Journal of Electronics and Information Science (2024)
Clausius Scientific Press, Canada
DOI: 10.23977/jeis.2024.090304
ISSN 2371-9524 Vol. 9 Num. 3

Research on Tennis Players' Momentum Calculation Model Based on Entropy Weight Method

Yanyan Lv[#,*], Zirui Yang[#], Tianming Zong[#]

Reading Academy, Nanjing University of Information Science and Technology, Nanjing, 210044, China

[*]Corresponding author: yylv@nuist.edu.cn

[#]These authors contributed equally.

Keywords: Entropy Weight Method, Momentum Calculation, Sports Performance

Abstract: Tennis, a globally revered sport, leverages the concept of momentum from physics, defined as "the force or strength gained through motion or a series of events." In the context of sports, momentum serves as a metric to delineate the performance trajectory of players over a given time frame, thereby reflecting the competitive edge of the athletes involved. This study presents the formulation of a computational model designed to estimate the momentum of tennis players during matches. The methodology commences with data preprocessing, encompassing the rectification of anomalous data points and the imputation of missing values. Subsequently, the paper elucidates the construction of a momentum estimation model, which encompasses the selection of pertinent indicators and a meticulous weight analysis. The authors have employed the entropy weight method to ascertain the relative importance of each indicator, subsequently devising a formula for momentum calculation grounded in these metrics. The paper culminates with a visual representation of the momentum dynamics, utilizing momentum change graphs and scatter plots to illustrate the fluctuations. The findings of this research offer valuable insights to tennis coaches and players, equipping them with a deeper comprehension of match dynamics and a strategic framework to enhance their competitive performance.

1. Introduction

Tennis, a sport adored globally, has not only captured the attention of audiences within the competitive arena but also enjoys widespread acclaim among amateur enthusiasts. Since its inception in mid-19th century Britain, tennis has evolved into a global phenomenon encompassing millions of participants and spectators. Beyond the realm of athletic competition, tennis has become a nexus for social, cultural, and economic activities. The sport of tennis captivates global sports enthusiasts with its unique competitive nature. It demands from athletes exceptional physical fitness, rapid reflexes, and considerable stamina, while also necessitating sound tactical awareness and psychological fortitude. Whether in singles or doubles competitions, tennis showcases the seamless integration of individual skill and teamwork. In the socio-cultural sphere, tennis embodies the societal values of fairness, competition, intelligence, and etiquette[1]. It transcends being merely a sport to become a hallmark of social interaction and lifestyle.

According to the previous research, the factors affecting the player's psychology in a tennis match mainly include subjective and objective aspects. Subjective factors include pre-match preparation, technical and tactical level of play, physical ability and will quality. Reasonable pre-match goal setting, adequate pre-match preparation, strong physical reserve and firm will quality are crucial to athletes' performance in the match[2]. Objective factors, on the other hand, include the competition environment, referees, coaches and competition opponents. These factors collectively affect the athletes' psychological state, which in turn affects their performance. Some studies have analyzed Djokovic's winning factors in matches using the entropy weight TOPSIS model[3]. The study shows that Djokovic's winning factors mainly include the success rate of serve retention, the first-serve percentage, the ACE rate, the winning point percentage, the first-serve percentage, and the break success rate. Through statistical modeling analysis, the study reveals the key role of Djokovic's superior performance in the serving phase and under pressure in winning his matches. This study provides tennis players with empirical insights to enhance their match performance, such as improving the threat of serving, ensuring the stability of hitting the ball, and adjusting the match mindset at the right time.

