Quest Journals
Journal of Research in Applied Mathematics
Volume 7 ~ Issue 9 (2021) pp: 34-42
ISSN(Online) : 2394-0743 ISSN (Print): 2394-0735
www.questjournals.org

Research Paper

An Analysis of400 Meter Final at IAAF World Athletics Championships 2019 Based on the Stepwise Regression Method

XiongJiangyuan[a]Huangqiaoxian[a]
[a]*Schoolof physical education & Sports science, South China normal university, China*

ABSTRACT:
The results of the men's 400m final at IAAF World Athletics Championships DOHA2019 were analyzed with the help of excel software by means of literature, video analysis and mathematical statistics. Through the acquisition of 50m results of men 400m athletes in each segment of the world athletics championships in 2019, a correlation study was conducted with 50m as the independent variable, the optimal indicator was screened out, and a regression model of 400m was established for regression analysis. It analyzes which section of 50m to 400m sports performance is the most significant, and finds a more effective training method, providing theoretical basis for China's sprint coaches and athletes in the future 400m training. The results showed that the 400 meters score is closely related to the 50m score in the eighth paragraph (x8), 50m score in the sixth paragraph (x6) and 50m score in the first paragraph (x1). The regression equation is: $\hat{y} = 131.456 + 1.895x_8 + 1.252x_6 - 21.197x_1$.

HIGHLIGHTS:
1.This article researches the performances of the eight athletes in the men's 400m final of the 2019 World Athletics Championships, trying to show the development trend of the 400m event, so as to accurately grasp the 400m training direction, and learn from the effective training methods abroad.
2.The 50m is an important predictor of the 400-metre race performance.

KEY WORDS: *athletics; 400 meters; 50meters; stepwise regression; segmentation*

Received 01 September, 2021; Revised: 12 September, 2021; Accepted 14 September, 2021 © The author(s) 2021. Published with open access at www.questjournals.org

I. INTRODUCTION

Track and field competition has a long history in the international sports arena, and track and field is also one of the most common sports. Therefore, track and field sports have the reputation of "the mother of sports". Track and field has always been in a high position in the long history of history. Among them, 400 meters is one of the most exciting and attractive track and field events. The 400-meter run not only requires the athletes to have excellent overall qualities, but also requires the participants to pay attention to the rationality of the speed distribution, the speed of reaction at the start, the consistency of the movements, and the proficiency of the curve running skills. Therefore, studying the speed distribution of the world's high-level 400-meter athletes has important reference significance for 400-meter running training. The World Athletics Championships are held every two years. Except for the Olympic Games, it is the most advanced track and field event in the world. This article researches the performances of the eight athletes in the men's 400m final of the 2019 World Athletics Championships, trying to show the development trend of the 400m event, so as to accurately grasp the 400m training direction, and learn from the effective training methods abroad. Girish and Darryl(2020) Providea theoretical basis for improving the performance of men's 400-meter race.

Sprint running is the shortest distance and the fastest speed in the track and field competition. The human organs and internal organs complete the extreme intensity periodic exercise under a large amount of hypoxia. The official sprint competitions include 100m, 200m, 400m, etc. In world-class track and field competitions, sprinting is an influential and widely loved sport among sports enthusiasts. It has a history of more than 100 years since the Athens Olympics were held in 1896. Sibel andAyşegül(2015) divide the sprint and training load of 400-meter athletes into 100-meter and 200-meter load, and their respective average and standard deviation values are calculated as 100 meters (11.96±0.17), 200 meters (25.0±0.38), 400 meters (56.07±0.91) and 400 meters hurdles (60.38±1.88). Liang Xia(2012)analyzed the current situation and existing problems of men's 400-meter hurdles, and explored the countermeasures to improve men's 400-meter hurdles. The scholar proposed to monitor the athlete's segmented speed to achieve the desired result.

Santos et al.(2020) connected the 400m and 100m performances and found that the top 10 athletes in the IAAF 100m and 400m rankings reached the personal record (PR) competition time Roughly the same. The research results of Krzysztof et al. (2010) found that the asymmetry of step length of sprinters of different groups exists, and the impact on the speed drop is significantly related to the shorter step length, while the rhythm change is not large. That is to say,for the 400-meter project, the matching relationship between the 200-meter stride and speed in each section will affect the 400-meter performance. Ralph et al. (1985) studied the sports variables of the gold medal, silver medalist and eighth runners in the men's 200m sprint finals in the 1984 Summer Olympics and found that the 50m and 100m results directly affect the speed of the 200m performance. 이장택(2010) think the best decathlon player performed relatively poorly in the 1500-meter race, but performed well in the long jump, 400-meter and 110-meter hurdles. This shows that the speed-driven 400-meter and 110-meter hurdles have a certain degree of influence.Roghaye et al. (2012) analyzed 16 young male athletes (age 20.58±3.25 years old, height 175.33±2.48 cm, body mass index 21.57±2.68 kg/m2), these subjects were taking sodium bicarbonate or placebo (carbonic acid). Calcium hydrogen) supplemented the balance sequence and participated in two 400-meter runs after one hour. It was found that after taking sodium bicarbonate or calcium bicarbonate, the last 100-meter sprint of the second 400-meter run can prevent blood sprinting during high-intensity exercise. The reduction in HCO-3. This article also uses a segmented method to observe the effect of sodium bicarbonate or calcium bicarbonate on the 400-meter performance.

