Unveiling the Hidden Momentum: A Comprehensive Analysis of Latent Dynamics in Tennis Matches

Summary

In the thrilling 2023 Wimbledon men's tennis final, the match dynamics were constantly shifting due to changes in Djokovic and Alcaraz's performance, making it difficult for either player to maintain control. Momentum plays a crucial role for coaching teams as it encapsulates various aspects of a player's performance and can accurately predict the course of the match, aiding athletes in improving their performance on the court. To better quantify momentum and explore its functions, we conducted an in-depth and detailed study from multiple perspectives and levels.

Firstly, to quantify momentum (M), we initially explored relevant data and found a strong correlation between players' scoring situations, serving conditions, fatigue levels, and momentum. Based on this, we developed a formula related to momentum and fitted it accordingly. We discovered that the first-order difference in the momentum difference (ΔM) and the progression of the match had a high correlation of 91%. Moreover, the positive or negative nature of (δM) determined the outcome of the match.

Next, we established an ARIMA forecasting model to predict fluctuations in matches. The ARIMA model, capable of capturing different autocorrelation structures in time series data, is particularly suitable for short-term forecasting. By referencing ACF and PACF plots, we selected the parameters (1,0,4), resulting in a fitted model with a 95% confidence interval and an R-squared goodness of fit of 0.905. Additionally, we employed a logistic regression model to predict binary response variable probabilities. The p-value for ΔM was calculated as 0.006, which is less than the commonly used significance level of 0.05.[1] This indicates that the coefficient of the ΔM variable is significantly non-zero.

Thirdly, we developed an analysis and evaluation model based on Bayesian spatiotemporal modeling to explore which factors significantly impact momentum. Using momentum (M) derived from the initial analysis as the dependent variable, we applied logistic regression and Markov Chain Monte Carlo (MCMC) to determine the impact magnitude of different factors on momentum. Finally, the weight coefficients obtained from WinBUGs14 approximately followed a normal distribution, aligning with the model's assumptions, indicating MCMC chain convergence and a satisfactory fit.

Fourthly, using a cross-categorization approach in our evaluation model, we utilized the $r_s^2 core$ as the evaluation metric. By designating each file as the test set sequentially, with the remaining files as the training set, we assessed the model's generalization capability across different datasets. We achieved an average accuracy of 93.17%, indicating our model's relative success in predicting match fluctuations.

Finally, we incorporated additional factors that could potentially influence the match's course to optimize the model for greater universality. Subsequently, we drafted a memorandum to the coach, outlining the role of momentum and providing targeted recommendations based on our findings.

Keywords: ARIMA forecasting model, Bayesian spatiotemporal model, logistic regression analysis, cross-validation model, momentum, match fluctuations.

Contents

1   Introduction

1.1   Problem Background

In today's professional tennis matches, pre-match preparation and tactical layout have become quite sophisticated. Players ensure they are in optimal condition on the court through pre-match physical training and carefully crafted game strategies. However, with the increasing level of competition and intensifying matches, relying solely on traditional preparation methods is no longer sufficient. Even top professional players experience fluctuations in their form, and there's always the possibility of being overturned despite seemingly having the advantage.[2]

In the thrilling Wimbledon men's tennis final of 2023, when Djokovic took the first set with a convincing score of 6-1, he seemed poised for an easy victory. However, as Alcaraz struggled to win the second set, Djokovic similarly won the third set with a 6-1 score. At this point, the outcome seemed almost certain, but the course of the match changed again as Djokovic lost control in the fourth set. As both players entered the fifth set, with the score tied at 1-1 in the first two games, Alcaraz broke serve in the third game. Djokovic's emotions fluctuated, and he struck the pillar beside the net with his racket. Subsequently, Djokovic's form declined, and Alcaraz ultimately won the match 6-4.

All these changes seem to be related to momentum, an intangible but influential factor in the course of a match. It prompts us to rethink how various variables in a tennis match interact to affect momentum, thereby influencing players' performance. By quantifying momentum, we can assist players in pre-match deployment and adjustments during the game from multiple perspectives, applying this comprehensive concept to more sports and influencing athletes' self-awareness of their performance states.

1.2   Restatement of the Problem

From the data collected after the second round of the 2023 Wimbledon Men's Tennis Tournament, we can understand various factors that influence the outcome of the matches as the game progresses, including detailed scoring for each player, number of victories, serving performance, distance covered on the court, quality of shots, and more. In order to develop a momentum-based model to predict the direction of the matches, provide recommendations for players during the game, and extend this model to other sports, we need to address the following questions:

1. Analyze momentum evaluation metrics and quantify momentum.

2. Calculate the momentum of players over time and predict pivotal moments in the match.

3. Validate the relationship between momentum fluctuations and players' match performance.

4. Analyze the varying impact of different factors on momentum within the dataset.

5. Collect and explore potential factors that influence momentum.

6. Explore additional information that could improve and promote the model.

1.3   Data Cleaning

Based on the observation of the provided data from the 2023 Wimbledon Men's Tennis Tournament after the second round, we have identified several issues and corresponding approaches for handling them.

The attachment contains some data that is challenging to process, including missing values, non-numeric data, and entries outside the expected range for certain columns. For instance, scores for players in rows 144-156 (which should be 0/15/30/40/AD) were incorrectly recorded as single-digit results. Additionally, there are numerous missing values for metrics such as hitting speed and return depth. Out of 7285 data points for hitting speed, 752 are missing, and out of 7285 data points for return depth, 1309 are missing. To address these issues, we conducted data cleaning to minimize errors and enhance the quality and accuracy of the data.

