ISSN : 0974 - 7435 Volume 8 Issue 1

BioTechnology

An Indian Journal

FULL PAPER

BTAIJ, 8(1), 2013 [44-49]

The application of multiple linear regressions in swimming fatigue degree assessment

Qicheng Wei*, Xiaona Zhu

Institute of Physical Education, Zhoukou Normal University, Zhoukou 466001, Henan, (CHINA)

E-mail: tiyuxi@qq.com

ABSTRACT

Fatigue degree assessment has a significant role in guiding for training strategies. This paper collected heart rate data during the eight weeks of training and subjective questionnaire data of the fatigue degree of 31 swimmers. Among them heart rate data is composed of five data: the heart rate value when getting up early in the morning, the heart rate value during the training preparation, the heart rate value when intensity of training is the largest, the heart rate value when finishing the training and the heart rate value when finishing the recovering. Subjective questionnaire data of the fatigue degree is composed of three data: the negative evaluation value, the physical strength overdraft value and the reduced value of confidence degree. Through heart rate data of the entire process, explore its relationship with the three items' data of the fatigue degree, and establish a multiple linear regression model. After data test, model has good fitting degree and significance. It is reasonable to use the model to evaluate the fatigue degree of athletes.

KEYWORDS

Training;
Regression model;
The degree of fatigue;
Functional recovery.

INTRODUCTION

We know that the human body will feel fatigue after some movement, through the rest the fatigue will be reduced as well as recover to the initial state or even exceed the original good state, this is why the recovery of body functions can improve the performance of athletes. The sport of swimming is a physical leading sport. In the training process of this movement, the athletes can easily feel fatigue. If athletes' fatigue degree is very high and reaches excessive fatigue, it will have an impact for the recovery of physical functions, which not only will not help improve performance, but also hurt the body healthy. In the Fifth International sports biochemistry meeting 1982, the definition of physical functions fatigue: physiological processes of the body cannot continue its function in a particular level or cannot maintain its intended exercise intensity. In competitive swimming training, the load capacity on athletes is large and continues for a long time. This kind of training sometimes makes athletes reach excessive fatigue, causes harm to mind and body. Moderate fatigue contributes to the improvement of athletic ability.

Many people have made efforts on research of sports fatigue and recovery, such as: Wu Shaoming and

Zhou Zhijie, from Sports Department of Wuhan University, in paper "Research on sports fatigue and recovery", expounded the relationship between fatigue and recovery, explained the influences to improve training results, explained and analyzed from the perspective of qualitative; Zhao Quan, from Pingdingshan Teachers College in paper "the assessment and recovery of the degree of fatigue in the movement", restated this issue with the results of the references. The former one is qualitative analysis with difficult operability, while the latter one just defined and restated some conceptual problems with poor operability too.

By analyzing the data of heart rate in five periods, this paper obtains three indexes of fatigue of body functions, and the operability is good. The use of multiple linear regression method well explained the process, and the goodness of fitting and significance are very good.

RESEARCH OBJECT AND METHOD

Research object

For certain Sports College of sports school, there are a total of 31 people of fourth grades students in swimming major, including eight national master, thirteen national first-level athletes and twelve national secondary athletes. The average age is 23.years old. Ranking in accordance with the level, the first 16 are a group and the last 15 are a group. Study the relation of the heart rate value when getting up early in the morning, the heart rate value during the training preparation, the heart rate value when intensity of training is the largest, the heart rate value when finishing the training, the heart rate value when finishing the recovering, the negative evaluation value of fatigue degree, the physical strength overdraft value and the reduced value of confidence degree in the eight weeks of training, as shown in TABLE 1:

TABLE 1: Data list of the two groups' athletes

Group	Weeks Number	the heart rate value when getting up early in the morning	the heart rate value during the training preparation	the heart rate value when intensity of training is the largest	the heart rate value when finishing the training	the heart rate value when finishing the recovering	the negative evaluation value	the physical strength overdraft value	the reduced value of confidence degree
the average value of 16 people of outstanding group	1	61.2	79.5	165.8	140	88	8.17	7.42	-0.67
	2	60.7	79.1	166.1	141.3	87.8	7.42	6.92	-1.08
	3	61.5	80.3	168.8	141.8	87.7	7.83	7.42	-0.67
	4	60.9	79.7	166.8	141.7	87	7.5	7.08	-1.17
	5	64.6	84.8	175.1	151.7	94.1	11.12	9.96	1.16
	6	63.3	84	175.2	152	72.9	9.96	9.46	0.81
	7	63.9	85.1	187.3	152.4	92.4	10	9.96	1.16
	8	63.4	84.2	176	152	91.8	9.65	9.29	0.36
the average value of the last 15 people	1	63.5	80	166.2	141.8	92.8	11.64	10.09	-0.67
	2	63.5	79.7	167	142.3	92.5	11.45	9.91	-1.08
	3	64.3	80.8	169.9	143.7	93.3	11.82	10.36	-0.67
	4	63.4	80.3	168.5	144.1	92.9	11.82	10.26	-1.17
	5	67.1	84.1	175.7	155.8	98.1	14.23	13.2	1.16
	6	67	85.6	175.9	155.6	97.4	14.84	12.89	0.81
	7	67.6	84.5	188.8	156.4	98	14.04	13.13	1.16
	8	66.8	83.7	177.9	155.7	97.4	17.87	12.68	0.36

Note: The units of heart rate value is (beats / min)

Research method

1) Document literature

Through CNKI and other data base, download related research papers, mathematical statistics tutorials and statistics courseware, which makes adequate preparation for the correctness of the theory and calculation.

FULL PAPER

2) Mathematical statistics

Use EXCEL spreadsheet soft and mathematical integrated computing system-FORLAB to conduct processing and computing on the statistical data.

THE MODEL ESTABLISHMENT OF MUL TIPLE LINEAR REGRESSIONS

Multiple linear regressions

In practical problems, the factors affecting a thing are often more than one or development trend influencing factors and things is not a linear relationship. To identify the functional relationship between these factors and things, multiple linear regression is a good method. This paper studies representative heart rate value of five stages of a cycle in swimming. So regard these five values as five variables, namely five-element linear regression model. If the random variable y is related to $p(p \geq 2)$ numbers of ordinary variable $x_1, x_2, x_3, \cdots, x_p$, and satisfies equation (1), equation (1) is the mathematical description way of the multiple linear regression.

$$\begin{cases} y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_p x_p + \varepsilon \\ E(\varepsilon) = 0, Var(\varepsilon) = \delta^2 < +\infty \end{cases} \quad (1)$$

Equation (1) is call p meta-theory linear regression model; $\beta_0, \beta_1, \beta_2, \beta_3, \cdots, \beta_p$ are the regression coefficients; $x_1, x_2, x_3, \cdots, x_p$ are called the regression factors or design factor, factors for short. Parameters $\beta_i (i = 1,2,3,\cdots, p)$ reflect the influence of factors $x_i (i = 1,2,3,\cdots, p)$ to the observation y, sois also known as the effects of factor.

Suppose there are n groups of sample observations $(x_{i1}, x_{i2}, x_{i3}, \cdots, x_{ip}; y_i)(i = 1,2,3,\cdots, n)$ which are not all the same, obtaining formula (2) by formula (1):

$$\begin{cases} y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_p x_{ip} + \varepsilon_i \\ E(\varepsilon_i) = 0, Var(\varepsilon_i) = \delta^2 < +\infty \end{cases} \quad (2)$$

Wherein $i = 1,2,3,\cdots, n$, moreover $\varepsilon_1, \varepsilon_2, \varepsilon_3, \cdots, \varepsilon_n$ is independent to each other.