In order to improve the competitiveness of the 400-meter event, the athletes' advantages in sprinting must be brought into play during the training process. However, most of the literature on the 400-meter full segment research focuses on the 400-meter research on the basis of 100-meter and 200-meter, and only a small number of scholars analyze the relationship between 100-meter, 200-meter and 400-meter sports performance. From the perspective of 50 meters to explore the impact of 400-meter sports performance, such as: Sibel (2015) in "Women's 400-meter hurdle athletes training load and its distribution" proposed that the study of the impact of the load of each segment on performance Important role. However, the research results of this article are about the relationship between the optimal running degree and the training load. It is found that there is only a significant relationship between the 50-60 meter sprint and the 150-450 meter sprint. Therefore, studying the characteristics of the 50-meter speed changes of the world's high-level athletes in the whole course, and discovering the most significant segments with the 400-meter performance, so as to guide the 400-meter training, which has reference significance for improving the 400-meter performance. In order to improve the performance of the 400-meter athletes, this article takes the results of the men’s 400-meter final at the 2019 World Athletics Championships as the research object based on the combination of sprint advantages. Step by step regression analysis, analyze which 50m and 400m performances are significantly related, establish a regression model to provide more scientific guidance for 400m training.

II. METHODS

2 Subjects

The model group was established for the eight athletes in the men's 400m final of the 2019 Doha World Championships in Men's 400m. The model group includes the second road BLOOM FIELD and the third road DEMISH GAYE, Fourth Road STEVEN GARDINER, Fifth RoadFRED KERLEY, Sixth Road MACHEL CEDENIO, Seventh Road KIRANI JAMES, the eighth roadANTHONY JOSE ZAMBRANO, the ninth road KORIR.

2.2 Proposal

Download the video of the men's 400m final of the 2019 World Athletics Championships on the official website of the CCTV5 Sports Events Channel, collect data according to the distance of each section of 50 meters, and combine with the IAAF official website about the 2019 Doha World Athletics Championships men's 400m final. Statistics provide data support for this article.

2.3Statistical analysis

With the help of SPSS software, the collected data are classified and processed, statistically analyzed, and mathematical support is provided for this article. The score data of the control group are expressed as mean±standard deviation, and two independent samples' t-test analysis is used. If $P<0.05$, it indicates a significant difference, and $P<0.01$ indicates a highly significant difference.

III. RESULTS AND ANALYSIS

3.1 Analysis of independent variables related to the 400-meter performance

3.1.1 Independent variable selection

Most scholars choose the independent variables that affect the 400-meter performance to be 100 meters, 200 meters, and 300 meters. Fewer scholars choose 50 meters as the independent variable for analysis to explore the impact of each section of 50 meters and 400 meters. After reading the relevant literature, I learned that the sprint performance has a considerable advantage in the first 30-50 meters. Therefore, this article divides the performance of the eight athletes in the men's 400m final of the 2019 World Athletics Championships into eight separate 50m as independent variables for regression analysis.

3.1.2 Data Sources

The data comes from relevant statistics on the men's 400m final of the 2019 Doha World Athletics Championships on the official website of the IAAF. Each pass represents each athlete. The 400-meter performance is represented by y/s, the first 50-meter performance is represented by x1/s, the second 50-meter performance is represented by x2/s, and the third 50-meter performance is represented by x3/s. In the fourth section, the 50-meter performance is represented by x4/s, the fifth section is represented by x5/s, the sixth section is represented by x6/s, and the seventh section is represented by x7/s. The 50-meter score in the eighth segment is expressed as x8/s. As shown in Table 1.

Table 1 Regression independent variables of 400 meter performance

Number	Track2	Track3	Track4	Track5	Track6	Track7	Track8	Track9
x_1/s	5.04	5.03	5.03	5.02	5.02	5.03	5.02	5.03
x_2/s	5.04	5.04	5.04	5.03	5.03	5.02	5.03	5.02
x_3/s	4.99	4.99	4.99	5.00	4.96	5.00	5.01	5.01
x_4/s	4.99	5.00	5.00	4.99	4.97	5.00	5.01	5.02
x_5/s	5.85	5.60	5.50	5.52	6.06	5.70	5.50	5.60
x_6/s	6.10	5.62	5.46	5.48	5.99	5.57	5.48	6.00
x_7/s	6.45	6.50	6.24	6.63	6.57	6.54	6.65	6.57
x_8/s	6.90	6.68	6.22	6.50	6.70	6.68	6.45	6.69
y/s	45.36	44.46	43.48	44.17	45.30	44.54	44.15	44.94

