Momentum Prediction of the Tennis Match Flow

Yuanhan Chen, Hanqi Wang
School of Optoelectronic Engineering and Instrumentation Science, Dalian University of Technology, Dalian, Liaoning, China

Abstract: Sports matches are characterized by dynamic momentum shifts that significantly influence performance and outcomes. There are many researches discussing about how momentum in sports games affects the match flow. This study aims to quantify, evaluate, and predict momentum dynamics in tennis to provide actionable insights for players and coaches. Using data from the 2023 Wimbledon men's singles, we developed a multi-method framework integrating the Analytic Hierarchy Process (AHP) and fuzzy comprehensive evaluation to assess player performance through weighted indicators such as advantage, skill, error, and emotion. Principal component analysis (PCA) was applied to identify critical momentum metrics, followed by a neural network model to predict momentum swings by analyzing factors like serve performance and error rates. Results revealed that points scored as the server (x_2) emerged as the most influential predictor of momentum shifts, with high correlation coefficients (Pearson >0.99) validating model accuracy. The framework was further tested on French Open data, demonstrating adaptability across venues and genders, with court surface and player gender identified as secondary influencing variables. Momentum fluctuations showed a direct correlation with performance trends, enabling visualization of game flow and strategic recommendations, such as prioritizing serve efficiency and managing errors during neutral phases. The model's modular design allows extension to other sports by adjusting sport-specific metrics. This study establishes a robust, data-driven approach to harnessing momentum in competitive settings, offering practical tools for optimizing decision-making during critical match phases.

Keywords: Sports Momentum; Tennis Match Flow Prediction; PCA; AHP

1. Introduction

1.1 Background

Dynamic momentum can influence match flow in various sporting events magically. [1-3] Tennis has some unique features compared to other sports. During a match, players must handle various problems and make decisions without the guidance of their coach. Additionally, tennis matches have long duration with short breaks between sets. [4] These factors can affect a player's momentum and the fluctuation of points, ultimately changing the outcome of the match. It is challenging to control the events that occur during a match and their effects, making it difficult to predict the impact on a player's momentum. Therefore, making full use of events and changes in players' momentum can be valuable in evaluating their performance and helping them win different matches.

1.2 Restatement of the Problem

We are provided with data regarding every point from all the men's matches held in Wimbledon 2023 after the first two rounds. The data includes information about the wins and losses in every game and set, points scored by both players, the order of serve, unforced errors, and the running distance of the players. To better assess players' performance and the role of momentum, the following paper will:

- Develop a mathematical model capable of describing a player's scoring situation in a game. Utilize this model to evaluate the player's performance over a given period during a match and generate visual representations of the evaluation results.
- Assess the player's momentum throughout the game while ensuring the model is reasonable and reliable to demonstrate the significance of momentum in a match.

- Identify the variables that may affect the game's result and use them to create a model predicting possible swings that may occur during the match. Based on past momentum fluctuations, offer advice to players on how to compete against new opponents.
- Test the model with data while considering potential factors that could be added to the model in the future. Research whether the model can apply to other types of tennis matches or matches of other sports.
- Provide coaches with a note emphasizing the importance of momentum in a match and advising them on how to prepare players for the various events that may occur during a match in their daily training.

1.3 Our Work

We used the Analytic Hierarchy Process to identify indicators measuring a player's performance and determine their respective weights for a comprehensive player evaluation system. We then used the fuzzy comprehensive evaluation method to combine the performance scores of players at a given time. Finally, we used line graphs to visually represent the performance trends of players across nine games.

Then we used principal component analysis to establish the evaluation system for momentum. Subsequently, we devised an assessment formula for momentum using the weights of the principal components and aligned it with the swings in momentum based on the evaluation rating. Finally, we applied the function to the same set of data. Through a comparison between the momentum score and the player evaluation score, we demonstrated the impact of momentum on the result of the game and its crucial role.

Next, we utilized a neural network to predict the swings in the match. Through analyzing various combinations of indicators and assessing their degree of it, we identified the most significant indicators of swings and momentum. Based on this factor and previous swings of momentum, we offered players recommendations for strategies under different momentum conditions and how to capitalize on momentum for optimal results.

Last, we used our model to analyze the men's and women's match data of the French Open to predict the swings. To make the model more complete and comprehensive, we took into account the influence of gender and venue factors on the prediction outcomes. Besides, we explored the feasibility of adapting this model to other sports, such as basketball and elite handball.

Figure 1 is the mind map of our model.

