Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 6
ISET 2024

Research on Quantifying and Evaluating Tennis Match Momentum Based on Comprehensive Analytical Methods

Shaoxukang Liu[1, *, #], Houjiang Liu[1, #], Hong Cao[2, #]

[1]School of Science, Southwest Petroleum University, Chengdu, China, 610500

[2]School of Electronic Engineering and Information, Southwest Petroleum University, Chengdu, China, 610500

*Corresponding author: 17828834170@163.com

#These authors contributed equally

Abstract. Momentum is crucial in tennis matches, influencing player performance and match outcomes. However, quantifying and evaluating momentum has long been a challenge in sports science research. This study aims to quantify momentum in tennis matches by establishing a hierarchical analysis-based model for momentum flow. The model analyzes factors affecting momentum through five criteria: Service Level, Catch Level, Competition Result, Finishing Ability, and Winning Streak, with respective weights of approximately (0.1, 0.1, 0.2, 0.4, 0.2). Subsequently, the study assigns scores to these criteria based on different actions within a match. By multiplying the weights and scores, the momentum of each player is quantified, and the ratio of the two players' momentum provides a measure of momentum flow during the match. To assess the impact of momentum on match outcomes, the study employs k-means++ clustering analysis, violin plot analysis, and point-biserial correlation analysis. The point-biserial correlation analysis reveals a positive correlation coefficient of 0.66 between a player's momentum and their victory. Similar results are derived from the violin plot and clustering analysis. Therefore, it can be concluded that there is a significant association between the direction of momentum flow and a player's success in a match.

Keywords: momentum, AHP, k-means++, violin plot analysis, clustering analysis.

1. Introduction

Tennis is a highly competitive sport, where the rivalry between players depends not only on skill level but also on mental state, the pace of the game[1], and changes in momentum. Momentum plays a crucial role in the game, it can significantly affect the performance of the players and the outcome of the game. However, the quantification and evaluation of the concept of momentum in practical applications has always been one of the difficult problems in sports science research. Momentum is usually defined as the sustained advantage gained by a player throughout a game, and it reflects not only the player's performance in the game but also their state of mind and confidence level. At a high level of tennis, momentum shifts tend to be a key factor in the outcome of the match. Quantifying and evaluating momentum effectively not only helps coaches develop more scientific tactical strategies but also helps players better adjust their status during the game.

Although the concept of momentum is well known, how to quantify and evaluate it scientifically remains a complex problem. Traditional methods, including the entropy weight method[2] are difficult to effectively quantify dynamic and complex indicators. For this study, this paper started to use the analytic hierarchy process, point double column correlation coefficient analysis.

2. Multi-Factor Analysis Method for Tennis Match Momentum Evaluation

2.1. Data Collection

The data collected in this study also includes the data set of the women's singles competition at the Wimbledon Open in 2023. The name of the database and the source of the data are as shown in Table 1 Data source:

Table 1 Data source

Database Names	Database Websites
Wimbledon women's singles statistics	https://www.wimbledon.com/index.html

The Analytic Hierarchy Process (AHP) is a systematic approach widely used in decision analysis and was developed by American operations researcher Thomas L. Saaty in the early 1970s. By decomposing a complex decision problem into smaller parts that are easier to analyze, the method constructs a hierarchical structure model of the problem, to make a choice or prioritize among multiple criteria and alternatives. The core of AHP is to use the pairwise comparison matrix to quantify the subjective judgment of decision-makers on the importance of criteria, and calculate the relative weight of criteria and alternatives through mathematical methods, and finally achieve the purpose of quantitative decision-making[3].

In the momentum analysis of tennis matches in this study, the AHP method can provide a systematic framework to evaluate and compare multiple factors that affect the momentum of a match. Given that momentum is determined by a multidimensional concept, including but not limited to factors such as serve results, catch ability, scoring ability, and streak ability, AHP allowed this study to incorporate these complex factors into a unified evaluation system [4].

