ISSN : 0974 - 7435 Volume 10 Issue 6

BioTechnology
An Indian Journal

FULL PAPER

BTAIJ, 10(6), 2014 [1447-1453]

The contrastive study of prediction of women's heptathlon performance based on grey theory and BP neural network prediction model

Zhongren Hou

Institute of Physical Education and Health, Zhaoqing University, ZhaoQing 526061, Guangdong, (CHINA)

ABSTRACT

By adopting literature consultation the data from the scores of successive women's heptathlon champions during the 23rd to 30th Olympic Games, this paper, based on the scores of the women' heptathlon champions during the 23rd to 29th Olympic Games, respectively establishes the Grey Forecasting Model and the BP Neural Network Prediction Model, and predicts the scores of the women' heptathlon champion of the 30th Olympic Games. Meanwhile, this paper also compares the prediction with the actual value of these two models of the scores of the women' heptathlon champion of the 30th Olympic Games, and analyzes these two models' forecast precision. According to the result, prediction of the Grey Forecasting Model was 6572, and its error was 5.5%. And prediction of the BP Neural Network Prediction Model was 6967, and its error was 0.18%. However, on the whole, the Grey Forecasting Model gives better fitting degree to data. The model precision was first grade, and the Grey Model is much fitter to prediction of scores of the women' heptathlon champion of the Olympic Games.

KEYWORDS

Grey theory;
BP neural network;
Prediction model;
Women' heptathlon.

INTRODUCTION

Quantitative prediction to competitive sports performance is very practical to research on athletics, policy-making, and arrangement of sportsmen, etc. With the development of sport culture, an increasing number of systematic sciences have been applied to prediction of athletic contest achievements and study of decision-making of athletic problems. Traditional prediction methods mainly include mathematical and statistical methods, which are based on a large number of data and require that problems to be dealt with must be static problems. Prediction problems of competitive sports performance usually are small sample problems, so it is hard for traditional methods to get comparatively precise results, especially with individual case prediction that traditional methods have nothing to do. Grey mathematics and mathematical neural network are emerging systematic science methods in last decades, and they have been widely used in the fields of pattern recognition and data excavation. The two methods' features make them own unique advantages in dealing with prediction problems of competitive sports. Their application in competitive sports problems is very important to the development of sports science.

Aiming at grey prediction model and neural network prediction model' application in prediction problems of competitive sports performance, a lot of schol-

FULL PAPER

ars at home and abroad have proceed systematic research on them. Among the researches, Song Ailing and others (2012) established prediction model of Chinese athlete Liu Jing by using artificial neural network function mapping method, based on MATLAB simulation platform. And they guided athletes' scientific train on the basis of this model. Sun Qun (2011) collected statistics and studied Liu Xiang 110 hurdles performance. And in view of the features of those statistics, they put forward a king of performance prediction model based on fuzzy-neural network, and then they got relatively precise prediction results[2]. Sun Qiang (2012) established gray equip-dimensional model GM (1.1) of best 400m performance of 20th to 29th Olympic Games, and predicted the best 400m performance of 30th Olympic Games. They provided principle evidence of the establishment of our country's 400m sport training plan. Wang Dao lin, Fan Xin sheng (2005) counted and studied China's amount of gold medals of 23th to 28th Olympic Games, and they set grey prediction model on the basis of the data features. In the same time, they used Grey Correlation to analyze the influence of events on amount of gold medals. Liu Jia jin and others added up number of papers that applied grey mathematics model to study competitive sports based on GNKI, and they deeply analyzed grey mathematics model' application prospect on competitive sports.

This paper collects statistics scores of the women' heptathlon champion of the 23rd to 30th Olympic Games, and respectively establishes grey prediction model GM (1.1) and BP neural network prediction model based on those statistics. After comparing prediction results of two models, the paper studies the two prediction models' application in the performance prediction of competitive sports.

