Journal of Theoretical and Applied Information Technology
31st March 2013. Vol. 49 No.3
© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

THE CORRELATION DEGREE ANALYSIS BETWEEN THE BASKETBALL TEAM ACHIEVEMENT AND TECHNICAL INDEX BASED ON THE GRAY SYSTEM THEORY

LIYA MA
Zhejiang Pharmaceutical College, Ningbo 315100, Zhejiang, China
E-mail: zb58@163.com

ABSTRACT

In this paper, the gray system theory is used to establish the comprehensive evaluation model. We take a team as an example and analyze it, and then the analysis and solution method is promoted to other teams. In the end, the correlative relationship between the technical index of each team and the achievement is solved. Firstly, the index data of each team is carried on statistical treatment and standardized processing (dimensionless method) and the margin list of each index is solved. The comprehensive evaluation model is established on the gray system theory and the relation grade between the technical index and achievement of each index is obtained by solving of Excel.

Keywords: *Gray, Correlation Degree, Standardization, Technical Index*

1. INTRODUCTION

The performance of the basketball player in the competition process is one of the main factors to decide the competition result. The on- the –spot technique statistics data of the basketball competition is not only the quantitative index to measure the level of player's skill development, but also the objective criterion to judge the competition result of a team[1-3]. At present, the main methods of studying the technical index of the player are: literature consultation; mathematical statistics method; the collected data is input the computer and the statistical software package is used and analyzed the result by mathematical statistics method; Comparative analysis, the correlation coefficient comparison is used to study and analyze the various technical index of the players in the competition [4-8].

In the paper, a basketball game in the university is taken as the research object. The game involves 12 college representative teams, which are divided into two groups. Each group has six teams. In each group, a team will battle with other teams in the same group, that is, each team will have five games, and then five groups of data of the same type will be produced[9,10]. We can take a team as an example, analyze and solve all the data, and then promote the analysis and solution methods to each team, and finally acquired the results of problem, which is convenient for us to analyze the data[11-13].

The specific statistics of data is made. As for any matches between two games, the performances of each player is given: playing time, two-score, three-scored, penalty shot, goal attempts, hit rate, offensive rebounds, ,defensive rebounds, total rebounds, assists, foul, turnover, stealing, blocking and scoring. We call the statistical data as the technical index of each team. As for large quantities of the data, we need to select the most valuable data by adopting the gray system theory.

2. HYPOTHESIS AND NOUN DEFINITION

2.1. Hypothesis

(1)In all the matches, both sides must try them all out, which means not cheating in the match, in other words, each competition result reflects the comparison of both sides' real strength.

(2) As for each team, the overall condition of each index should be considered, but the performance of each player should not be considered.

(3)Each basketball team is a system.

2.2. Definition

(1)Integral Rate: the team scoring in each match divided by the total scores of both sides.

(2)Total Integral Rate: The integral sum rate of five groups.

(3) Win rate: the integral average rate.

© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

(4) Losing weight: If A group and B group have different strength, the winning ratio is different.

2.3 The Sign and Variable

(1) $x_i\left(i=1,2,\cdots,17\right)$: various technical index of the basketball team(many factors of the system);

(2) $x_0=\left\{x_0(k)\middle|k=1,2,\cdots,5\right\}$: The competition results of the five games of each team, which is regarded as the comparison datum;

(3) $x_i=\left\{x_i(k)\middle|k=1,2,\cdots,5\right\}(i=1,2,\cdots,17)$: the ith technique statistics in the kth matches of the team;

(4) $\xi_i(k)$: The correlation coefficient between technical statistics and comparison datum. The index reflects the correlation degree between compared sequence and reference sequence at a certain time. And

$$\xi_i(k)=\frac{\min\limits_i\min\limits_k\left|x_0(k)-x_i(k)\right|+\rho\times\max\limits_i\max\limits_k\left|x_0(k)-x_i(k)\right|}{\left|x_0(k)-x_i(k)\right|+\rho\times\max\limits_i\max\limits_k\left|x_0(k)-x_i(k)\right|}.$$

(5) $r_i=\frac{1}{n}\sum_{k=1}^{n}\xi_i(k)$: Correlation degree refers to the correlation degree between technical index and comparison datum, which is the only index to measure the correlation degree between the compared sequence and the reference sequence.