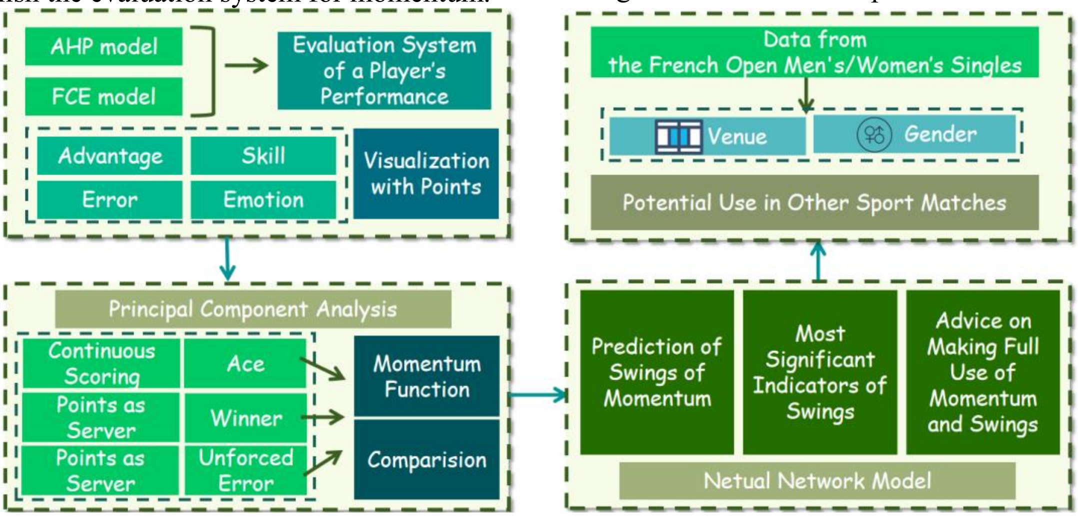

Figure 1. Model Overview

2. Models and Results

2.1 Evaluation Model of Player's Performance

2.1.1 Analytic hierarchy process

First of all, we are looking for indicators to evaluate players' performance. Scores in games within a set and the point score within a game can influence the players' performance in tennis matches. In addition, the probability of winning a point also depends on whether the point is a breaking point.[5,6] Furthermore, when tennis players become fatigued after running for a long distance, their hitting accuracy can be sharply decreased. [7] To develop a comprehensive and reliable model for evaluating players' performances, a two-layer criteria Analytic Hierarchy Model has been developed. This model includes 12 factors that can effectively reflect players' performance and provide a clear understanding of the factors that influence it. The indicators and their categorization are presented below as Figure 2 shows for further clarity:

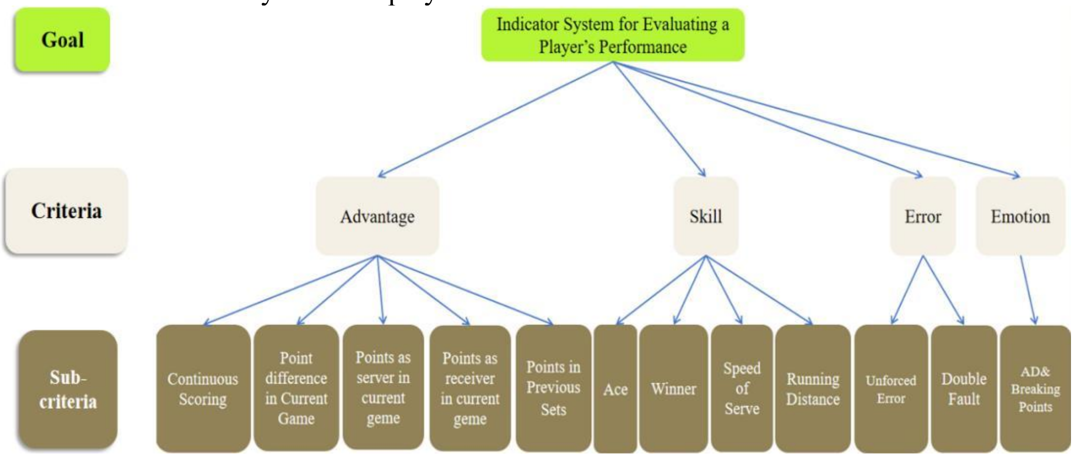

Figure 2. Indicators for the Evaluation System

The calculated weights are shown in Table 1 below:

Table 1. Weights of Indicators

Indicator	Weight
Continuous Scoring	0.1142
Points Difference in Current Game	0.0406
Points as Server in Current Game	0.0644
Points as Receiver in Current Game	0.1409
Points in Previous Sets	0.0568
Ace	0.0467
Winner	0.0299
Speed of Serve	0.0159
Running Distance	0.0179
Unforced Error	0.0741
Double Fault	0.1483
Ad & Breaking Points	0.2502

2.1.2 Fuzzy comprehensive evaluation

The fuzzy comprehensive evaluation method is an application of fuzzy mathematics, which is suitable for solving various non-deterministic problems. [8] Based on the indicators and weights above, we use a fuzzy comprehensive evaluation model to evaluate the players' performance in a game through the scores they get in the model.

