values from 2007 to 2013, and in order to combine with the results of GM (1,1) grey prediction model conducted for best performance sequence of World Men's 200 Meters from 2003 to 2013,it needs to be processed. Assumed that

$$\varepsilon_2^{(1)}(k) = \begin{cases} 0 & k \le 4 \\ \varepsilon_1^{(1)}(k-4) & k > 4 \end{cases}$$

Assumed that $y^{(0)}(k)$ is the prediction value of GM(1,1) grey prediction model based on residual correction, where $k=1,2,\cdots,13$. Then $y^{(1)}(k)$ is cumulative data sequence of prediction value of residual grey prediction model, namely it can be obtained that:

$$y^{(1)}(k)=\hat{x}^{(1)}(k)+\varepsilon_2^{(1)}(k-4)$$

Restore the obtained prediction sequence $y^{(0)}(k)$ with cumulative data sequence $y^{(1)}(k)$, and prediction value and the error value can be obtained as shown in Table 9:

TABLE 10. ERROR VALUE OF GREY PREDICTION SEQUENCE GM(1,1) MODEL

Year	Original value	Model value	Residual	Relative error
2003	20.13	21.3000	0	0
2004	20.58	20.7342	-0.1542	0.0075
2005	20.69	20.7377	-0.0477	0.0023
2006	20.91	20.7412	0.1688	0.0081
2007	20.58	20.5800	0	0
2008	20.91	20.9360	-0.0260	0.0012
2009	20.94	20.8772	0.0628	0.0030
2010	20.74	20.8188	-0.0788	0.0038
2011	20.82	20.7605	0.0595	0.0029
2012	20.70	20.7025	-0.0025	0.0001
2013	20.63	20.6447	-0.0147	0.0007

Error value and residual value of obtained model in Table 10 are used to calculate GM(1,1) grey prediction model based on residual correction to get all inspection indexes for calculation and the results obtained are shown in Table 8:

TABLE 11. PRECISION TABLE OF GM(1,1) PREDICTION MODEL BASED ON RESIDUAL CORRECTION

Precision class	Relative error	Absolute association degree	Mean square deviation ratio
First class	0.0027	0.9980	0.3328

From Table 8, it can be known that relative error of model is 0.0027 with first-class precision and absolute association degree is 0.9980 with first-class precision and mean square deviation ratio is 0.3328 with first-class precision. Namely obtained prediction precision class of GM (1, 1) grey prediction model based on residual correction is at the first class, thus, it can be used for exact prediction of annual best performance of World Men's 200 Meters.

Base on prediction of annual best performance of World Men's 200 Meters in 2014 with obtained model and from formula

$$\hat{x}^{(1)}(k+1)=122821e^{-0.000168803k}-122801$$

It can be obtained that when $k=10$, $k=11$, $\hat{x}^{(1)}(11)=227.5006$, $\hat{x}^{(1)}(12)=248.2699$.

Namely it can be obtained that the prediction value of GM(1, 1) grey prediction model in 2014 is $\hat{x}^{(0)}(12)=20.7639$.

Similarly, it can be obtained that residual prediction value in 2014 is $\hat{\varepsilon}_2(8)=-0.1821$. Then it can be obtained that annual best performance value of World Men's 200 Meters in 2014 is 20.59 with GM (1, 1) model based on residual correction.

VI. CONCLUSION

In this paper, grey prediction modeling is conducted for annual best performance of World Men's 200 Meters and at the same time of prediction of performance in 2014, application methods of grey model in competitive sports performance prediction is discussed and application of grey model in competitive sports is promoted. GM (1, 1) grey model is suitable for modeling process of which sample data distribution shows index change, and through modeling analysis of statistical data in this paper, grey prediction model is expanded and GM (1, 1) grey prediction model based on residual correction is used to predict annual best performance of World Men's 200 Meters and it is compared with GM (1, 1) grey model. From the results, it can be found that residual GM (1, 1) grey prediction model has low complexity and high prediction precision, which is suitable for competitive sports performance prediction.

