Quantification of momentum in tennis matches and its impact: a study based on AHP-EWM method and data analysis

Wei Wu*

College of Engineering, China University of Petroleum-Beijing at Karamay, Karamay, China, 834000

* Corresponding Author Email: 2021016113@st.cupk.edu.cn

Abstract. Tennis is a globally popular sport that tests players' abilities in all aspects, including mental quality. In tennis, momentum is a psychological phenomenon that reflects the strengths and weaknesses of players. It affects players' mentality, performance, and match results. However, there are still some inadequacies and controversies in the current research on momentum, such as its concept, formation mechanism, quantification method, and actual effect. This study deeply analyzes the concept and formation mechanism of momentum in tennis matches and its influence on the result of the match. Through a combination of data analysis and quantitative methods, this study reveals the importance of momentum changes on player performance and match results. Through the application of AHP-EWM, Welch's T-Test, and Run Test, this paper not only quantifies momentum but also confirms the significance and non-randomness of momentum, refuting the traditional views on the randomness of competition performance fluctuation and success rate. The results of this study not only provide a new perspective for understanding the strategy and dynamics of tennis matches, but also provide practical guidance for athletes' training and competition preparation. In addition, the methodology and findings of this study also have reference value for other sports that need to analyze the impact of momentum, opening up a new path for future sports science research.

Keywords: momentum, AHP-EWM, Welch's T-test, Run Test, Athlete performance.

1. Introduction

In the realm of competitive athletics, particularly within a sport as intricate and strategic as tennis, the influence of momentum on match outcomes is widely acknowledged. Although its impact is commonly recognized in practical situations, there remains considerable discourse within the scholarly community regarding the precise definition and measurement of momentum, as well as its specific ramifications on athlete performance and match results.

Drawing upon data gathered from www.comap.com, this study delves into the very essence, formation mechanisms, and consequential effects of momentum in tennis matches. Employing the AHP-EWM method, this study amalgamates expert knowledge with data-driven characteristics[1] to objectively assess the various factors that contribute to the momentum of a match and to chart the potential consequences of momentum fluctuations on match outcomes. Welch's T-test and the Run Test have also been employed to scrutinize the influence of momentum on the progression of a match. The findings confirm that momentum in tennis matches is significantly distinguishable, non-random, and predictive, with players exhibiting greater momentum more prone to emerge victorious. Furthermore, this paper explores the fluctuations in momentum throughout a match and the profound shifts in momentum resulting from consecutive points or losses. These discoveries vehemently refute the notion that momentum changes occur randomly and hold important implications for dissecting the concept of arbitrary fluctuations in competitors' performance and success rates. Moreover, they provide empirical validation for future investigations into competition strategy, athlete psychology, and other factors that may impact competition outcomes.

2. Tennis Momentum Quantification

2.1. Data pre-processing

During the process of gathering and transmitting measurement data, certain data points may be unsuitable or lost due to environmental disturbances or human factors. To restore the objective authenticity of the data, this paper initially purifies the competition data and eliminates any anomalous data present. If there are missing values in the competition speed, this paper employs the KNN interpolation method based on Gower distance to address this issue. Gower distance serves as a means of gauging the similarity between data points and computes the distance using the data points' local density and diffusion[2]. Considering the existence of missing values in qualitative variables such as serve_width, serve_depth, and return_depth, this paper first translates these missing values into an ordered variable and subsequently applies KNN interpolation. Additionally, due to the variation in dimensions caused by the competition's scoring rules, subsequent calculations might be affected. In this paper, the original score is substituted with an integral value for this reason.