On the basis of the previous research, this paper establishes a momentum index system based on the entropy weight method. For the 2023 Wimbledon men's singles match data, five aspects, including serve, error, technical ability, mentality and physical fitness, are selected as the first-level indexes, which are further refined into 13 second-level indexes. Through the entropy weighting method, the relative importance of each indicator was determined, and the momentum calculation formula was constructed accordingly. This formula combines the weight of each indicator with the performance value of the athlete at each time point to calculate the momentum score of the athlete during the game. This quantitative method not only provides coaches and athletes with a real-time basis for tactical decision-making, but also intuitively demonstrates the fluctuation of momentum during the game by visualizing the momentum change graph and scatter plot. In addition, at critical moments, significant changes in momentum are often closely associated with decisive scores or unforced errors, which provides important reference information for strategic planning. Through the analysis of these data, it is possible to better understand the performance of the athletes during the game and to use the momentum data for further game strategy optimization and athlete performance evaluation.

2. Data Source and Preprocessing

2.1. Data Source

This study's data is sourced from the official website of the Consortium for Mathematics and Its Applications (COMAP) Undergraduate Competition. The dataset used is from the 2023 Wimbledon Men's Singles matches after the second round. It includes 31 matches between Carlos Alcaraz and Nicolas Jarry, totaling 7284 records, including match scores and statistics for both players.

2.2. Outlier Handling

The dataset was first preprocessed. It was found by manual subjective judgment that the match numbered 1303 was longer than 24 hours, the match numbered 1304 had an error in the win/loss relationship; and the data of the match numbered 1403 was incomplete. In order to ensure the continuity of the analysis and the accuracy of the data results, it was decided to remove the data of these three matches from the dataset when processing the data.

2.3. Missing Value Handling

Using the pandas library in Python to handle missing values is considered a common and efficient method. After conducting a missing value inspection, it was found that the columns ‘speed_mph’, ‘serve_width’, ‘serve_depth’, and ‘return_depth’ contain missing values. We employed cubic spline interpolation and forward fill methods to address these issues, respectively. This resulted in a complete dataset.

3. Establishment of Momentum Calculation Model Based on Entropy Weight Method

3.1. Establishment of Momentum Index System for Tennis Players

In order to effectively measure the athletic performance of tennis players, five aspects of tennis players' serve, error, technical ability, mentality, and physical fitness were constructed as Primary indicators, as shown in Table 1. Serving is a crucial technical aspect of the game; the error rate reflects the player's stability during the match; technical capabilities ensure the player's optimal performance; a positive mindset can enhance overall performance; physical fitness directly impacts the player's athletic performance. These five primary indicators are further refined into 13 secondary indicators. Among these, the secondary indicators under the primary categories of physical fitness and errors are negative indicators, while all other indicators are positive. Higher values of positive indicators indicate better performance, whereas higher values of negative indicators indicate poorer performance.

Table 1: Indicator selection

Primary indicators	Secondary indicators	Description
Serve(A1)	server(B1)	server of the point
	ace(B2)	player hit an untouchable winning serve
Mentality(A2)	sets(B3)	Sets won by player
	games(B4)	Games won by player
	points(B5)	number of points won by player in match
Technical capabilities(A3)	break_pt_won(B6)	player won the other player is serving
	net_pt_won(B7)	player won the point while at the net
	winner(B8)	player hit an untouchable winning shot
Physical fitness(A4)	distance_run(B9)	Player’s distance ran during point (meters)
	rally_count(B10)	number of shots during the point
Error(A5)	double fault(B11)	player missed both serves and lost the point
	unforced error(B12)	Player made an unforced error
	break_pt_missed(B13)	player missed an opportunity to win a other player is serving

3.2. Dimensionless Processing

The primary purpose of dimensionless processing is to eliminate the effects caused by different measurement units [4], ensuring consistency and comparability of data during the analysis process.

This provides a solid foundation for subsequent statistical analysis and decision support[5]. In the construction of an evaluation index system, since the indicators have different dimensions and magnitudes, it is necessary to perform dimensionless processing on the raw data. This not only removes dimensions but also standardizes the data. Dimensionless processing has different methods for positive and negative indicators.