2.2. The Establishment of Model

Step 1: Establish the hierarchy

Goal Level 1: Assesses the momentum of the server and receiver in the game.

Criteria Level 2: These include Service Level, Catch Level, Finishing Ability, Competition result, and Winning Streak. These criteria were selected for this study for the following reasons:

Service Level: The service is very important in tennis because it is the start of every point. A player with a good serve can build momentum by dominating the game with a powerful serve. The evaluation of service level takes into account not only the winning serve but also the service errors, to fully reflect the influence of the player's performance in the service session on the momentum of the match.

Catch Level: The catch level reflects the player's ability to respond to the opponent's serve and baseline play. Technical points such as volleys at the net and backhand winning returns can significantly change the momentum of a match. In addition, pressure turnovers may weaken momentum, so the level of receiving is directly related to the control of the game.

Competition Result: The result of the competition reflects the actual score and is the final determinant of the performance of the contestants. Momentum is often reflected in rapid changes in scores, so the magnitude of momentum can be directly assessed by observing how often and how well a player scores during a match.

Finishing Ability: Performance on key points is particularly important as they are often the key moments in a game where momentum shifts. How players perform on match points is directly related to whether they can capitalize on the momentum they have built or turn the tide against them.

Winning Streak: The ability to win multiple points in a row or streak effectively builds and maintains momentum. When a player wins a few points in a row, it not only boosts their confidence but may also affect the mentality of the opponent, thus creating a psychological advantage for the player.

Level 3: By consulting information[5, 6] and comparing the difficulty and benefits of different behaviors in the game, this study obtained the detailed performance scores of the server and receiver as shown in Table 2 Performance scores:

Table 2 Performance scores

Criterion	Symbol	Scoring
Service Level	a_s	Winning serve +1; Every missed serve -1
Catch Level	a_c	Net +1 (if win +2); Winning return forehand +1(if backhand +2); Oppressive blunders -1
Competition Result	a_r	The server loses -1 and the receiver +1
Finishing Ability	a_f	Match point score +1, match point loss -1
Winning Streak	a_w	Win two games in a row +1, win three games in a row +2

It is worth noting that due to the natural advantage of the server, no extra points are given in this study for the cases where the server scores or the receiver loses points.

The criterion layer matrix is obtained through the above study $A=\{a_s,a_c,a_r,a_f,a_w\}$.

Step 2: Construct the judgment matrix

Each factor of the criterion layer is compared pairwise to determine the relative importance of their influence on the target layer. A 1-9 scale is used, where 1 means that two factors are equally important and 9 means that one factor is much more important than the other. In this study, by consulting relevant information[7, 8] and evaluating the relative importance of each factor, the judgment matrix is obtained as follows:

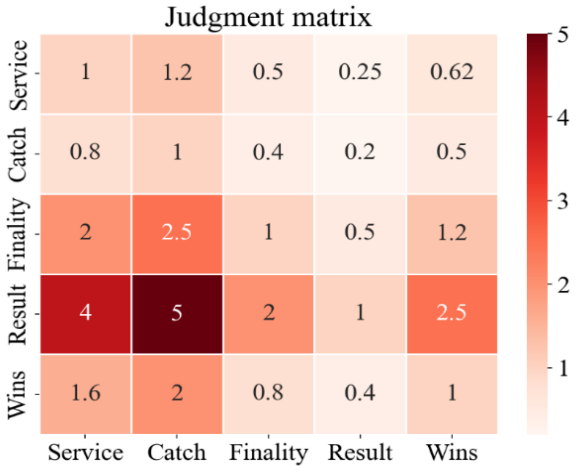

Figure 1 Judgment matrix

Step 3: Hierarchical single sorting and consistency check

In this study, the maximum eigenvalue and its corresponding eigenvector are calculated for each judgment matrix, and the eigenvector is normalized and used as the weight of each factor. Then a

consistency check is performed to ensure the consistency of the judgment. If the Consistency ratio (CR) is less than 0.1, the judgment matrix is acceptable.