2.2. Establishment and solution of quantization model

In tennis matches, the server usually has the advantage, which is due to the server having more initiative and control in the service game[3]. To measure momentum more accurately, this article considers giving a higher weight to the server. However, the subjectivity of weight assignment has been an important issue [4]for AHP. Therefore, this paper proposes to use AHP-EWM to improve. AHP-EWM not only considers expert opinions, but also reflects the relative importance [5]of each indicator more objectively from the point of view of the data itself. In this way, the importance of each indicator can be measured more accurately and momentum can be further quantified. The process of constructing the model is illustrated in Figure 1:

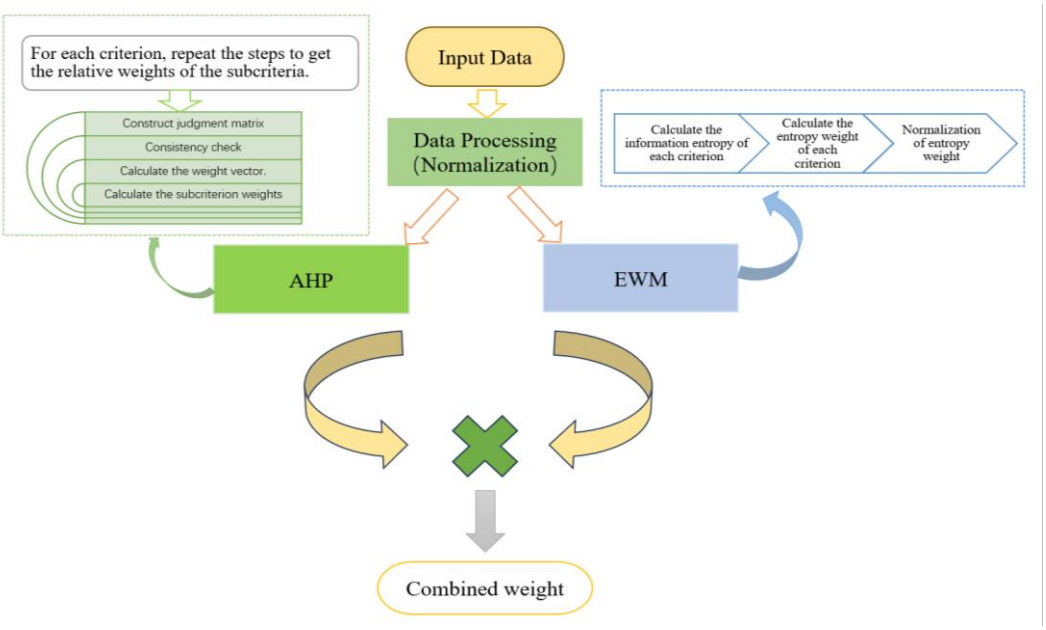


Figure 1. Flow chart of momentum quantization algorithm

After finishing the data processing, the first step is to establish the hierarchy, build the decision matrix, conduct the consistency test, and calculate the weights of the sub-criteria. Next, calculate the information entropy and entropy weight, and normalize the entropy weight to ensure the total sum of entropy weights for each criterion is 1[6]. Finally, multiply the entropy weight of the sub-criterion by the relative weight of its corresponding criterion to obtain the final comprehensive weight.

2.2.1 Indicator selection

During a tennis match, the live circumstances, external disruptions, and psychological elements may cause players to sense a certain momentum, which can influence the game's trajectory. To further

investigate the impact of momentum, factors such as player scores, server advantage, fatigue, and confidence levels are taken into comprehensive consideration. These indicators (for instance, the case of players p1 and p2) are selected for quantifying the presence of momentum:

- **p1_score_difference:** The score difference of an athlete intuitively reveals his scoring ability and is an important standard for evaluating his competitive state. The change in score difference can often reflect the momentum fluctuation and competition situation of an athlete in real-time, which is expressed by the score difference between p1 and p2.
- **server:** In the service game, the server has more time and space to prepare his serve, and can make corresponding adjustments according to the reaction of the opponent, which enables the server to better control the rhythm of the game. The server can take advantage of precise drop control, and speed changes make it difficult for the receiving side to cope. To ensure that subsequent calculations are not affected by how the ordering variable is encoded, we do something with the server indicator:

When calculating p1,

$$server = \begin{cases} 1, p1\ serve \\ 0, p2\ serve \end{cases} \tag{1}$$

In calculating p2,

$$server = \begin{cases} 0, p1\ serve \\ 1, p2\ serve \end{cases} \tag{2}$$

- **p1_d:** We consider that the longer the distance traversed by an athlete, the greater their level of exhaustion consequently.
- **p1_break_pt_won:** Due to the advantage of the server, the receiving team is often in a passive position during the game. Winning a serve in the opponent's service game can greatly boost the player's confidence and morale.
- **p1_double_fault:** An increase in the number of double faults in a player's serve game can lead to poor form and increased psychological stress[7].

In conclusion, these five metrics assess the state and fluctuations of players in tennis matches, examining their performance in terms of score, serve, physical prowess, technique, and mental strength, correspondingly. By continuously monitoring these indicators, we can gain profound insights into the athletes' competitive state. Utilizing these measures to quantify momentum enables us to grasp the extent of their enhanced performance.

2.2.2 Evaluation matrix construction and consistency test

According to the expert scoring results, the following evaluation matrix is constructed:

$$criteria\ matrix = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ \frac{1}{2} & 1 & 3 & 2 & 4 \\ \frac{1}{4} & \frac{1}{3} & 1 & \frac{1}{2} & 2 \\ \frac{1}{3} & \frac{1}{2} & 2 & 1 & 3 \\ \frac{1}{5} & \frac{1}{4} & \frac{1}{2} & \frac{1}{3} & 1 \end{bmatrix} \tag{3}$$

The formula for defining the consistency index is as follows:

$$CI = \frac{\lambda - n}{n - 1} \tag{4}$$

Wherein, λ refers to the maximum eigenvalue of the evaluation matrix, and the calculated value of CI is 0.017. Meanwhile, The consistency index RI for matrices with different n values is shown in Table 1:

Table 1. RI value result

n	1	2	3	4	5	6
RI	0	0	0.58	0.90	1.12	1.24

The consistency test of the matrix, then RI takes the value of 1.12, yields: $CR = \frac{0.017}{1.12} = 0.0152 < 0.1$ indicating a successful passing of the consistency test.

2.2.3 Result

The final weights for AHP are presented in Table 2:

Table 2. AHP weight result

Indicators	score_difference	server	distance	break_pt_won	double_fault
Weights	0.4185	0.2625	0.0973	0.1599	0.0618

The final weights for EWM are presented in Table 3:

Table 3. EWM weight result

Indicators	score_difference	server	distance	break_pt_won	double_fault
Weights	0.5651	0.2410	0.0424	0.1188	0.0327

momentum can be calculated from the weight value by the following formula:

$$momentum = 0.5651 \times score\ difference + 0.2410 \times server + 0.0424 \times distance + 0.1188 \times break\ pt\ won + 0.0327 \times double\ fault \quad (5)$$

The results of momentum calculated by formula 5 are shown in Table 4:

Table 4. momentum result

match_id	elapsed_time	momentum_p1	momentum_p2
2023-wimbledon-1301	00:00:00	0.204999	0.24792
2023-wimbledon-1301	00:00:38	0.22726	0.175711
2023-wimbledon-1301	00:01:01	0.194737	0.27766
2023-wimbledon-1301	00:01:31	0.24587	0.126407
2023-wimbledon-1301	00:02:21	0.354694	0.177596
2023-wimbledon-1301	00:02:50	0.642274	0.62476
2023-wimbledon-1301	00:03:33	0.202823	0.27549
2023-wimbledon-1301	00:04:01	0.627909	0.69424
2023-wimbledon-1301	00:04:48	0.190002	0.28145
2023-wimbledon-1301	00:05:32	0.626815	0.70153
2023-wimbledon-1301	00:07:05	0.2764	0.246247
2023-wimbledon-1301	00:07:28	0.72455	0.607326
2023-wimbledon-1301	00:08:06	0.29104	0.188441
2023-wimbledon-1301	00:08:41	0.149828	0.19646
2023-wimbledon-1301	00:09:11	0.26318	0.340257
2023-wimbledon-1301	00:09:37	0.70951	0.675784
2023-wimbledon-1301	00:10:20	0.27446	0.237908
2023-wimbledon-1301	00:10:45	0.7229	0.60943
2023-wimbledon-1301	00:11:24	0.308017	0.062561
2023-wimbledon-1301	00:11:47	0.627708	0.70449
2023-wimbledon-1301	00:12:17	1.071474	1.09837
2023-wimbledon-1301	00:12:44	0.624269	0.70388
2023-wimbledon-1301	00:13:26	1.08533	1.04788
2023-wimbledon-1301	00:15:31	0.28658	0.195865

