

IJACT 20-9-27

## A model of predicting performance of Olympic female weightlifters using time series analysis

Jin-hee Won, In-ho Cho\*

Human Performance Laboratory, Korea National Sport University, Seoul 05541,  
Republic of Korea, Korea National Sport University, Seoul 05541,

Republic of Korea\*

E-mail : judo69@knsu.ac.kr\*

### Abstract

*The purpose of this study was to predict the performance of female weightlifters using time series analysis. Based on this purpose, a time series analysis was used to calculate the performance prediction model for women(58kg) among the domestic women weightlifters who participated in the Olympics. As a result of creating time series data based on 10 years of record and then evaluating the sequential charts of each athlete group, the female athletes' records did not show any seasonality or difference. In addition, after examining the independence of the data through the creation of a time series model, it was shown that the models produced conformed to the criteria for compliance and that there was no difference in the data, but there was a trend. Accordingly, Holt linear trend analysis of the exponential smoothing model was applied. As a result of deriving the prediction model of the athletes through this process, it was found that the women (58kg) who participated in the Olympics continued to improve within the range of 166.1kg to 184.1kg.*

**Keywords:** time series analysis, prediction model, performance, female weightlifter

### 1. Introduction

Along with the identification of athletes' performance, interest and research on prediction have become the subject of major interest in the sports field. In particular, it is a very important ability for sports leaders who are making various efforts to improve their performance through accurate predictions about the players' performance.

Initial studies related to performance evaluated and predicted performance based on players' physical strength and skills, but since the recent study found that performance was determined by the interaction of various variables, research has been conducted to identify or predict performance based on analysis methods to identify the mutual influence of various variables related to the characteristics of the athlete.

Researches to predict the athletes' performance are based on the core skills, physical strength, and psychological factors for each sport, centering on physical strength[1,2,3,4], skill[5], and psychological factors[6,7,8], and present the criteria for identifying and predicting the performance.

However, these studies are meaningful in that they identified the factors that determine the athletes' performance in detail and provided a criterion[9] for predicting their performance, but the performance of the athletes is not a one-off performance or result, but the passage of time. It is true that there is a limit to empirically predicting athletic performance based on the changes according to the situation.

---

Manuscript Received: August 31, 2020/Revised: September 09, 2020 /Accepted: September 17, 2020

Corresponding Author: judo69@knsu.ac.kr

Tel:+82-10-8939-3670, Fax: +82-2-410-6990

Author's affiliation : Department of Health and Exercise Science, Korea National Sport University

In the early days, time series analysis is used as a method of inferring social changes or predicting economic conditions centered on social science and economic fields. Time series analysis is divided into univariate time series analysis and multivariate time series analysis by analyzing the future based on time series data, which are values that change over time. The time series analysis method is divided into simple trend analysis, moving average method, decomposition method, index smoothing method, and box-genkins analysis method according to the characteristics and prediction period of the data. In addition, it has the advantage of being highly useful in the field through analysis because it is possible to predict trends or seasonal changes, cyclical changes and irregular changes[10].

If you look at studies using time series analysis at sports sites, most of them focus on forecasting demand related to professional sports and sports facilities[11,12,13] or predicting the economic value of sports[14], and there is no research related to athletes' performance. In fact, it seems theoretically reasonable to analyze and utilize the characteristics of each player to predict their performance. However, predicting performance through a time series analysis may be a more empirical alternative, given that a player's performance is made up of various variables and has variable characteristics that vary with time fluctuations.

So far, most of the research on weightlifters' performance has identified or predicted athletes' performance[4,15], mostly based on technical and psychological variables or leader's behavior, and it is true that weightlifters' level of performance has been determined or predicted at sports sites based on their training, experience and subjective judgment. This method can have a positive effect in that it identifies the performance of weightlifters, but it is true that there is a limit to establishing objective and reasonable performance improvement strategies as they have identified or predicted their performance through subjective judgment and compliance. Accordingly, this study aims to present a comparative evaluation of performance and performance prediction model and a performance prediction model using time series data related to the performance of weightlifters through the competition records, unlike the regression analysis or path analysis applied in the sports field so far. The results of this study according to this purpose will provide an empirical basis for establishing a strategy for improving the performance of weightlifting leaders who are striving to improve the performance of weightlifters in the sports field.