ANALYSIS OF THE SCORES OF THE WOMEN' HEPTATHLON CHAMPION OF THE 23RD TO 30TH OLYMPIC GAMES

A modeling method has its own advantages as well has its limitations. To comparative study the effect of grey prediction model GM (1.1) and BP neural network dealing with performance prediction of competitive sports, this paper uses literature consultation, counted scores of the women' heptathlon champion of the 23rd to 30th Olympic Games, and respectively establishes grey prediction model GM (1.1) and BP neural network prediction model based on the scores. The statistics data are in TABLE 1:

It is necessary to use different methods to establish models of different time series of variation tendency to achieve content prediction effect. one of the examples is that after adopting accumulation to the grey prediction model GM (1.1) of the grey model, the new time series gets strong index law, and so we can get relatively more precise results in the grey prediction model of the time series which has grown into a series with non-monotone swing development.

From Figure 1, we can find that the change of scores of the women' heptathlon champion of the 23rd to 30th Olympic Games is in the swinging tendency. On the whole, its variation range is between 6390 and 7291. Among them, the score of the women' heptathlon champion of the 23rd Olympic Games is the lowest, 6390. And the score of the women' heptathlon champion of the 24th Olympic Games is highest, 7291. Among 24th to 27th Olympic Games, the scores of the women' heptathlon champion kept descending, but scores of the women' heptathlon champion of the 27th to 30th Olympic Games were in the swinging tendency. Traditional time series forecasting methods, like probabilistic method and so on, need a large number of data, and the change of data must have some regulation. We can see from Figure 1 that the data of scores of the women'

TABLE 1: Scores of the women' heptathlon champion of the 23rd to 30th Olympic Games

Olympic Games	23rd	24th	25th	26th	27th	28th	29th	30th
Score	6390	7291	7044	6780	6584	6952	6733	6955

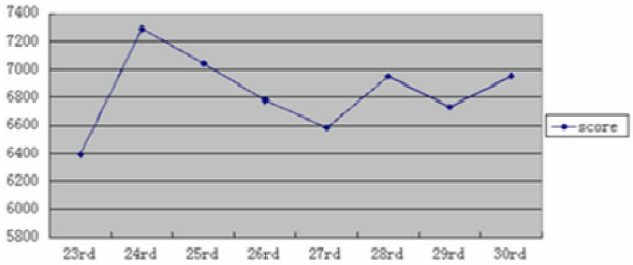

Figure 1: The change of scores of the women' heptathlon champion of the 23rd to 30th Olympic Games

heptathlon champion of the 23rd to 30th Olympic Games are less and have no clear regulations. Thus, it is hard to use traditional methods to do function fitting. In that, these texts adopt grey prediction model and BP neural network prediction model to analyze prediction problems of scores of the women' heptathlon champion of the Olympic Games. At the same time, according to the features of data changing, the grey prediction model this paper adopting is grey prediction model GM (1.1).

Based on the scores of the women' heptathlon champion of the 23th to 29th Olympic Games, referring to scores of the women's heptathlon champion of the 30th Olympic Games, this paper uses the grey prediction model GM (1.1) and BP neural network prediction model respectively to forecast the scores of the women' heptathlon champion of the 30th Olympic Games. And this text compared the prediction results with the scores of the women' heptathlon champion of the 30th Olympic Games so that we can analyze the application of individual cases prediction problems of grey prediction model GM (1.1) and BP neural network prediction model.

THE ESTABLISHMENT OF THE GREY PREDICTION MODEL GM (1.1)

Grey System Theory has formed the theoretical system that bases on Grey relational space and takes grey prediction model GM (1.1) as its main body since it was established by Professor Deng Ju long in 1982. It has been widely used in the fields of system analysis, modeling, controlling, decision-making etc. Grey prediction model GM (1.1), by proceeding differential function fitting with generation of spatial series of time series, ensures the developmental quotient and cooperation index, and then, by whitening the grey differential equation, confirms the response type winterization equation at last. By that response type winterization equation, we can get prediction data that we need, and check its degree of accuracy. This text is based on scores of the women' heptathlon champion of the 23rd to 30th Olympic Games and establishes grey prediction model GM (1.1).