Table 2 400-meter correlation coefficient matrix R

Number	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	y
0-50m	1	0.303	0.095	0.204	0.104	0.430	-0.564	0.446	0.267
51-100m		1	-0.362	-0.316	-0.014	-0.065	-0.578	-0.143	-0.197
101-150m			1	0.886	-0.793	-0.363	0.237	-0.194	-0.374
151-200m				1	-0.734	-0.229	0.028	-0.216	-0.347
201-250m					1	0.740	0.080	0.659	0.820
251-300m						1	0.061	0.783	0.912**
301-350m							1	0.344	0.359
351-400m								1	0.926**
400m									1

**. When the confidence level (double test) is 0.01, the correlation is significant.

3.2 Two linear regression equation models to predict the 400-meter performance

From the simple correlation coefficient of 400 meters, see Table 2. The correlation coefficient between 50 meters and 400 meters in the eighth section is 0.926. When α=0.01, the degree of freedom n'=n-2=8-2=6, check t value table$t_{0.01/2}(6)=3.707 < |t_r| = 6.008$，$p < 0.01$.

The correlation coefficient between 50 meters and 400 meters in the sixth section is 0.912, when α=0.01, degrees of freedom n’=n-2=8-2=6, check the t value table$t_{0.01/2}(6)=3.707 < |t_r| = 5.446$，p＜

0.01.Therefore, when the confidence level is 0.01, the eighth segment 50 meters and the sixth segment 50 meters have a very significant linear correlation with 400 meters.

In conclusion, a one-variable linear regression equation with two independent variables x8 and x6 and the 400-meter performance of the dependent variable can be established, which can be compared with the regression equation to find the pros and cons, thereby establishing the optimal equation for predicting the 400-meter performance.

3.2.1The independent variable " eighth segment 50 meters (351-400m)" predicts the 400-meter unary linear regression equation

The unary linear regression equation for predicting the 400-meter performance of the dependent variable in the eighth segment of the independent variable is: $\hat{y} = 25.708 + 2.854x_8$ .Analysis of varianceF=36.068，$F > F_{0.01}(1,6)=13.7$，$P < 0.01$，The regression equation is very significant. In the eighth 50m, the remaining standard deviation of the 400-meter univariate linear regression equation for the 50-meter prediction is 0.259 seconds.As shown in Table 3.

Table 3 Variance analysis table of the eighth segment 50-meter prediction 400-meter unary linear regression equation

Variance source	Sum of squares	Degree of freedom	Mean square	F value	Probability judgment
Return	2.425	1	2.425	36.068	P＜0.01
Rest	0.403	6	0.067		
Sum	2.828	7			

3.2.2 The independent variable "sixth segment 50 meters (251-300m)" predicts the 400-meter linear regression equation

The unary linear regression equation for predicting the 400-meter performance of the dependent variable in the sixth paragraph of the independent variable is： $\hat{y} = 32.294 + 2.146x_6$ .Analysis of varianceF=29.501，$F > F_{0.01}(1,6)=13.7$，$P < 0.01$，The regression equation is very significant. In the sixth paragraph, the remaining standard deviation of the 400-meter univariate linear regression equation for the 50-meter prediction is 0.282 seconds.As shown in Table 4.

Table 4 Variance analysis table of the sixth segment 50-meter prediction 400-meter unary linear regression equation

Variance source	Sum of squares	Degree of freedom	Mean square	F value	Probability judgment
Return	2.350	1	2.350	29.501	P＜0.01
Rest	0.478	6	0.080		
Sum	2.828	7			

3.3 Stepwise regression analysis of independent variables of 400 meter performance

3.3.1The basic idea of gradual return

One of the methods in multiple regression analysis is stepwise regression analysis. Regression analysis is mainly used to explore the relationship between multiple variables. When a variable changes with another variable, the relationship between them can be expressed by an equation .Therefore, the stepwise regression analysis method can establish the optimal regression model, and then explore the relationship between various variables more deeply(Zhang,2020).

The principle of stepwise regression analysis is actually that when a dependent variable y is affected by multiple independent variables x1, x2, x3,..., xk, it is necessary to screen the independent variables, select the appropriate independent variable into the regression equation, and establish the regression equation When the first independent variable x1 is introduced, if the first independent variable x1 is significant, the influence of the independent variable x1 on the dependent variable y is retained; then the second variable x2 is introduced. If the second independent variable x2 is significant, then Keep the influence of the independent variable x2 on the dependent variable y; then introduce the third variable x3. If the third independent variable x3 is not significant, then remove the introduced variables until all the independent variables are introduced for testing. This purpose is to make the regression equation contain only significant independent variables. This process of introducing independent variables is repeated until the equation results neither eliminate insignificant variables nor introduce significant variables into the regression equation and stop. This allows the resulting regression model to better reflect the influence of the independent variable on the dependent variable, and to understand the correlation between the two variables.