For positive indicators,

$$x_{ij} = \frac{x_{ij} - \min x_{ij}}{\max x_{ij} - \min x_{ij}} \tag{1}$$

For negative indicators,

$$x_{ij} = \frac{\max x_{ij} - x_{ij}}{\max x_{ij} - \min x_{ij}} \tag{2}$$

Table. 2 below is a demonstration of dimensionless processing data, showcasing a selection identified by the code 2023-wimbledon-1304.

Table 2: Data table after dimensionless processing

elapsed_time	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13
0:00:00	0	0	0	0	0	0	0	0	0.91	0.92	1	1	1
0:00:38	0	0	0	0	0.0059	0	1	0	0.90	0.92	1	1	1
0:01:01	0	0	0	0	0.0119	0	0	0	0.78	0.76	1	1	1
0:01:31	0	0	0	0	0.0179	0	0	1	0.86	0.46	1	1	1
0:02:21	1	0	0	0.17	0.0179	0	0	0	0.92	0.85	1	1	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...

3.3. Entropy Weighting Method for Determining Indicator Weights

Entropy weight method is an objective assignment method, which can avoid the interference of human factors in evaluating the weights of indicators[6], and is widely used in the field of comprehensive evaluation of multiple indicators[7][8]. Entropy comes from the thermodynamic concept in physics, which mainly reflects the degree of chaos of the system. In information theory, entropy is used to reflect the size of information, the more information an indicator carries, the greater its role in decision-making, and the greater the weight. Entropy weighting method utilizes this information entropy theory to determine the weight of each indicator by calculating the information entropy of each indicator, so as to evaluate the system in a scientific and reasonable way. The following are the steps of entropy weight method:

Let there be a total of n evaluation objects and m evaluation indicator variables, and the value of the i^{th} evaluation object with respect to the j^{th} indicator variable is $x_{ij}(i = 1,2,\dots,n;\ j = 1,2,\dots,m)$. Construct the data matrix $A = (a_{ij})_{n \times m}$.

Calculation of the weight of the ith object with respect to the value of the j^{th} indicator

$$p_{ij} = \frac{y_{ij}}{\sum_{i=1}^{n} y_{ij}} \tag{3}$$

y_{ij} denotes the indicator value. p_{ij} denotes the entropy value of the j^{th} indicator. The entropy value reflects the information confusion degree of the indicator, the larger the entropy value is, the more dispersed the information of the indicator is, and the smaller the influence on the comprehensive evaluation is; the smaller the entropy value is, the more concentrated the information of the indicator is, and the larger the influence on the comprehensive evaluation is. The entropy value calculation formula is as follows:

$$e_j = -\frac{1}{\ln(n)} \sum_{i=1}^{n} p_{ij} \ln(p_{ij}) (e_j > 0) \tag{4}$$

Compute the information entropy redundancy:

$$g_j = 1 - e_j \tag{5}$$

Calculate the weight of the j^{th} indicator. The weight reflects the contribution of each indicator to the system evaluation, and the entropy weight method determines the weight of each indicator through its information entropy. The weight calculation formula is as follows:

$$w_j = \frac{g_j}{\sum_{j=1}^{m} g_j} \tag{6}$$

In equation (6), m is the total number of indicators. This formula calculates the weight of each indicator by normalizing the coefficient of variation.

The weights of the indicators calculated by the entropy weighting method above are shown in Table 3.

Table 3: Weights of each indicator obtained by entropy weight method

Primary indicators	Secondary indicators	Weight	Order
A1	B1	0.2241	1
	B2	0.0645	7
A2	B3	0.1629	3
	B4	0.0924	5
	B5	0.0580	8
A3	B6	0.0203	9
	B7	0.0955	4
	B8	0.1752	2
A4	B9	0.0046	13
	B10	0.0049	12
A5	B11	0.0100	11
	B12	0.0726	6
	B13	0.0144	10

In terms of the weight distribution of momentum indicators affecting tennis players, the indicators that have the most important impact on momentum are serve (server), winner (winner) and the number of sets (sets) that players have won. This is followed by the success of the net catch (net_pt_won), the number of games the player has won (points_won), unforced errors(unf_err), ACE balls(ACE) and the number of points the player has won(points_won). The metrics with relatively weak effects are breaking serve points (break_pt_won), breaking serve errors (break_pt_missed), double faults (double_fault), number of shots hit (rally_count) and distance run (distance_run).