3. THE ESTABLISHMENT AND SOLUTION OF MODEL

According to analysis of the problems, we can take the Information College as research object and analyze:

Information College is a representative team in the second group. There are five competitions and each competition gives a group of on-the-spot technique statistics data. We put these data together so as to find out the relationship between them. We give the number of the five competitions respectively. (The number of the competitions can be changed)

As for the on-the-spot technique statistics data recorded in each match, we accumulate and sum up the performance of all the teams of each index, and the results are regarded as the technique parameters of the index of each team. Thus, five groups of data are obtained as follows:

Table 1：Information College Table

G	1	2	3	4	5
scoring	78	98	98	94	100
2-point shot(FG)	14	14	20	15	27
2-point shot	30	25	31	27	40
2-PointFG(S)	46.67	56	64.52	55.56	67.5
3-point shot(FG)	6	16	13	13	10
3-point shot	20	27	24	27	28
3-point FG(S)	30	59.26	54.17	48.15	35.71
Free Throw	32	22	19	25	16
Free Throw(FG)	50	35	26	31	24
FT%	64	62.86	73.08	80.65	66.67
Rebound(offensive)	18	8	10	9	9
Rebound(defensive)	29	28	18	23	22
Rebound	47	36	28	32	31
Assist	12	16	17	7	13
Foul	20	34	32	27	14
Turnovers	16	19	15	21	9
Steal	5	12	9	4	8
Block shot	3	5	4	2	3

© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

According to the Table 1, our task is to find out the interrelations between each index and the score of the competition. However, it seems that the relationship between the each data or statistical data and the final scores is not obvious. In order to find out the interrelationship, we should draw on these data so as to find out whether there is a little regularity or not.

As can be seen from the Figure 1, if the curve geometry is closer, the change of slope is closer, the correlation degree is greater. Due to the curve of the scoring and the 2-point FG(s) is the closed, it can be concluded that the interrelationship between the 2-point FG(s) and scoring is the greatest. However, the close curve can be seen and several indexes which are related to the scoring can be obtained. The interrelationship between all the indexes and the scoring can not be obtained, either the correlative relationship of quantization can not be obtained, and the issue can not be solved fundamentally.

Based on this kind of situation, we adopt the gray system method to analyze the data so as to get the correlative relationship between each index and scoring. Above all, we should give the definition of the correlation coefficient, and then give the correlation degree—the only index to measure the correlation degree between each index and the scoring.

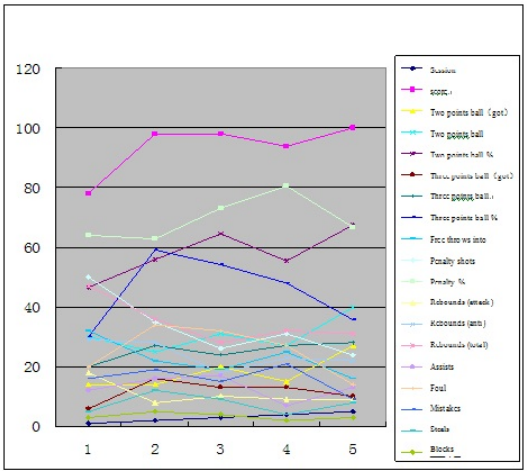

Figure 1: Information College Analysis

Take $x_0 = \{x_0(k) | k = 1,2,\cdots,5\}$ as the competition result of the fifth match of the basketball team, and we regard it as the comparison datum, x_0 can be expressed as the sequence (benchmark sequence):

$$x_0 = \{x_0(k) | k = 1,2,\cdots,5\} = (x_0(1), x_0(2), x_0(3), x_0(4), x_0(5))$$

Among them, k expressed games, and $x_0(k)$ is the statistics value which x_0 obtained in the K game. Take $x_i (i = 1,2,\cdots,17)$ as various technical indexes of the team (many factors of the gray system), and we need to compare it with the competition result directly. First of all, the compare sequence is constructed:

$$x_i = \{x_i(k) | k = 1,2,\cdots,5\} = (x_i(1), x_i(2), x_i(3), x_i(4), x_i(5)) \quad (i = 1,2,\cdots,17)$$