During the data inspection process, we identified numerous factors that may influence momentum and encountered some non-numeric data, which hindered the assessment of corresponding metrics. To resolve this issue, we quantified and normalized these data to reduce the impact of different units and scales on the analysis results, thereby improving the accuracy of data analysis.
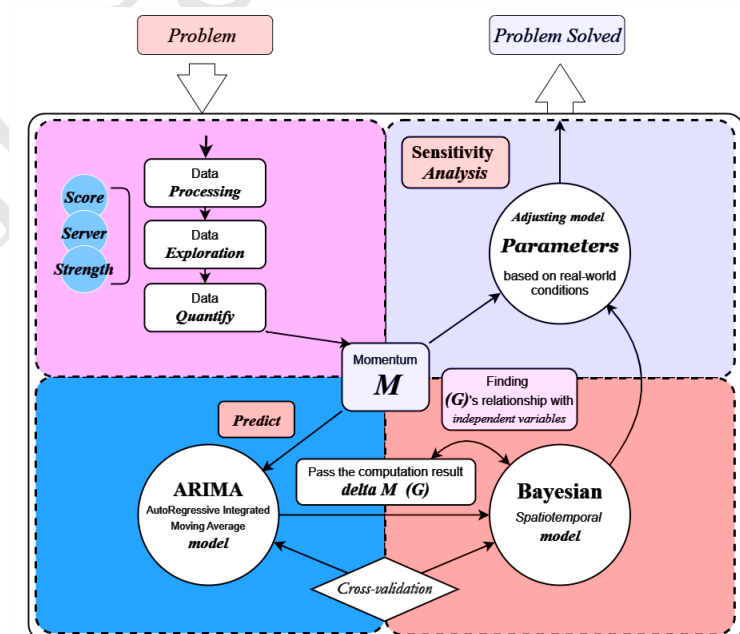
1.4   Our work

Figure 1: Work Flow

2 Preparation of the Models

2.1 Assumptions

Based on the simplifications made to the model, we have established the following reasonable assumptions:

1. Momentum influences the progression of the match, thereby affecting the outcome, while other aspects of players' performance also impact the result. This assumption effectively connects player performance, match outcomes, and momentum, thereby better defining and quantifying momentum.

2. The difference in skill level among participating players is relatively small, and match outcomes are not directly determined by a significant gap in absolute strength but are related to changes in momentum for both sides. This assumption ensures that momentum's influence on the match is not affected by external factors.

3. Players' momentum fluctuates, allowing momentum to be correlated with pivotal moments in the match. This assumption enables the association of momentum with critical turning points in the game.

4. Momentum is related to the psychological state of the players. When players are on a winning streak or serving, they have greater confidence in their next shot, leading to a noticeable increase in momentum. This assumption clearly identifies factors that have a significant impact on momentum, facilitating the quantification of momentum.

By adopting these assumptions, the model simplifies the complex dynamics of the game while still capturing essential aspects related to momentum and player performance.

2.2 Notations

The primary notations used in this paper are listed in Table 1.

Table 1: Notations

Symbol	Definition
ΔM	the players' momentum difference at a given moment
δM	the ΔM_t difference between adjacent time points
G	δM
P	the probability of one player gaining an advantage

3　Model 1:Momentum Quantification and ARIMA Integration

We integrated momentum, a comprehensive assessment derived from various aspects of a player's performance, with ARIMA forecasting model. Quantifying momentum allows us to reflect players' competitive status to facilitate rapid adjustments. By analyzing momentum changes and trends, we can make relatively accurate predictions about the progression of the match. Therefore, we first considered factors that significantly influence momentum and derived a criterion to evaluate momentum. Subsequently, we used the ARIMA model to forecast the match's progression and ultimately identify the turning points where one side transitions to the other.

3.1　Construction of Momentum Evaluation Formula

Firstly, we consider that players gain confidence during winning streaks, allowing them to utilize their full potential. However, this effect diminishes as their confidence reaches a certain level. Moreover, the outcome of the previous point directly influences a player's subsequent performance. Additionally, we observe that players have a higher winning rate when serving compared to when receiving. However, if a player makes an error on their first serve, they face the risk of losing the point directly on their second serve, leading to a more cautious approach. Furthermore, facing break points or multiple failed serves increases psychological pressure on players, negatively affecting their momentum. Also, we notice a strong correlation between players' fatigue levels and their performance; as fatigue increases, performance declines. The fluctuation in momentum at different moments reflects the fluctuation in player states, which often impacts the match's progression. Consequently, we find a strong correlation between the difference in momentum at different times and the match outcome.

To investigate the impact of changes in momentum for both sides on match fluctuations, we denote the difference in momentum between two players at the same moment as ΔM_t, which reflects differences in player states. Then, we define the difference in ΔM_t at different time points as δM. Since the momentum of both players is not necessarily equal at a specific time point, δM reflects changes in the difference in momentum between the two players. If δM is positive, it indicates a higher likelihood of winning the match. After fitting operations, we observe a strong correlation between δM and subsequent victories.

$$K_1 = 1.3 - 0.3 \cdot \exp\left(1 - x\right) \tag{1}$$

$$K_2 = victouy?1 : -0.9 \tag{2}$$

$$P_1 = \begin{cases} 1.2 \to is\ server \\ 0 \to not\ server \end{cases} \tag{3}$$

$$P_2 = \begin{cases} 0.75 \rightarrow fist\ fault \\ 1 \rightarrow fist\ succeed \end{cases} \tag{4}$$

$$\delta Q = -\frac{\sum\limits_{i=1}^{n}(distan_1 - distan_2)}{sum} \tag{5}$$

$$preM_{t+1} = preM_t + \frac{\prod\limits_{i=1}^{n} K_i + \prod\limits_{j=1}^{m} P_j}{2} \tag{6}$$

$$\Delta M_t = preM_{1t} - preM_{2t} + \delta Q \tag{7}$$

$$\delta M = \Delta M_{t+1} - \Delta M_t \tag{8}$$

As for the weights of $\prod\limits_{i=1}^{n} K_i$ and $\prod\limits_{j=1}^{m} P_j$, we later utilized sensitivity analysis.

Going forward, we will use MDELM to represent ΔM and dDELM for δM for brevity.