(1) Pearson correlation using orthogonal screening method is also called simple correlation coefficient. Select the first independent variable into the equation, and calculate its partial regression sum of squares:

$$V_1^{(0)} = \left(r_{1y}^{(0)}\right)^2 \div r_{11}^{(0)} = (0.2670)^2 \div 1 = 0.0713$$

$$V_2^{(0)} = \left(r_{2y}^{(0)}\right)^2 \div r_{22}^{(0)} = (-0.1970)^2 \div 1 = 0.0388$$

$$V_3^{(0)} = \left(r_{3y}^{(0)}\right)^2 \div r_{33}^{(0)} = (-0.3740)^2 \div 1 = 0.1399$$

$$V_4^{(0)} = \left(r_{4y}^{(0)}\right)^2 \div r_{44}^{(0)} = (-0.3470)^2 \div 1 = 0.1204$$

$$V_5^{(0)} = \left(r_{5y}^{(0)}\right)^2 \div r_{55}^{(0)} = (0.8200)^2 \div 1 = 0.6724$$

$$V_6^{(0)} = \left(r_{6y}^{(0)}\right)^2 \div r_{66}^{(0)} = (0.9120)^2 \div 1 = 0.8318$$

$$V_7^{(0)} = \left(r_{7y}^{(0)}\right)^2 \div r_{77}^{(0)} = (0.3590)^2 \div 1 = 0.1289$$

$$V_8^{(0)} = \left(r_{8y}^{(0)}\right)^2 \div r_{88}^{(0)} = (0.9260)^2 \div 1 = 0.8575$$

The maximum parameter of $V_i^{(0)}$ is $V_8^{(0)} = 0.8575$ .Therefore, the independent variable x8 is introduced into the equation for F test，$F_{(0.05)}=6.59$，$F=(8-2)\times0.8575/(1-0.8575)=36.1053 > F_{\alpha}$，the independent variable x8 has a significant effect and is introduced into the regression equation. $R^{(0)}$ changed to $R^{(1)}$ by using Gauss-Jordan elimination method.

$$R^{(1)} = \begin{pmatrix} 0.8011 & 0.3668 & 0.1815 & 0.3003 & -0.1899 & 0.0808 & -0.7174 & -0.4460 & -0.1460 \\ 0.3668 & 0.9796 & -0.3897 & -0.3469 & 0.0802 & 0.0470 & -0.5288 & 0.1430 & -0.0646 \\ 0.1815 & -0.3879 & 0.9624 & 0.8441 & -0.6652 & -0.2111 & 0.3037 & 0.1940 & -0.1944 \\ 0.3003 & -0.3469 & 0.8441 & 0.9533 & -0.5917 & -0.0599 & 0.1023 & 0.2160 & -0.1470 \\ -0.1899 & 0.0802 & -0.6652 & -0.5917 & 0.5657 & 0.2240 & -0.1467 & -0.6590 & 0.2098 \\ 0.0808 & 0.0470 & -0.2111 & -0.0599 & 0.2240 & 0.3869 & -0.2084 & -0.7830 & 0.1869 \\ -0.7174 & -0.5288 & 0.3037 & 0.1023 & -0.1467 & -0.2084 & 0.8817 & -0.3440 & 0.0405 \\ 0.4460 & -0.1430 & -0.1940 & -0.2160 & 0.6590 & 0.7830 & 0.3440 & 1 & 0.9260 \\ -0.1460 & -0.0646 & -0.1944 & -0.1470 & 0.2098 & 0.1869 & 0.0405 & -0.9260 & 0.1425 \end{pmatrix}$$

（2）Select the second independent variable to enter the regression equation, calculate the partial regression sum of squares:

$$V_1^{(1)} = \left(r_{1y}^{(1)}\right)^2 \div r_{11}^{(1)} = (-0.1460)^2 \div 0.8011 = 0.0266$$

$$V_2^{(1)} = \left(r_{2y}^{(1)}\right)^2 \div r_{22}^{(1)} = (-0.0646)^2 \div 0.9796 = 0.0043$$

$$V_3^{(1)} = \left(r_{3y}^{(1)}\right)^2 \div r_{33}^{(1)} = (-0.1944)^2 \div 0.9624 = 0.0393$$

$$V_4^{(1)} = \left(r_{4y}^{(1)}\right)^2 \div r_{44}^{(1)} = (-0.1470)^2 \div 0.9533 = 0.0227$$

$$V_5^{(1)} = \left(r_{5y}^{(1)}\right)^2 \div r_{55}^{(1)} = (0.2098)^2 \div 0.5657 = 0.0778$$

$$V_6^{(1)} = \left(r_{6y}^{(1)}\right)^2 \div r_{66}^{(1)} = (0.1869)^2 \div 0.3869 = 0.0903$$

$$V_7^{(1)} = \left(r_{7y}^{(1)}\right)^2 \div r_{77}^{(1)} = (0.0405)^2 \div 0.8817 = 0.0019$$