The CR value of the judgment matrix in this study is 0, which means that the agreement between pairwise comparisons is perfect and passes the consistency test.

The weight vectors calculated by the eigenvalue method in this study are shown in Table 3 Weight vectors:

Table 3 Weight vectors

Criterion	Weights
Service Level	0.10638298
Catch Level	0.08510638
Competition result	0.21276596
Finishing Ability	0.42553191
Winning Streak	0.17021277

Hence the set of weights corresponding to the criterion layer.

$W = \{0.10638298, 0.08510638, 0.21276596, 0.42553191, 0.1702127\,7\}$

Step 4: Comprehensive score calculation

In this study, the comprehensive score of momentum between the server and the receiver is calculated based on the weights of the criterion layer and the scores of the corresponding factors in the scheme layer. Weighted sum of serving level, receiving level, match result, scoring ability, and winning streak ability $M = A \cdot W^T$. Since tennis usually involves two players playing against each other, the weighted scoring results in this study are divided into 1 M_1 for player 1 and M_2 player 2.

Step 5: Ratio calculation reflects the momentum flow

Since the game is a confrontation process, the flow of momentum in the game is also determined by the momentum of the two players themselves. By calculating the ratio of the comprehensive score of the momentum of the server and the receiver $R = \frac{M_1}{M_2}$. This study can quantify the momentum flow of the game. A ratio greater than 1 indicates that the server is dominant, while a ratio less than 1 indicates that the receiver is dominant.

2.3. The Solution of Model

Through the above steps, this study obtained the visualization of the momentum flow of match_id 1303, where the momentum chart shows the change of the momentum score of the players in the game, and the momentum ratio is the logarithm of the ratio to 10, showing the momentum flow in the game as shown below:

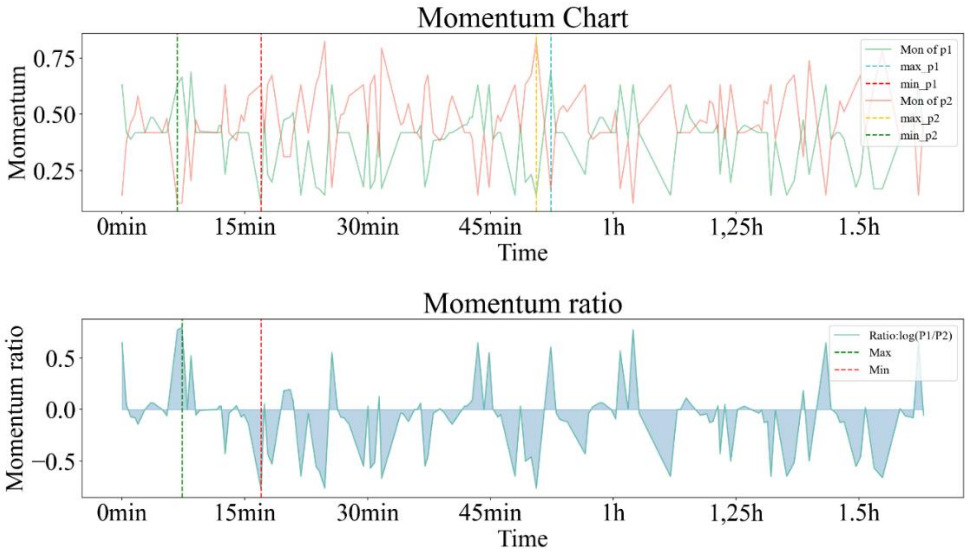

Figure 2 Momentum flow in the game

For a Momentum Chart:

Player momentum score change: The top of the chart shows the change in momentum score for both players throughout the game. The momentum of p1 is shown in green and the momentum of p2 is shown in red. In this study, it can be seen that the momentum scores of the two players fluctuate significantly during the game, and the peak and valley values of momentum alternate, showing the constant conversion of the game momentum between the two players; Fluctuating momentum score: A rapid change in momentum may indicate multiple tense moments in the match, such as break, save, or key point contests, which can cause a rapid shift in momentum.

