Quantification Analysis of Sports Momentum Based on a Multi-Index Approach and TOPSIS

Zehao Wu[#], Weishi Xu[#], Jiayu Huang[#, *]

School of Mathematics, South China University of Technology, Guangzhou, China

* Corresponding Author Email: ma202030520392@mail.scut.edu.cn

#These authors contributed equally.

Abstract. This study quantifies the concept of momentum in sports, inspired by Novak Djokovic's unexpected loss in 2023, and evaluates its impact on game outcomes. Utilizing a structured multi-index approach, we categorized performance into fourteen primary and three secondary indexes. The TOPSIS model enhanced by entropy weighting, along with the Delphi model, was applied for weight assessments of these indexes. Our analysis, incorporating the Mann-Kendall trend test and Spearman correlation coefficient, revealed a significant correlation of 0.51 between momentum and match results, underscoring its pivotal role. The results contribute to a deeper understanding of momentum, enabling coaches to refine strategies and optimize performance based on dynamic changes in momentum. This study provides a comprehensive framework for quantifying momentum, with potential applications extending beyond tennis to other sports where strategic momentum management could influence outcomes.

Keywords: Momentum Quantification, Multi-index Approach, Correlation Analysis, TOPSIS Model.

1. Introduction

Tennis has become one of the most fascinating and widespread sport in the world. Its intense matches, skilled athletes and randomness captivated a large amount of audiences around the world. The professional tennis circuit, including prestigious tournaments like the Grand Slam events, ATP Tour, and WTA Tour, attracts millions of fans and generates substantial revenue [1].

Because of the prevalence of football and basketball, more people play sports betting in football and basketball. As a result, the models and methods of factor analysis and match result prediction for basketball and football have been very mature.[2]Compared with basketball and football, the research of factor analysis and match result prediction for tennis is much later. And the study in this field now are still focus in the traditional statistical approach. [3]. The need of this kind of study not only lies in the extension of sports betting market, but also in the demand of the coach team helping athletes to make progress. More advanced method are required to find the key factor of winning the match and the prominence technique of the sports in order to improve the athletes' skills and performance. By building a visualized and robust model to show the game's flow and evaluate a player's performance as well as predict the outcome of the match, the coach can learn how to guide the player to perform better in the coming matches. Some previous studies have done by implementing regression model using factors such as ATP ranking, age to predict the outcome of the match.[4].Others found that Elo ratings-based methods, including the recently proposed Weighted Elo method, are performing well when forecasting tennis match results.[5] More advanced studies have used the machine learning models to give more accurate results.[6].But the research about visualizing the game flow and quantify the performance and condition during a single match are seldom studied. The aim of this article is to delve into the realm of momentum modeling in tennis matches, by building momentum model which represent a player's state, we can visualize the turns and twists of a game. By studying the key factors that shape the game flow, we seek to unravel the intricate relationship between momentum and strategic decision-making. Through insightful analysis and comprehensive research, we aim to empower tennis enthusiasts, players, and coaches with valuable tools to enhance their understanding and approach to the game.

Understanding the dynamics of momentum in tennis can provide invaluable insights into the ebb and flow of a match. It enables us to identify pivotal moments, analyze player performance patterns, and anticipate shifts in momentum that can ultimately sway the outcome of a game. Armed with this knowledge, players and coaches can develop effective strategies to capitalize on momentum swings and gain a competitive edge on the court.[7]

2. Quantification of momentum and visualization of game flow

In the process of quantifying momentum, a review of the literature revealed that multiple metrics in the game have an impact on momentum. [8-9]These indexes are called first grade indexes, and they are respectively number of double faults(DF(N)), number of ACE balls(ACEs), success rate in saving break points(BPSR), first delivery success rate(FSSR), first serve scoring rate(FSScR), second serve scoring rate(SSScR), percentage of first serve returns(FSRScR), breakthrough rate(BPCR), number of break points(BP), percentage of second serve returns(SSRScR), unforced error(UE(N)), round running distance(RRD), score at the net(NPS) and three consecutive scores(CTPS). Then, these first grade indexes are grouped into three categories, each forming three second grade indexes.The three second grade indexes are serve momentum(SMI), receive momentum(RMI) and other momentum(OMI). The final quantified values of three different indexes are calculated by using same model. Then, according to Delphi method, three values are weighted to get the final momentum quantization result.