Then, the definition of the correlation coefficient of the compared sequence x_i is the benchmark sequence x_0 in the k:

$$\xi_i(k) = \frac{\min\limits_i \min\limits_k |x_0(k) - x_i(k)| + \rho \max\limits_i \max\limits_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \max\limits_i \max\limits_k |x_0(k) - x_i(k)|}$$

Among them $\rho \in [0,+\infty)$ is called the resolution coefficient, $\min\limits_i \min\limits_k |x_0(k) - x_i(k)|$ and $\max\limits_i \max\limits_k |x_0(k) - x_i(k)|$ is called the minimum difference between two poles and the maximum difference between two poles. In general, the resolution coefficient ρ is 0.5.

The correlation coefficient expresses the correlation degree between the compared sequence and the benchmark sequence in a certain games. But the correlation coefficient in each game is too scattered to be compared comprehensively. Thus, the correlation degree between the compared sequence x_i and the benchmark sequence x_0 is defined as: $r_i = \frac{1}{n}\sum_{k=1}^{n} \xi_i(k)$, which is the only index to measure the correlation degree among indexes in the system.

© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

As can be seen from the expression of the correlation degree, the correlation coefficient of each games are integrated to an average and realize the concentration of the scattered information in order to deal with on the whole.

Look at the Table 1 again, due to the different dimension of the data in the Table1 and the dimension is required to keep consistent in the calculation process, we need to carry on the initial processing of these data. Mainly the dimensionless form of all the data is obtained and the public intersection is obtained among all the sequences so as to be compared conveniently. We adopt the following methods:

Suppose the original sequence is: $X = (x(1), x(2), \cdots, x(n))$ can construct the initialization sequence:

$$\overline{X} = \left(1, \frac{x(2)}{x(1)}, \frac{x(3)}{x(1)}, \cdots, \frac{x(n)}{x(1)}\right)$$

, then the initialization sequence is satisfied to the requirement of the dimensionless form, and if all the original sequence can be constructed the initialization sequence, there must be the public intersection point 1. Thus, the table was formed by the initialization sequence is obtained, just as the follow Table 2:

Table 2: Information College initialization Sequence

G	1	2	3	4	5
scoring	1	1.25641	1.25641	1.205128	1.282051
2-point shot(FG)	1	1	1.428571	1.071429	1.928571
2-point shot	1	0.833333	1.033333	0.9	1.333333
2-PointFG(S)	1	1.199914	1.382473	1.190486	1.446325
3-point shot(FG)	1	2.666667	2.166667	2.166667	1.666667
3-point shot	1	1.35	1.2	1.35	1.4
3-point FG(S)	1	1.975333	1.805667	1.605	1.190333
Free Throw	1	0.6875	0.59375	0.78125	0.5
Free Throw(FG)	1	0.7	0.52	0.62	0.48
FT%	1	0.982188	1.141875	1.260156	1.041719
Rebound(offensive)	1	0.444444	0.555556	0.5	0.5
Rebound(defensive)	1	0.965517	0.62069	0.793104	0.758621
Rebound	1	0.765958	0.595745	0.680851	0.659574
Assist	1	1.333333	1.416667	0.583333	1.083333
Foul	1	1.7	1.6	1.35	0.7
Turnovers	1	1.1875	0.9375	1.3125	0.5625
Steal	1	2.4	1.8	0.8	1.6
Block shot	1	1.666667	1.333333	0.666667	1

To calculate the correlation degree, we should solve $|x_0 - x_i(k)|$ and the minimum difference between two polar and the maximum difference between the two polar. Firstly, we solve $|x_0 - x_i(k)|$, which is shown in the Table 3:

Table 3: $|x_0 - x_i(k)|$ table

		1	2	3	4	5
$\|x_0 - x_i(k)\|$	1	0	0.25641	0.172161	0.133699	0.64652
	2	0	0.423077	0.223077	0.305128	0.0513
	3	0	0.0565	0.126063	0.014642	0.164274
	4	0	1.410257	0.910257	0.961539	0.384616

© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5	0	0.0936	0.0564	0.144872	0.117949
6	0	0.718923	0.549257	0.399872	0.0917
7	0	0.56891	0.66266	0.423878	0.782051
8	0	0.55641	0.73641	0.585128	0.802051
9	0	0.274223	0.114535	0.0550	0.240332
10	0	0.811966	0.700854	0.705128	0.782051
11	0	0.290893	0.63572	0.412024	0.52343
12	0	0.490453	0.660665	0.524277	0.622477
13	0	0.0769	0.160257	0.621795	0.198718
14	0	0.44359	0.34359	0.144872	0.582051
15	0	0.06891	0.31891	0.107372	0.719551
16	0	1.14359	0.54359	0.405128	0.317949
17	0	0.410257	0.0769	0.538461	0.282051

And:

$$\min_i \min_k |x_0(k) - x_i(k)| = 0$$

$$\max_i \max_k |x_0(k) - x_i(k)| = 1.14359$$

So,

$$\xi_i(k) = \frac{0.5 \times 1.14359}{|x_0(k) - x_i(k)| + 0.5 \times 01.14359}$$

,

and from :

$$r_i = \frac{1}{5}\sum_{i=1}^{5} \xi_i(k)$$

, the correlation degree between each technique indexes and the competition results are:

$r_1 = 0.77988$

$r_2 = 0.80297$

$r_3 = 0.91297$

$r_4 = 0.568$

$r_5 = 0.899$

$r_6 = 0.71606$

$r_7 = 0.63354$

$r_8 = 0.61249$

$r_9 = 0.85074$

$r_{10} = 0.58809$

$r_{11} = 0.68779$

$r_{12} = 0.64215$

$r_{13} = 0.8056$

$r_{14} = 0.73272$

$r_{15} = 0.79247$

$r_{16} = 0.65408$

$r_{17} = 0.76302$

So far, we obtain the result of the Information College.

Obviously, the above computing process is too troubled, and we need to calculate the correlation degree between each technique index and the competition results in the 12 colleges so that the computational complexity is too big, thus, we consider to use the computer to solve.

Due to the initialization sequence is solved in the computing process, the minimum difference between two polar must be 0, that is: $\min_i \min_k |x_0(k) - x_i(k)| = 0$.

Thus, the formula of the correlation coefficient can be simplified as:

© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$\xi_i(k) = \frac{0.5 \times \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + 0.5 \times \max_i \max_k |x_0(k) - x_i(k)|}$$

The formula of the correlation degree is:

$$r_i = \frac{1}{n} \sum_{k=1}^{n} \xi_i(k)$$

The above two formulas are the main formulas in the program.

According to the program, firstly, the results of the information college that are verified are the same as the calculating results, which can be expressed the workability of the program. Then we use the program to carry on the statistics and solving of the other eleven colleges to obtain the correlation degree between the technique indexes of each team and the competition result, as the Table 4.

Now look back and discuss further on the correlation degree between each technique index of the Information College and the competition results. According to the calculating correlation degree, we can find that the 2-point FG(s) is related to the final competition result and its correlation degree is 0.91297; the second is 3-point shot, and its correlation degree is 0.899; the third is the free throw percent, and its correlation degree is 0.85074. Then followed by: Assist, 2-point shot, Turnovers and so on.

In order to prove that ranking of correlation degree is consistent with the similarity of the graph, we draw the line graph according to the correlation degree between the several main factors and the final achievement, as shown in the Figure 2. From the Figure 2, we can see the consistency between each index and the competition result is great.

However, some figures in the original data are large, some are small, and they are quite different, so the consistency between them can not be seen obviously. Thus, we draw the same graph according to the corresponding initialization sequence, as shown in the Figure 3.

In addition, as can be seen in the graph in the Figure 3, the negative correlation is observed between the free throw percentage and the competition result, that is, when the curve of the competition result rises, the curve of the free throw percentage falls; when the curve of the competition result falls, the curve of the free throw percentage falls. It is contrary to the common sense. We consider that it caused by the accidental factors of the data, that is, if the data quantity is large enough, the positive correlation is observed between them.

As for the negative correlation, we have more concrete analysis in the further discussion of the model.