For the Momentum Ratio plot:

Change in momentum flow: The bottom of the chart shows the change in the momentum ratio, which is the log value of 10 for Mon_p1 and Mon_p2. Positive values indicate that p1 has stronger momentum relative to p2, while negative values indicate that p2 has stronger momentum relative to p1; Momentum ratio fluctuation: The momentum ratio fluctuation (alternating positive and negative values) shows how momentum flows throughout the game. In this study, it can be observed that the momentum ratio sometimes increases and sometimes decreases, which means that the dominance of the momentum in the game switches between the two players.

In conclusion, this study can draw the following conclusions:

There are obvious ups and downs in the momentum of the two players at different stages of the game, which may be related to their game status, psychological state, and key events in the game; The momentum ratio graph shows the dynamics of the momentum flow and provides an intuitive representation of the momentum change during the game. The fluctuation of the ratio indicates that the momentum of the game is not dominated by one side, but the result of both sides; This dynamic change in momentum is crucial information for coaches and players, as it helps them to identify key moments in the game and adjust their tactics accordingly.

In this study, the momentum distribution in the game can also be visually displayed through the histogram as follows:

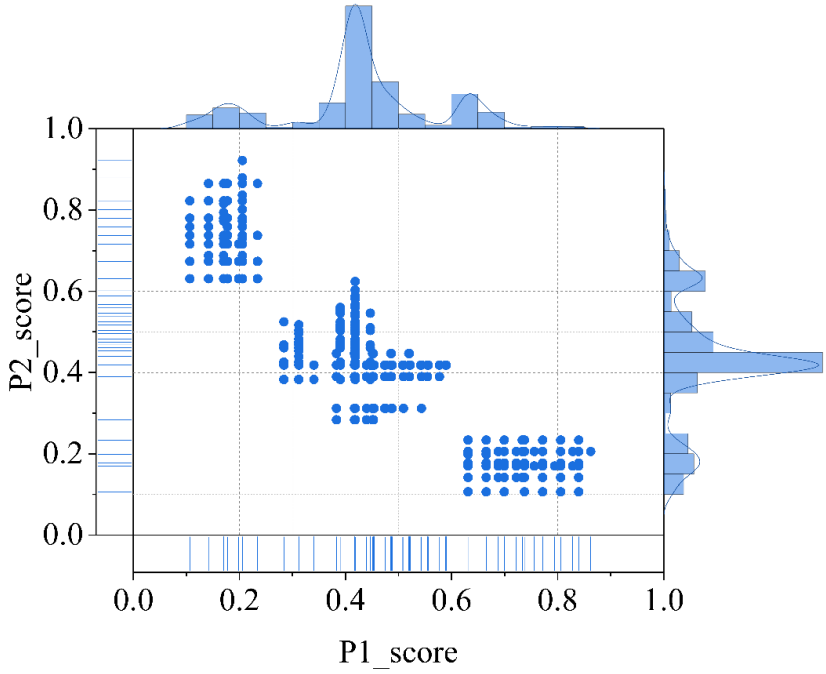

Figure 3 Momentum distribution

As shown here, most of the data points are concentrated in a few specific areas, which may indicate that there are several key moments or stages in the game where the momentum of the players tends to concentrate on a certain level of scoring. The histogram also shows a higher frequency in areas with higher and lower scores, possibly pointing to clear momentum swings in the game, where players were able to take or lose full control of the game at certain points.

3. Clustering and Correlation Evaluation of Tennis Match Momentum

3.1. K-means ++ cluster analysis

The K-means++ algorithm is a classical unsupervised learning algorithm used to partition data points into a predetermined number of clusters (clusters), where each data point belongs to the cluster corresponding to its nearest cluster center (centroid). The goal of the algorithm is to minimize the sum of the distances between data points within a cluster and their cluster centers [9]. At the same time, the optimal number of classes is determined by the contour coefficient map.