2.1. Entropy Weighting-Based TOPSIS Method for Measuring Secondary Indicators

The TOPSIS method based on entropy weighting (TEW) is a widely used method with high feasibility and acceptance, so it was chosen to measure Secondary indicators[10].

TEW synthesizes the entropy weight method with the TOPSIS method. Entropy weight method is a weight allocation method based on information entropy, which can be used to deal with multiple indicators in the decision-making problem. TOPSIS method is a multi-indicator decision-making method can evaluate the degree of superiority of multiple programs. The combination of the two can achieve an objective decision-making optimal program.This method can consider multiple indicators at the same time, avoid the limitation of single indicator decision-making, and make full use of the information between indicators to improve the accuracy and reliability of evaluation.Obviously, this method can solve the problem very well.The following model is based on TEW.

Firstly, Harmonize indicator types and do indicator normalization. Very small indicator to very large indicator formula:

$$\max x - x \tag{1}$$

The intermediate index is transformed into the very large index formula:

$$M = \max\{|x_i - x_{best}|\} \tag{2}$$

$$\widetilde{x_i} = 1 - \frac{|x_i - x_{best}|}{M} \tag{3}$$

The interval type index is transformed into an extremely large index formula:

$$M = \max\{a - \min\{x_i\}, \max\{x_i\} - b\} \tag{4}$$

$$\widetilde{x_i} = \begin{cases} 1 - \dfrac{a - x_i}{M}, & x_i < a \\ 1, & a \le x_i \le b \\ 1 - \dfrac{x_i - b}{M}, & x_i > b \end{cases} \tag{5}$$

Then, Normalization of the normalization matrix. Suppose there are n objects to be evaluated, and the positive matrix composed of m evaluation indicators (already positive) is as follows:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \tag{6}$$

Then the normalized matrix is denoted by Z, where Z is each element

$$z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{n} x_{ij}^2}} \tag{7}$$

The normalized matrix Z is obtained

$$Z = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1m} \\ z_{21} & z_{22} & \cdots & z_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nm} \end{bmatrix} \tag{8}$$

The score is calculated and normalized. Defining the maximum value:

$$Z^+ = (\max\{z_{11}, z_{21}, \cdots, z_{n1}\}, \max\{z_{12}, z_{22}, \cdots, z_{n2}\}, \cdots, \max\{z_{1m}, z_{2m}, \cdots, z_{nm}\}) \tag{9}$$

Defining the minimum value:

$$Z^- = (\min\{z_{11}, z_{21}, \cdots, z_{n1}\}, \min\{z_{12}, z_{22}, \cdots, z_{n2}\}, \cdots, \min\{z_{1m}, z_{2m}, \cdots, z_{nm}\}) \tag{10}$$

Define i (i=1,2,... ,n) the distance between the evaluation objects and the maximum value:

$$D_i^+ = \sqrt{\sum_{j=1}^{m} (Z_j^+ - z_{ij})^2} \tag{11}$$

Define i (i=1,2,... ,n) the distance between the evaluation objects and the minimum value:

$$D_i^- = \sqrt{\sum_{j=1}^{m} (Z_j^- - z_{ij})^2} \tag{12}$$

It can be calculated that i (i=1,2,... ,n) non-normalized scores of evaluation objects:

$$S_i = \frac{D_i^-}{D_i^+ + D_i^-} \tag{13}$$

According to the model and data, the following calculation results are obtained:

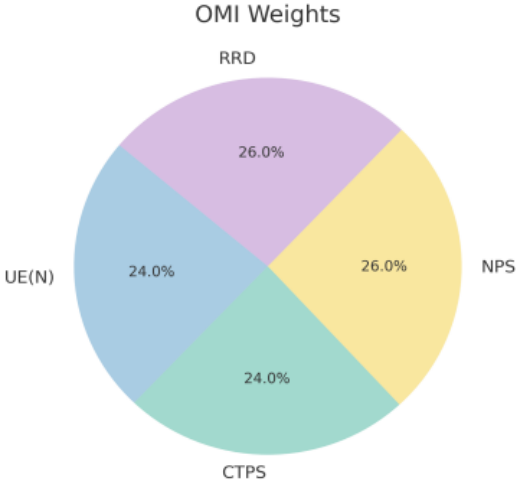


Figure 1. Waiting Chart for OMI

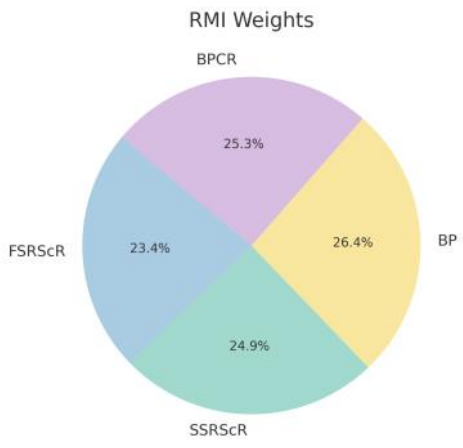


Figure 2. Waiting Chart for RMI

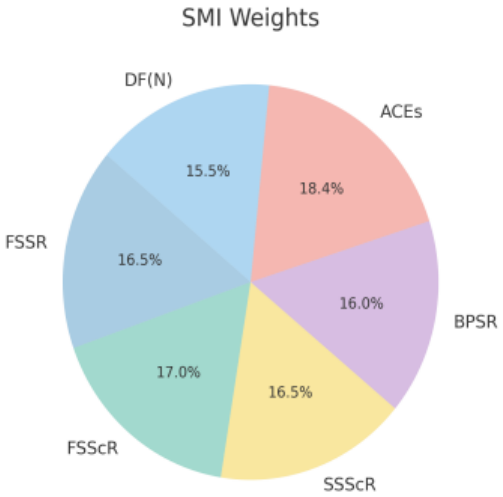


Figure 3. Waiting Chart for SMI

Observing Figure.1-3, it can be seen that the weighting of each first grade index is relatively even across the three second grade indexes. So it can be obtained that in each second grade index, the influence size of each first grade index on the second grade index is similar.

2.2. Calculating the momentum

The Delphi technique is an intuitive forecasting technique that combines the experience of multiple experts with subjective judgment.The Delphi method was chosen to calculate the value of the final momentum because Delphi method can avoid some possible disadvantages of group decision making. In addition, managers can ensure that important ideas are not overlooked when seeking input to make decisions.

Within each dimension, the expert compares the magnitude between each two metrics and scores them, resulting in a complementary judgment matrix. By solving the eigenvectors of these matrices, the weights of each indicator are obtained. Eventually, the momentum can be expressed as

$$M = \beta_S \cdot SMI + \beta_R \cdot RMI + \beta_O \cdot OMI \tag{14}$$

In the above formula β_S means weights for serve momentum, β_R means weights for receive momentum, and β_O means weights for other momentum.

The final result of the calculation is obtained:

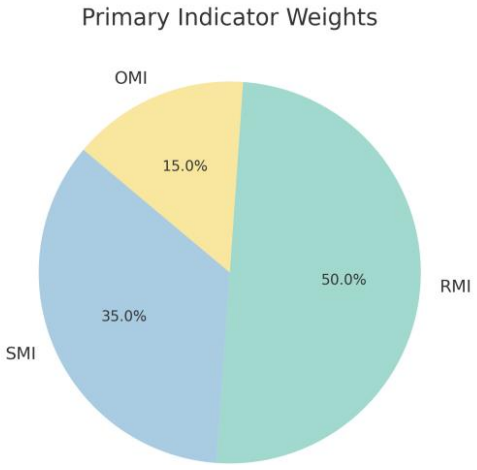


Figure 4. Chart of weights of secondary indicators

By looking at Figure.4, it can be concluded that in momentum, RMI has the highest percentage, followed by SMI and finally OMI . bringing in the data the value of momentum can be derived through the equation.