Test of data series

Grey prediction model GM (1.1) has limitations that its application to establish model must be applied in some condition. Whether a set of time series data can establish prediction model by grey prediction mode GM (1, 1) can be confirmed by comparing its grade ratio.

After collecting statistics, we can do series generation to scores of the women' heptathlon champion of the 23th to 30th Olympic Games, and the series generation is as following:

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, (x)^{(0)}(7))$$

To ensure whether prediction problems of scores of the women' heptathlon champion of the Olympic Games can apply the grey prediction model or not, we need to calculate and check the class radio series of the time series that we need to establish. When the class radio series values $\lambda(k)$ all set in block $(e^{-\frac{2}{n+1}}, e^{\frac{2}{n+2}})$, we can use grey prediction model to forecast. In this paper, the question of prediction about the scores of the Women's Heptathlon champion of the Olympics, $n = 7$, we can know that its magnitude ratio belongs to the interval (0.7788, 1.2488). And its magnitude ratio formula of time series is

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, k = 2,3,\cdots,7$$

From the formula above, the ratio is as TABLE 2:

From TABLE 2, all the magnitude ratios belong to the interval (0.7788, 1.2488). Therefore, on the basis of it, we can build up the grey prediction model of the data forecast.

GM (1, 1) the establishment of the grey prediction model

(1) The accumulation and mean method of time series

After the AGO accumulation and weakening the randomness of $x^{(0)}$, we can get a new series:

$$x^{(1)} = (6390,13681,20725,27505,34089,41041,47774)$$

Meanwhile, we apply mean method to $x^{(0)}$, and the formula is:

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k)$$

And the time series come from the means is as following:

$$z^{(1)} = (10035,17203,24115,30797,37565,44408)$$

(2) The equalization of development quotient a and co-

FULL PAPER

operation index b

From the mean time series above, we can get data matrix B and data vector Y of the grey model:

$$B=\begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(7) & 1 \end{bmatrix}, Y=\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(7) \end{bmatrix}$$

With the matrix equation $Y = Bu$, we can get the development quotient a and the cooperation index b in the GM (1, 1) grey prediction model. With the matrix equation and least square method, we can know:

$$\hat{u} = (\hat{a}, \hat{b}) = (B^T B)^{-1} B^T Y$$

From the equation above, we can know that the development quotient a is 0.0137 and the cooperation index b is 7272.4.

(3) The establishment of the grey model

From the development quotient and the cooperation index above, we can develop a grey differential equation of the model:

$$x^{(0)}(k) + 0.0137 z^{(1)}(k) = 7272.4$$

And from this equation, we can develop another albinism differential equation and finally we can develop the following equation:

$$x^{(1)}(k+1) = 530361 - 523971 * e^{-0.0137k}$$

(4) The calculation of the predicted value

We can know the predicted value of the accumulated value $x^{(1)}(8)$ from the functions and apply inverse accumulated generating operation to the predicted value. After these, we can know the that the score of the Women's Heptathlon champion of the 30th Olympics is 6572, and the predicted values of the scores of the Women's Heptathlon champions of the 24th—29th Olympics are showed in TABLE 3:

GM (1, 1) the inspection of the grey prediction model

All the standards of inspection are showed below:

Now we can know that the max of the relative error is 4.01% and the minimum of the class ratio deviation is 0.1355.

(1) Residual qualified model

The series of relative error is:

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

And the average relative error is:

$$\overline{\Delta} = \frac{1}{n} \sum \Delta_k = 0.0178$$

(2) Correlation qualified model

Absolute correlation degree g is the absolute correlation degree between raw series $x^{(0)}$ and the corresponding grey predicted series $\hat{x}^{(0)}$.

$$g = 0.9941$$

(3) mean variance ratio qualified model

S_1^2 and S_2^2 are the variances of raw series $x^{(0)}$ and residual series $\varepsilon(k)$. And we can get the mean variance ratio below.

$$C = S_2 / S_1 = 0.3218$$

From all the inspection data, we can conclude that the precision of the GM (1, 1) grey model we have built up is A-level.