Due to space constraints, only parts are shown

3. Quantitative Analysis of Match Influences

With the guidance of momentum, we can discern the ebb and flow of an athlete's force and the magnitude of their enhanced performances. To provide a more visually captivating portrayal of the players' prowess, we have graphically represented the progression across four matches as shown in Figure 2:

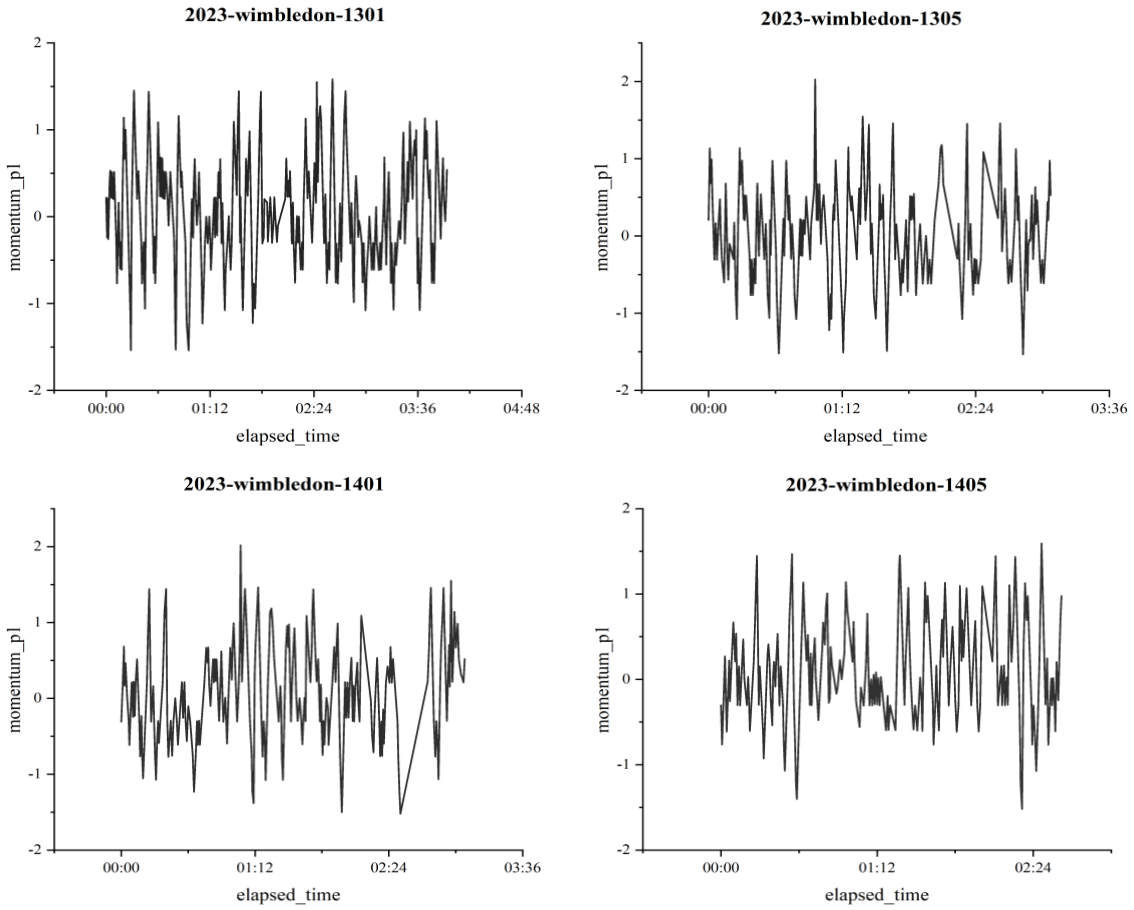


Figure 2. Momentum trend chart in four matches of p1

Based on the constructed model, we visualized the match status between Carlos Alcaraz and Novak Djokovic in 2023 Wimbledon-1301 as shown in Figure 3, and conducted an analysis based on the provided score data in the question:

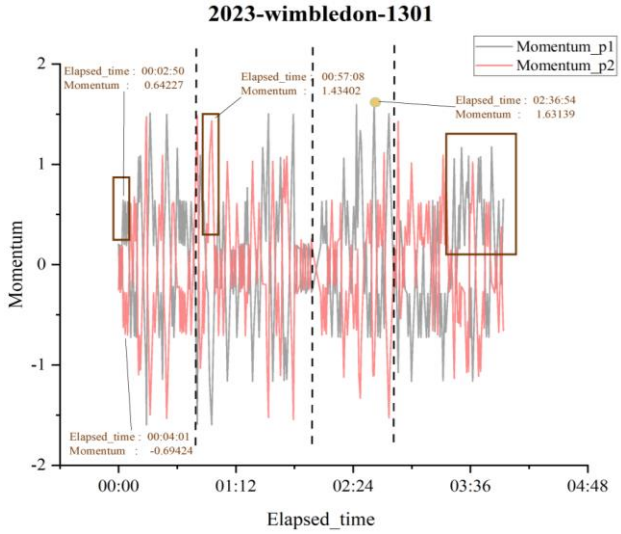


Figure 3. Momentum trend of players in 2023-Wimbledon-1301

As you can see from the figure, the change in momentum in the game is very noticeable.

◆ In the first set, p1 dominated in the service game. At 00:02:50, p1's momentum reached 0.64227 and continued at a high level for the next few points. At this point, p2's momentum is -0.62476. At 00:04:01, p2 momentum is down to -0.69424. Then the two sides tied each other continuously. Finally, in game 5, p1 went on the offensive again, scoring consecutive points to win the first set.

◆ In the second set, p2 starts to fight back. p2 has momentum as high as 1.43402 at 00:57:08, at which point the score is 0:40. However, p1 broke p2 in the next game to keep up. The two sides went toe-to-toe until p2 won in the tie-break.

◆ In the third set, p1 regained his form and held firm on serve. With 02:36:54 on the clock, p1 has a momentum of 1.63139, the highest momentum of the match. Eventually p1 won the third set easily.

◆ In the fourth set, p2 dominated the service game, then the two sides were tied, and finally, p1 maintained a high level of momentum to win the final set and the match.

3.1. Welch's T-test

To evaluate whether "momentum" in tennis matches affects the play, that is, on the match result, this paper analyzes the difference between the momentum of players after each serve and the match result. Welch's T-test was applied to match results and momentum because momentum in different matches did not meet the homogeneity[8] of variance.

Welch's T-test calculated the T-value according to the following formula:

$$t = \frac{x_1 - x_2}{\sqrt{\frac{s_1^2 + s_2^2}{n_1 + n_2}}} \tag{6}$$

Where x_1 and x_2 represent the means of the two samples, s_1^2 and s_2^2 denote the variances of the two samples, n_1 and n_2 indicate the sizes of the two samples.

The P-value of Welch's s T-test falls significantly below the standard level of significance at t=-22.958. This signals a substantial disparity between the victorious and defeated groups in the momentum index, with players exhibiting higher momentum tending to emerge triumphant.