First, we processed the data, eliminated outliers, and split it by games. Then we established a Weight Vector A for each criterion, using the weights of each indicator obtained from model 2.1.1. The elements of represent the weights of all factors in their criterion.

To better evaluate the performance of the players, we formulated three grades: good, medium, and poor, with scores of 90, 70, and 50 respectively. Based on the trapezoidal distribution, we separately established a membership function for each grade, describing the likelihood of each factor belonging to each grade. α1, α2, and α3 are the lower quartile, mean, and upper quartile of each factor based on the processed data respectively.

$$u_1(x)=\begin{cases}1, & 0\le x\le \alpha_1\\ \dfrac{\alpha_2-x}{\alpha_2-\alpha_1}, & \alpha_1\le x\le \alpha_2\\ 0, & x>\alpha_2\end{cases} \tag{1}$$

$$u_2(x)=\begin{cases}0, & 0\le x\le \alpha_1\\ \dfrac{x-\alpha_1}{\alpha_2-\alpha_1}, & \alpha_1\le x\le \alpha_2\\ \dfrac{\alpha_3-x}{\alpha_3-\alpha_2}, & \alpha_2\le x\le \alpha_3\\ 0, & x>\alpha_3\end{cases} \tag{2}$$

$$u_3(x)=\begin{cases}0, & x\le \alpha_2\\ \dfrac{x-\alpha_2}{\alpha_3-\alpha_2}, & \alpha_2\le x\le \alpha_3\\ 1, & x>\alpha_3\end{cases} \tag{3}$$

Next, we built the Fuzzy Relation Matrix R.

$$R=\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & r_{n3} \end{bmatrix} \tag{4}$$

Then, we calculated the Result Vector B.

$$B=A\times R \tag{5}$$

After calculating each of the four criteria, four row vectors can be obtained: B-advantage, B-skill, B-error, and B-emotion. Based on the data of each criterion, we defined R' = [B-advantage B-skill B-error B-emotion]. Besides, the total weight vector A' was defined based on the weight of each criterion's elements. Therefore, the Synthetic Result Vector B' of a player's performance in a given time can be expressed as:

$$B'=A'\times R'=[T_1 \quad T_2 \quad T_3] \tag{6}$$

Then, we calculated the player's evaluation score using the following method:

$$t=90\times sin^2(T_3\times\frac{\pi}{2})+70\times sin^2(T_2\times\frac{\pi}{2})+50\times sin^2(T_1\times\frac{\pi}{2}) \tag{7}$$

Based on the process above, we obtained the evaluation points of player 1 in nine games of the first set. The player's performance is illustrated in Figure 3:

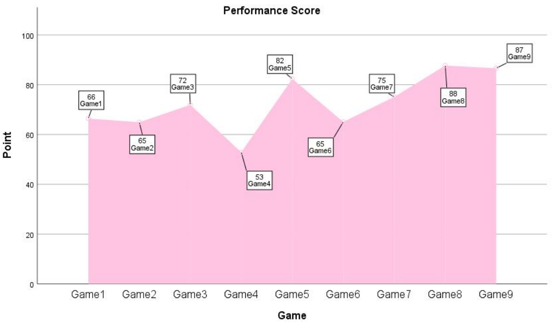

Figure 3. Performance Scores of Player 1

2.2 Momentum Evaluation Model

2.2.1 principal component analysis

Momentum can be defined as "strength or force gained by motion or by a series of events". It can affect a player's cognition, actions, and emotions, leading to changes in their strategies and the eventual result of the match. To better understand the role of momentum in the flow of a match, we developed the following indicators in Table 2 to measure momentum.

