ISSN : 0974 - 7435 Volume 10 Issue 11

2014

BioTechnology

An Indian Journal

FULL PAPER

BTAIJ, 10(11), 2014 [5004-5012]

# Grey BP neural network-based women heptathlon performance prediction application

**Xiaoyan Han**

**Physical Education Department, Xi'an University of Architecture And Technology, Xi'an 710055, Shaanxi, (CHINA)**

## ABSTRACT

The paper makes statistics of world women heptathlon annual best performances during 2005～2013, takes prediction on heptathlon best performance in 2013 as research objects. According to statistical world women heptathlon annual best performance during 2005～2012, the paper establishes sports performance prediction model that is series combined by GM(1, 1) grey prediction model and BP neural network prediction model, and applies established model to predict world women heptathlon annual best performances in 2013. By comparing world women heptathlon annual best performance actual value with model predicted value in 2013, study and analyze grey BP neural network prediction model application in sports performance prediction problem. Research results show that for sports performance prediction problem, grey BP neural network prediction model's prediction accuracy is high, application is simple, and generalizability is strong, it has more advantages over single GM(1, 1) grey prediction model and BP neural network model by comparing.

## KEYWORDS

Sports performance; Women heptathlon; BP neural network, Prediction model.

## INTRODUCTION

Competitive sports performance prediction has great practical value, such prediction can provide training targets for athletes, and can also provide references for sports development research. Traditional prediction method mainly has time sequence method, analogy, regression analysis method and so on. Most of these aspects take lots of data that can be analyzed as basis, their research range also tends to limit in static problems. Sports performance prediction problems possess less data for analysis, data randomness is larger, influence factors are more, each factor mutual affects and relative complicated relationships these features, which lets traditional prediction method's prediction result on sports performance not to be ideal.

With constantly development of system science, lots of scholars carry on analysis and study on how to apply less data to make effective prediction on sports performance such problem. Among them, Yuan Jian-Guo (1992) applied grey model and posterior difference checking method to study individual refinement sports performance prediction problem, established men short track speed skating national record grey prediction model[1] ; Liu Jia-Jin and others （2005）applied mathematical statistics, documents literature and other methods to make statistics of Tong fang full text journal library's number of scientific research papers that applied grey theory to research sports performance prediction problems during 1994 to 2003 , and got a conclusion that grey theory application in Chinese sports performance prediction problems was still in the preliminary stage that provided references for its application in Chinese sports performance prediction[2] ; Wang Guo-Fan and others （2010）applied economics, fuzzy C mean value clustering analysis and other theories, they put forward economics and competitive sports strength discrepancy theory combinative improved prediction model, and verified established model's scientificity and feasibility[3] ; Zhang Yu and others（2013）took 23～30$^{th}$ Olympic Games men 200m and others five events competition champions performances as research objects, established BP neural network prediction model, verified its practicability and accuracy in sports performance prediction problems[4] ; Zhuang Chong（2006）analyzed and compared grey prediction model with neural network prediction model's precise on sports performance prediction, which provided reference for sports performance prediction methods selection[5].

The paper makes statistics of 2005～2013 world women heptathlon annual best performances, and on the basis of the data, established grey BP neural network prediction model. By analyzing established grey BP neural network prediction model, study its application in sports performance prediction problems.

## GREY BP NEURAL NETWORK MODEL

The paper established grey BP neural network model is series combined by GM (1, 1) grey prediction model and BP neural network prediction model that has both grey prediction model processing functions on small samples, poor data information and BP neural network model's robustness and fault tolerance as well as other features, it fully reflects sports grey information excavation's discipline intersectionality.

### BP neural network prediction model establishment steps

In artificial neural network model, BP neural network model is one of neural network models that people thoroughly research. According to statistics, BP neural network model application occupies 80%～90% of artificial neural network model. BP neural network model possess a three layers neural network structure that are input layer, hidden layer and output layer. Between layers, it adopts full joining ways; its joining weights carry out error back propagation and repeatedly learning training according to each training mode, till its weights meet each training mode requirement. BP neural network model can draw near any non-linear functions by existing training modes' training, it can fast,

accurate process with non-linear system problems, so is widely applied in functional approximation, prediction, system identification, classification and data compression and other fields.