Among them, the biggest value is $V_6^{(1)} = 0.0903$，Check whether x6 should be introduced：

$F=(8-3)\times0.0903/(0.1425-0.0903)=8.6494 > F_{\alpha}=6.59$,Therefore, the independent variable x6 has a significant effect and is introduced into the regression equation，$R^{(1)}$ was changed into $R^{(2)}$.

$$R^{(2)} = \begin{pmatrix} 0.7842 & 0.3570 & 0.2256 & 0.3128 & -0.2367 & -0.2088 & -0.6739 & -0.2825 & -0.1850 \\ 0.3570 & 0.9738 & -0.3641 & -0.3396 & 0.0530 & -0.1214 & -0.5035 & 0.2381 & -0.0873 \\ 0.2256 & -0.3641 & 0.8472 & 0.8114 & -0.5429 & 0.5456 & 0.1901 & -0.2332 & -0.0924 \\ 0.3128 & -0.3396 & 0.8114 & 0.9441 & -0.5570 & 0.1547 & 0.0701 & 0.0948 & -0.1181 \\ -0.2367 & 0.0530 & -0.5429 & -0.5570 & 0.4360 & -0.5790 & -0.0261 & -0.2057 & 0.1015 \\ 0.2088 & 0.1214 & -0.5456 & -0.1547 & 0.5790 & 2.5846 & -0.5385 & -2.0237 & 0.4832 \\ -0.6739 & -0.5035 & 0.1901 & 0.0701 & -0.0261 & 0.5385 & 0.7695 & -0.7656 & 0.1411 \\ 0.2825 & -0.2381 & 0.2332 & -0.0948 & 0.2057 & -2.0237 & 0.7656 & 2.5846 & 0.5477 \\ -0.1850 & -0.0873 & -0.0924 & -0.1181 & 0.1015 & -0.4832 & 0.1411 & -0.5477 & 0.0522 \end{pmatrix}$$

（3）Select the third independent variable to enter the regression equation, and calculate the partial regression sum of squares:

$$V_1^{(2)} = \left(r_{1y}^{(2)}\right)^2 \div r_{11}^{(2)} = (-0.1850)^2 \div 0.7842 = 0.0436$$

$$V_2^{(2)} = \left(r_{2y}^{(2)}\right)^2 \div r_{22}^{(2)} = (-0.0873)^2 \div 0.9738 = 0.0078$$

$$V_3^{(2)} = \left(r_{3y}^{(2)}\right)^2 \div r_{33}^{(2)} = (-0.0924)^2 \div 0.8472 = 0.0101$$

$$V_4^{(2)} = \left(r_{4y}^{(2)}\right)^2 \div r_{44}^{(2)} = (-0.1181)^2 \div 0.9441 = 0.0148$$

$$V_5^{(2)} = \left(r_{5y}^{(2)}\right)^2 \div r_{55}^{(2)} = (0.1015)^2 \div 0.4360 = 0.0236$$

$$V_7^{(2)} = \left(r_{7y}^{(2)}\right)^2 \div r_{77}^{(2)} = (0.1411)^2 \div 0.7695 = 0.0259$$

Among them $V_1^{(2)} = 0.0436$ is the biggest，Check whether to introduce x1：

F=（8-4）×0.0436/（0.0522-0.0436）=20.2791＞F_α=6.59，The independent variable x1 has a significant effect, which is introduced into the regression equation and $R^{(2)}$ transformed into $R^{(3)}$.

$$R^{(3)} = \begin{pmatrix} 1.2752 & 0.4552 & 0.2877 & 0.3989 & -0.3018 & -0.2662 & -0.8594 & -0.3603 & -0.2359 \\ -0.4552 & 0.8114 & -0.4668 & -0.4820 & 0.1608 & -0.0264 & -0.1967 & 0.3667 & -0.0031 \\ -0.2877 & -0.4668 & 0.7823 & 0.7214 & -0.4749 & 0.6057 & 0.3839 & -0.1519 & -0.0391 \\ -0.3989 & -0.4820 & 0.7214 & 0.8193 & -0.4626 & 0.2380 & 0.3389 & 0.2075 & -0.0442 \\ 0.3018 & 0.1608 & -0.4749 & -0.4626 & 0.3646 & -0.6420 & -0.2295 & -0.2909 & 0.0457 \\ -0.2662 & 0.0264 & -0.6057 & -0.2380 & 0.6420 & 2.6402 & -0.3591 & -1.9485 & 0.5324 \\ 0.8594 & -0.1967 & 0.3839 & 0.3389 & -0.2292 & 0.3591 & 0.1903 & -1.0084 & -0.0179 \\ -0.3603 & -0.3667 & 0.1519 & -0.2075 & 0.2909 & -1.9485 & 1.0084 & 2.6864 & 0.6143 \\ 0.2359 & -0.0031 & -0.0391 & -0.0442 & 0.0457 & -0.5324 & -0.0179 & -0.6143 & 0.0085 \end{pmatrix}$$