## 2. Research Method

### 2.1 Research Subject and Data Collection

The range of data used in this study is the average data of the total number of weightlifters who participated in the National Sports Festival between 1998 and 2018, and was collected and utilized separately from 78 to 298 athletes divided by gender and weight class. The data collected for the time series analysis were classified by year, gender, and weight class, and were organized using Microsoft's Excel office 365 program, and IBM SPSS 20.0 was used.

**Table 1. Definition of variable**

Data	Response Variable	Forecast Period
Woman 58kg	Total(kg)	long-term (over 10years)

### 2.2 Research Procedure

In the first step process, the data was converted into normal time series data through the stabilization process, and in the second step, independence and suitability of the data were evaluated through sequential charts and autocorrelation analysis. In the third step, a time series model was created based on the time series data that secured independence and suitability, and the analysis method according to the generated time series model was determined. Finally, in step 4, the results of predicting the performance of weightlifters were derived by evaluating the validity of the prediction model determined through the time series analysis method.

### 2.3. Data Processing Method

Data collected according to the research purpose were converted into time series data using Microsoft's Excel office 365 and IBM SPSS 20.0 programs, and then the presence or absence of normal and abnormal time series was evaluated. If the collected data were abnormal time series, it was converted into normal time series data through the stabilization process of data using the difference. In addition, the independence of time series data was secured through the verification of Ljung-Box statistics by complying with the autocorrelation function(ACF) and partial autocorrelation function(PACF) of normal time series data, and then the suitability of the data was evaluated. After obtaining reliability through conformity assessment, a time series model was created and the analysis method suitable for the model was determined by following the time series model generated. The predictive model was derived based on the model produced according to the last analysis method, and the performance of the weightlifters was predicted after the feasibility assessment of the derived predictive model.

## 3. Result

### 3.1. Women's(58kg) Performance Prediction Model

#### 3.1.1 Data Stabilization and Time Series Analysis

**Table 2. Time Series Model Creation**

Model	Ljung-Box Q(18)			Outlier
	Statistic	<i>df</i>	<i>p</i>	
LAGS Model	14.122	17	.772	0

The data collected prior to the analysis were converted into time series data and then the time series model was deduced by the model generation. As a result of the analysis, as shown in Table 2, the significance probability of the Ljung-Box Q value of the generated time series model was greater than .05, indicating that the independence of the generated time series model satisfies the fit criterion, and there were no outliers and missing values. In addition, as a result of trend analysis based on the Spike model, the time series model was stable and there was no seasonal period, but there was a trend.

**Table 3. Autocorrelation Analysis Result**

Time Difference	Autocorrelation	Standard error	Box-Ljung Statistic		
			Value	<i>df</i>	<i>p</i>
1	.088	.208	1.854	1	.173
2	.052	.202	1.979	2	.372
3	.210	.197	2.152	3	.542
4	-.099	.191	2.053	4	.403
5	-.027	.112	2.529	5	.332
6	-.052	.178	6.049	6	.543
7	.045	.172	6.292	7	.330
8	.053	.165	6.383	8	.233
9	-.018	.158	7.999	9	.993
10	.000	.151	10.223	10	.772
11	-.165	.143	11.223	11	.677

12	-.174	.135	10.543	12	.773
13	-.062	.126	8.443	13	.443
14	-.239	.117	7.994	14	.403
15	-.093	.107	9.884	15	.499
16	-.170	.095	10.322	16	.332

**Table 4. Analysis Result of Partial Correlation**

Time Difference	Partial Autocorrelation	SE
1	-.268	.224
2	-.163	.224
3	.024	.224
4	-.408	.224
5	-.078	.224
6	-.176	.224
7	-.101	.224
8	-.053	.224
9	-.035	.224
10	-.210	.224
11	-.051	.224
12	.131	.224
13	.003	.224
14	-.090	.224
15	-.025	.224
16	-.053	.224

The autocorrelation function(ACF) and partial autocorrelation function(PACF) analysis were conducted to evaluate the suitability of the time series model. As a result of the analysis, it was found that the significance value was greater than .05, indicating that the suitability of the data was good, and the difference and seasonal difference were not necessary.

### 3.2. Creating and Evaluating Predictive Models

A model was created based on the time series data. In the model, there were no differences and seasonal differences, but a trend appeared, and a prediction model was generated by applying the Holt linear trend.

As a result of model generation, the significance probability of Ljung-Box was greater than .05, and it was independent from white noise, and the generated model was found to be a suitable model for inferring the prediction model.