In this study, the results of the competition can be divided into four categories according to K-means ++: p1 wins and the momentum gap is large (purple), p1 wins and the momentum gap is small (green), p2 wins and the momentum gap is large (blue), p2 wins and the momentum gap is small (yellow), as shown in the figure, and their centroid coordinates are respectively as shown in Table 4 Centroid coordinates:

Table 4 Centroid coordinates

Color	P1 Momentum	P2 momentum	Scorer
Purple	0.65914311	0.17485127	1
Green	0.45352701	0.40849849	1
Blue	0.17300512	0.66188977	2
Yellow	0.40725887	0.46250626	2

Where the X-axis represents p1 momentum, the Y-axis represents p2 momentum, and the z-axis represents the winner (1 is p1 score, 2 is p2 score).

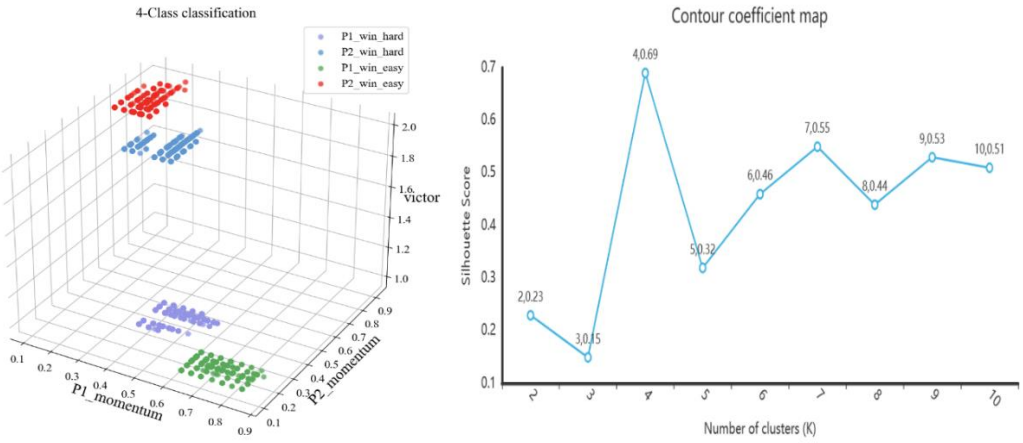

Figure 4 K-means cluster map and Contour coefficient map

From the centroid coordinates of the graph and classification, this study can analyze the following:

A player wins when their momentum is greater than the momentum of their opponent.

When the momentum of the two is similar, the momentum value distribution of the winner also tends to become larger and the momentum value of the other side becomes smaller, which indicates that the trend of the game is related to the relative momentum value between the players.

Overall, when p1 has much more momentum than p2, it is almost impossible for p2 to win.

3.2. Point bi-column correlation coefficient analysis and violin plot analysis

To further determine the role of momentum in the game, this study obtained the momentum score of p1 when p1 and p2 were the opponents in a game through the point-double column correlation coefficient [10], and the point-double column correlation coefficient matrix between p2's momentum score and p2's victory was as follows. At the same time, the momentum distribution is observed by violin plot according to the classification of players' wins and losses.

