



Field events performance prediction applied research based on data curve fitting

Kelei Guo

Department of Physical Education, Xianyang Normal University, Xianyang, China

ABSTRACT

It is particularly important to seize and predict the total changing trend of sporting performance and achieve reasonable suggestions through exploring of athlete's qualities in modern times. The sole measurement criteria to athlete is their results achieved in Olympic Games or other large scaled games by fundamentally improved their integrated qualities. This study utilized 28 sessions field events performances collected from "Chinese Sports Almanac", made respectively analysis of 28th Olympic games athletics performance as high jump, long jump, triple jump and pole vault, made curve regression of such performance with generalized least squares, and got changing curve so as to describe performance changing trend as well as made prediction to future sessions of Olympic games performance. Regression curve is the simple and feasible method to predict performance.

Key words: Athletics, generalized least squares, curve fitting, performance prediction

INTRODUCTION

With the social progress and development, sports have already become the indispensable part in human civilized life. Athletics as the most widespread sport is well-received by people. Olympic Games-led several kinds of sports competitions have sprung up to improve people's athletic level. Olympic Games is a sporting event that attracts worldwide attention, from which athletics is the more widespread and well-received sport. Performance in every Olympic Games would be set. To do fully seize of athletics sport performance total changing trend, make prediction accordingly so as to give athletes a general training objects and improve our country's competitive sports quality.

For curve fitting method as well as prediction research of sports performance, lots of predecessors have made great efforts on it, based on that science and sports are making constant progress and fast development. Some of them put forward their own thoughts, from which: Zhuang Chong(2006)Through contrast research on gray system model and BP neural network model, get so far the relative classic, mature and with better effective mathematic method time in prediction is time series model, GM(1, 1) grey system model together with recently popular BP neural network model [1-3]. Liu Jia-JIN(2006) Explore new method of competitive sports area's grey enveloped prediction model, put forward the prediction results is a possible point that cannot acknowledge existing area of prediction value [2]. Especially when original data has excessive fluctuate ranges, the modeling precision of GM (1, 1) grey model is too low and cannot fulfill prediction function [3, 4].

This paper based on previous studies, utilizing crowns Asia third place performance in 28 sessions high jump, long jump, triple jump and pole vault from "Chinese Sports Almanac", get average performance to represent their items performance on such sessions Olympic Games. Apply generalized least squares to make multivariate regression analysis and work out curve graph, so that describe the total changing trend of past, present and future athletic performance. Then use regression functions to predict future some sessions Olympic Games performance, and present performance prediction range with the error value from prediction to make more accuracy performance prediction.

PREDICTION ON 4 FIELD EVENTS PERFORMANCE

Track events, field events and combined events are collectively called as athletics. Athletics has very long history, is one of the most widespread sports in the world and also an enjoyable sport in competition. Athletics features as closely relationship with life, simple and feasible, wideness as well helpful for physical and psychological health make it well-received by people and being passed so far. To better research of athletics performance, and provide training objects that can refer to athletes through performance prediction. By data statistics and handling of every athletic in 28 sessions Olympic Games, get Table 1 data with average value of top three performances.

From Table 1, it can be seen that the 3rd 12th 13th Olympic Games data missed because of two world wars influence made Olympic Games broken off. On a whole, each sports performance is in the rising trend. Generally speaking, due to modern human physical quality improvement as well as people's high emphasis on athletics and scientific rational development of athletics training methods, athletics has made leap progress and always keep the rising trend. However, the rising extent is different at times, this paper respectively make curve multivariate regression analysis of high jump, long jump, triple jump, pole vault data with SPSS, and provide curve review model so as to make scientific rational prediction of each athletics performance [5-7].

Table 1: Summary of 28 sessions' athletics performance

Session	High jump	Long jump	Triple jump	Pole vault
1	1.703	6.213	12.976	3.133
2	1.810	7.103	14.026	3.250
3	1.786	7.036	14.041	3.455
4	1.888	7.218	14.686	3.666
5	1.910	7.330	14.480	3.883
7	1.910	7.108	14.418	3.796
8	1.950	7.326	15.440	3.933
9	1.920	7.570	15.163	4.083
10	1.970	7.563	15.386	4.271
11	2.010	7.890	15.720	4.283
14	1.960	7.641	15.263	4.233
15	2.100	7.466	15.906	4.493
16	1.960	7.663	16.210	4.530
17	2.153	8.090	16.623	4.616
18	2.173	8.030	16.756	5.050
19	2.220	8.416	17.293	5.400
20	2.216	8.150	17.236	5.416
21	2.230	8.160	17.123	5.500
22	2.326	8.310	17.270	5.693
23	2.330	8.340	17.103	5.666
24	2.366	8.493	17.516	5.850
25	2.340	8.550	17.710	5.783
26	2.370	8.343	17.803	5.920
27	2.339	8.450	17.546	5.900
28	2.340	8.460	17.606	5.900