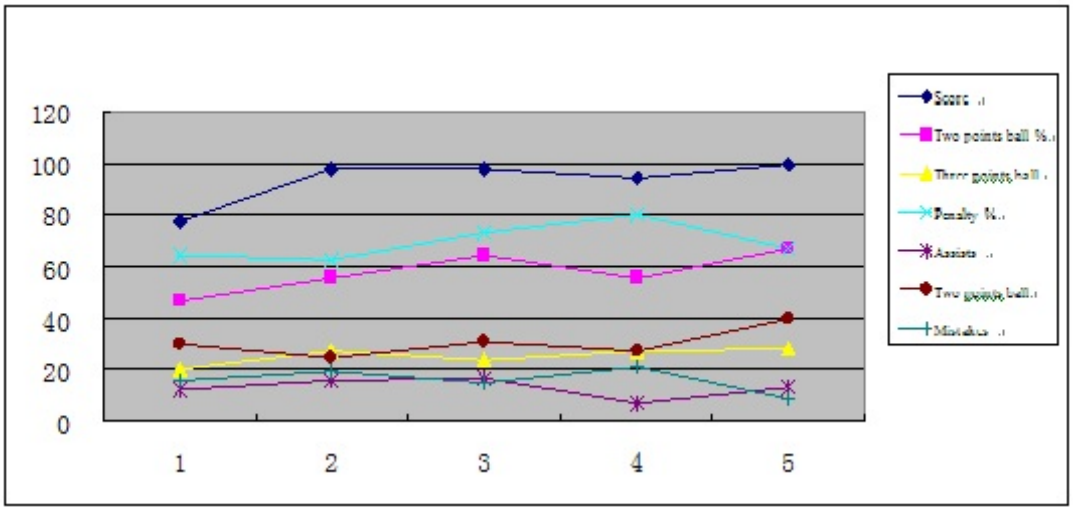

Figure 2: Correlation Degree Analysis

Journal of Theoretical and Applied Information Technology
31st March 2013. Vol. 49 No.3
© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

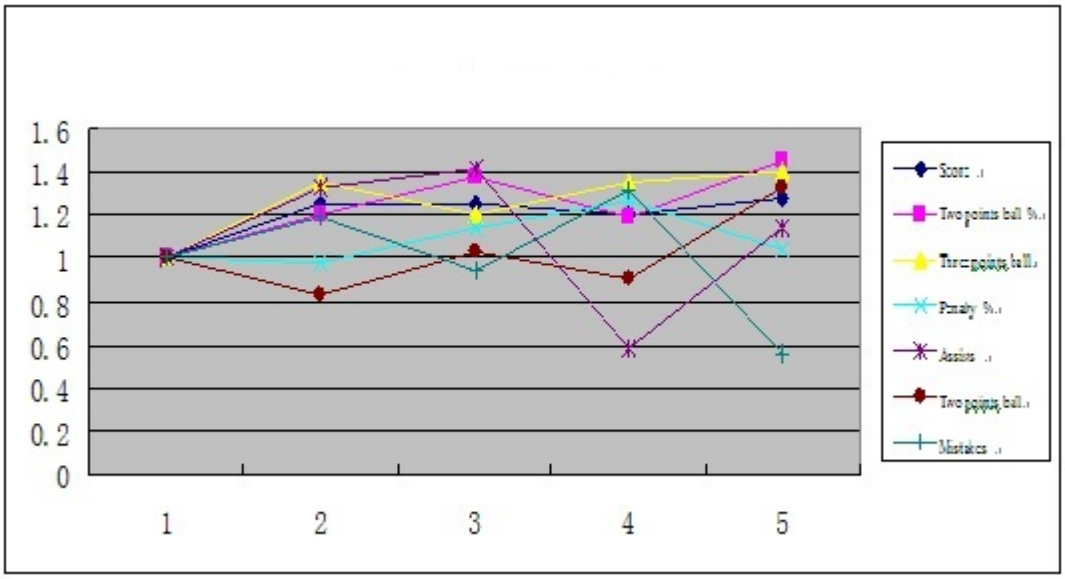

Figure 3: Initialization Sequence Correlation Degree Analysis

According to the analysis, the correlation degree between various technique index of basketball team and scoring is obtained.