BP neural network model establishment should complete following four steps：

（1）Research problems, elements refinement and training mode collection ；（2）According to problem, set BP neural network each parameter ；（3）Establish training samples, input training samples' input data into established BP neural network system, and compare system output data with expected mode, if it has errors, then execute（4）, otherwise, return back（3）；（4）By back propagation, carry on system each layer joining weights correction.

**GM (1, 1) grey prediction model establishment steps**

Grey system theory is on the basis of relevance space, smooth discrete functions and others, it defines grey derivative and grey differential equation, and then use discrete data sequence to establish differential equation dynamic model[6]. Grey prediction model applies corresponding grey system theory, fully excavates small samples data explicit information and hidden information, looks for each factor relations. Compare with traditional probability statistical method, grey prediction model has certain superiorities in handling with small samples and poor data information data problems. GM (1, 1) grey prediction model is simple and easily to understand, applied range is wide, prediction accuracy is high, is one of most widely used models in grey models.

Apply GM(1, 1) grey prediction model to solve small samples data prediction problem, firstly it needs to extract data from problems, establish original sequence, and then does smoothness test and handling with established original sequence, carries on grey modeling with original sequence or sequence after handling that conforms to requirements, finally test model predicted value. Its specific steps are as following:

(1) According to handled problems, establish corresponding original sequence:

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

(2) Test original sequence, original sequence ultimate ratio:

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

If original sequence ultimate ratio values all cover in the interval $(e^{-\frac{2}{n+1}}, e^{\frac{2}{n+2}})$, then original sequence meets smoothness requirements, it can directly establish prediction model, otherwise, it can do translation conversion handling.

(3) For original sequence, carry on AGO accumulated generating operation and mean value generating handling, respectively generate sequence $x^{(1)}$ and sequence $z^{(1)}$, and establish corresponding grey differential equation:

$$x^{(0)}(k) + az^{(1)}(k) = b, k = 2,3,\cdots,n$$

Among them, $a$ is model development coefficient, $b$ is model coordination coefficient.

(4) Set $u = (a,b)^T$, it can use least square method $\hat{u} = (a,b)^T = (B^T B)^{-1} B^T Y$ to solve development coefficient $a$ and coordination coefficient $b$. Among them $Y$ and $B$ are respectively :

$$Y = (x^{(0)}(2), x^{(0)}(3), \cdots, x^{(0)}(n))^T$$

$$B = \begin{bmatrix} -z^{(1)}(2) & -z^{(1)}(3) & \cdots & -z^{(1)}(n) \\ 1 & 1 & \cdots & 1 \end{bmatrix}^T$$

Finally, it can use grey differential equation to solve:

$$x^{(1)}(k+1) = (x^{(0)}(1) - \frac{b}{a})e^{-ak} + \frac{b}{a}$$

(5) For solved prediction model, it can carry out testing whether prediction model solved sequence and original sequence residual, absolute correlation degree as well as mean square error ratio conform to precise requirements.

**Grey BP neural network model**

In the paper, grey BP neural network prediction model adopts GM (1, 1)grey prediction model and BP neural network prediction model series combination forms, that is taking GM(1, 1) grey prediction model prediction result as input, actual performance as output, establishing sample data, carrying out weights training on established BP neural network. The paper makes statistics of year 2003～2013 world women heptathlon events annual best performances, takes performance in year 2003～2011 as basis, establishes grey BP neural network prediction model sample data, trains established neural network model. And meanwhile, apply established prediction model to predict year 2012 and 2013 two years' world women heptathlon annual best performances, make comparison of them with such two years' women heptathlon best performances actual values, and tests established prediction model's accuracy.

## GREY BP NEURAL NETWORK PREDICTION MODEL ESTABLISHMENT

Women heptathlon event seven individual events are mutual connected, their each event performance are mutual affected. That is when making performances prediction on them; it needs to consider each event relations. The paper established grey BP neural network prediction model, its BP neural network model fully researches each event performances mutual relations, its GM (1, 1)grey prediction model fully utilizes statistical data time sequence, and excavates its data time sequence relations.