2.3. Visualization of competition flow

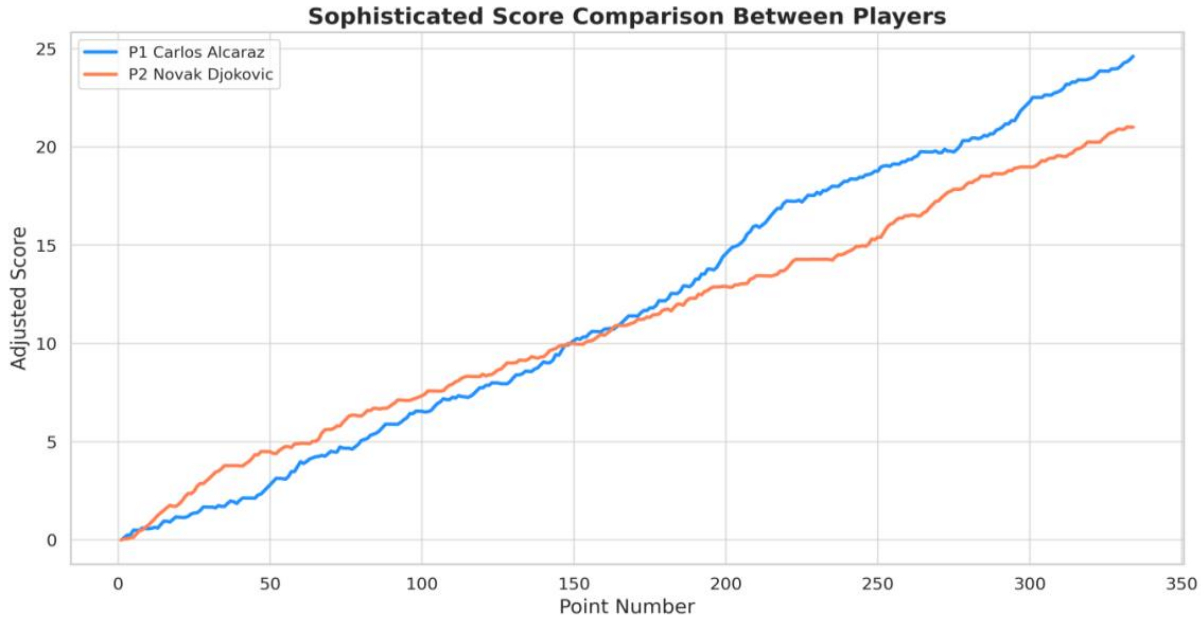


Figure 5. Plot of player momentum change over time

Figure 5. shows the momentum change of the two players in the final. It can be seen that the momentum of player 2 is high at the beginning, but the momentum of player 1 is higher than that of player 2. You can see that the change in momentum is similar to the change in the result.

The momentum of both players in the graph increases with the number of swings. This is because in the model we set the momentum not to zero after each round, so the momentum generated in the next round will be directly added to the previous momentum value. In other words, in the model, the momentum values are accumulated over time, and therefore will generally trend upwards as the match progresses.

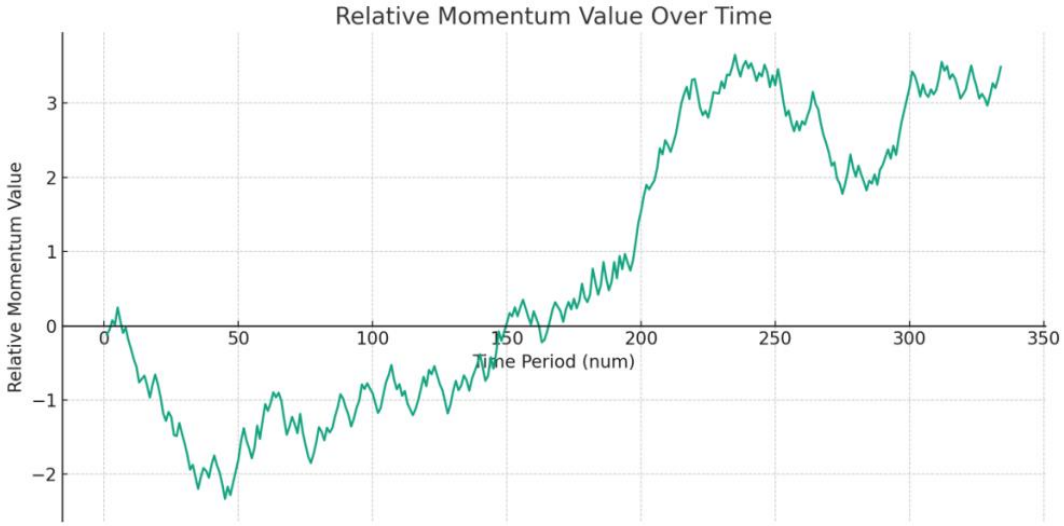


Figure 6. Diagram of relative momentum

It can be seen from Figure.6 that both sequences of momentum are accumulated over time and have increasing trend. So the flow and result of the match are more likely to be determined by the relative value of two players' momentum. Thus, the relative momentum curve is obtained by subtracting the momentum of the two players over time. If the value of relative momentum is above 0, the match is now more controlled by player 1. By analyzing the trend of the curve, the flow of the match can be described. If the curve has increasing trend, it means player 1 is taking control.Using this relative momentum curve to solve problem.

3. Evaluation of the correlation between momentum and game outcome

One of the tennis coaches in the question wondered if "momentum" played a role in the match. Therefore, we modeled the correlation between momentum and match results. In this way, it is possible to specify the extent to which momentum affects the outcome of a match.