TABLE 2 : The magnitude ratios of the scores in the Women's Heptathlon champions in the 24th—29th Olympics

Olympics	24th	25th	26th	27th	28th	29th
ratio	0.8764	1.0351	1.0389	1.0298	0.9478	1.0325

TABLE 3 : The predicted values of the scores of the Women's Heptathlon champions of the 24th —29th Olympics based on the grey model

Olympics	24th	25th	26th	27th	28th	29th
Predicted value	7136	7039	6943	6848	6755	6663

THE ESTABLISHMENT OF THE BP NEURAL NETWORK PREDICTION MODEL

Artificial neural network is built up with the research and analysis of the structure and function of biologic neural network, which can make the artificial intelligence people need a reality. Since 1980s, artificial neural network has developed rapidly and been widely applied in the area of pattern recognition, signal processing, decisions recognizing, and real-time optimal control. With the advantages of easy operation, reliable and stable system, and convenient data save, weight cor-

TABLE 4 : GM (1, 1) inspection table of the model

Year	Raw Value	Model Value	Residual	Relative Error	Class Ratio Deviation
23rd	6390	6390.0	0	0	
24th	7291	7135.8	155.2	0.0213	0.1355
25th	7044	7038.6	5.4	0.0008	-0.0210
26th	6780	6942.7	-162.7	0.0240	-0.0248
27th	6584	6848.2	-264.2	0.0401	-0.0157
28th	6952	6754.9	197.1	0.0284	0.0658
29th	6733	6662.9	70.1	0.0104	-0.0185

recting algorithm of the BP neural network prediction model is a kind of back propagation algorithm. And models coming from it can approximate a function in any precision. On the basis of the scores of the Women's Heptathlon champions of the 23rd —30th Olympics, we build up a BP neural network prediction model

The confirmation of the training sample of the BP neural network prediction model

In this paper, the predication model of BP neural network forecasts The Olympic Games women's heptathlon champion scores still in accordance with time series. The model of BP neural network is built up by taking the scores of the first three Olympic Games women's heptathlon champion as an input and taking the fourth Olympic Games women's heptathlon champion scores as an output to build up the predication model of neural network as well, which means we take the 23rd ~25th Olympic Games women's heptathlon champion scores as an input and the 26th Olympic Games women's heptathlon champion scores as output. At the meanwhile, we take the 24th ~26th Olympic Games women's heptathlon champion scores as an input and take the 27th Olympic Games women's heptathlon champion scores as output and so on. We altogether establish 4 kinds of mapping relations to be the training samples of the predication model of neural network. At the same time, we take the 27th~29th Olympic Games scores as an input to forecast the 30th Olympic Games scores so that we can build up the samples of input matrix P:

$$P=\begin{bmatrix}6390 & 7291 & 7044 & 6780\\ 7291 & 7044 & 6780 & 6584\\ 7044 & 6780 & 6584 & 6952\end{bmatrix}$$

The established output matrix $T=[6780 \quad 6584 \quad 6952 \quad 6733]$.

Because of the model of BP neural network's own characteristics, we need to deal with the data to some degree and make every figure in input matrix and output matrix all locates between (0, 1), so that the predication model will have a better convergence effect in the training. The paper normalizes the statistical data, which means the established input matrix P and the established output matrix T can become the matrix P1:

$$P1=\begin{bmatrix}0.8764 & 1.0000 & 0.9661 & 0.9299\\ 1.0000 & 0.9661 & 0.9299 & 0.9030\\ 0.9661 & 0.9299 & 0.9030 & 0.9535\end{bmatrix}$$

output matrix $T1=[0.9299 \quad 0.9030 \quad 0.9535 \quad 0.9235]$.