High jump performance curve regression model:

High jump is one of the field events in athletics that is composed of rhymed run-up, one-foot takeoff, clear the bar and landing completion so on other motions, a competition item with its performance calculated with the cross-bar edge height that athletes jumped. High jump is a human height conquest sports item, also a human individual gravity and earth gravity overcoming physical items, and a symbol of human incommutability, challenge the nature with great courage and enable to scale new heights. High jump is a sports item that widely spread in worldwide, is one of the young children most favorite sports activities. There are many postures in high jump, it including leap-type, scissors type, straddle and back layout etc., from which back layout is the most popular one. High jump is an ancient sport and strict with athlete's qualities, therefore high jumper should be trained from multiple areas. Physical training in back layout training is the base to improve jumpers' technical motion quality. Only with strong and high quality body, base can be made for performance improving. Technical training is the root for improving players' performance. Only high technology can get excellent performance. It is necessary to take psychological training, players psychological conditions directly effect on competition performance, only with good psychological quality can maintain higher competitive state. With the increasingly training intensity, players' recovery training also becomes more and more important.

Table 2: High jump model summarizing and parameter estimation

equation	model summarizing					parameter estimation value		
	R square	F	df1	df2	Sig.	constant	b1	b2
linear	.937	342.599	1	23	.000	1.735	.024	
logarithm	.813	100.125	1	23	.000	1.572	.213	
quadratic	.937	164.352	2	22	.000	1.741	.022	-0.125

Apply curve regression method, fit all previous Olympic Games high jump performance and sessions with above three regression curve model(Fig.1), point represents observed actual high jump performance, from which linear and quadratic curve model make the best fitting with high jump performance.

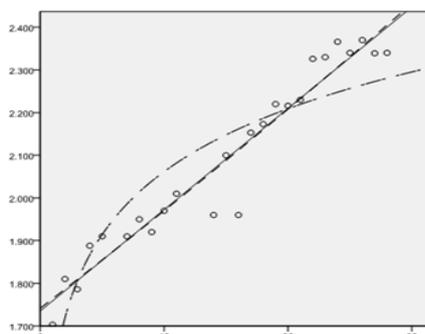


Fig.1: curve regression of high jump performance and sessions

Similar results can also get from Table 2, from which both linear model and quadratic model's R square is 0.937, the next is logarithm model. From value of F, linear and quadratic model get the best since its value being the maximum as 342.599. Besides, all three model's probability is 0.000, therefore three models all has significant correlations. To sum up, no matter fitting or F value, probability, all current models have best fitting levels, and linear model is the most significant correlation fitting model, so choose linear model. According to constant and coefficient got from linear fitting, its regression equation as formula (1) shows:

$$Y = 0.024x + 1.735 \quad (1)$$

Long Jump performance's curve regression model:

Long jump also is a long history athletics item. It is composed of run-up, takeoff, soar and land four parts, and such four motions be done at one go without interruption, closely connects as a unified entirety. Therefore correctly fulfils and makes dynamic integration of each part motion in long jump technique is essential to long jump technique. Many factors influence jumpers' long jump performance, from which normalization of technical motions is the main factor, to the exclusion it also including jumpers' adaptation degree to spots as well as their psychological state and so on. Coaches should specified jumpers motions techniques and make proper arrangement and plan to jumpers training time and intensity in daily training so as to achieve good results in games. Besides, coaches should also pay attention to jumpers' psychological states neither too excited nor too negative enable them to remain best psychological states to take the challenge.