3.2   Testing of the Momentum Evaluation Formula

We plotted the curve of ΔM against time using the data from the match 2023-wimbledon-1701. It is easy to observe that if the subsequent point is a red dot, ΔM between the previous and subsequent points decreases; if the subsequent point is a blue dot, ΔM increases. That is, when δM is positive, Player 1 has a high probability of winning the match. When the sign of δM changes, the trend of winning the match shifts from one player to the other.
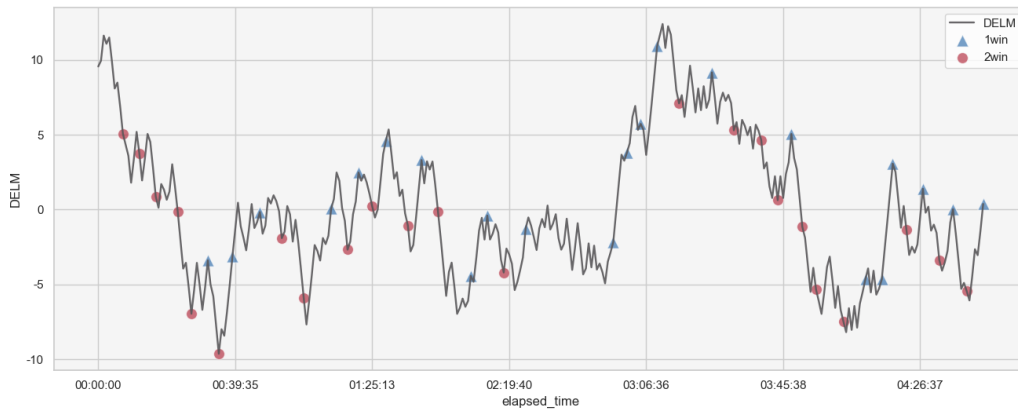

Figure 2: The result of ΔM

We conducted a Pearson correlation analysis between the DELM, dDELM data, and whether the player won, resulting in a heatmap.
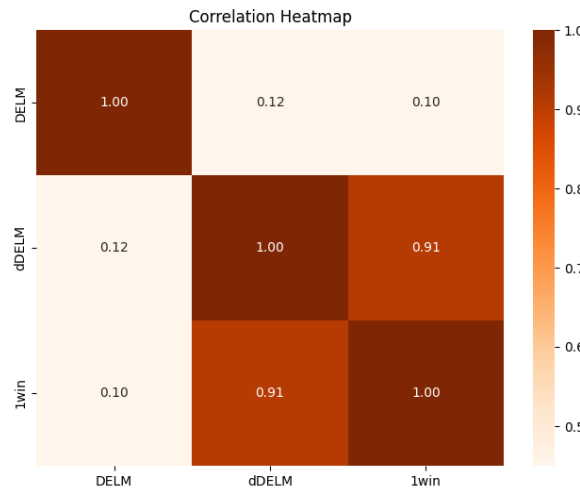
Figure 3: HeatMap

We found a correlation coefficient of 0.91 between δM and the outcome of the match. Therefore, we can conclude that the fitting of dDELM, representing δM, is effective.[3]

3.3   Predicting Player Performance with ARIMA and δM

We utilize the momentum evaluation formula to calculate the value of δM. Given δM's strong correlation with match results and player performance, we employ the ARIMA model to process δM and predict the progression of the match.

3.3.1   Explanation of the ARIMA Model

The ARIMA model (Autoregressive Integrated Moving Average) is a widely used method for time series forecasting analysis. To enhance prediction accuracy, the ARIMA model combines autoregression (AR) and moving average (MA) methods and integrates differencing (I) to handle the characteristics of non-stationary time series. Because the ARIMA model can capture various autocorrelation structures in time series, it is particularly suitable for short-term forecasting.[4] Researchers can simulate and forecast corresponding time series data by adjusting the parameters (p, d, q), where p represents the order of autoregression, d represents the differencing level, and q represents the order of the moving average.

3.3.2   Fitting with the ARIMA Model

We first define the 'check' class, which receives the differencing flag and handles it. Additionally, it conducts unit root (ADF test) and white noise tests to evaluate the stationarity and randomness of the data. Then, we use the arma_order_select_ic method to select the orders of the ARIMA model that best fit the data based on the AIC and BIC criteria.

Finally, we fit the data with the ARIMA model, plot the autocorrelation and partial auto-correlation graphs, make predictions, and compare the predicted results with the actual values. We calculate the mean absolute error of the predictions and analyze the normal-ity of the residuals to evaluate the model. We also define the 'generate' class, which first splits the dataset into training and testing sets. Then, it analyzes trends and changes by computing and comparing the differences between actual and predicted data.

Below are the results of the ADF test:

Table 2: ADF Test Results

Variable	Differencing Order	t	P	AIC	Critical Values		
					%1	%5	%10
	0	-3.413	0.011**	1040.585	-3.451	-2.871	-2.572
DELM	1	-16.172	0.000***	1044.606	-3.451	-2.871	-2.572
	2	-9.439	0.000***	1073.071	-3.452	-2.871	-2.572

Note: ***, **, and * respectively represent significance levels of 1%, 5%, and 10%.

The result of this sequence test shows that, based on the variable DELM: when the differencing order is 0, the significance level P-value is 0.011**, indicating significance at the 1% level. We reject the null hypothesis, indicating that the sequence is a stationary time series. Therefore, we determine our d parameter to be 0.