The installation of the model of BP neural network parameter

The artificial model of BP neural network consists of output layer, input layer and hidden layer as well, among which the numbers of nerve cells of output layer and input layer are decided by the training sample. The output layer of the established predication model of BP neural network in this passage has three nerve cells and the input layer has a nerve cell according to the installation training sample. Supposing that the hidden layer of the established predication model of BP neural network is 1, the numbers of the nerve cells of the hidden layer can be calculated by the empirical formula $i=\sqrt{n+m}+a$, among which n is the numbers of nerve cells of the input layer, m is the numbers of nerve cells the output layer, a is between 0 and 1. We know from the passage that $n=3$, $m=1$, the model $a=2$. So that we can calculate the hidden layer has 4 nerve cells by the empirical formula. Meanwhile, the model chooses

FULL PAPER

S tangent function, transfer function, linear function, gradient descent with momentum BP well function, gradient descent with momentum weight learning function and mean square error of standardized performance function as the study function of model weights. The maximum number of training times of BP neural network is 5000. The training accuracy is 0.005, the display of interval training is 500 times.

The solution of the model of BP neural network

As for the first p raining mode, supposing that the sum of the inputs of the neural network unit j is a_{pj}, outputs id o_{pj}, then:

$$a_{pj} = \sum_{j=1}^{N} W_{ji} o_{pi}$$

$$o_{pj} = f(a_{pj}) = \frac{1}{1+e^{-a_{pj}}}$$

Meanwhile, if we set arbitrary numbers about the different layers neural net of the link weight, then to any training mode, when the deviations of the corresponding output o_{pj} of the input and the expecting output can not meet the qualification accuracy, we need to calculate the network error:

$$E = \sum_{p} E_p$$

$$E_p = \frac{1}{2}\sum_{i}(d_{pj} - o_{pj})$$

We carry out the weight training to the established model of BP neural network by using the setting training function. When the weighted value of the mode can not meet the qualification accuracy, its weighted value begins modifying as follow.

$$W_{ij}(t+1) = W_{ij}(t) + \eta E_p o_{pj} + a[W_{ij}(t) - W_{ij}(t-1)]$$

Among this, a is momentum factor, and the value in this passage is 0.5. We continuously amend the weight value of the established neural network model with the formula till it meets the requirements.

According to the degree of simulation of MATLAB, we get the solution of the prediction model. Figure 2 is the error analysis of curve of the predication model of BP neural network according to the MATLAB software. After nearly 1000 times of the network may meet the requirements of error. Among which the Y-axis is error accuracy and x-coordinate represents the learning cycle.

The solution of the model of BP neural network

According to the above established model of BP neural network, we forecast the 30th Olympic Games women's heptathlon champion scores. its input vector is (0.9030, 0.9535, 0.9235), and the output is 0.9556. Restoring the output, we can calculate that the 30th Olympic Games women's heptathlon champion scores is 6967. At the same time, emulating the 26~29th Olympic Games women's heptathlon champion scores according to the established model of BP neural network, their scores are 6063, 7152, 7351 and 6474. The re-

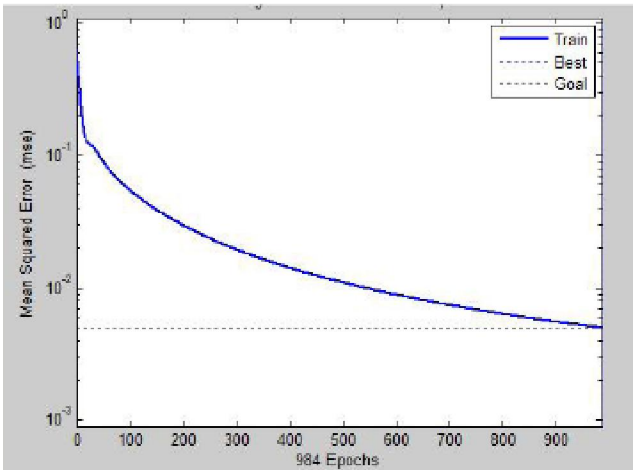

Figure 2 : The network training error curve

TABLE 5 : The comparison between the prediction results of grey model and BP neural network model