The matrix expression of the formula (2) is shown in formula (3):

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix}, \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}, \begin{cases} Y = X\beta + \varepsilon \\ E(\varepsilon) = 0, Var(\varepsilon) = \delta^2 I_n \end{cases} \quad (3)$$

Whereinis I_n the n-order unit matrix, 1_n is the n-dimensional column vector that the elements are all 1. Y is called the observation vector of random variables, β is the unknown parameter vector, is design matrix, X is the -dimensional random error vector. In the regression analysis, generally assume rank(X)=P+1, namely require that X is the column satisfaction. So it has $E(Y) = X\beta, Var(\varepsilon) = \delta^2 I_n$. Generally take $\varepsilon \sim N(0, \delta^2)$, so the expression form of formula (3) is shown as formula (4) below:

$$\begin{cases} Y = X\beta + \varepsilon \\ \varepsilon \sim N(0, \delta^2 I_n) \end{cases} \quad (4)$$

Regression coefficient calculation method

Using the least squares method to solve the estimated vector $\hat{\beta}$ of the regression coefficients column vector β in model (4), the regression coefficients error sum of squares is shown in equation (5):

$$Q(\beta_0, \beta_1, \beta_2, \cdots, \beta_p) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - \beta_0 x_{i2} - \cdots - \beta_p x_{ip})^2 \quad (5)$$

Based on the idea of the least squares method, choose the group with the smallest error sum of square as the parameters estimation of the regression coefficients, that is to say, if $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_p$ exists and it satisfies the equation (6):

$$Q\left(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_p\right) = \min\{Q(\beta_0, \beta_1, \beta_2, \cdots, \beta_p)\} \quad (6)$$

$\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_p$ is the least squares estimation of parameters .

Regression effect and regression coefficient significance test

The linear relationship closeness of y and $x_1, x_2, x_3, \cdots, x_p$ is related to the proportion of regression sum of square in the total sum of squares. Wherein the expression of regression sum of squares U and the total sum of squares L_{yy} is shown in formula (7) below:

$$\begin{cases} L_{yy} = Q + U \\ L_{yy} = \sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2 \\ Q = \sum_{i=1}^{n} \left(y_i - \hat{y}\right)^2 \\ U = \sum_{i=1}^{n} \left(\hat{y}_i - \bar{y}\right)^2 \end{cases} \quad (7)$$

$R=\sqrt{\frac{U}{L_{yy}}}$ is called the sample multiple correlation coefficient y of $x_1, x_2, x_3, \cdots, x_p$ and, denoting the goodness of fitting. In the practical application of the multiple linear regressions, the negative correlation coefficient indicates the good and bad of the fitting degree of regression equation to the sample data, and its range is $0 \le R^2 \le 1$. The closer to 1 its value is, the higher the goodness of fitting becomes. But when the number of independent variables and sample size is close to each other, R is vulnerable to approach 1. At this moment the good and bad of the model is determined by R which should be carefully considered.

When the regression effects are significant, it can explains that $\beta_0, \beta_1, \beta_2, \cdots, \beta_p$ are not all 0. But one or $m(m<p)$ numbers of β_i are zero, which means that y and x_i are irrelevant or that the role of x_i is replaced by $x_j (j \ne i)$. Thus x_i should be removed from the regression equation

Fatigue feasibility calculations

The regression equation obtained by achievements sample $(x_{i1}, x_{i2}, x_{i3}, \cdots, x_{ip}; y_i)(i=1,2,\cdots,n)$ is shown in formula (8) below:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3 + \cdots + \hat{\beta}_p x_p \quad \textbf{(8)}$$

After test the regression effect and regression coefficient are both significant. When given a set of values $(x_{01}, x_{02}, x_{03}, \cdots, x_{0p})$, it is reasonable to take $\hat{y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_{01} + \hat{\beta}_2 x_{02} + \hat{\beta}_3 x_{03} + \cdots + \hat{\beta}_p x_{0p}$ as the point estimation of $E(y_0)$. Because according to formula (4), $\hat{y}$ is the unbiased estimation of $E(y_0)$.

RESULT ANALYSIS

Changes of the fatigue index

According to the data in TABLE 1, draw scattered trend chart, as shown in Figure 1.