**Women heptathlon annual best performance statistics**

Women heptathlon is composed of 100m hurdle, high jump, shot put, 200m, long jump, javelin throw and 800m seven sports events, it involves running, jumping and throwing events, is a comprehensive women event, which has very high requests on athletes' physical quality, psychological quality and technical levels. The paper makes statistics of year 2003～2013 world women heptathlon event annual best performances, their performances statistics are as TABLE 1.

**TABLE 1 : Year 2003～2013 world women heptathlon event annual best performance**

| Year | 100m hurdle | High jump | Shot put | 200m | Long jump | Javelin throw | 800m |
|---|---|---|---|---|---|---|---|
| 2003 | 13.18 | 1.94 | 14.19 | 22.98 | 6.68 | 49.90 | 132.12 |
| 2004 | 13.21 | 1.91 | 14.77 | 23.27 | 6.78 | 48.89 | 134.15 |
| 2005 | 12.62 | 1.91 | 12.61 | 24.12 | 6.78 | 53.07 | 134.66 |
| 2006 | 13.35 | 1.89 | 14.56 | 23.86 | 6.65 | 46.94 | 134.95 |
| 2007 | 13.15 | 1.95 | 14.81 | 23.38 | 6.85 | 47.98 | 132.98 |
| 2008 | 13.44 | 1.80 | 17.29 | 24.39 | 6.63 | 48.60 | 137.72 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2009 | 12.93 | 1.92 | 14.14 | 23.25 | 6.29 | 43.54 | 132.22 |
| 2010 | 12.95 | 1.89 | 14.05 | 23.21 | 6.43 | 46.71 | 130.18 |
| 2011 | 13.32 | 1.83 | 14.17 | 23.50 | 6.61 | 52.95 | 128.04 |
| 2012 | 12.54 | 1.86 | 14.28 | 22.83 | 6.48 | 47.49 | 128.65 |
| 2013 | 13.63 | 1.86 | 13.92 | 23.93 | 6.48 | 47.90 | 131.58 |

## Original data processing

When carrying on grey BP neural network prediction model training learning, it needs to make normalization processing with statistical women heptathlon total performance and each individual performance, let their samples values to be within interval$^{[0, 1]}$. That is selecting year 2003～2013women heptathlon annual best values' each event sports performance maximum value, and takes it as denominator, and regards its corresponding event other years' performances as numerator, handles with each performance, solved data after processing is as TABLE 2.

**TABLE 2 : Year 2003～2013 world women heptathlon event annual best performance normalization data**

| Year | 100m hurdle | High jump | Shot put | 200m | Long jump | Javelin throw | 800m |
|---|---|---|---|---|---|---|---|
| 2003 | 0.9670 | 0.9949 | 0.8207 | 0.9422 | 0.9752 | 0.9403 | 0.9593 |
| 2004 | 0.9692 | 0.9795 | 0.8587 | 0.9541 | 0.9898 | 0.9212 | 0.9741 |
| 2005 | 0.9259 | 0.9795 | 0.7293 | 0.9889 | 0.9898 | 1.0000 | 0.9778 |
| 2006 | 0.9795 | 0.9692 | 0.8421 | 0.9783 | 0.9708 | 0.8845 | 0.9799 |
| 2007 | 0.9648 | 1.0000 | 0.8566 | 0.9586 | 1.0000 | 0.9041 | 0.9656 |
| 2008 | 0.9861 | 0.9231 | 1.0000 | 1.0000 | 0.9679 | 0.9158 | 1.0000 |
| 2009 | 0.9486 | 0.9846 | 0.8178 | 0.9533 | 0.9182 | 0.8204 | 0.9601 |
| 2010 | 0.9501 | 0.9692 | 0.8126 | 0.9516 | 0.9387 | 0.8802 | 0.9453 |
| 2011 | 0.9773 | 0.9385 | 0.8195 | 0.9635 | 0.9650 | 0.9977 | 0.9300 |
| 2012 | 0.9200 | 0.9538 | 0.8259 | 0.9360 | 0.9460 | 0.8949 | 0.9341 |
| 2013 | 1.0000 | 0.9538 | 0.8051 | 0.9811 | 0.9460 | 0.9026 | 0.9554 |