3.1. Analysis M sequence

Since momentum varies with match time and the final value of momentum is related to duration and number of beats, we want to detect the change in momentum and the magnitude of relative momentum. In this section, we use the Mann-Kendall trend test.

The following is the process of analyzing the sequence:

For time series M, Let's say the sample size is n, calculate S first:

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(M_j - M_k) \tag{15}$$

$$\operatorname{sgn}(x) = \begin{cases} 1 & if \quad x > 0 \\ 0 & if \quad x = 0 \\ -1 & if \quad x < 0 \end{cases} \tag{16}$$

The Mann-Kendall Trend Test analyzes difference in signs between earlier and later data points. The idea is that if a trend is present, the sign values will tend to increase constantly, or decrease constantly. Every value is compared to every value preceding it in the time series, which gives a total of n(n-1) / 2 pairs of data, where “n” is the number of observations in the set. Then calculate the variance:

$$Var(S) = \frac{1}{18}(n(n-1)(2n+5) - \sum_{k=1}^{p} q_k(q_k-1)(2q_k+5)) \tag{17}$$

By transforming S transformed to the test statistic Z_{MK}, such that for large samples of data, the Z_{MK} can approximately satisfy the standard normal distribution:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{Var(S)} & if \quad S > 0 \\ 0 & if \quad S = 0 \\ \dfrac{S+1}{Var(S)} & if \quad S < 0 \end{cases} \tag{18}$$

Finally, a hypothesis test is enough. Figure.7 is the result of Mann-Kendall trend test:

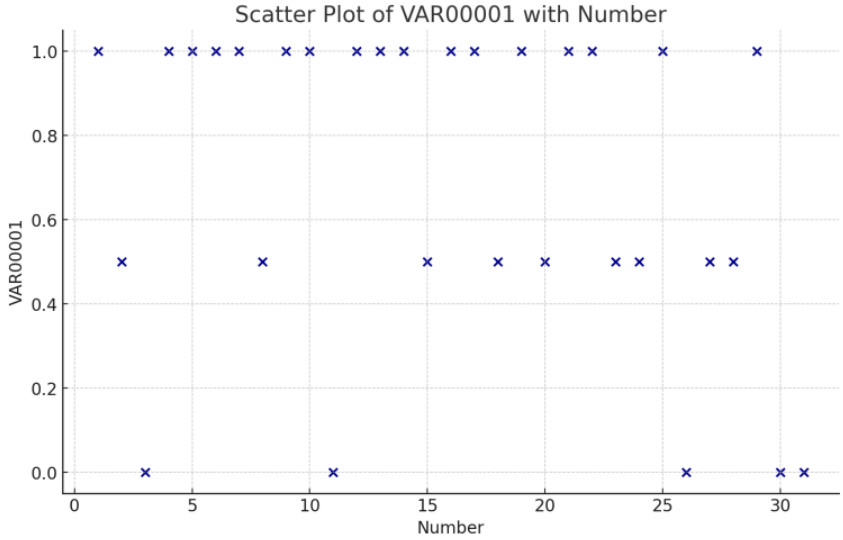


Figure 7. Map of Mann-Kendall result

3.2. Correlation analysis

The Spearman correlation coefficient is a rank correlation coefficient that reflects the association between the direction and strength of the trend of two random variables. Since the context in which it is applied matches the problem, it is chosen to be used to calculate the degree of correlation between momentum and the outcome of the match.

$$\rho = \frac{\frac{1}{n}\sum_{i=1}^{n}((R(x_i)-\overline{R(x)})\cdot(R(y_i)-\overline{R(y)}))}{\sqrt{(\frac{1}{n}\sum_{i=1}^{n}(R(x_i)-\overline{R(x)})^2)\cdot(\frac{1}{n}\sum_{i=1}^{n}(R(y_i)-\overline{R(y)})^2)}} \tag{19}$$

where $R(x)$ and $R(y)$ are the rank of x and y, respectively, and $\overline{R(x)}$ and $\overline{R(y)}$ denote the average rank, respectively.