From the perspective of the first-level indicators, the greater impact on the player's momentum is the serve, mentality and competition performance, while the impact of physical fitness and errors is relatively small. As a key technique in tennis, the serve directly affects the rhythm of the game and the player's self-confidence. Therefore, the ability to serve is often an important factor in determining whether a player can take the initiative in a match. Mentality is the player's mental state when facing pressure. A good mentality will help the player to play well at critical moments, but on the contrary, it may lead to an increase in errors and affect the overall performance. Competition performance is a combination of the player's technical, tactical and psychological qualities in the game, and is an important indicator of the player's overall strength [9]. On the other hand, although physical fitness has a certain impact on the player's ability to sustain combat, professional athletes usually undergo systematic training, have a strong physical fitness base, and are able to maintain a high level of athleticism during a long period of competition[10]. Although errors are an unavoidable part of the game, professional players are usually able to effectively adjust their mindset and reduce the negative impact of errors on their overall performance through long-term psychological training and accumulation of game experience. This is also in line with the theoretical performance of professional tennis players, indicating that professional tennis players are able to afford a certain amount of exercise, and can also overcome the psychological pressure caused by errors to a certain extent[9][10].

3.4. Momentum Calculation Model

Equation (7) performs a weighted summing operation between the indicator weights w_j and the player's value y_{ij} after the dimensionless processing of each indicator to obtain the player's momentum for each time period:

$$S_i = \sum_{j=1}^{m} w_j y_{ij} \tag{7}$$

Table 4: The momentum of player1 and player2 at 2023-wimbledon-1308

match_id	player1	player2	elapsed_time	p1_momentum	p2_momentum
2023-wimbledon-1308	Laslo Djere	Stefanos Tsitsipas	0:00:00	0.482	0.4091
2023-wimbledon-1308	Laslo Djere	Stefanos Tsitsipas	0:00:16	0.3116	0.0226
2023-wimbledon-1308	Laslo Djere	Stefanos Tsitsipas	0:00:39	0.0981	0.101
2023-wimbledon-1308	Laslo Djere	Stefanos Tsitsipas	0:01:08	0.3981	0.1549
2023-wimbledon-1308	Laslo Djere	Stefanos Tsitsipas	0:02:20	0.2555	0.535
...	...	...	...	...	...
2023-wimbledon-1308	Laslo Djere	Stefanos Tsitsipas	2:07:24	0.7891	-0.2816
2023-wimbledon-1308	Laslo Djere	Stefanos Tsitsipas	2:08:59	0.4111	0.1037
2023-wimbledon-1308	Laslo Djere	Stefanos Tsitsipas	2:09:31	-0.1372	0.5183
2023-wimbledon-1308	Laslo Djere	Stefanos Tsitsipas	2:10:16	-0.1229	0.4013
2023-wimbledon-1308	Laslo Djere	Stefanos Tsitsipas	2:10:56	0.0522	0.2447

This approach is able to combine the effects of all the metrics to provide a composite score for each sample, which can be used for further analysis and comparison. As an example, for the 2023-wimbledon-1308 game, the momentum calculation model was used to calculate the momentum of the two batsmen for each ball in that game is shown in Table 4.