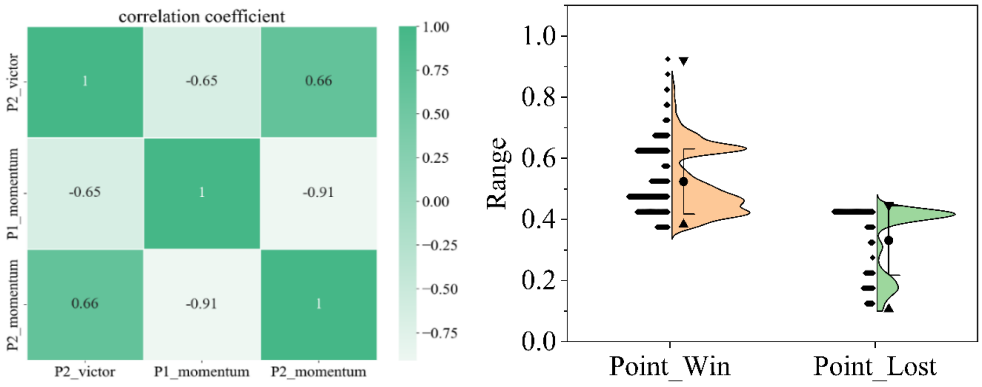

Figure 5 Point bi-column correlation coefficient and Violin plot

In this study, it can be observed that the correlation coefficient between P2 win (P2_victor) and P1 momentum (P1_momentum) is -0.65: this means that P1 momentum is negatively correlated with P2 win. When P1's momentum score increases, the probability of P2 winning decreases, indicating that P1's high momentum is unfavorable for P2 winning. The correlation coefficient between P2 win (P2_victor) and P2 momentum (P2_momentum) is 0.66: this indicates that P2 momentum is positively correlated with its likelihood of winning. That is, the higher P2's momentum score is, the more likely it is to win the game. However, the correlation coefficient between P1 momentum (P1_momentum) and P2 momentum (P2_momentum) is -0.91: this is a very strong negative

correlation, meaning that when P1's momentum increases, P2's momentum decreases, and vice versa. This reflects the zero-sum nature of the game, where one side's gain is often the other's loss.

From these data combined with box plots, this study can conclude that momentum has a significant impact on the outcome of the game, which can be confirmed by the positive correlation between P2's momentum and its winning probability. The momentum transition in a match is real-time and dynamic, as reflected by the strong negative correlation between P1 and P2 momentum; When a player shows strong momentum, it usually weakens the momentum of his opponent. These relevance metrics provide a quantitative basis for understanding the importance of momentum in a game and help guide coaches and players in developing strategies to exploit or combat momentum shifts in a game.

Overall, momentum plays an important role in tennis, and its effects are visible during the match. Coaches and players should pay attention to this dynamic change in momentum, as it can be one of the decisive factors affecting the final game outcome.

3.3. Analysis of the factors affecting the momentum transition

As for the factors that affect the momentum conversion, this study makes the five-dimensional graph of the correlation coefficient of p1 and p2's momentum converted into when they are opponents as follows:

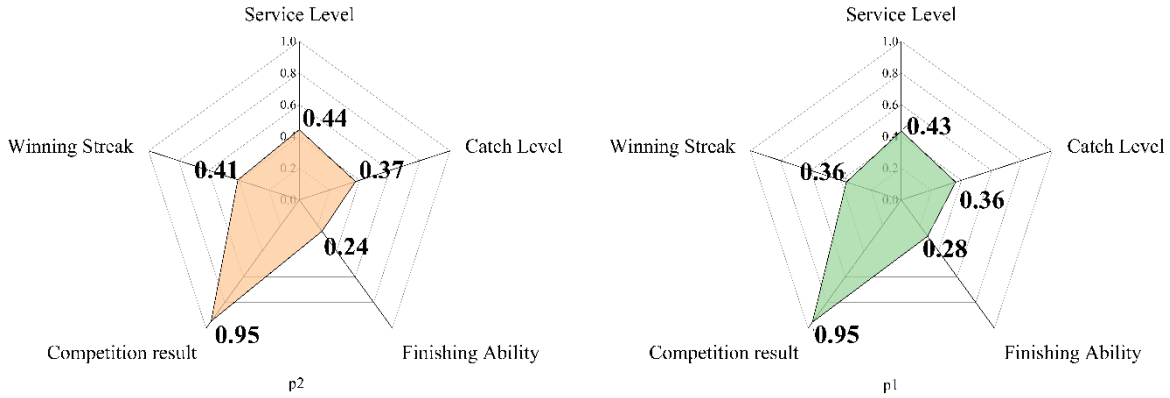

Figure 6 Correlation coefficient five-dimensional chart

By observing this study, it can be found that for both players, the strong positive correlation between momentum and game results (0.95) highlights the importance of maintaining momentum in the game to achieve the final victory. Serving (0.43, 0.44) and receiving (0.36, 0.37) are important components of the momentum of the two players; So the result of the serve is the most important to influencing the momentum. Improving your performance in match points (0.28, 0.24) also significantly increases momentum; The amount of points scored in a match (0.36, 0.41) also has an important impact on momentum.