（4）Determine whether variables should be excluded from the regression equation, and calculate the sum of squares of the partial regression of the selected independent variables x8, x6, and x1.

$$V_8^{(3)} = \left(r_{8y}^{(3)}\right)^2 \div r_{88}^{(3)} = (0.6143)^2 \div 2.6864 = 0.1405$$

$$V_6^{(3)} = \left(r_{6y}^{(3)}\right)^2 \div r_{66}^{(3)} = (0.5324)^2 \div 2.6402 = 0.1074$$

$$V_1^{(3)} = \left(r_{1y}^{(3)}\right)^2 \div r_{11}^{(3)} = (-0.2359)^2 \div 1.2752 = 0.0436$$

Among them，$V_1^{(3)}$ is the smallest，and the following is $V_6^{(3)}$ ， it can be seen from the foregoing that since x1 and x6 are introduced into the equation at the latest, they should not be excluded from the regression equation, so we only consider whether x8 should be eliminated, and whether x8 should be eliminated for testing below：F=（8-4）×0.1405/0.0085=66.1176＞F_α=6.59。Therefore, x8 has a significant effect and cannot be excluded from the regression equation.

（5）Introduce the fourth independent variable, and calculate the sum of squared partial regression of the remaining independent variables x2, x3, x4, x5, x7.

$$V_2^{(3)} = \left(r_{2y}^{(3)}\right)^2 \div r_{22}^{(3)} = (-0.0031)^2 \div 0.8114 = 0$$

$$V_3^{(3)} = \left(r_{3y}^{(3)}\right)^2 \div r_{33}^{(3)} = (-0.0391)^2 \div 0.7823 = 0.0020$$

$$V_4^{(3)} = \left(r_{4y}^{(3)}\right)^2 \div r_{44}^{(3)} = (-0.0442)^2 \div 0.8193 = 0.0024$$

$$V_5^{(3)} = \left(r_{5y}^{(3)}\right)^2 \div r_{55}^{(3)} = (0.0457)^2 \div 0.3646 = 0.0057$$

$$V_7^{(3)} = \left(r_{7y}^{(3)}\right)^2 \div r_{77}^{(3)} = (-0.0179)^2 \div 0.1903 = 0.0017$$

Among them, $V_5^{(3)} = 0.0057$ is the biggest， Check whether the equation can be introduced to x5.

F=0.0057×（8-5）/0.0085-0.0057=6.1071＜F_α=6.59，The role of the independent variable x5 is not obvious and cannot be introduced into the regression equation. The stepwise regression calculation ends here. Among the eight factors that affect the 400-meter run, there are three significant factors: x8 50 meters in the eighth section, x6 50 meters in the sixth section, and x1 50 meters in the first section (although the first section 50 meters is simply related The coefficient is low, but the stepwise regression calculation process shows that the first stage of 50 meters and 400 meters are also related).

3.3.2 The independent variables "50 meters in the eighth section" and "50 meters in the sixth section" predict the binary linear regression equation for 400 meters

Use the stepwise regression calculation process $R^{(2)}$ to establish a binary linear regression equation： $\hat{y} = 26.905 + 1.690x_8 + 1.135x_6$, F=45.142，F＞$F_{0.01}$(2,5)=13.3，p＜0.01.The independent variables "50 meters in the eighth segment" and "50 meters in the sixth segment" predict that the remaining standard deviation of the 400-meter binary linear regression equation is 0.172 seconds.As shown in Table 5.

Table 5 Analysis of variance of the binary linear regression equation

Variance source	Sum of squares	Degree of freedom	Mean square	F value	Probability judgment
Return	2.680	2	1.340	45.142	P＜0.01
Rest	0.148	5	0.030		
Sum	2.828	7			

3.3.3The independent variables "50 meters in the eighth section", "50 meters in the sixth section" and "50 meters in the first section" predict the 400-meter stepwise regression equation

The stepwise regression equation is established as: $\hat{y} = 131.456 + 1.895x_8 + 1.252x_6 - 21.197x_1$, F=149.0 45，F＞$F_{0.01}$(3,4)=16.7，P＜0.01.The remaining standard deviation of the stepwise regression equation is 0.078 seconds.see Table 6.