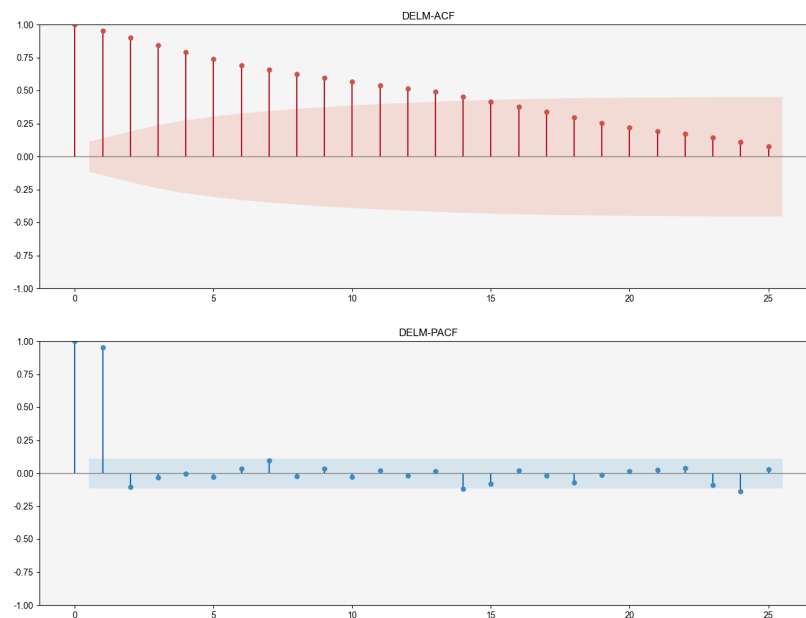Next, we use ACF and PACF plots to select the p, d, and q parameters for ARIMA.

Figure 4: ACF & PACF

Referring to the images and using the AIC, BIC selection criteria, along with some adjustments, we have chosen the (1,0,4) parameters.

The formula for the model obtained using ARIMA(1,0,4) is as follows:

$$y(t) = 0.027 + 0.938 \cdot y(t-1) + 0.138 \cdot \varepsilon(t-1) + 0.036 \cdot \varepsilon(t-2)$$
$$+0.033 \cdot \varepsilon(t-3) + 0.035 \cdot \varepsilon(t-4) \tag{9}$$

Subsequently, we obtained the fitting effect of the predicted values and actual values, and found that they closely matched the original data.
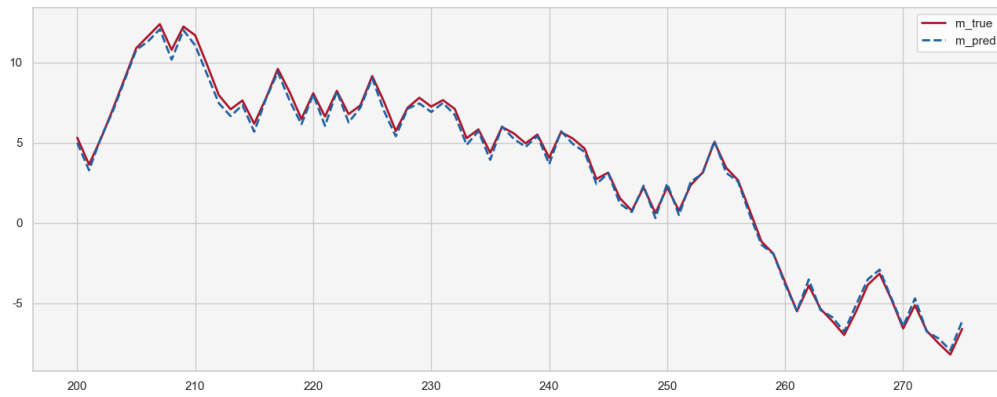

Figure 5: ARIMA

3.3.3 The forecast results using the ARIMA model

The visualization of the forecast results is shown in the following figure, with the 95% confidence interval shaded in gray. We have achieved short-term forecasting, which can be used to determine the trend of winning or losing using δM.
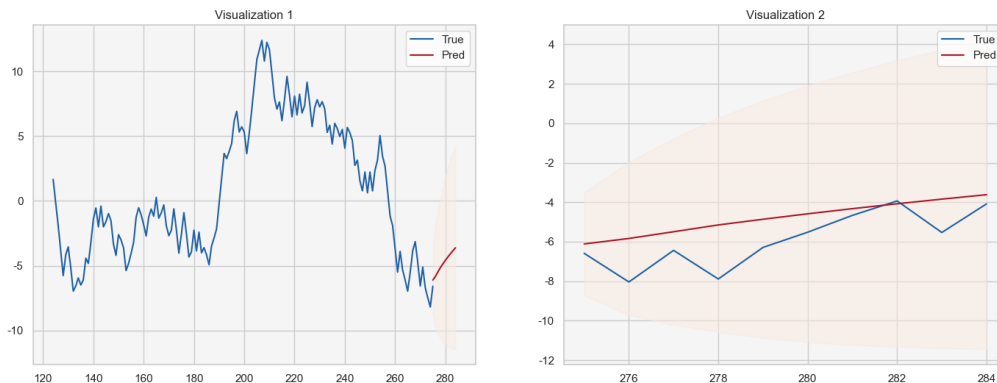

Figure 6: ARIMA test

Seeing the satisfactory results, we proceeded to forecast the next 10 sets.

Figure 7: ARIMA predict

There is a downward trend, and we anticipate that the momentum of player 1 will be poor in the next match, leading to a loss. We will use dDELM to assess the degree of decrease in their momentum.

3.4 The impact of momentum on the match results

To demonstrate that the "momentum" we fitted has an impact on the results, we use ACF, PACF plots as references, along with a logistic regression model to analyze whether the previous δM has an effect on the outcome.

3.4.1 Logistic Regression[5] Model Explanation

Logistic Regression is a statistical method widely used for predicting the probability of binary response variables. Its essence lies in mapping the output of a linear regression model to probabilities between 0 and 1 using a logistic or log-odds model.

The model is typically represented as:

$$\log\left(\frac{p}{1-p}\right) = \beta\_0 + \beta\_1X\_1 + \beta\_2X\_2 + \cdots + \beta\_nX\_n \tag{10}$$

In the equation, p represents the probability of the dependent variable (such as the probability of winning) taking the value 1. X_1, X_2, ..., X_n are the independent variables (explanatory variables), and β_0, β_1, ..., β_n are the model parameters.