Moreover, the average momentum index of athletes in the winning group stands at 0.178, while the mean for the losing group is -0.186. The disparity in mean values between the two groups confirms that, within our dataset, winning players possess greater momentum while their losing counterparts have less, thus solidifying the credibility and soundness of our model.

3.2. Run Test

Tennis coaches generally consider fluctuations in scores and the likelihood of success in a match to be random and therefore changes in momentum. To test this hypothesis, using the 2023 Wimbledon-1301 tournament as a reference, this paper conducted a Run Test on the Momentum_p1 data collected by Carlos Alcaraz. Since the data do not obey binary classification, this paper uses the median as the cutting point and processes the raw data to divide it into two categories (n_1, n_2) [9].

For data with a large sample size, the test statistic Z follows an approximate normal distribution, and its calculation formula is as follows:

$$Z = \frac{R - \mu_R}{\sigma_R} \tag{7}$$

Where *R* represents the number of runs, μ_R represents the expected value of the number of runs, and σ_R represents the standard deviation of the number of runs[10], they are calculated as follows:

$$\mu_R = \frac{2n_1n_2}{n_1+n_2} + 1 \tag{8}$$

$$\sigma_R = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}} \tag{9}$$

Where, n_1 and n_2 respectively represent the number of two types of data, while $n_1 + n_2 = n$ represens the size of the sample.

The P-value denotes the probability of observing a standard normal variable with an absolute value greater than Z[11], which is calculated using the following formula:

$$P = 2(1 - \Phi|Z|) \tag{10}$$

The results of the Run Test show that the data are non-random data based on the variable Momentum_p1, and the level is significant. $Z = -7.402$

3.3. Momentum and swings in play

The results of the Run Test show that the momentum of the athletes in the game is not a random quantity, which indicates that some systematic factors may affect the momentum of the athletes in the game. To have a more intuitive understanding, this paper visually shows the change of momentum and cumulative points of athletes in a single round of competition, as illustrated in Figure 4:

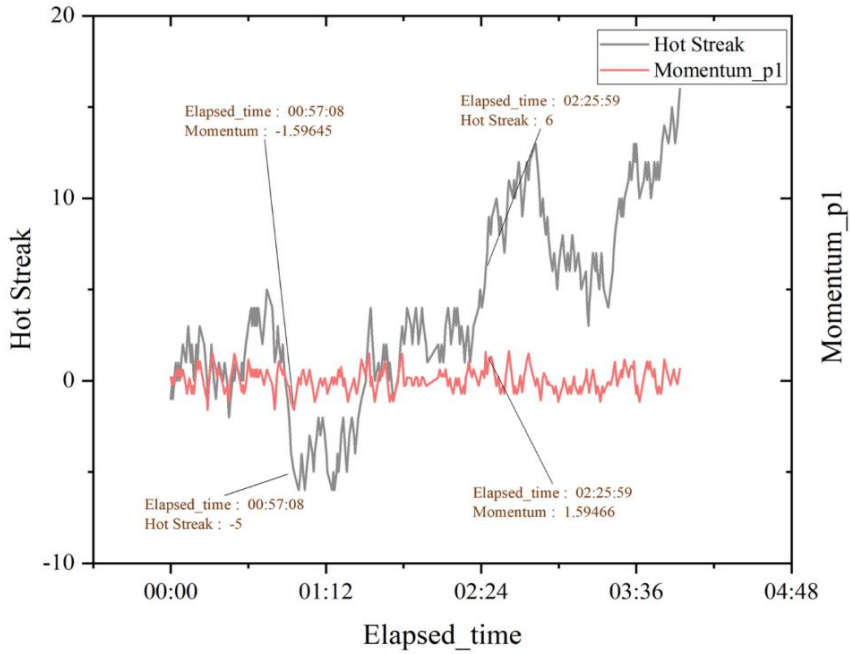


Figure 4. Momentum and runs of success

As you can see, at 00:57:08, Carlos Alcaraz lost six points in a row during the race, and momentum shows a lower value of -1.59645; On the contrary, at 02:25:59, Carlos Alcaraz won 6 points in a row, and momentum shows a high value of 1.59466, which effectively verifies that "the fluctuation of results and the possibility of success" is not random.