Table 6 Analysis of variance of stepwise regression equation

Variance source	Sum of squares	Degree of freedom	Mean square	F value	Probability judgment
Return	2.803	3	0.934	149.045	P＜0.01
Rest	0.025	4	0.006		
Sum	2.828	7			

IV. DISCUSSION

4.1A Comparative Study of Four Regression Equation Models for Predicting 400-meterPerformance

According to the regression equation obtained in the above stepwise regression calculation process, the actual performance data (44.6686±0.4232) of the men's 400m final of the 2017 World Athletics Championships is compared with the predicted data of the regression equation. The correlation coefficient between the predicted performance of the four models and the actual game performance is P<0.01, indicating that there is a very significant difference between the predicted performance of the model and the actual performance. The result of the paired t-test is P>0.05, which indicates that there is no significant difference between the predicted performance of the model and the actual performance. It can be seen that the four models have better actual results in predicting the 400-meter performance. The consistent correlation coefficient between the results predicted by the stepwise regression equation and the actual game results is 0.9357, and the remaining standard deviation of the stepwise regression equation is the smallest 0.078 seconds. The prediction effect is good and the accuracy is high.As shown in Table 7

Table 7 Comparison of the four regression equation models for predicting the 400-meter capacity

Variables	Predictive regression equation	400m predicted value	Coefficient of agreement r	Paired t test
Model 1：Eighth paragraph 50m prediction	$\hat{y} = 25.708 + 2.854x_8$	42.3957±0.5040	0.9267	1.0183
Model 2：The sixth stage 50m prediction	$\hat{y} = 32.294 + 2.146x_6$	44.6826±0.3710	0.8855	0.9644
Model 3：Binary regression prediction	$\hat{y} = 26.905 + 1.690x_8 + 1.135x_6$	43.3359±0.4871	0.9061	0.9912
Model 4：Stepwise regression prediction	$\hat{y} = 131.456 + 1.895x_8 + 1.252x_6 - 21.197x_1$	44.6571±1.6998	0.9357	1.0302

（Note：$r_{0.05/2}(5) = 0.755$，$r_{0.01/2}(5) = 0.875$。$t_{0.05/2} = 2.447$。）

4.2Theoretical basis for the results of stepwise regression analysis

The first segment of the 400m is the 50m after the start. This stage is as critical as the final sprint in the 400m race. The performance required for the start has a pivotal effect on the overall performance. The ability to react quickly at the start can enable athletes to gain a favorable speed in the game, and at the same time cause psychological pressure on the opponents of the same group, and then Destroy the running tactics and rhythm developed by the opponent before the game, and ultimately suppress the strength of the opponent in the game. Leeet al.(2017) scholars found in their research that the weak foot at the back of the 400-meter race has an impact on the 400-meter performance. Therefore, based on the results of the stepwise regression analysis, the important role of the starting stage in the 400-meter training is also obvious.

The sixth segment of 400 meters is the last 50 meters on the way. This stage is the key to entering the final sprint. For a standard track and field track, as far as its innermost track (track 1) is concerned, the straight track of the 400m track is only about 171.92m, while the curve is about 228.08m. Obviously the curve is much longer than the straight, which means It takes a lot more time for an athlete to finish a curve than it takes to finish a straight, and it consumes more physical energy. Zhang etal.(2020) concluded that if the athlete’s curve running technique can be effectively improved, it should be strengthened in training so that the athlete can run in the curve. Part of the time saved by the above will be of great help to the overall results of sprint events. In the 400-meter curve acceleration process, effective use of curve running technology, moderate increase in speed in the game, does not consume too much physical energy to achieve satisfactory results, in order to leave sufficient energy to produce extremely strong.Bowtell(2007)consider that the explosive force of the speed into the straight sprint accelerates, which will be one of the important factors for obtaining an excellent 400m result. Therefore, based on the results of the previous stepwise regression analysis, the distance of 50 meters in the sixth segment, that is, the final down corner, has an important influence on the 400-meter performance.

The eighth segment of the 400m is the last 50m of the sprint. This stage is usually the decisive factor affecting the 400m performance. When the 400 meters run to the critical 50 meters of the final sprint, the human body consumes excessive oxygen, a large amount of lactic acid accumulates in the muscles, the skeletal muscles become fatigued, the raising of the thigh is restricted, and the idea of slowing down the sprint speed is generated, resulting in a decrease in speed. Therefore, 400-meter athletes, coaches and sports researchers have always attached great importance to the cultivation of athletes’ final sprint ability, and have long explored various effective training methods and methods to improve athletes’ speed endurance and skeletal muscle resistance to lactic acid. Through continuous training to improve the performance of the 400-meter athletes. Because the final 50-meter sprint stage is the final link to complete the 400-meter race, and the energy consumption of the athlete's body makes the body extremely fatigued, the final sprint stage becomes the key to affecting the 400-meter performance(Wlodzimierz 2013).

V. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusion

Through a stepwise regression analysis of the impact on the 400-meter performance from the 50-meter perspective, the results show:

(1) For the multiple correlation coefficient R=0.996, after the analysis of variance test, P<0.01, the difference is highly significant, and further regression analysis can be used. This shows that the 400-meter regression model has a good fitting effect. Through stepwise regression analysis, it is found that the 400-meter performance y excludes 5 from the 8 indicators, which are the second distance x2, the third distance x3, the fourth distance x4, the fifth distance x5 and the seventh distance x7, the main factors affecting the athlete's 400-meter performance are the eighth distance x8, the sixth distance x6, and the first distance x1. Each of these factors has very significant significance. Then the regression equation of the influence of the independent variable 50 meters on the 400 meter performance of the dependent variable in the men's 400 meters final at the 2019 World Athletics Championships is: $\hat{y} = 131.456 + 1.895x_8 + 1.252x_6 - 21.197x_1$

(2) The first stage of the 50-meter performance (x1) has a significant negative impact on the 400-meter performance. The first stage performance (x1) changes by one unit, it will cause the 400m score to drop by 21.197 units; the eighth stage 50m score (x8) has a significant impact on the 400m score, the eighth stage 50m score (x8) changes by one unit , The 400m score will increase by 1.895 units; the sixth stage 50m score (x6) has a significant impact on the 400m score, the sixth stage 50m score (x6) changes by one unit, the 400m score will increase by 1.252 units. Under certain conditions, the 400-meter performance can be improved by improving the sixth stage 50-meter performance and the eighth stage 50-meter performance.

5.2 Suggestions

(1) The relevant variables that have a significant impact on the 400-meter performance can be further studied in depth.

(2) Research on the influencing factors of 400m performance can choose other variables (such as strength, endurance, speed indicators) instead of simply dividing the 400m into segments for research.

Disclosure statement

No potential conflict of interest was reported by the author(s)

Acknowledgements

We thank the men's 400m finalists at the 2019 World Athletics Championships.

We appreciate the public data provided by the IAAF.

REFERENCES

[1]. Girish R,Darryl W(2020). Does a home advantage effect exist in international track and field athletics?.Managing Sport and Leisure,25(5).

[2]. Sibel D, Ayşegül Y, Eylem Ç, TansuY(2015). A Record Holder Female Athlete's Training Loads and its Distribution in Women 400 meter Hurdle Running[J].International Journal of Science Culture and Sport,3(Special Issue 3).

[3]. Liang Xia, Sports Department of Heilongjiang Institute of Science and Technology Harbin,Hei Longjiang,150027 China.The Present Situation of the Chinese Men's 400-meter Hurdles and the Countermeasure[A].IEEE Beijing Section、Zhengzhou Institute of Aeronautical Industry Management.Proceedings of 2012 IEEE International Conference on Engineering Technology and Economic Management（ICETEM2012）[C].IEEE Beijing Section、Zhengzhou Institute of Aeronautical Industry Management:IEEE BEIJING SECTION,2012:5.

[4]. Santos PA,SousaCV,Barbosa LP,Aguiar SS, et al(2020). Breaking the athletics world record in the 100 and 400 meters: an alternative method for assessment.[J]. The Journal of sports medicine and physical fitness.

[5]. Krzysztof M,Ryszard M,Milan Ć(2010). Asymmetry of Step Length in Relationship to Leg Strength in 200 meters Sprint of different Performance Levels[J]. Journal of Human Kinetics,25(-1).

[6]. Ralph M,John H(1985,). Kinematic Analysis of Olympic Sprint Performance: Men's 200 Meters[J]. International Journal of Sport Biomechanics,1(2).

[7]. 이장택. A Study on the performance pattern in the elite decathlon[J]. 한국데이터정보과학회지,2010,21(6).

[8]. RoghayeP,Mohammad A, Jadidi P, Karim A, Hoseyn F(2012). The Effect of Sodium Bicarbonate Supplement on Lactic Acid, Ammonia and the Performance of 400 Meters Male Runners[J]. Baltic Journal of Health and Physical Activity,4(2).

[9]. Zhang PY, Zhou ZJ, Wang F(2020). Analysis of Influencing Factors of Shanxi Economic Development based on Multiple Linear Regression Analysis[J]. World Scientific Research Journal,6(10).

[10]. Lee K, Hong W, Kim EG(2017). Analysis of Changes in Stride Length, Time, and Electromyography Finding Depending on Athletic Crouch Start Method. 27(1):75-81.

[11]. Zhang Y, Li F, Peng NJ ,Peng LH(2020). Environmental impact assessment of air-permeable plastic runway production in China. 730.

[12]. BowtellM, Spence A, Wilson A, et al(2007). Limitation to maximal speed human sprinting — Insights from bend running and fatigue. 146(4):S109-S109

[13]. Wlodzimierz SE, Patrycja L(2013). Kinematics of marathon running tactics. 32(6):1379-1392.