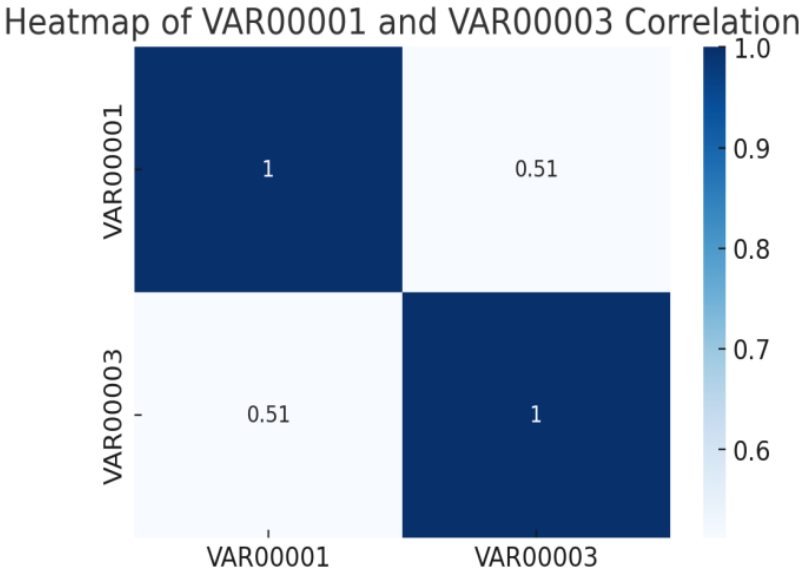


Figure 8. Heatmap of VAR00001 and VAR00003 Correlation

Figure.8 is one of the graphs showing the results of Spearman's correlation coefficient. It can be seen that the correlation coefficient is 0.51, 0.51 and 0 is still a big difference, so it can be seen that the momentum and the results of the game are more relevant.

4. Conclusions

The main purpose of this article is to identify which players are better at the game and how well they are doing it, using a variable to measure the level of player performance, the variable. There are variables that need to be quantified for momentum, and for this purpose, all indicators that affect the size of momentum are classified and stratified. Among them, there are 3 secondary indicators and 14 primary indicators. By using TOPSIS model and the enhancement of entropy weight method, the significance of each first-level index's contribution to the second-level index was determined. The Delphi model is used to evaluate the weights of the three secondary indexes. Finally, the momentum is quantified successfully and reliable visualization results are obtained. Using entropy weighting ensures that metrics that provide more unique information are given higher weights. And TOPSIS portfolio optimization model based on the entropy weight method is more objective than the ordinary TOPSIS model. So the result is much more accurate.

Next, the Mann-Kendall trend test was used to detect the change in momentum and the magnitude of relative momentum, and the Spearman correlation coefficient was used to examine the degree of correlation between momentum and race outcomes. The results show that there is a large correlation coefficient (0.51) between the two.

References

[1] Chen H, Li C, Meng X, et al. Development of youth tennis players: A study based on the ranking history of top ATP/WTA players worldwide and China [J]. Plos one, 2023, 18(11): e0289848.

[2] Etuk R, Xu T, Abarbanel B, et al. Sports betting around the world: A systematic review[J]. Journal of Behavioral Addictions, 2022, 11(3): 689-715.

[3] Wilkens, Sascha. 'Sports Prediction and Betting Models in the Machine Learning Age: The Case of Tennis'. 1 Jan. 2021: 99 – 117.

[4] Buhamra N, Groll A, Brunner S. Modeling and prediction of tennis matches at Grand Slam tournaments[J]. Journal of Sports Analytics, 2024, 10(1): 17-33.

[5] Angelini G, Candila V, De Angelis L. Weighted Elo rating for tennis match predictions [J]. European Journal of Operational Research, 2022, 297(1): 120-132. Read N.Forecasting the Outcomes of Professional Tennis Matches [D]. UCLA, 2024

[6] Read N. Forecasting the Outcomes of Professional Tennis Matches[D]. UCLA, 2024.

[7] Pluim B M, Jansen M G T, Williamson S, et al. Physical demands of tennis across the different court surfaces, performance levels and sexes: a systematic review with meta-analysis [J]. Sports medicine, 2023, 53(4): 807-836.

[8] Rong Zhang. Tennis match result forecast and player analysis[D]. Yunnan University, 2022.

[9] Zou Xin. Comprehensive Strength Evaluation and Regression Model Forecast Analysis of World Women's Hard Court Singles [J], Sports Science and technology literature Bulletin, 2022, 30(1): 67-70

[10] Qing-Fang Liu, Ying-Ying Lan, Zhao-Ying Zuo, et al. Research on comprehensive service quality evaluation of gas station based on entropy weight -Topsis method [J], China Inspection and Testing, 2022, 6: 87-90