Table 4: Correlation Degree between the technique index and the scoring

	2-point shot (FG)	2-point shot	2-point shot (FGs)	3-point shot (FG)	3-point shot	3-point shot (FGs)	Free Throw (FG)	Free Throw	Free Throw (FGs)
Correlation degree	0.763	0.819	0.867	0.792	0.855	0.821	0.708	0.702	0.917
	Rebound Offense	Rebound Defense	rebound	Assist	Foul	Turnovers	Steal	Block shot	
Correlation degree	0.765	0.657	0.734	0.792	0.797	0.697	0.732	0.768	

4. CONCLUSIONS

(1) The comprehensive evaluation model of the gray system is compatible to the data given in the title, and the correlation analysis method is adopted according to the similarity of the development and the dissimilarity among the factors can measure the correlation degree among the factors. Thus it can be promoted to evaluation of the unrelated index and the different dimensional index.

(2) For example: it can be promoted to the football match, the other various competitions and building index evaluation. The comprehensive evaluation model of the comprehensive index method can be promoted to the comprehensive evaluation in the other industries, such as worker's performance ranking, the comprehensive raking of each month in the hospital.

(3) The concept of the ratio is introduced and promoted to the other various competitions and evaluations. In the daily life and during the process of the competition, the same scores or achievements are often occurred. In that case, the ratio is considered to reduce the situation where there is the same scores and where it is difficult to rank.

(4)In accordance with the average compared solution method can be applied to the other aspects of evaluation, and the graph can be used to express which index are lower than the average, then it can be concluded that which indexes need to be strengthen. This method can be adapted to learning, work and all kinds of competitions.

© 2005 - 2013 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REFERENCES:

[1] Ning Dehuang, "Gray System GM(1,1) Model and its Application to Market Forecasting", *Journal of KunMing University of Science and Technology*,Vol.03, 2012, pp.56-63.

[2] Wang Zhongxian, Wu Chundu, "The Establishment and Application of a Improved GM(1,1) Model", *Mathematics In Practice and Theory*, Vol.09, 2003, pp.89-95.

[3] Li Zuoyong , Zhang Ming, Deng Xinming, "Optimal Gray GM(1,1) Model Based on Genetic Algorithm and Its Verification", *Systems Engineering-theory and Practice*,Vol.08, 2002, pp.45-52.

[4] Liu Changhong, Chen Qiu, Shen Mengyue, "A Kind of Simulation Method for Data Processing in Fatigue Test", *Machine Design*, Vol.06, 2000, pp.78-83.

[5] Wen Pinren, Chen Xiaohong, "The Establishment of the gray model of the GDP Index", *China Science and Technology Information*, Vol.17, 2005, pp.89-93.

[6] Wang wei, Jiang Yan, Zhao Yong, "Preliminary Study of Mine Tremor Prediction in Terms ofGray System Theory", *Engineering of surveying and Mapping*, Vol.01, 2007, pp.56-63.

[7] Guo Hongsheng, Teng Jinlong, Ma Xiaoguang, "Application of the GM(1, 1) Model in Building Deformation Prediction", *Dong Bei Surveying and Mapping*, Vol.01, 2011, pp.78-83.

[8] Zhou Ming, Wang Hongfa, "Gray Model GM(1,1) Based on Genetic Algorithm", *Journal of Nan Chang University(Natural Science)*, Vol.04, 2002, pp.58-63.

[9] Huang Weisong, Ji Peirong,Hu Xiangyong, "Experimental Study on Error Features of Gray Model GM(1,1)", *Journal of university Hydraulic and Electric Engineering/Yi Chang*, Vol.01, 2000, pp.85-89.

[10] Wang Jinzhu, "Study on Estimation of Time-varying Parameter of Gray Model", *Journal of Qing Dao University Engineering and Technology Edition*, Vol.01, 2000, pp.88-94.

[11] Robert Vallee, "Cybernetics and Systems, From Past to Future Precedings of the Joint Conference of WOSC and IIGSS", *Pittsburgh, Pennsylvania*, Vol.8, 2002, pp.24-26.

[12] Wang Xuliang, Nie Hong, "Miners rule co nsidering the fuzziness of stress damage under differ ent load sequences", *Tr ansactions of Nanjing Univ ersity of Aeronautics & Astronautics*, Vol.25, No.3, 2008, pp.182-186.

[13] Nie Hong, "Biaxial stress fatigue life pr edict ion by the local st rain method", *Inter national Journal of Fatigue*, Vol.19, No.6, 1997, pp.517-522.