Table 3: Long jump model summarizing and parameters estimation

equation	model summarizing					parameter estimation value	
	R square	F	R square	F	R square	F	
linear	.868	151.705	1	23	.000	6.812	.066
logarithm	.868	151.462	1	23	.000	6.255	.631
quadratic	.883	83.013	2	22	.000	6.645	-.001

Apply curve regression method, make fitting to every session Olympic Games' long jump performances, get three regression curve model as Fig.2 shows, from which we can see that quadratic curve model has the best fitting degree and its R square is 0.883 by referencing Table 3. Logarithm and linear model get the second fitting level, whose R square is the same and be 0.868. From value of F, linear and logarithm model are the best, because their value is the maximum as 151.705 and 151.462. Besides, probability value of three models is 0.000; therefore three models have correlated significance. To sum up, the best fitting degree as well as best correlated fitting model quadratic model is chosen. According to constant and coefficient got from linear fitting, its regression equation as formula (2) shows:

$$Y = -0.001x^2 + 0.1x + 6.645 \quad (2)$$

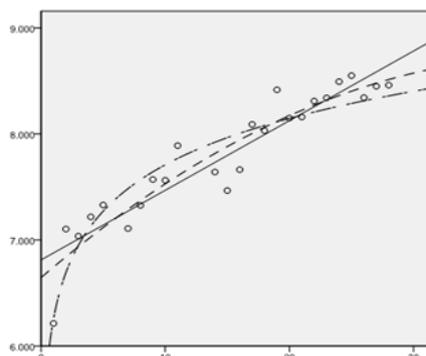


Fig.2: curve regression of long jump performance and sessions

Triple jump performance curve regression model:

Triple jump is a sport item that requires jumpers make dynamic combination of speed, strength and balance ability. Due to triple jump's higher requests with jumpers' physical quality and its own relative complicated techniques, its technique is constantly changing since it entered into competitive sports circle. It requires jumpers with fast run-up ability; excellent jump ability. Trip jumpers' leg strength should be stronger enough to better concentrate speed and strength and then complete trip jump exercises as well as achieve good results. Therefore, jumpers should take explosive force training, jumping training and waist abdomen muscle training in triple jump. Main factors that influence on triple jump performance are speed, jumpers not only should have good run-up speed, but also good movements speed and jump speed. Jumpers soaring angle after running-up every time, their own initial speed as well as balance degree in air posture and so on would affect on triple jump performance. Triple jump motions connection and its distance ratio as well as rhythm play a very important role in performance improving.

Table 4: triple jump model summarizing and parameters estimation

equation	Model summarizing					parameters estimation value		
	R square	F	df1	df2	Sig.	constant	b1	b2
linear	.938	345.304	1	23	.000	13.662	.159	
logarithm	.894	194.431	1	23	.000	12.399	1.496	
quadratic	.951	212.849	2	22	.000	13.291	.237	-.003

Similarly curve regression method is adopted, regression curve model as Fig.3 is got by fitting triple jump Olympic Games performance with sessions, from which quadratic model has the best fitting degree and its R square is 0.951 as Table 4 shows, the next is logarithm and linear model. From value of F, linear model is the best because it has the maximum value. Besides, all three models' probability are 0.000, therefore the three models have significant correlations. According to above, quadratic function model is chosen.

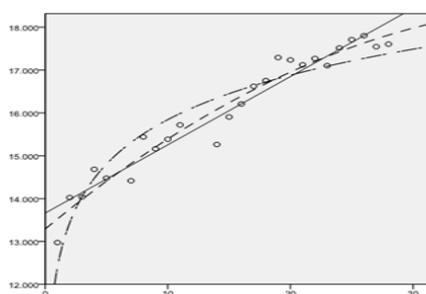


Fig.3: curve regression of triple jump performance and sessions

According to constant and coefficient that achieved, its regression equation as formula (3) shows.

$$Y = -0.003x^2 + 0.237x + 13.291 \quad (3)$$

Pole vault performance curve regression model:

Pole vault requires athletes enable body across a certain height with pole motions as overhang, swing, leg raise, pull up etc. by pole supporting and its elastic force so as to calculate performance. By far woman pole vault records holder is Russian athlete Yelena Isinbayeva, man pole vault records holder is Italian Bubka. Pole vault is very strict with athletes' physical conditions, speed strength and specific technique. So far, America is one of the most

developed countries in pole vault, though our country's pole vault level ranks in the front of world, it still has a long way to go to catch up with America. So as to constantly shorten the gap, both coaches and athletes have made great efforts. Many scholars as well as sports researches also provide lots of scientific research on pole vault. Adding scientific training method together with continuously efforts of athletes and coaches, our country's pole vault level improvement can be expected soon.