3.4.2 Using the Logistic Regression model to analyze the impact of δM

We referred to the ACF and PACF plots and observed autocorrelation showing a "tail" pattern. Since future values in a time series may be influenced by past values, the tail

pattern implies strong correlation. Thus, we can preliminarily conclude that δM has an impact on the outcome.

We first instantiated the 'judge' class and then selected DELM and dDELM from the dataframe as independent variables (explanatory variables) and 1ifwin as the dependent variable (response variable). By adding a constant term to the independent variables, we fitted the logistic regression model.

Referencing the results of the Logistic Regression model:

Table 3: Logit Regression Results

Variable	Coefficient	Standard Error	z	P>|z|
const	-0.0142	0.113	-0.126	0.900
DELM	-0.0679	0.025	-2.742	0.006
dDELM	-0.0018	0.084	-0.022	0.983

In this model, the p-value for the variable DELM is 0.006, which is less than the commonly used significance level of 0.05. This indicates that the coefficient of the DELM variable is significantly different from zero. Therefore, it can be concluded that the DELM variable has a significant impact on the dependent variable.
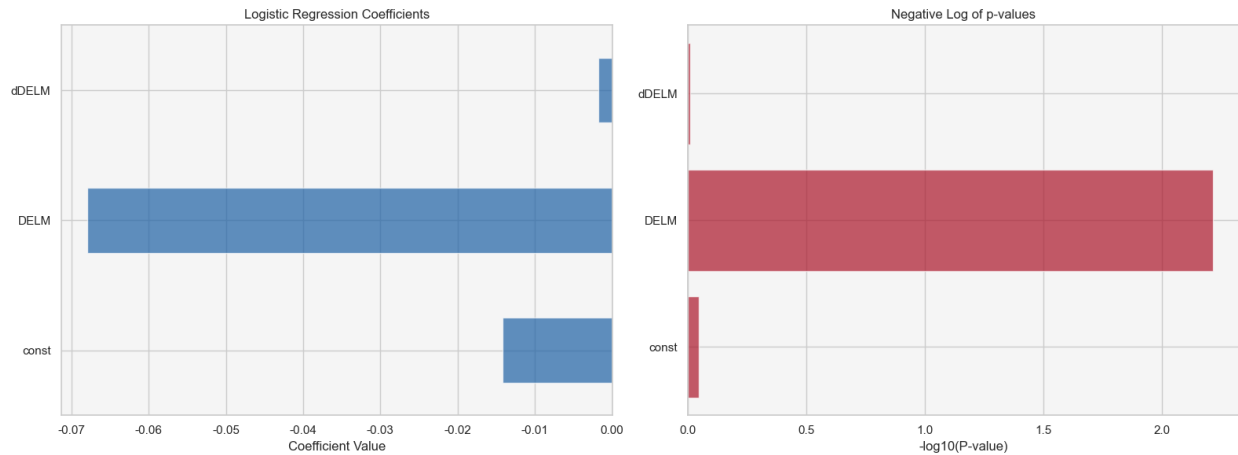

Figure 8: p-value

Therefore, we can conclude that the fitted DELM has a significant impact on the outcome of the match. This can refute the coach's idea.

4 Model 2:Analysis and Evaluation Model Based on Bayesian Spatiotemporal Model

4.1 Characterization of the fluctuation probability in tennis matches

Characterizing Factors Influencing Fluctuations in Tennis Matches

For both players and coaches, understanding the factors influencing fluctuations can better capture signs of momentum reversal. By changing strategies, namely the factors affecting fluctuations, one can shift the momentum of the tennis match in their favor.

Based on the conclusions of the momentum model described in the previous section, as shown in Figure 3.2-1, our momentum model exhibits strong fluctuations over time. Meanwhile, dDELM, the rate of change in momentum difference between both sides, implies the final outcome of the match. As depicted in Figure 3.2-2, the high correlation between dDELM and match results suggests that using dDELM as a criterion for assessing fluctuations in tennis matches is highly appropriate. Therefore, our aim to identify factors influencing fluctuations essentially involves determining the determinants of dDELM. For convenience, we refer to dDELM as the reversal factor, denoted as G.

Based on the dataset, we derive the distribution of all G, denoted as Ω. We observe that G is roughly concentrated between four values, with a limited distribution range that is approximately symmetrical about the origin. Combining this with empirical assumptions, we hypothesize that the possible values of G for each player during the match are concentrated in the [-2, 2] interval, and roughly follow a "U"-shaped distribution about 0.
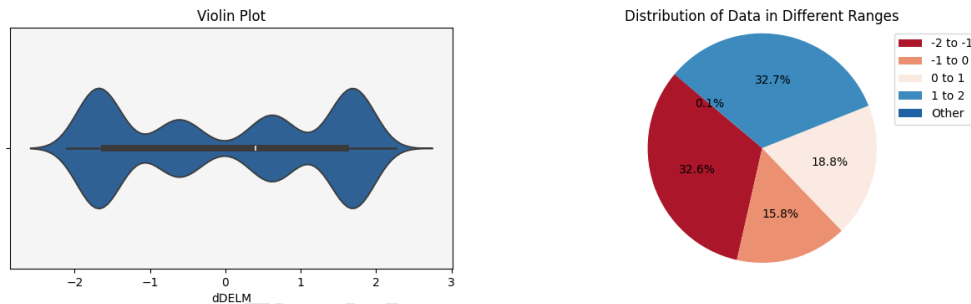
Figure 9: Violin plot & Pie chart

The fluctuations in tennis matches exhibit both temporal and spatial correlations, with players acting as spatial factors and match time serving as the temporal dimension, jointly determining the ups and downs of tennis matches. The Bayesian spatio-temporal model is a mathematical model established within the framework of Bayesian statistics to analyze the temporal and spatial information inherent in spatio-temporal data. It utilizes prior knowledge and conditional probabilities to compute the posterior distribution of parameters, with the most common method being Markov Chain Monte Carlo (MCMC) for computation. Considering the Bayesian spatio-temporal model's suitability for fitting time, space, and probability distributions (i.e., the probability of an event occurring),[6] [7] we use the log-logistic regression under the Bayesian spatio-temporal model to fit and analyze other variables.