the Olympic Games		26th	27th	28th	29th	30th
Actual result		6780	6584	6952	6733	6955
Gray model	Predicted value	6942.7	6848.2	6754.9	6662.9	6572
	Fractiona-1 error	0.0240	0.0401	0.0284	0.0104	0.0550
Neural net	Predicted value	6063	7152	7351	6474	6967
	Fractiona-l error	0.1058	0.0863	0.0574	0.0385	0.0018

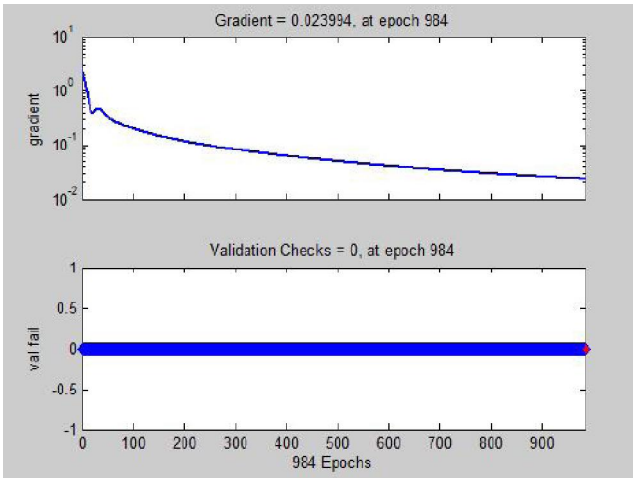

Figure 3 : The results of the model of BP neural network

sults of the model of BP neural network are as following:

CONCLUSION

According to the calculations above, we can get the simulation error of the 26~30th Olympic Games women's heptathlon champion scores between the GM (1, 1) gray prediction model and the predication model of BP neural network:

From the above, we can know that GM (1, 1) the gray predicted model and the predication model of BP neural network forecast the 30th Olympic Games women's heptathlon champion scores are 6572 and 6967 respectively. Their deviation are 5.5% and 0.18%, which means the predication model of BP neural network has the higher accuracy to forecast the 30th Olympic Games women's heptathlon champion scores. But comparing with the fractional error of fitted value of the two model of the 26th ~30th Olympic Games women's heptathlon champion scores, gray model can accurate the approximation of time series which is made up of the scores effectively, which means in the matter of forecasting the Olympic Games women's heptathlon champion scores, GM (1, 1) gray predicted model has better effect of prediction comparing with the model of BP neural network.

REFERENCES

[1] Chen Xue-mei; Comprehensive Evaluation and Recovery on Exercise Fatigue of Athlete[J]. Bulletin of Sport Science & Technology, **15(6)**, 22-23, 43 **(2007)**.

[2] Hu Qi; The Producing of the Fatigue and Recovery[J]. Journal of Physical Education Institute of Shanxi Teachers University, **19(1)**, 90-92 **(2004)**.

[3] Lan Zheng; The Analysis on the Main Influencing Factors of the Football Players' Running Ability[J]. Sport Science And Technology, **29(1)**, 58-60 **(2008)**.

[4] Wu Shao-ming, ZHOU Zhi-jie; The research on fatigue and recovery in sports training[J]. Journal of Wuhan Institute of Physical Education, **35(4)**, 93-94 **(2001)**.

[5] Zhong Da peng, ZHANG Xin yun; Occurrence and elimination of sports fatigue[J]. Journal of Wuhan Institute of Physical Education, **36(5)**, 50-51 **(2002)**.

[6] Zhao Zhong-hua; Research Progress of Neuro-muscular Fatigue in Sport[J]. Journal of Shenyang Sport University, **3**, 44-46 **(2003)**.