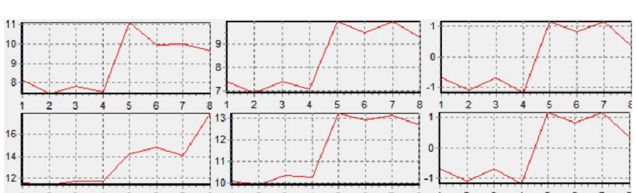

Figure 1 : Line chart of two groups' fatigue index

The upper left in Figure 1 is the negative evaluation value of first sixteen people in outstanding group; the upper middle is the physical strength exhaustion value of first sixteen people in outstanding group; the upper right is the reduced value of confidence degree of first sixteen people in outstanding group; the lower left is the negative evaluation value of last fifteen people; the lower middle is the physical strength exhaustion value of last fifteen people; the lower right is the reduced value of confidence degree of last fifteen people.

Variable data

Use vector $X_1 = (x_{1,1} \ x_{1,2} \ \ldots \ x_{1,16})$ to present the heart rate value when getting up early in the morning.

Use vector $X_2 = (x_{2,1} \ x_{2,2} \ \cdots \ x_{2,16})$ to present the heart rate value during the training preparation.

Use vector $X_3 = (x_{3,1} \ x_{3,2} \ \cdots \ x_{3,16})$ to present the heart rate value when intensity of training is the largest.

Use vector $X_4 = (x_{4,1} \ x_{4,2} \ \cdots \ x_{4,16})$ to present the heart rate value when finishing the training.

Use vector $X_5 = (x_{5,1} \ x_{5,2} \ \cdots \ x_{5,16})$ to present the heart rate value when finishing the recovering.

Use vector $Y_1 = (y_{1,1} \ y_{1,2} \ \cdots \ y_{1,16})$ to present the negative evaluation value.

Use vector $Y_2 = (y_{2,1} \ y_{2,2} \ \cdots \ y_{2,16})$ to present the physical strength overdraft value.

Use vector $Y_3 = (y_{3,1} \ y_{3,2} \ \cdots \ y_{3,16})$ to present the reduced value of confidence degree.

The establishment of regression equation

In accordance with the least squares method, the data in TABLE 1 and the variable data in 3.2, establish the regression equation of Y_1 and $X_1, \ldots, X_5$, as shown in formula (9) below:

$$\hat{Y}_1 = -43.4409 + 1.3277 \hat{X}_1 - 0.5698 \hat{X}_2 - 0.0923 \hat{X}_3 + 0.2149 \hat{X}_4 + 0.0088 \hat{X}_5 \quad \textbf{(9)}$$

Establish the regression equation of Y_2 and $X_1, \ldots, X_5$, as shown in formula (10) below:

$$\hat{Y}_2 = -46.4417 + 1.0666 \hat{X}_1 - 0.1054 \hat{X}_2 - 0.0079 \hat{X}_3 - 0.0025 \hat{X}_4 + 0.0144 \hat{X}_5 \quad \textbf{(10)}$$

Establish the regression equation Y_3 of and $X_1, \ldots, X_5$, as shown in formula (11) below:

$$\hat{Y}_3 = -30.7866 + 0.0464 \hat{X}_1 + 0.3284 \hat{X}_2 + 0.0215 \hat{X}_3 - 0.0108 \hat{X}_4 - 0.0144 \hat{X}_5 \quad \textbf{(11)}$$

Model test

The statistical parameters of model (9) (10) (11) is shown in TABLE 2:

TABLE 2 : List of model parameters

Category	The residual sum of squares	The sum of deviation squares	Multiple correlation coefficient	Variance
Model(9)	11.6173	129.911	0.9542	1.9678
Model(10)	1.0845	67.9092	0.992	0.6012
Model(11)	1.0735	13.8175	0.9604	0.5982

The goodness of fitting and significant is good. The model can make good use of heart rate value in five training period, to determine the fatigue degree of the athletes.