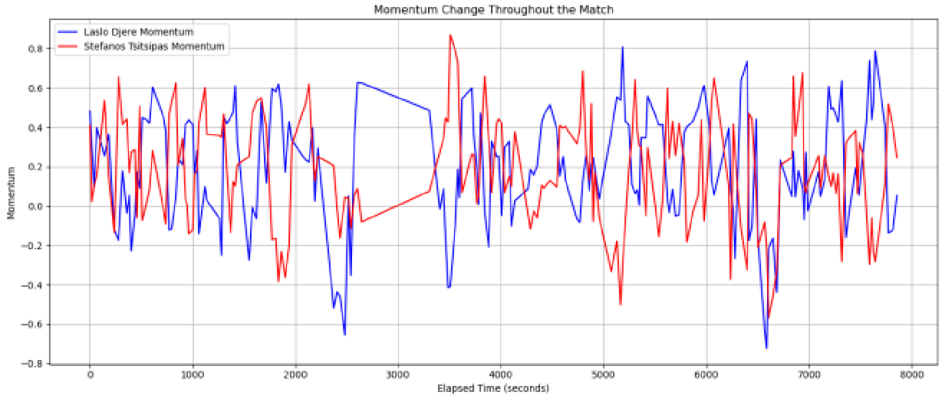

Figure 1: Momentum of player1 and player2 at 2023-wimbledon-1308 (line chart)

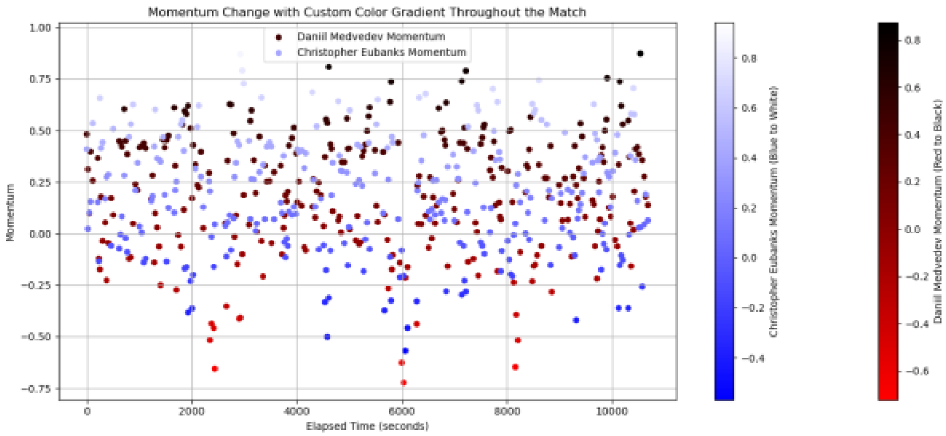

Figure 2: The momentum of player1 and player2 at 2023-wimbledon-1308 (Scatter plot)

Figure 1 and Figure 2 show the change in momentum of both players in 2023-wimbledon-1308 in seconds, the blue line represents the momentum of Laslo Djere and the red line represents the momentum of Stefanos Tsitsipas. It is not difficult to see from the information in the figures that the momentum of the two players, except for the similarity of the first ball, show a clear symmetry in the later balls. This is very much in line with the theory that there is a significant change in momentum after a ball is won or lost. The momentum of the two players alternated several times throughout the match, indicating the intensity of the match. At critical moments (e.g., 0:02:20), large changes in momentum may be related to key scores or errors made by the players during the game, and this information is an important reference for coaches and players to formulate their game strategies. Through the analysis of these data, we can better understand the players' performance in the game, and the momentum data can be used for further game strategy optimization and player performance evaluation.

4. Conclusions

This study presents an exhaustive analysis of the dataset from the 2023 Wimbledon Men's Singles Championship. The methodology encompassed data preprocessing, the construction of an evaluative index system, normalization, and the application of the entropy weight method to ascertain the indices' weights. A momentum calculation model predicated on the entropy weight approach has been developed to quantitatively gauge the performance of tennis players at discrete temporal junctures, thereby offering a framework for strategic adjustments. This model has been applied to perform an integrated assessment of the performance of both competitors in a given match.

