A Study of Momentum Change in Sports Competitions Based on Evaluation and Correlation Analysis

Haonan Zhang [1,*], Yu Liu [2]

[1] College of Management, Ocean University of China, Qingdao, China, 266000

[2] Faculty of Information Science and Engineering, Ocean University of China, Qingdao, China, 266000

* Corresponding Author Email: zhanghaonan1207@stu.ouc.edu.cn

Abstract. In sports competitions, momentum has an important impact on the outcome of the game, and how to accurately capture and analyze momentum changes has become a key issue in research. But traditional methods often lack comprehensiveness and objectivity, so a new assessment model is needed to better understand and predict momentum changes. In this paper, an assessment model of momentum change is proposed, aiming to establish a comprehensive and objective momentum assessment model through data transformation, correlation analysis and weight calculation. The assessment model of momentum change is based on the data transformation technique, which converts score data into game data and generates 17 indicators to reflect various factors in the game. Meanwhile, using Spearman correlation coefficient and the Criteria Importance Through Intercriteria Correlation (CRITIC) algorithm, the weights of these indicators were assigned and correlation analysis was carried out to establish the Momentum Assessment Model (MAM). This paper takes the 2023 Wimbledon Men's Singles Final as an example to demonstrate the effectiveness of MAM in analyzing momentum changes in the match flow. The results show that the model can help to understand and predict the relationship between player performance and match results during a match, providing an important research tool for the field of sports and athletics.

Keywords: Spearman Correlation Coefficient, CRITIC algorithm, Momentum Assessment.

1. Introduction

The 36-year-old Servia tennis star Djokovic lost the men's singles final at Wimbledon 2023. The failure ended a remarkable run for one of the all-time great players in Grand Slams. In this exciting match, the momentum, which indicates the strength or force gained by motion or a series of events, influences the game's situation. Tennis is an intensely competitive sport often played with long periods of confrontation. Some incredible swings often exist in tennis, from favoring one player to another. These changes can be attributed to the momentum often possessed by the player with the advantage.

All indications are that momentum is crucial for the players. However, quantifying the perception of momentum or a dynamic edge during a game is challenging. Additionally, the way in which different occurrences within a game contribute to or alter this momentum is not easily discerned. This paper expects to unravel the mystery of momentum by establishing a mathematical model which can also be generalizable to other matches.

For a comprehensive and objective assessment of momentum to better understand and predict changes in momentum in sports competitions, this paper first converts the data corresponding to each score into the data corresponding to each game and generates 17 indicators that apply to the game. Then this paper further selected 5 of the 17 variables through variable selection and Spearman correlation coefficient. At last the weight of each variable is calculated through CRITIC algorithm to form the formula of MAM.

2. Momentum Assessment Model

This paper establishes the Momentum Assessment Model (MAM) to capture the flow of play (momentum) as points occur and apply it to one or more matches. MAM is based on the Data converting [1], Spearman correlation coefficient (SCC) and Criteria Importance Through Intercriteria Correlation (CRITIC) algorithm. The advantage of the CRITIC algorithm lies in its ability to effectively identify and evaluate the importance of different factors and their interrelations, thereby offering a more accurate and objective basis for decision-making [2, 3]. The details of MAM are plotted in Figure 1.

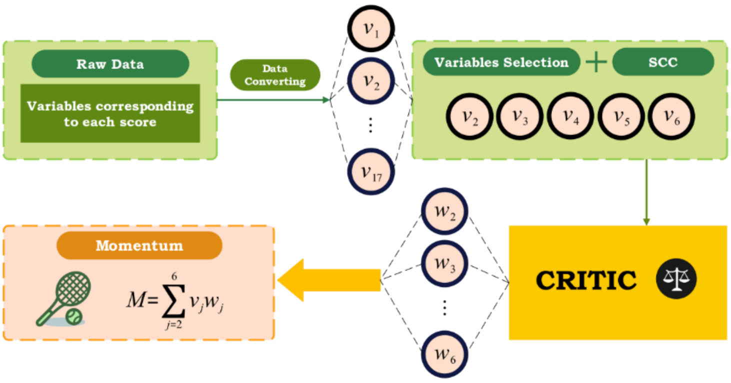

Figure 1. The details of MAM

2.1. Data Converting

Before establishing MAM, this paper converts the data corresponding to each score into the data corresponding to each game. Then this paper tweaks the variables and generates 17 indicators that apply to the game [4]. With these conversions, this paper turns integer variables into continuum variables. The turning facilitates interpolation and extrapolation to improve analysis accuracy [5, 6]. The 17 indicators $(v_j, j = 1,2,\ldots,17)$ based on player 1 as an example are shown in Table 1 (n is the number of times of scores in a game).