4.2   Correlating factors affecting match fluctuations
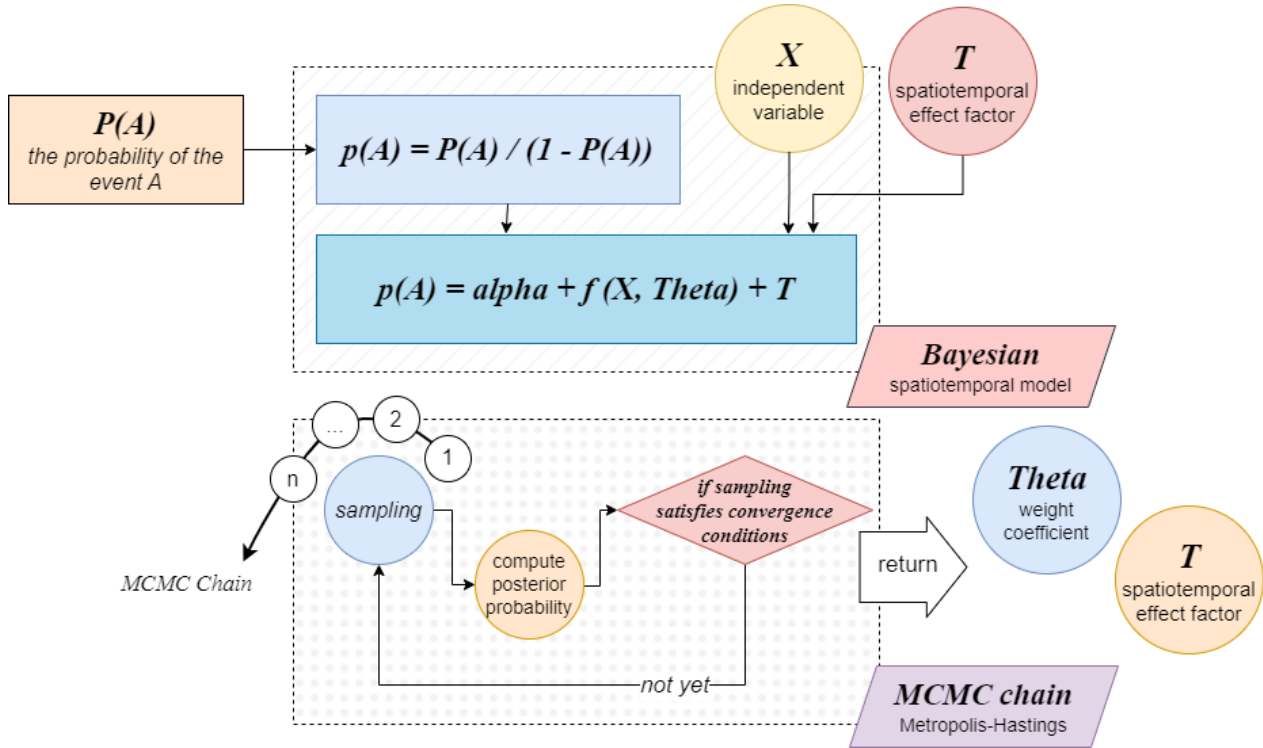
Figure 10: Model Logic Flowchart

Considering a model with random time effects in a single match, we can transform players' fluctuations during the game into the probability of one player gaining an advantage by the following means.

$$G_t < 0 \tag{11}$$

Obviously, when player 1 is at a disadvantage, and vice versa when

$$G_t \cdot G_{t+1} < 0 \tag{12}$$

we consider that a reversal of fortune has occurred. Let Ω represent the sample space composed of all G, and we use the relative distribution of G to represent the probability of player 1 having an advantage.

$$P(t) = \frac{n(i|G_i \in [G_{min}, G_t])}{n(\Omega)} \tag{13}$$

We found that in the bipolar distribution of G, when a player is at the lower (upper) bound, the probability of a comeback (being counterattacked) is also lower. This aligns with our everyday understanding and the sample space formed by the G depicted in the above figure.

To explore the relationship between G and other variables we are interested in, we can, based on log-logistic and Bayesian spatiotemporal model theories, calculate the probability distribution column P for all G. Subsequently, we can use this P value in regression analysis with the other variables we want to explore and introduce.

Based on the log-logistic and Bayesian spatiotemporal models, we obtain the equation:

$$logitP(i) = \alpha + f(X(i), \Lambda) + \tau_i \tag{14}$$

to fit probability values In this equation:

i — data point τ_i — Time random effect factor X — Matrix of independent variables f — Quadratic polynomial Λ — Weighting matrix

For variables representing two players, such as P1_Distance_Run and P2_Distance_Run, we can eliminate the influence of spatial random variables on the coefficients by taking the difference.

For categorical variables, such as the 5 different serving methods in Serve_Width, which are not numerical, to account for the impact of different serving methods on P, we will use the method of dummy variables. For each sample point, we use a matrix [x_0, x_1, x_2, x_3, x_4], where x_i $\in$ 0,1 represents Serve_Width, and x_i indicates whether the i-th serving method is used. To assess the effects of different serving methods, i.e., when measuring the effects under different values of x_i, we construct a coefficient matrix Θ = [k0, k1, k2, k3, k4] to evaluate the individual effects of x_i. The advantage of this approach is that it can reduce errors in the analysis of the causes and effects of independent variables due to data processing issues.