4. Conclusions

To improve coaches' and players' understanding and application of momentum in matches, the goal of this study is to develop a model to capture momentum flow in tennis matches and explore its role in matches and how it can be used to influence match outcomes.

In this study, a momentum quantification model based on the Analytic Hierarchy Process (AHP) was established, and the correlation coefficient matrix was constructed by combining five criteria: serving level, receiving level, shooting ability, game result, and streak, and Consistency Ratio (CR) of the matrix was tested. Then the weights of the five indicators are determined using the eigenvalue method. This study then assigns scores to the different behaviors in the matching of the standard layer. The momentum of each player is quantified by multiplying the weight by the score. Finally, the

quantification of the momentum flow in the game is obtained by the ratio of the individual momenta of the two players. When the gap between the momentum values of the two players is too large, the probability of the team with the larger momentum value winning is almost.

To evaluate the role of momentum in the competition, this study developed a momentum evaluation model of the influence of momentum on the victory of the competition, which combines clustering analysis based on the K-means++ algorithm, violin plot analysis, and pairwise sequence correlation coefficient analysis. This study found that there was a significant correlation between the outcome of the match and the momentum index (0.95), and the result of the serve had the greatest impact on the momentum, so it was concluded that the momentum had a significant impact on the match result, in which the effect of the serve on the momentum was the most significant, and the reliability of the AHP quantitative index was verified.

This paper provides a research idea and framework applied to sports analysis-related fields, K-means++ clustering analysis results prove the feasibility of the AHP momentum quantification model.

References

[1] FERNANDEZ J, MENDEZ-VILLANUEVA A, PLUIM B J B J O S M. Intensity of tennis match play [J]. 2006, 40(5): 387-91.

[2] WANG C, WANG K, HU Y J H I S, ENGINEERING, et al. Momentum analysis of tennis players based on entropy weight method and personal's correlation coefficient [J]. 2024, 101: 612-22.

[3] KHAN A U, ALI Y J I J O T A H P. Analytical hierarchy process (AHP) and analytic network process methods and their applications: a twenty year review from 2000-2019: AHP & ANP techniques and their applications: Twenty years review from 2000 to 2019 [J]. 2020, 12(3).

[4] ÖZKAN B, KARASAN A, KAYA I J J O M-V L, et al. A Fuzzy Based Performance Model for the Assessment of Individual Sport Branches: A Case Study for Tennis Players [J]. 2021, 37.

[5] O'DONOGHUE P G J R Q F E, SPORT. The most important points in grand slam singles tennis [J]. 2001, 72(2): 125-31.

[6] ANGELINI G, CANDILA V, DE ANGELIS L J E J O O R. Weighted Elo rating for tennis match predictions [J]. 2022, 297(1): 120-32.

[7] ZHANG W. The Technology of Serving the Ballin the Tennisand its Trend [J]. Journal of Anyang Normal University, 2004, (5): 100-2.

[8] XIAO L. Technical Analysis of Women's Singles Finals in the 2019 Australian Open [J]. Physical Education Revivew, 2019, 38(12): 138-9.

[9] AHMED M, SERAJ R, ISLAM S M S J E. The k-means algorithm: A comprehensive survey and performance evaluation [J]. 2020, 9(8): 1295.

[10] BONETT D G J B J O M, PSYCHOLOGY S. Point-biserial correlation: Interval estimation, hypothesis testing, meta-analysis, and sample size determination [J]. 2020, 73: 113-44.