REFERENCES

[1] Wang J. BP neural network-based sports performance prediction model applied research[J]. Journal of Chemical & Pharmaceutical Research, 2014.

[2] Milosevic M B, Milosevic M M. Model for Assessing the Physical Status, as Well as Prediction and Programming of Training and Sports Performance of a Soccer Player[J]. Journal of Physical Education & Sport, 2013, 13(3).

[3] Jinyu Hu and Zhiwei Gao. Distinction immune genes of hepatitis-induced heptatocellular carcinoma[J]. Bioinformatics, 2012, 28(24): 3191-3194.

[4] Liu, Y., Yang, J., Meng, Q., Lv, Z., Song, Z., & Gao, Z. (2016). Stereoscopic image quality assessment method based on binocular combination saliency model. Signal Processing, 125, 237-248.

[5] Jinyu Hu, Zhiwei Gao and Weisen Pan. Multiangle Social Network Recommendation Algorithms and Similarity Network Evaluation[J]. Journal of Applied Mathematics, 2013 (2013).

[6] Yishuang Geng, Jin Chen, Ruijun Fu, Guanqun Bao, Kaveh Pahlavan, Enlighten wearable physiological monitoring systems: On-body rf characteristics based human motion classification using a support vector machine, IEEE transactions on mobile computing, 1(1), 1-15, Apr. 2015

[7] Lv, Z., Halawani, A., Feng, S., Ur Réhman, S., & Li, H. (2015). Touch-less interactive augmented reality game on vision-based

wearable device. Personal and Ubiquitous Computing, 19(3-4), 551-567.

[8] Jinyu Hu and Zhiwei Gao. Modules identification in gene positive networks of hepatocellular carcinoma using Pearson agglomerative method and Pearson cohesion coupling modularity[J]. Journal of Applied Mathematics, 2012 (2012).

[9] Jiang, D., Ying, X., Han, Y., & Lv, Z. (2016). Collaborative multi-hop routing in cognitive wireless networks. Wireless personal communications, 86(2), 901-923.

[10] Jin S, Guo T. Grey prediction model in world women's pentathlon performance prediction applied research[J]. Journal of Chemical & Pharmaceutical Research, 2014.

[11] Song L, Shen M, Chen X, et al. Neural network-based athletics performance prediction optimization model applied research[J]. Journal of Chemical & Pharmaceutical Research, 2014.

[12] Tartaruga M P, Mota C B, Peyré-Tartaruga L A, et al. Scale Model on Performance Prediction in Recreational and Elite Endurance Runners[J]. International Journal of Sports Physiology & Performance, 2014, 9:650-655.

[13] Tartaruga M P, Mota C B, Peyré-Tartaruga L A, et al. Scale Model on Performance Prediction in Recreational and Elite Endurance Runners[J]. International Journal of Sports Physiology & Performance, 2014, 9:650-655.

[14] Jackson P S. Performance prediction for Olympic kayaks[J]. Journal of Sports Sciences, 1995, 13(13):239-45.

[15] Tian B, Qin L M, Jie-Ling L I. Relations between target orientation,perfectionism and sports performance with respect to intelligence handicapped athletes[J]. Journal of Physical Education, 2010.

[16] Baltzopoulos V. NEUROMUSCULAR DIFEERENCES BETWEEN ISOKINETIC AND SPORT MOVEMENTS: IMPLICATIONS FOR TRAINING - PERFORMANCE PREDICTION[C]// 1997.

[17] Harry Prapavessis. The POMS and sports performance: A review[J]. Journal of Applied Sport Psychology, 2000, 12(1):34-48.

[18] Cooper S, Baker J, Eaton Z, et al. A simple multistage field test for the prediction of anaerobic capacity in female games players.[J]. British Journal of Sports Medicine, 2004, 38(6):784-9.

[19] Whitehead R, Butz J W, Kozar B, et al. Stress and performance: an application of Gray's three-factor arousal theory to basketball free-throw shooting.[J]. Journal of Sports Sciences, 1996, 14(5):393-401.

.