This paper defines the momentum of a match to be related only to the serve player, the real time score and the number of breaks of serve. Consequently, this study selects first six of these variables (v_1, v_2, v_3, v_4, v_5 and v_6) as factors affecting momentum. Through Spearman correlation coefficient plotted in Figure 2 this paper gets the correlation between v_1 and v_2 is 0.73 which is large, so this paper randomly removes v_1 and reserve v_2 [7, 8]. Consequently, v_2, v_3, v_4, v_5 and v_6 are brought into the CRITIC algorithm to calculate weights.

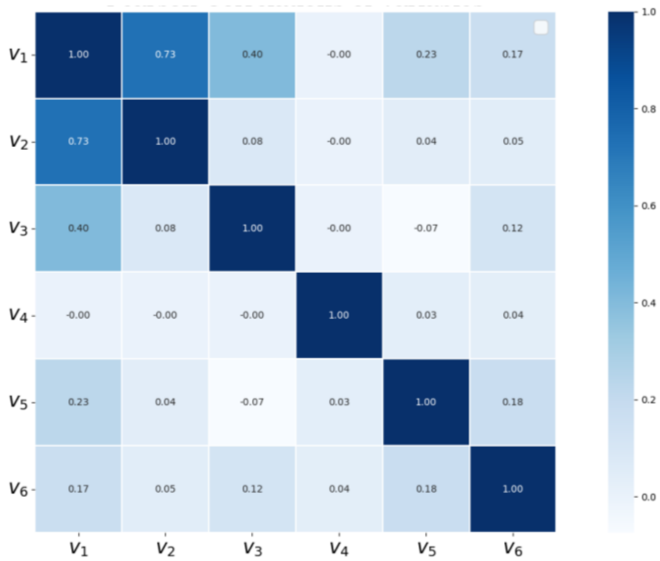

Figure 2. Spearman correlation coefficient results

Table 1. The 17 indicators based on player 1

Variable	Formula	Definition
v_1	$Max(P1\_points\_won) - Max(P2\_points\_won)$	Difference between the total number of scores of player1 & player2 up to the current game
v_2	$\sum_{i=1}^{n} P1\_sets - \sum_{i=1}^{n} P2\_sets$	Difference in sets won by player 1 & player 2
v_3	$\sum_{i=1}^{n} P1\_games - \sum_{i=1}^{n} P2\_games$	Difference in games won by player 1 & player 2
v_4	$server\ (1\ or\ 2)$	The server of a game
v_5	$\sum_{i=1}^{n} 1_{\{point\_victor=1\}}$	Total number of points won by player 1 in a game
v_6	$\frac{\sum_{i=1}^{n} P1\_break\_pt\_won}{n}$	The probability that player 1 won the break point
v_7	$\frac{\sum_{i=1}^{n} P1\_winner}{n}$	The probability that player 1 hit an untouchable winning shot
v_8	$\frac{\sum_{i=1}^{n} P1\_double\_fault}{n}$	The probability that player 1 missed both serves and lost the point
v_9	$\frac{\sum_{i=1}^{n} P1\_unf\_err}{n}$	The probability that player 1 made an unforced error
v_{10}	$\frac{\sum_{i=1}^{n} P1\_net\_pt\_won}{\sum_{i=1}^{n} P1\_net\_pt}$	The probability that player 1 won the point while at the net
v_{11}	$\frac{\sum_{i=1}^{n} P1\_ace}{n}$	The probability that player 1 hit an untouchable winning serve
v_{12}	$\sum_{i=1}^{n} P1\_break\_pt$	The total number of break point
v_{13}	$\mathrm{var}(serve\_width)$	Variance of direction of serve
v_{14}	$\mathrm{var}(serve\_depth)$	Variance of depth of serve
v_{15}	$\mathrm{var}(return\_depth)$	Variance of depth of return
v_{16}	$\sum_{i=1}^{n} P1\_distance\_run$	Player 1's total distance ran during a game
v_{17}	$\frac{\sum_{i=1}^{n} speed\_mph}{n}$	Player 1's average speed of serve

2.2. Mathematical Principles of CRITIC Algorithm

The Criteria Importance Through Intercriteria Correlation (CRITIC) method is mainly used to determine the weight of attributes [9]. In the present method, the attributes aren't in contradiction with each other, and the attributes' weights are determined using the decision matrix. In the CRITIC algorithm, there is no need for attribute independence, and the qualitative attributes are transformed into quantitative attributes [10].