Furthermore, variables like Serve_Width can exhibit fluctuations in their impact on P depending on the Server. Therefore, we first appropriately handle the values of Server, mapping them to -1,1. Subsequently, we construct interaction terms between Server and variables affected by its fluctuations, such as Serve_Width, to reduce this influence.

To estimate the coefficients in the logistic regression expression, i.e., the coefficient matrix in the expression, we use an MCMC chain for regression analysis. After properly processing the dataset, it is imported into WinBUGS14 for iteration. When the number of iterations is set to 10,000, the obtained coefficients are as follows:

The weight coefficients obtained from WinBUGS14 roughly follow a normal distribution, aligning with the model's assumptions. The MCMC chain converges, indicating a good fit.

Some of the parameter confidence intervals include zero, indicating that the effect of those independent variables on the model is not significantly pronounced, such as the intercept; or it could be that the variable itself exhibits variability, as in the case of Return_Depth.

On the other hand, variables with confidence intervals (CI) excluding zero, especially those with narrow upper and lower bounds, exhibit a stronger and more pronounced impact on the model, such as the serving side. According to our model, the serving side plays a significant role in numerous interaction terms, influencing variables like the serving technique to a considerable extent.

Node			mean	sd	MC error	95% CI		
						2.50%	median	97.50%
Intercept			0.008	0.074	0.002	-0.140	≈ 0	0.151
G			0.206	0.055	0.002	0.094	0.206	0.319
Delta Score			0.008	0.005	<0.001	-0.001113	0.007	0.01644
Delta Run_distance			-0.002	0.002	<0.001	-0.006109	-0.002	0.002417
Speed_mph			-1.794	0.394	0.011	-2.59	-1.794	-1.027
Net			0.082	0.129	0.003	-0.171	0.082	0.335
BreakPoint			-0.208	0.223	0.007	-0.651	-0.203	0.215
Server			-7.398	1.631	0.163	-9.952	-7.583	-4.534
Interaction term between Server and G			0.049	0.046	0.001	-0.042	0.049	0.141
Interaction term between serving/receiving style and serving side	Serve_Width	B	6.681	1.988	0.184	3.115	6.734	9.937
		BC	7.374	1.854	0.176	3.579	7.409	10.36
		BW	7.749	1.621	0.156	5.004	7.738	10.47
		C	8.265	1.586	0.154	5.475	8.295	10.89
		W	8.445	1.637	0.163	5.526	8.619	11.02
	Serve_Depth	CTL / NCTL	-0.151	2.424	0.243	-6.41	0.519	3.455
	Return_Depth	D / ND	-0.205	2.433	0.244	-3.833	-0.853	6.107

Figure 11: Correlation Coefficient

4.3   Discussion on Model Fit Accuracy

We analyzed the distribution of the time random effects factor, represented by the mean (red line), and its 95% confidence interval (yellow line), as depicted in the graph below. From the plot, we observe a significant fluctuation in the time factor over time, suggesting the rationality of the time random effects factor we have set.
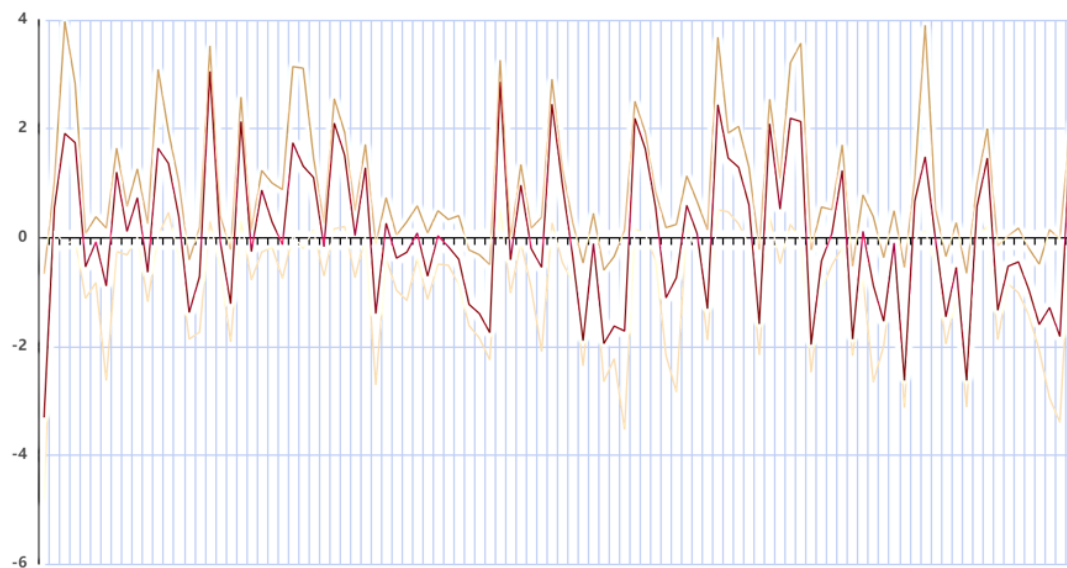

Figure 12: Time Factor

After obtaining the coefficient matrix, intercept, and time random effects factor, we proceeded to predict P, resulting in the predicted values (pre). The fitted graph is shown

below, with Point_no on the horizontal axis and probability values on the vertical axis.
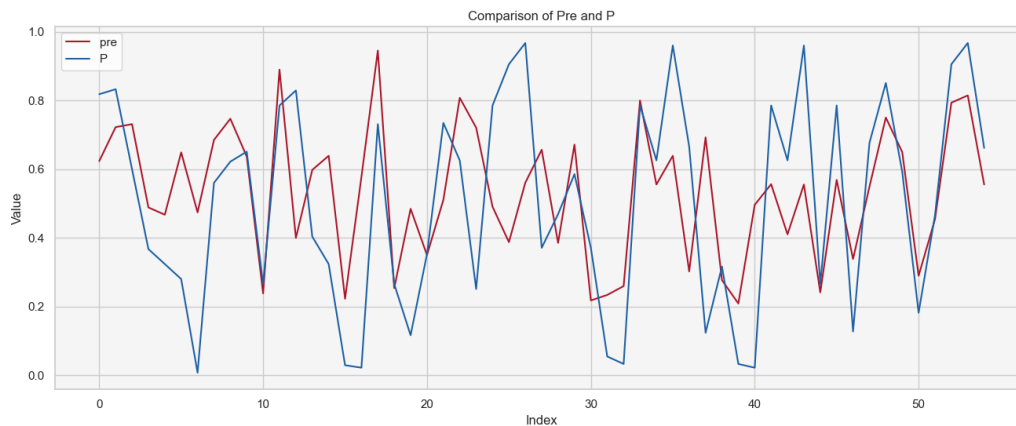