The momentum calculation model formulated in this research integrates multiple facets of a player's performance, including serving proficiency, mental fortitude, technical acumen, physical stamina, and error rates. By computing a composite evaluation metric, a singular score is assigned to each instance, reflecting the dynamic shifts in a player's momentum throughout the match. This score serves as an intuitive indicator for coaches and players to make real-time tactical decisions. The momentum analysis conducted in this paper elucidates the nexus between the fluctuation of momentum during a match and its outcome. Notably, at pivotal junctures, pronounced shifts in momentum are often intricately linked to decisive points or unforced errors, offering significant insights for strategic planning.

As big data and artificial intelligence technologies advance, the momentum calculation model introduced herein is poised for further refinement. The integration of an expanded dataset and more granular player analytics will enhance the model's precision in tracking performance trajectories, thereby affording more nuanced tactical insights to coaches and athletes. The model's forecasting capabilities are also anticipated to be augmented, enabling preemptive strategic planning by predicting match dynamics and thereby increasing the likelihood of victory.

Furthermore, the methodologies and models presented in this paper are amenable to extrapolation to other athletic disciplines, offering novel analytical tools and perspectives for performance assessment and tactical analysis across various sports. With the evolution of sports science, the prospect of achieving a higher degree of personalization and intelligence in athlete training and match strategy development is within reach. This advancement is expected to elevate the overall standard of competitive sports performance.

References

[1] Yang Fang, Lei Shaohua. A Sociocultural Examination of Tennis [J]. Sports and Science, 2005(02):71-72+80.

[2] Xu Jiawei, Zhang Zhiyuan. Analysis of Factors Affecting Athletes' Psychology in Tennis Matches and Countermeasures [J]. Ice and Snow Sports Innovation Research, 2023, (15): 164-166.

[3] Zhang Penghui. Analysis of Novak Djokovic's Winning Factors under the Entropy Weight TOPSIS Model and Experience Inspiration for Chinese Tennis Players[C]// Shaanxi Provincial Sports Science Society, Shaanxi Student Sports Association. Proceedings of the Second Shaanxi Provincial Sports Science Conference (Special Topic 2). School of Physical Education, North University of China; 2024: 11.

[4] Guo Yajun, Yi Pingtao. Analysis of the properties of the linear dimensionless method [J]. Statistical Research, 2008, 25(2): 93-100.

[5] Li Qing, Chen Pengyu. The impact of indicator dimensionless methods on entropy weight evaluation results: A case study of journal academic influence evaluation [J]. Science and Technology Information, 2019, 17(7): 184-186.

[6] Zhang Sui, Zhang Mei, Chi Guotai. A scientific and technological evaluation model based on the entropy weight method and its empirical study [J]. Journal of Management, 2010, 7(01): 34-42.

[7] Zhao Teng, Yang Shizhong. Application of the entropy weight TOPSIS method in the evaluation of enterprise financial risk: A case study of JGJ Company [J]. Finance and Accounting Monthly, 2019(03): 9-16.

[8] Tang Cheng. Dynamic evaluation of sustainable agricultural ecological development in the Yangtze River Economic Belt: An evaluation method based on the entropy weight method [J]. Modern Agriculture, 2024(06): 40-44.

[9] Xia Lei. Research on the Special Physical Characteristics and Training Strategies of Competitive Tennis [J]. Bulletin of Sports Science and Technology Literature, 2024, 32(04): 91-94+201.

[10] He Bohan, Wang Wenzeng. Exploring the Relationship between Competitive Pressure and Unforced Errors in Tennis Players and Their Psychological Intervention Strategies [C]// Chinese Society of Sports Science. Abstracts of the 13th National Sports Science Conference—Poster Session (Sports Psychology Division) (2). Graduate Affairs Department, Party Committee of Shenyang Sport University; School of Sports Health, Shenyang Sport University; 2023: 3.