The paper takes statistical each individual event performance as basis, establishes equivalent dimensions additional GM(1, 1) grey model on it that firstly respectively takes year 2003～2006 data to carry on GM(1, 1) grey modeling on each event, and predicts year 2007 performance corresponding normalization processing data. Then take year 2004～2007 data to predict year 2008 performance corresponding normalization data, by parity of reasoning, it gets year 2007～2011 performance corresponding normalization data predicted value. Solved prediction data is as TABLE 3.

**TABLE 3 : Year 2007～2013 world women heptathlon event annual best performances prediction data**

| Year | 100m hurdle | High jump | Shot put | 200m | Long jump | Javelin throw | 800m |
|---|---|---|---|---|---|---|---|
| 2007 | 0.9687 | 0.9658 | 0.7958 | 0.9980 | 0.9647 | 0.9003 | 0.9831 |
| 2008 | 0.9958 | 1.0037 | 0.9419 | 0.9454 | 0.9972 | 0.8353 | 0.9623 |
| 2009 | 0.9835 | 0.9197 | 1.0728 | 1.0011 | 0.9767 | 0.9331 | 1.0023 |
| 2010 | 0.9506 | 0.9536 | 0.8555 | 0.9654 | 0.8832 | 0.8011 | 0.9698 |
| 2011 | 0.9259 | 1.0053 | 0.6994 | 0.9205 | 0.9124 | 0.8361 | 0.9148 |
| 2012 | 0.9878 | 0.9190 | 0.8183 | 0.9664 | 0.9884 | 1.0928 | 0.9154 |
| 2013 | 0.9198 | 0.9384 | 0.8372 | 0.9350 | 0.9572 | 0.9385 | 0.9253 |

**Training samples establishment**

In the paper, grey BP neural network prediction model takes GM(1, 1) grey prediction model's women heptathlon seven individual events performances' predicted values as input, and actual performances as output. In training samples, there are seven samples that are respectively year 2007～2011 grey prediction model predication data and actual data as input and output. Set its input and output matrix respectively as P and T, then it has:

$$P=\begin{bmatrix}0.9687 & 0.9658 & 0.7958 & 0.9980 & 0.9647 & 0.9003 & 0.9831\\ 0.9958 & 1.0037 & 0.9419 & 0.9454 & 0.9972 & 0.8353 & 0.9623\\ 0.9835 & 0.9197 & 1.0728 & 1.0011 & 0.9767 & 0.9331 & 1.0023\\ 0.9506 & 0.9536 & 0.8555 & 0.9654 & 0.8832 & 0.8011 & 0.9698\\ 0.9259 & 1.0053 & 0.6994 & 0.9205 & 0.9124 & 0.8361 & 0.9148\end{bmatrix}^T$$

$$T=\begin{bmatrix}0.9648 & 1.0000 & 0.8566 & 0.9586 & 1.0000 & 0.9401 & 0.9656\\ 0.9861 & 0.9231 & 1.0000 & 1.0000 & 0.9679 & 0.9158 & 1.0000\\ 0.9486 & 0.9846 & 0.8178 & 0.9533 & 0.9182 & 0.8204 & 0.9601\\ 0.9501 & 0.9692 & 0.8126 & 0.9516 & 0.9387 & 0.8802 & 0.9453\\ 0.9773 & 0.9385 & 0.8195 & 0.9635 & 0.9650 & 0.9977 & 0.9300\end{bmatrix}^T$$

**BP neural network prediction model establishment**

In the paper's neural network prediction model, hidden layer is one layer. By above, it is clear that neural network has seven inputs and seven outputs, that the input layer and output layer's nerve cells numbers are both seven. Among them, hidden layer nerve cells numbers research can apply empirical formula：

$$i=\sqrt{n+m}+a \qquad (1)$$

Among them, $n$ is number of input layer nerve cells, $m$ is number of output layer nerve cells, $a$ value is between 0and 1， by above empirical formula, the model takes hidden nerve cells numbers as 7.