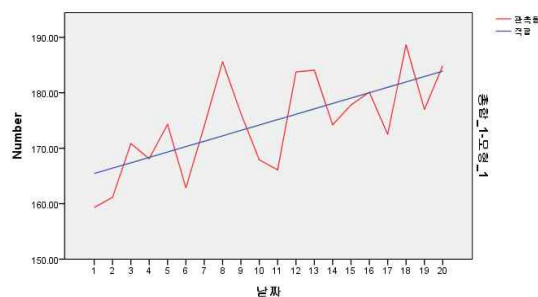
**Table 5. Model Statistics**

Model	Predictor Variable	Model Fit Statistic			Ljung-Box Q(18)			Number of Outliers
		Normal R Square	R Square	Normalized BIC	Statistic	Degree of Freedom	p	
LAGS	0	.522	-.193	4.778	14.552	17	.881	0

As a result of model generation, the significance probability of Ljung-Box was greater than .05, indicating that it was a suitable model for inferring the prediction model.

**Table 6. Holt linear Trend Parameter**

Model	Estimated Value	SE	t	p
LAGS	174.47	1.771	88.202	.000



As shown in Table 6, the difference between the observed value and the predicted value was found to be large, but the weight of the short-term predicted total was increased within the range of 166.11 kg to 184.1 kg.

#### 4. Discussion

In this study, the following results were derived through time series analysis to predict the performance of weightlifters. Up to now, athletes' performance has been identified and predicted by a causal relationship analysis method that examines the mutual influence between various variables related to the athlete's characteristics in terms of physical strength, skill, and psychology based on a multi-dimensional structural model.

However, this method has contributed positively in that it has detailed the components of performance and provided a direction to predict the formation and level of performance by each player, but it is true that there are limits to predicting performance by team or event from a macroscopic perspective by limiting performance to individual characteristics. Therefore, in order to predict a player's performance, it is reasonable to consider both aspects of micro and macro perspectives.

Until now, the time series analysis method is a type of big data analysis method that faithfully fulfills the ultimate purpose of the big data analysis method to predict the future through information collection and analysis. Therefore, time series analysis is being used as a universal method to predict the future and establish strategic planning in various fields today.

Up to now, time series analysis in the sports field has been widely used in studies to develop a model for predicting the win or loss of a game through the construction of an artificial neural network[16] or to predict demand in the sports industry[17]. Their studies were mainly conducted for the purpose of generating profits or maximizing profits centering on professional sports, and it is true that studies focusing on amateur sports are insufficient.

Therefore, the purpose of this study was to extract the performance prediction model of domestic

weightlifters, focusing on weight lifting. As a result, time series data were extracted based on records of female athletes (58 kg) who participated in the Olympics for 10 years. Prior to the first analysis, the time series data were generated and then the characteristics of the data were analyzed. For women (58kg), the record of 10 years of increase and total clean and jerk was shown to be upward trend among the fluctuations. The results indicate that the women participating in the Olympics have been improving their performance for 10 years. Therefore, the prediction model was derived by applying a Holt linear trend analysis. As a result of this procedure, the performance-related prediction model of the competitors was derived, and in the case of women (58 kg), it was predicted that they would rise within the range of 166.11 kg to 184.1 kg in proportion to changes in timing.

Based on this prediction model, it is judged that the women (58kg) athletes participating in the Olympics have a good strategy to improve their performance.

## **5. Conclusion and Suggestions**

The purpose of predicting the performance of weightlifters, this study derives a performance prediction model through a time series analysis method based on the records of weightlifters for 10 years. As a result of analysis according to this procedure, the following results were derived.

In the case of female (58kg) athletes, the performance for 10 years showed an upward trend, and as a result of predicting the performance through the prediction model based on this, the total was predicted to rise within the range of 166.11kg to 184.1kg.

During this study, suggestions for follow-up research are as follows.

First, in this study, a study was conducted on 58kg female athletes among weightlifters. In the case of weightlifting, it is true that there are various weight classes for men and women in addition to women. Therefore, in the follow-up study, a follow-up study to derive a model for predicting the performance of domestic athletes of different weight classes should be conducted.

Second, in time series analysis, there are comparative time series analysis methods other than alone. Therefore, in order to better understand the level of performance of domestic athletes and to provide reference data for strategies to improve domestic athletes' performance at world competitions, comparative time series analysis should be conducted with time series data of domestic and foreign weightlifters selected.