The decision matrix is based on entering the method and expressing the alternatives and attributes are based on the information received from the decision maker, as shown in Equation (1).

$$X=\begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}_{m\times n} \quad ;\ i=1,\dots,m,\ j=1,\dots,n \tag{1}$$

The a_{ij} indicates the element of the decision matrix for ith alternative in jth attribute.

2.2.1 The Normalized Decision Matrix

In order to normalize the positive and negative attributes of the decision matrix, Equation (2) and Equation (3) are used respectively.

$$x_{ij}=\frac{a_{ij}-a_i^-}{a_i^+-a_i^-} \quad i=1,\dots,m, \quad j=1,\dots,n \tag{2}$$

$$x_{ij}=\frac{a_{ij}-a_i^+}{a_i^--a_i^+} \quad i=1,\dots,m, \quad j=1,\dots,n \tag{3}$$

The xij represents a normalized value of the decision matrix for ith alternative for jth attribute and $a_i^+ = \mathrm{Max}(r_1, r_2, \dots, r_m)$; $a_i^- = \mathrm{Min}(r_1, r_2, \dots, r_m)$.

2.2.2 The Correlation Coefficient

Then the correlation coefficient among attributes is determined by the following equation (Equation (4)).

$$\rho_{jk}=\frac{\sum_{i=1}^{m}(x_{ij}-\bar{x}_j)(x_{ik}-\bar{x}_k)}{\sqrt{\sum_{i=1}^{m}(x_{ij}-\bar{x}_j)^2\sum_{i=1}^{m}(x_{ik}-\bar{x}_k)^2}} \tag{4}$$

The $\overline{x_j}$ and $\overline{x_k}$ display the mean of *j*th and *k*th attributes. $\overline{x_j}$ is computed from Equation (5). Similarly, it is obtained for $\overline{x_k}$. Also, ρ_{jk} is the correlation coefficient between *j*th and *k*th attributes.

$$\bar{x}_j=\frac{1}{n}\sum_{j=1}^{n}x_{ij} \quad i=1,\dots,m \tag{5}$$

2.2.3 The Index C

First, the standard deviation of each attribute is estimated by Equation (6).

$$\sigma_j=\sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(x_{ij}-\bar{x}_j)^2} \quad i=1,\dots,m \tag{6}$$

Then, the index C is calculated using Equation (7).

$$C_j=\sigma_j\sum_{k=1}^{n}(1-\rho_{jk}) \quad j=1,\dots,n \tag{7}$$

2.2.4 The Weight of Attributes

The weights of attributes are determined by Equation (8).

$$w_j = \frac{C_j}{\sum_{j=1}^{n} C_j} \quad j = 1, \ldots, n \tag{8}$$

2.2.5 The Final Ranking of Attributes

The weights of attributes are arranged in descending order for the final ranking of attributes.

2.3. The Establishment and Application of MAM

After utilizing data converting and CRITIC algorithm, this paper figures out the weights of five selected new variables ($w_j, j = 2,3,\ldots,6$), which is plotted in Table 2. The sum of these weights is 1.

Table 2. The weights of five selected new variables based on player 1

Selected new variable	Weight	Selected new variable	Weight
v_2	$w_2 = 0.085$	v_5	$w_5 = 0.102$
v_3	$w_3 = 0.044$	v_6	$w_6 = 0.593$
v_4	$w_4 = 0.176$	-	-

According to the five selected new variables and the weight of each new variable sought, this paper quantifies the momentum in Equation (9). By using Min-Max Scaling method this paper scales the momentum to the range 0-1 and establish the MAM model which is shown below in Equation (10).

$$M' = \sum_{j=2}^{6} v_j w_j \tag{9}$$

$$M = \frac{M' - M'_{\min}}{M'_{\max} - M'_{\min}} \tag{10}$$

M' is the quantified momentum of a player. *M* is the scaled quantified momentum of a player (ranging from 0-1).

In order to verify the validity of the MAM model, this paper turns the attention back to the remarkable 2023 Wimbledon men's singles final and quantify the momentum of the two players and changes of match flow over the course of the match by means of the model, as shown in Figure 3.