BP neural network model's transfer function, training function, learning function and performance function are various, as transfer function has S type logarithmic function 、S type tangent function and pure linear function and so on, training function has BFGS quasi-Newton BP algorithm function、 gradient BP descent algorithm function and gradient descent momentum BP algorithm function and so on, learning function has gradient descent weights learning function, gradient descent momentum weights learning function and so on, performance function has mean square error performance function and mean square error normalization performance function and so on. The BP neural network model selects S type tangent function transfer function and pure linear function, gradient descent momentum BP gain function, gradient descent momentum weight learning function and mean square error normalization performance function. And meanwhile, this BP neural network model selected maximum training times are 10000times, training accuracy is 0.005, training indication interval is 500times.

## MODEL SOLUTION AND ANALYSIS

Apply MATLAB software as simulation platform and setup training samples and training parameters, carry out network weights training on above established grey BP neural network prediction model. By repeatedly training BP neural network, finally it trains BP neural network weights, establishes corresponding grey BP neural network prediction model. Apply test samples data to test established BP neural network then can get that established sample is accurate.

## Grey BP neural network training performance analysis

Apply established sample data, BP neural network training parameters to carry out weights training on established BP neural network. In order to analyze established grey BP neural network training process and model accuracy, the paper studies grey BP neural network training process and training results.

Figure 1 is model training process curve figure; Figure 2 is model training result curve figure. In Figure1, horizontal coordinate is BP neural network training times, vertical coordinate is BP neural network training precise. It is clear that BP neural network model weights training has circulated 1518 times, and its precise has arrived at 0.00492.

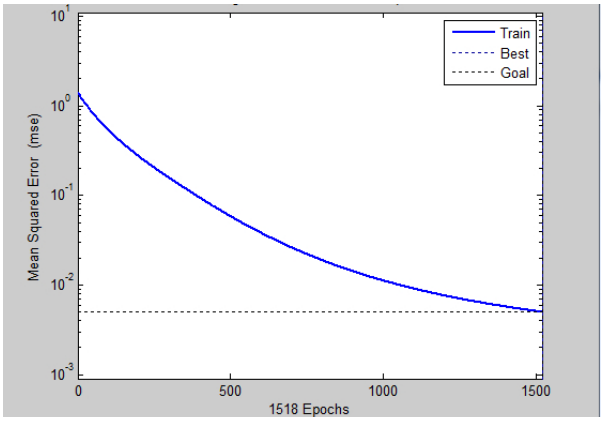


**Figure 1 : Training process curve**

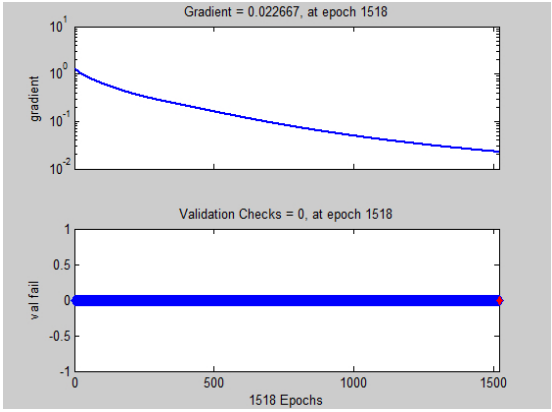


**Figure 2 : Training result curve**

## Grey BP neural network precise test

The paper takes year 2012 and 2013 data as test samples, use them to test established grey BP neural network prediction model accuracy. Grey BP neural network prediction model input is annual best values of year 2012 and 2013 world women heptathlon obtained by applying GM(1, 1) grey prediction model. By comparing grey BP neural network prediction model output values with year 2012

and 2013 world women heptathlon annual best performances, study and analyze grey BP neural network prediction model prediction precise. Set year 2012 and 2013 GM(1, 1) grey model prediction data that are grey BP neural network inputs are respectively P1and P2, grey BP neural network outputs are respectively T1 and T2, then it has:

$$P1 = [0.9878 \quad 0.9190 \quad 0.8183 \quad 0.9664 \quad 0.9884 \quad 1.0928 \quad 0.9154]^T$$

$$P2 = [0.9198 \quad 0.9384 \quad 0.8372 \quad 0.9350 \quad 0.9572 \quad 0.9385 \quad 0.9253]^T$$