Table 2. Indicators of momentum

Indicator	Parameter
Continuous Scoring	x1
Points as Server in Current Game	x2
Points as Receiver in Current Game	x3
Ace	x4
Winner	x5
Unforced Error	x6

To reduce the dimensionality of those indicators, we utilized Principal Component Analysis PCA). It produces linear combinations of the original variables to generate the axes, also known as principal components. [9]

- Develop a data matrix X = (x1 x2 x3 x4 x5 x6).
- Standardize the samples and calculate the covariance matrix R.

$$R=\frac{\sum_{k=1}^{n}\left(x_{ki}-\overline{x}_i\right)\left(x_{kj}-\overline{x}_j\right)}{\sqrt{\sum_{i=1}^{n}(x_{ki}-\overline{x}_i)^2\sum_{i=1}^{n}(x_{kj}-\overline{x}_j)^2}} \tag{8}$$

Calculate the eigenvalues and eigenvectors of R and calculate the principal component contribution and cumulative contribution.

- After calculating, the feature vectors involved with the 4 main components are:

$$\begin{cases} F_1=\left(-0.30 \quad -0.62 \quad 0.72 \quad 0.07 \quad -0.04 \quad 0.06\right)^T \\ F_2=\left(-0.59 \quad 0.30 \quad 0.06 \quad -0.33 \quad 0.61 \quad 0.29\right)^T \\ F_3=\left(-0.50 \quad 0.34 \quad -0.02 \quad 0.71 \quad -0.33 \quad 0.14\right)^T \\ F_4=\left(0.31 \quad -0.10 \quad -0.04 \quad 0.04 \quad -0.08 \quad 0.94\right)^T \end{cases} \tag{9}$$

F1, F2, F3, and F4 inherit 79% possible variance from X, which is more than 75%. Thus, it's effective and properly reflects the original factors. Therefore, (F1, F2, F3, F4) is defined as the metrics system for the evaluation model of a player's momentum.

2.2.2 Principal component analysis

Based on the principal component weights from, the momentum function was fit:

$$F=\frac{0.27}{0.79}F_1+\frac{0.18}{0.79}F_2+\frac{0.17}{0.79}F_3+\frac{0.16}{0.79}F_4 \tag{10}$$

This function helps to evaluate a player's momentum in the game and to measure any changes in their momentum throughout the game. It converts the player's momentum into a score, enabling us to assess how high or low their momentum is. To demonstrate the significance of momentum in a game, we calculated the momentum scores of player 1 for each game in the first set. We then compared these scores to the player's performance scores. The results of this comparison, including the change in momentum and the relationship between the two scores, are shown in Figure 4 and Figure 5 below.

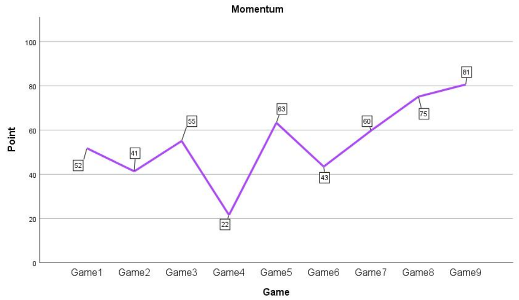

Figure 4. Momentum Scores

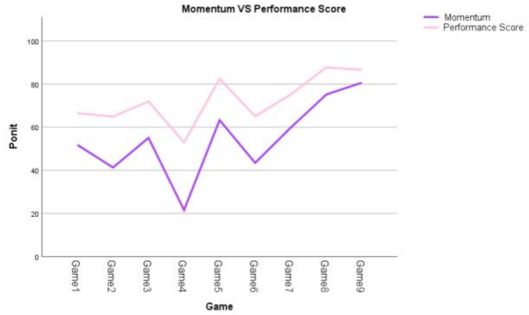

Figure 5. Comparison of Scores of Momentum and Performance

As the figures show, swings and points in a game are not random. Positive momentum typically leads to a good performance and victory, while negative momentum often results in a lackluster performance and defeat. In summary, a player's momentum has a direct correlation with their performance and the result of games.

2.3 Prediction Model of Swings in Matches

2.3.1 Neural network-based prediction of momentum swings

Momentum is influenced by six factors. To better predict the flow of play, we employed a neural network model to anticipate shifts and swings in momentum during the match. We used a training set of 1000 data points to train the neural network. Then, we used the trained model to predict the momentum swings of 500 additional data. Following the prediction, we conducted a correlation analysis using the Pearson correlation coefficient to obtain the correlation coefficients when all six influencing factors were taken into account. We observed that the correlation coefficients of the fitted results showed less variation when the less influential indicators were not taken into consideration. As a result, we delved deeper into the primary factors by examining various combinations of factors and their associated correlation coefficients of fitted results.