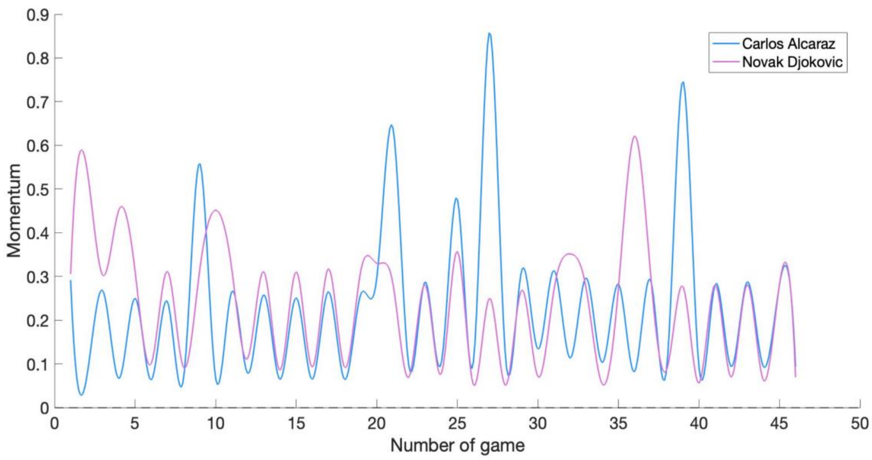

Figure 3. The visualization of match flow in the 2023 Wimbledon men's singles final

By reviewing the whole match process, this paper finds that players' momentum is closely linked to the match situation. In the first set (Game 1-7), the momentum of Djokovic was much higher than his opponent, so he easily dominated and won the set. In the tense second set (Game 8-20), the pair's momentum was neck-and-neck, with Carlos narrowly edging out Djokovic by a single point. After the first victory over his component, the momentum of Carlos seemed much higher, so he easily won the third set (Game 21-27). However, Djokovic's momentum grew over Carlos in the fourth set (Game 28-36) and won the set. In the last set (Game 37-46), perhaps out of a desire to win, Carlos' momentum became extremely high, and he eventually took the match. This match is a strong test to the effectiveness of MAM in capturing players' momentum and identifying the details of their performance in a match. The model can also be applied in analyzing the situation of other matches.

3. Conclusion

In conclusion, this research marks a significant contribution to sports analytics by combining advanced statistical and machine learning methods to develop sophisticated models for capturing and predicting momentum in tennis. The study first converts score data into match data and generate 17 indicators to reflect various factors in the match. At the same time, using the Spearman correlation coefficient and through CRITIC algorithm, the weights of these indicators are assigned and correlation analysis is performed to establish a momentum assessment model (MAM). This paper demonstrates the effectiveness of the MAM in analyzing momentum changes in the flow of a match, using the 2023 Wimbledon Men's Singles Final as an example. The results show that the model helps to understand and predict the relationship between player performance and match results during the match process, providing an important research tool for the field of sports and athletics.

References

[1] Noy O, Shamir R. Time-dependent Iterative Imputation for Multivariate Longitudinal Clinical Data [J]. arXiv preprint arXiv: 2304.07821, 2023.

[2] Žižović M, Miljković B, Marinković D. Objective methods for determining criteria weight coefficients: A modification of the CRITIC method [J]. Decision Making: Applications in Management and Engineering, 2020, 3 (2): 149-161.

[3] Pamucar D, Žižović M, Đuričić D. Modification of the CRITIC method using fuzzy rough numbers [J]. Decision Making: Applications in Management and Engineering, 2022, 5 (2): 362-371.

[4] Moss B, O'Donoghue P. Momentum in US Open men's singles tennis [J]. International Journal of Performance Analysis in Sport, 2015, 15 (3): 884-896.

[5] Zhang R. Result prediction and player analysis of tennis matches [D]. Yunnan University, 2022.

[6] Zhang Q. Analysis of win-loss factors of Federer vs Nadal in 2010-2017 season [D]. Nanjing Sports Institute, 2018.

[7] May J O, Looney S W. On sample size determination when comparing two independent Spearman or Kendall coefficients [J]. Open Journal of Statistics, 2022, 12 (2): 291-302.

[8] Chok N S. Pearson's versus Spearman's and Kendall's correlation coefficients for continuous data [D]. University of Pittsburgh, 2010.

[9] Krishnan A R, Kasim M M, Hamid R, et al. A modified CRITIC method to estimate the objective weights of decision criteria [J]. Symmetry, 2021, 13 (6): 973.

[10] Mishra A R, Rani P, Pandey K. Fermatean fuzzy CRITIC-EDAS approach for the selection of sustainable third-party reverse logistics providers using improved generalized score function [J]. Journal of ambient intelligence and humanized computing, 2022: 1-17.