## **References**

- [1] Kim, K. H., Jeong, C. K., & Kang, H. S. (2013). Physical Fitness of Collegiate Soccer Athletes and Prediction of Performance Abilities. *The Korean Journal of Sport*, 11(3), 67-74.
- [2] Kim, E. J. & Cho, I. H. (2017). The Validation for Estimation of Performance Required for Ice hockey players on the Basis of Field test. *The Korean Journal of Sport*, 15(2), 893-906.
- [3] Park, S. H., Kim, K. H., & Huh, Y. (2009). College Ssireum Player's Strength Profile and Performance Predictions. *The Journal of Korean Alliance of Martial Arts*, 11(1), 213-225. DOI: 10.35277/kama.2009.11.1.213
- [4] Cho, Y. H., & Seo, T. B. (2020). Vertical jump and hand grip strength are indicators for predicting performance of Korea national weightlifters. *The Korea Journal of Sports Science*, 29(2), 1025-1035. DOI : 10.35159/kjss.2020.04.29.2.1025
- [5] Lee, S. H., & Kim, J. H. (2020). Seasonal Variation in Physiological Fitness of a Semi-Professional K3 League Soccer Team. *Journal of Digital Convergence*, 18(1), 357-368. DOI: 10.14400/JDC.2020.18.1.357

- [6] Kim, J. S., Lee, M. K. & Park, J. S. (2015). Effects of self-regulation with EEG neuro-feedback training on high school bowling player's cognitive and performance abilities. *The Korean Journal of Physical Education*, 54(6), 495-503.
- [7] So, Y. H. (2016). Relationship among Coach-Athlete Relationship, Exercise Passion, and Perceived Performance of High School Athletes. *The Korean Journal of Physical Education*, 55(5), 105-115.
- [8] Jung, J. H. (2005). The prediction of Psychological Factors for canoe players' performance. *The Korean Journal of Physical Education*, 44(6), 347-357.
- [9] Chang, J. K., & Lee, D. H. (2009). Predicted Performance Based on the Performance Level of Sprinter and Coaches' Professional Experience. *Korean Journal of Sport Science*, 20(2), 426-433. DOI: 10.24985/kjss.2009.20.2.426
- [10] Choi, Y. M., & Kim, S. H. (1998). Forecasting Accuracy of Single Variable Time - series Model for Tourism Demand ; Comparison of Actual and Forecasted Data of Korean Travelers Overseas. *The Tourism Sciences Society of Korea*, 21(2), 111-128.
- [11] Kim, M. C. (2013). A Study on the Relationship between Service Quality and Customers in the Sports Industry Using KS-SQI : Focused on Professional Baseball Teams, Golf Clubs, and Ski Resorts. *Journal of sport and Leisure Studies*, 51(1), 343-352.
- [12] Park, D. Y., & Kim, J. W. (2016). Analytic Study on the Demand Forecast of Pro-Baseball Spectators of KBO(2016-2020). *The Korean Journal of Sport*, 14(4), 415-426.
- [13] Park, D. Y., Jung, J. H., & Han, J. K. (2011). An Analysis of Demand Forecasting for Jejudo Leisure Sports Tourists by Using Time Series Analysis. *The Korean Journal of Sport*, 9(4), 147-159.
- [14] Yoon, H. J., & Ahn, J. K. (2013). Predicting the Horse Racing Revenue in Korea Using an ARIMA Model. *Korean Journal Sport Management*, 18(5), 1-10.
- [15] Nam, H. J., Lee, J. H., & Hong, D. S. (2018). The Effect of Using Mouth Guard on Athletic performance in Weight-Lifting Athletes. *The Korea Journal of sports Science*, 27(4), 833-844. DOI: 10.35159/kjss.2018.08.27.4.833
- [16] Chea, J. S., & Song, J. K. (2014). Performance Analysis and a Forecasting Model for The Short-Term Series in The Korean Professional Baseball League. *The Journal of Sport Science*, 25(1), 92-107. DOI: 10.24985/kjss.2014.25.1.92
- [17] Kim, M. C. (2009). Mediator effect analysis of teenagers' impulse buying of sporting goods in the on-Line purchase decision-making process. *The Korean Journal of Physical Education*, 48(4), 189-202.