Take P1and P2 as outputs, it gets simulation value:

$$T1 = [0.9221 \quad 0.9481 \quad 0.8267 \quad 0.9411 \quad 0.9411 \quad 0.8992 \quad 0.9377]^T$$

$$T2 = [0.9238 \quad 0.9327 \quad 0.8401 \quad 0.9355 \quad 0.9591 \quad 0.9450 \quad 0.9268]^T$$

Compare women heptathlon best performances normalization data simulation value and actual value errors with their actual values, as TABLE 4 shows.

**TABLE 4 : Comparison between predicted value and actual value**

| | | 100m hurdle | High jump | Shot put | 200m | Long jump | Javelin throw | 800m |
|---|---|---|---|---|---|---|---|---|
| 2012 | Actual value | 0.9200 | 0.9538 | 0.8259 | 0.9360 | 0.9460 | 0.8949 | 0.9341 |
| | Prediction deviation | 0.0021 | -0.0057 | 0.0008 | 0.0051 | 0.0019 | 0.0043 | 0.0036 |
| | Deviation ratio | 0.23% | 0.60% | 0.10% | 0.54% | 0.21% | 0.48% | 0.39% |
| 2013 | Actual value | 0.9198 | 0.9384 | 0.8372 | 0.9350 | 0.9572 | 0.9385 | 0.9253 |
| | Prediction deviation | 0.0040 | -0.0057 | 0.0029 | 0.0005 | 0.0019 | 0.0065 | 0.0015 |
| | Deviation ratio | 0.44% | 0.61% | 0.03% | 0.06% | 0.19% | 0.69% | 0.02% |

By TABLE 4, it can get that apply grey BP neural network model to predict year 2012 and 2013 world women heptathlon each event performance, the predicted values deviation ratios are less than 0.7%. In order to compare grey BP neural network with GM (1, 1) grey prediction model prediction precise on world women heptathlon, the paper establishes GM (1, 1) grey prediction model, its predicted values are as TABLE 5.

**TABLE 5 : Grey model predicted value**

| Year | 100m hurdle | High jump | Shot put | 200m | Long jump | Javelin throw | 800m |
|---|---|---|---|---|---|---|---|
| 2012 | 0.9648 | 0.9484 | 0.8543 | 0.9004 | 0.9347 | 0.9021 | 0.9405 |
| 2013 | 0.9697 | 0.9441 | 0.8571 | 0.9586 | 0.9276 | 0.8992 | 0.9349 |

By comparing established grey BP neural network prediction model simulation value with GM (1, 1) grey prediction model's predicted value, it is clear that grey BP neural network prediction model precise is far higher than GM(1, 1) grey model's precise.

Restore grey BP neural network prediction model's year 2012 and 2013 world women heptathlon simulation value, it can solve model predicted data values on year 2012 and 2013 world women heptathlon performance as TABLE 6 shows.

**TABLE 6 : World women heptathlon annual best performance predicted value**

| Year | 100m hurdle | High jump | Shot put | 200m | Long jump | Javelin throw | 800m |
|---|---|---|---|---|---|---|---|
| 2012 | 13.75 | 1.68 | 15.26 | 21.19 | 6.90 | 47.49 | 128.64 |
| 2013 | 13.10 | 1.72 | 15.32 | 22.93 | 6.62 | 53.33 | 129.50 |

## CONCLUSION

For women heptathlon and other comprehensive sports events, their each event individual sport are mutual affected, sports performances also have certain mapping relations. In order to better predict each event sports performances, it should make full use of each event known data. Grey BP neural network model has both BP neural network prediction model robustness, error tolerance and other features, and can make full use of GM(1, 1) grey prediction model to excavate data time sequence relations, for multiple factors prediction problems, its predicted result precise is not second to GM(1, 1) grey prediction model.

## REFERENCES

**[1]** 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)**.

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

**[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]** 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)**.

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

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