In brief, this paper demonstrates the pivotal role of momentum in a match and presents a visual exploration of noteworthy momentum shifts in scenarios of consecutive point scoring or losses. These discoveries hold profound significance in dispelling the notion of random fluctuations in a player's performance and how they achieve success in a game. Furthermore, they serve as a valuable reservoir of insights for future research into game strategy, athlete psychology, and other determinants that can impact game outcomes.

4. Conclusions

Through in-depth analysis and statistical examination of tennis match data, this study has revealed the non-randomness of momentum in matches and its significant impact on the outcome of matches. It is found that momentum changes with obvious regularity when players score or lose consecutive points in a match. These findings provide a strong refutation to the traditional view that the performance fluctuations and results of athletes in a match are random. The quantitative model in this paper not only combines AHP and EWM methods to improve the objectivity of evaluating the importance of each evaluation index but also verifies the statistical characteristics of momentum through Welch's T-test and Run Test. The empirical results show that athletes with higher momentum are more likely to win in competitions, which highlights the importance of factors such as psychology and strategy in competitions, provides a new perspective and data basis for subsequent research, and helps to optimize training and competition strategies and improve athletes' competitive performance. In addition, this study also provides a research idea, framework, and empirical basis for further exploration of competition strategy, athlete psychology, and other factors that affect competition results, demonstrating the potential and value of quantitative analysis in understanding and optimizing sports competition.

References

[1] Wu Yang, Liu Haiyu, Kou Mengnan, et al. Comparative analysis of methanol and traditional boiler fuel based on AHP-EWM [J]. Thermal Power Generation,2023,52(10):113-121.

[2] Bektas, A., & Schumann, R. (2019). How to optimize Gower distance weights for the k-medoids clustering algorithm to obtain mobility profiles of the Swiss population. In 2019 6th Swiss Conference on Data Science (SDS) (pp. 51-56). IEEE.

[3] Xie Zhongya. Kinematic Analysis of Serving Technique of Female Tennis Players [J]. Youth,2016(5):135.

[4] Luo Yi, Zhou Jianchi, Liu Xiangjie. Application of multi-level grey correlation Analysis Method in Operation Evaluation of thermal power Units [J]. Proceedings of the CSEE, 2012, 32(17): 97-103.

[5] Bai Yu, Xu Xuewen, Lei Yao. Fuzzy Comprehensive Evaluation of Missile Equipment Support Capability Based on Improved AHP-EWM [J]. Journal of Naval Aeronautical University,2023,38(1):151-157.

[6] Yang Zhao, Zheng Zhibo, Zhou Haoming, et al. Study on positive quality markers of blue cloth based on analytic Hierarchy Analysis (AHP) -entropy weight method and network pharmacology [J]. Chinese Journal of Chinese Materia Medica,2023,48(20):5450-5459.

[7] Xu Jia-Wei, ZHANG Zhi-Yuan. Analysis of the factors affecting athletes' psychology and countermeasures in Tennis Match [J]. Innovation Research of Ice and Snow Sports,2023(15):164-166.

[8] Chen Ying, Chen Changsong, Hu Honggang. Research on Power Consumption Analysis of SM4 Hardware Circuit [J]. Information Network Security,2018(5):52-58.

[9] Zhu K, Li Y. Application of RPT method in multivariate run test [J]. China Health Statistics,2016,33(2):362-363. (in Chinese)

[10] Pei De-Zhao, Li Yan-Ting. Multivariate nonparametric control chart based on Run Test [J]. Industrial Engineering,2020,23(2):124-132,149.

[11] Jump, lian-yong feng. China pilot carbon emissions trading market effectiveness analysis [J]. Journal of management and management, 2022, 31 (8): 195-202.