Table 3 Pearson Correlation Coefficient of Different Sets of Indicators

Sets of Indicators	Pearson Correlation Coefficient
x2, x3, x4, x5, x6	0.9949
x1, x3, x4, x5, x6	0.9967
x1, x2, x4, x5, x6	0.9960
x1, x2, x3, x5, x6	0.9998
x1, x2, x3, x4, x6	0.9990
x1, x2, x3, x4, x5	0.9986
x1, x2, x3, x4, x5, x6	0.9988

According to Table 3, the correlation coefficient undergoes the most significant change when x2 is not taken into account. This leads us to the conclusion that x2, which represents points scored as the server in the current game, is the most critical indicator that affects a player's momentum and the swing of a game. Scoring points on serve, as well as the number of points scored, can contribute to swings in a match. When a player scores more points, there is a higher chance of gaining momentum and winning.

Therefore, the number of points scored as the server and whether they are scored contribute to changes in momentum and swings in the match. The more points a player gets, the higher the probability of gaining momentum and winning will be.

2.3.2 Advice for players

- Score as many points as possible on the serve set.

Points scored on the serve set can greatly impact the swings. Therefore, players need to make full use of the opportunity and focus on their serving technique as well as observing the performance of their opponents in their serve set.

- Control the pace of a match when there is a big difference in momentum.

When the momentum is completely in favor of a player, his opponent may become aggressive. When the momentum is completely in favor of the player’s opponent, the player may slow down and just follow rituals. Therefore, when the momentum is transformed, it will favor the side that plays more aggressively. To take advantage of momentum shifts, players should predict their opponent'space, maintain their current momentum, and capitalize on opportunities to change the momentum in their favor.

● Make fewer mistakes when momentum is neutral.

When two players have neutral momentum, they generally play at the same pace. At this point, players should pay attention to basic serve and receive maneuvers,

keep a steady mindset, and capitalize on points during their serve set.

2.4 Additions and Extensions to the Model

2.4.1 Model testing based on data from other tennis matches

To ensure the precision and consistency of the model, we carried out testing on the model, sing data from other tennis matches. We selected the data of men's singles and women's singles of the French Open tennis Tournament. The data source is https: // github. com /JeffSackmann /tennis_ slam_ pointbypoint /blob / master /2021-frenchopen points.csv. Then, we calculated the momentum score and predicted the momentum swings according to our model. Figure 6 and Figure 7 are the prediction results which are shown below:

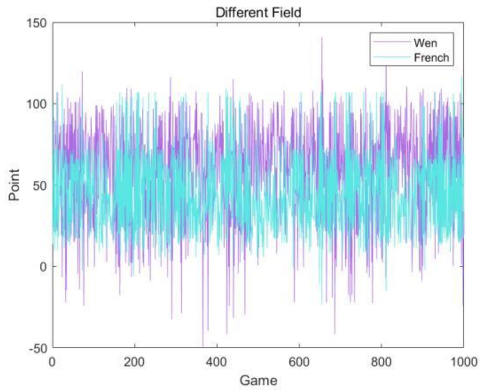

Figure 6. Momentum Points in Different Fields

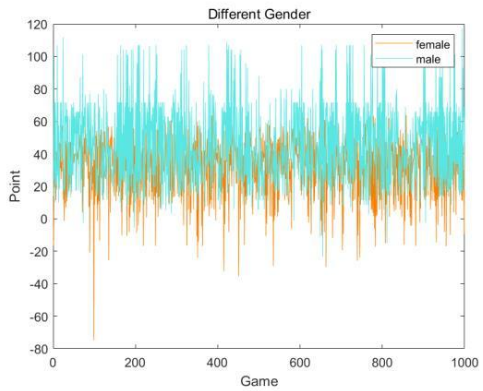

Figure 7. Momentum Points of Male Players and Famale Players

It is evident that the model possesses a considerable level of accuracy, indicating its potential to be implemented in various tennis matches. However, the scores and swings in players' omentum in different tournaments are significantly different. Thus, to search for the impact of other variables on the projected results, we conducted a comparative analysis between the French Open Men's Singles and the Wimbledon Men's Singles, as well as between the French Open Men's Singles and the French Open Women's Singles.

In our analysis, we found that both the venue and the gender of the players have an impact on the momentum scores and swings.