Evaluation of the fatigue degree

The index evaluation of the fatigue degree of model (9) (10) (11) is shown as formula (12) below:

$$\begin{cases} Z^* = \sum_{i=1}^{3} Y_i \\ Z^* \in [13,16],\ Z^* = Z_1^* \\ Z^* \in [16,23],\ Z^* = Z_2^* \\ Z^* \in [23,+\infty],\ Z^* = Z_3^* \end{cases} \quad (12)$$

In Formula (12) Z^* is a concentrated expression of the degree of fatigue, the greater its value is, the greater the fatigue degree becomes; Z_1^* represents that the training process is of the best fatigue degree, Z_2^* represents that the training process is of strong fatigue degree, Z_3^* represents that the training process is of the excessive fatigue, less than Z_1^* represents that the training process is still lack of the fatigue and can also increase the load amount.

In summary, the changes in the human heart rate can reflect the fatigue degree of physical function. In the whole motor process, we can use the five heart rate values to measure the fatigue degree of the body, making the appropriate response for the training intensity adjustment.

CONCLUSIONS

Five-element linear regression model well explains the degree of fatigue during exercise; the regression model established has high goodness of fitting, very good significance and very high accuracy; the intensity adjustment of swimming training can be determined based on the evaluation level of fatigue degree introduced in this model.

ACKNOWLEDGEMENTS

"And the concept of athletic ability, approach - avoidance goal orientation related cognitive evaluation model," Henan Social Science Planning Project, Project Number: 2011FTY011.

REFERENCES

[1] Chen Bi-jiang; Occurrence and elimination of sports fatigue. Examination Weekly, **24**, 98-99 **(2007)**.
[2] Chen Jing, Tian Fu-jun; Research progress on sports fatigue and recovery. Journal of Tonghua Normal University, **30(8)**, 82-86 **(2009)**.
[3] Chen Wu-san, Study on the Intensity of Warm-down Exercise after Loading and the Blood Lactate Obliteration in Swimming. Sport Science, **18(5)**, 58-60 **(1998)**.
[4] Dai Yan; Research on Mechanism of the Sports Fatigue and Recovery Measures. Journal of Shenyang Institute of Physical Education, **23(3)**, 333-335 **(2004)**.
[5] Feng Yan; Emotion and Coping Way of Female Swimming during Intensive Training. China Sport Science and Technology, **40(4)**, 27-29 **(2004)**.
[6] Feng Yan; Psychological and Physiological Fatigue of the Swimming Athletes During Intensive Training and Their Coping Ways. Journal of Tianjin Institute of Physical Education, **19(3)**, 54-56 **(2004)**.
[7] Fu Ke-cui; On the Outline of the Exercise Fatigue. Sichuan Sports Science, **(6)**,30-33 **(2004)**.
[8] Guo Zhe-bin, Miao Miao; Effects of the supplement of carbohydrate-electrolyte drinks on body fluid balance and sports capacity during hypoxic training. Journal of Physical Education, **15(3)**, 108-122 **(2008)**.
[9] Hou Ya-ping, Liu Cheng-yi, Liu Song-hao; Cellular Rehabilitation Mechanism of Photobiomodulation. Acta Laser Biology Sinica, **17(4)**, 559-564 **(2008)**.
[10] Wang Hui-Min, Zhang Xin-Jiang, Tian Fu-Jun; Some Doubts on Research of Sport Fatigue. Bulletin of Sport Science & Technology, **18(4)**, 124-126

(2010).

[11] Wang Rong-shen; The Research of Athletics Fatigue and the Recovering Process. Journal Of Physical Education, **4**, 119-120 **(1997)**.

[12] Yuan Yuan; How to understand the fatigue in swimming training. Swimming, **20(4)**, 14-15 **(2001)**.

[13] Zhang Jian-chun; Recovery from the Fatigue after Track and Field Training. Journal of Pengcheng Occupation University, **19(5)**, 86-87 **(2004)**.

[14] Zhang Yin; Exploring and Analysis of Resuming Means after Swimming Trainings. Journal of Jiaying University, **20(3)**, 114-115 **(2002)**.

[15] Zhao Jia-gui; Body recovery issues research on university athletic training. Journal Of Anhui Sports Science, **3**, 37-39 **(1998)**.

[16] Zhao Yuan-ji; Review of Causes of Gymnasts' Psychological Fatigue and Its Rehabilitation Methods. Journal of Beijing University of Physical Education, **27(10)**, 1345-1347 **(2004)**.