Table 5: Pole vault model summarizing and parameters estimate

equation	Model summarizing				parameters estimation value		
	R square	F	R square	F	R square	F	
linear	.965	633.839	1	23	.000	3.096	.107
logarithm	.827	109.681	1	23	.000	2.372	.956
quadratic	.965	305.256	2	22	.000	3.129	.101

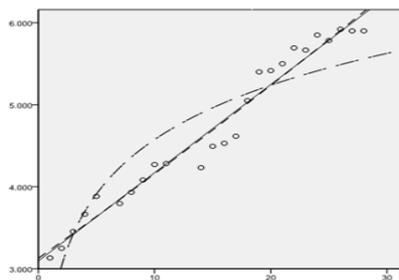


Fig.4: curve regression of pole vault performance and sessions

Achieve curve as Fig.4 shows by fitting pole vault performance with sessions. In the above three regression curve models, we can see that linear model and quadratic curve model have the best fitting degree, and their R square is 0.937 as Table 5 shows, the next is logarithm model. From value F, linear model is the best since it gets maximum value as 633.839. Besides, all three models' probability is 0.000, therefore they have relative correlation. For above, linear model is chosen. According to constant and coefficient that achieved, its regression equation as formula (4) shows.

$$Y = 0.0107x + 3.096 \quad (4)$$

Field events performance prediction:

According to above distribution conditions of each athletic items performance points, makes curve fitting respectively to each competition items, and use SPSS testify their fitting levels. Through measurement and research of each aspect, each athletic game development trend followed by increasing of sessions can be achieved. Among them, both high jump and pole vault have linear relations with sessions, while long jump and triple jump have quadratic function relations with sessions. With the changing of curve, all can describe performance point locations, therefore better prediction of future sessions performance can be done with fitting function.

To sum up, high jump, long jump, triple jump and pole vault jump's fitting curve are respectively as formula(1), (2), (3), (4)shows, make prediction of 29th and 30th Olympic Games athletic items performance by utilizing fitted curve model, its error value and actual value status as Table 6 shows.

Table 6: Curve model prediction value to the 29th and 30th Olympic Games athletic items performance

Sessions item	High jump			long jump			Triple jump			Pole vault		
	predict	actual	error	predict	actual	error	predict	actual	error	predict	actual	error
29 th	2.43	2.33	4.15	8.70	8.26	5.37	17.6	17.6	0.42	6.20	6.18	0.31
30 th	2.46	2.36	4.03	7.75	8.20	5.64	17.3	17.4	0.83	6.31	5.96	5.81

Note: Unit of error values is %, as 4.15 in Table represents 0.0415=4.15%.

CONCLUSION

This paper collected and handled of 28 sessions Olympic Games performances in high jump, long jump, triple jump, pole vault these four athletics, got performance distribution features through data researching. According to research on 28 sessions sports data, utilized curve regression model, made fitting and prediction of high jump, long jump and others Olympic Games performances and sessions, got performance distribution ranges. According to each regression curve equation, future two sessions Olympic Games the 29th Beijing Olympic Games and the 30th London Olympic Games competition results can be predicted, their error analysis can be done. It can be known from "China sports almanac" data, the 29th Olympic Games' long jump performance prediction value is 8.704m, while

actual is 8.275m, error between them is 5.18%. High jump performance in the 30th Olympic Games is 2.63m, while prediction shows 2.455m, error between them is 4.02%. From the above data, it can be seen that it's feasible to make prediction of athletics with fitting function, no great difference between predicted performance and actual performance that approximately in the range of 0.4% to 5%, such error might have connection with athlete's body and psychological states, therefore make athletic performance prediction with curve as this paper suggestion is one of the simple and feasible methods.

REFERENCES

- [1] Liu Jia-Jin. *Ningbo University sports institute*, **2006**, 3.
- [2] Zhang B.. *Int. J. Appl. Math. Stat.*, **2013**, 44(14), 422-430.
- [3] Zhang Yuan-Zhang. *Journal of north China institute of water conservancy electric power*, **2010**, 31(1).
- [4] Zhuang Chong. *Journal of Nanjing Sports Institution*, **2006**, 06.
- [5] Zhang B., Zhang S. and Lu G.. *Journal of Chemical and Pharmaceutical Research*, **2013**, 5(9), 256-262.
- [6] Zhang B., Yue H.. *Int. J. Appl. Math. Stat.*, **2013**, 40(10), 469-476.
- [7] Zhang B., Feng Y.. *Int. J. Appl. Math. Stat.*, **2013**, 40(10), 136-143.