Venue

The French Open tends to have lower average momentum scores for men's singles compared to Wimbledon. The unique court characteristics of the French Open, such as its red soil surface, can create challenges for players, including unstable landing points for incoming balls and unfavorable conditions for sliding movements. This can result in slower movement speeds, increased physical exertion, and a higher likelihood of spinning balls. Additionally, the absence of the Hawk-Eye Challenge System can further compound these challenges. All of these factors can have a negative impact on a player's momentum.

Gender

Women's momentum scores at the French Open are lower on average than men's, which means gender has an impact on momentum. Women's momentum is slightly lower than men's on average. Female players may have lower physical fitness passion and ability to read the game than male players, which may also indirectly contribute to the disparities in momentum between male and female players in sports.

Therefore, we can introduce gender and venue as two dummy variables in our model. We can further refine our system and apply it to various tennis matches. With this model, we can accurately measure the momentum of tennis matches across different genders and venues. Ultimately, our model can be used to quantify player momentum in men's and women's singles matches at any open tennis tournament worldwide, delivering a unique tennis experience for players and coaches.

2.4.2 Application to other sport matches

The model has high potential to be applied to other ball games, but it needs to be adapted to the rules of different games and the characteristics of different sports. For example, in women's basketball, there is a positive relation between the rate of reinforcement before an adversity and a team's positive response to that This means teams who were performing well before an adversity generally responded better to that adversity than teams who were performing poorly. [10] Furthermore, in elite handball matches, the ability of a team to limit the amount of opposition attacks helps to create swings in momentum in favor of that team.[11] As long as the metrics for evaluating momentum are identified, the present model can calculate the momentum score and predict swings of momentum using neural networks. Therefore, our model holds significant application value for different sports.

3. Conclusion

This study successfully developed a comprehensive framework to quantify, evaluate, and predict momentum dynamics in tennis matches by integrating the Analytic Hierarchy Process (AHP), fuzzy comprehensive evaluation, principal component analysis (PCA), and neural networks. Key findings revealed that momentum fluctuations are strongly correlated with critical in-game factors, particularly points scored as the server, which emerged as the most influential predictor of momentum shifts. The models demonstrated high accuracy in visualizing performance trends, assessing momentum scores, and forecasting swings, validated through applications to Wimbledon and French Open datasets. By incorporating gender and venue as dummy variables, the framework showcased adaptability across diverse tennis tournaments, while its modular design suggests potential applicability to other sports through tailored adjustments. These insights provide actionable strategies for players and coaches to harness momentum during critical phases, though future research should expand datasets and test generalizability across broader contexts to further refine predictive capabilities.

References

[1] Iso-Ahola S E, Blanchard W J. Psychological momentum and competitive sport performance: A field study. Perceptual and Motor Skills, 1986, 62(3): 763-768.

[2] Crust L, Nesti M. A review of psychological momentum in sports: Why qualitative research is needed. Athletic Insight, 2006, 8(1): 1-15.

[3] Kimiecik J C, Jackson S A. Optimal experience in sport: A flow perspective. 2002. Kimiecik J C, Jackson S A. Optimal experience in sport: A flow perspective. 2002.

[4] M. Zhao, H. Jiang, B. Qi and J. Yang. Chinese tennis players emotional intelligence theory model construction and scale preparation. Psychological Research, 2024, 17(01): 42-53. DOI: 10.19988/j.cnki.issn.2095-1159.2024.01.005.

[5] Knight G, O' donoghue P. The probability of winning break points in Grand Slam men's singles tennis. European Journal of Sport Science, 2012, 12(6): 462-468.

[6] Newton PK, Keller J B. Probability of winning at tennis I. Theory and data. Studies in Applied Mathematics, 2005, 114(3): 241-269.

[7] Polly R. Davey, Rod D. Thorpe & Clyde Williams (2002) Fatigue decreases skilled tennis performance, Journal of Sports Sciences, 20:4, 311-318, DOI: 10.1080/026404102753576080

[8] DaiZ, Sun C, Zhao L, et al. Assessment of smart learning environments in higher educational institutions: a study using AHP-FCE and GA-BP methods. IEEE Access, 2021, 9: 35487-35500.

[9] Holland S M. Principal components analysis (PCA). Department of Geology, University of Georgia, Athens, GA, 2008, 30602: 2501.

[10]Roane H S, Kelley M E, Trosclair N M, et al. Behavioral momentum in sports: A partial replication with women's basketball. Journal of applied behavior analysis, 2004, 37(3): 385-390.

[11]Mortimer P, Burt W E. Does momentum exist in elite handball?. International Journal of Performance Analysis in Sport, 2014, 14(3): 788-800.