Figure 13: Level of Fit

4.4 Recommendations Based on the Impact of Various Factors on Matches

Given the numerous uncontrollable factors on the court, such as the opponent's serve direction and intensity, to provide players with actionable recommendations for altering the situation, we should focus on measures they can effectively execute in response to the next ball. For instance, adjustments in serving posture or receiving stance. Utilizing the model, we calculate the change in the probability variable (ΔP) for each specific scenario. By determining the distribution proportion where $\Delta P > 0$, we obtain a probability distribution table indicating their likelihood of changing the game situation based on altering different factors.

When player1 is the server (with 'Return_Depth' column indicating the opponent as the returner), the calculation yields:

Serve Width					Serve Depth		Return Depth	
B	BC	BW	C	W	CTL	NCTL	D	ND
0.228	0.345	0.367	0.533	0.545	0.843	0.272	0.414	0.249

Figure 14: Suggestions

This table illustrates that when the player is serving, choosing strategies C, W, and CTL can increase the probability of scoring. On the other hand, when receiving, opting for return strategy D can enhance the likelihood of scoring.

5   Model 3: Momentum Evaluation Based on Cross-Validation Model

5.1   Testing the Universality of the Model

With a relatively good fit obtained in the third question, to assess how well the model can predict fluctuations in matches, we employ cross-validation techniques. Leave-One-Out (LOO) cross-validation is a straightforward yet effective method. In LOO cross-validation, one sample is retained as the test set each time, while the remaining samples serve as the training set. This means that for a dataset containing n samples, the model is trained and evaluated n times, using n-1 samples for training in each iteration and testing with the remaining one sample.

The primary advantage of LOO cross-validation is its ability to make the most use of the data for model evaluation, as each sample is used for testing. However, its drawback lies in the relatively high computational cost, especially for large datasets, as it requires training a substantial number of models.

First, we compute differences for data processing and preparation. Then, we use r2_score as the evaluation metric. Finally, by sequentially using each file as the test set and the rest as the training set, we assess the model's generalization ability on different datasets.

The accuracy of the prediction model is illustrated in the following figure, with an average accuracy value of 93.17%. This indicates that our model performs relatively well in predicting fluctuations in matches.
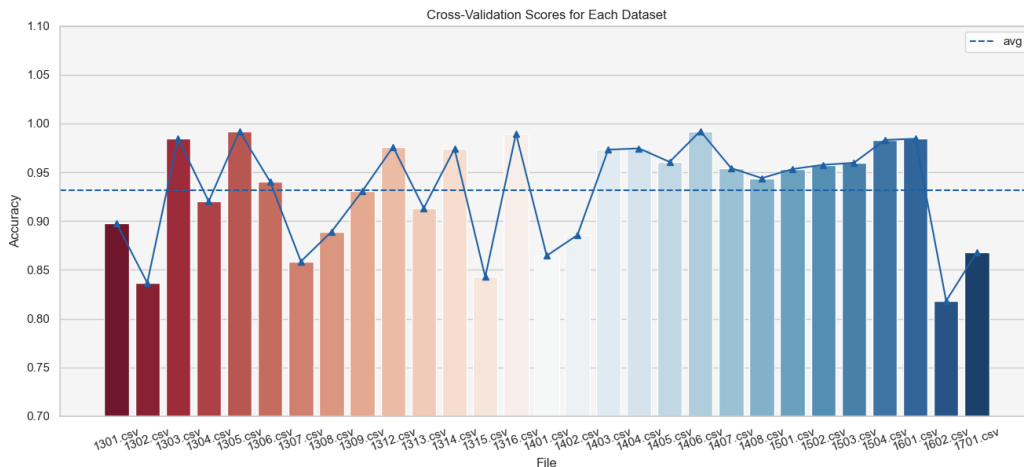

Figure 15: Accuracy

5.2   Enhancing the Model by Incorporating Relevant Factors

Based on our model and the observed fitting results, there are areas for improvement. The model exhibits a lag in predicting certain data and appears insensitive to data with

low volatility. For a more comprehensive model in future development, we can consider incorporating spatiotemporal random effects for various match time points, especially when dealing with multiple events. I suggest including the following factors into consideration.

1. The world rankings and historical head-to-head records of the two players: From the first model, we observe a significant correlation between changes in momentum and consecutive victories, as well as the outcome of the previous match. Both factors play a crucial role in boosting a player's confidence and maintaining a good competitive state.

2. Surface conditions: Different playing surfaces affect the ball's bounce speed and height, as well as the players' movement on the court. Players adapt differently to surfaces like hardcourt, grass, and clay, which can influence the speed of their adaptation to the playing conditions.

3. Audience and home court advantage: When players compete on their home turf, there is often more cheering and encouragement. These positive stimuli can quickly excite players and help them get into the groove more rapidly.

4. Climate conditions: Temperature can alter the elasticity of the ball, humidity affects the ball's flight speed and bounce height, and both factors can contribute to increased fatigue for the athletes. These elements collectively influence a player's performance.

5. Psychological expectations and pressure: While victory is a common goal for every player, each match holds different significance for individual participants. When a particular match is imbued with the potential to impact a player's career, the expectations and tension surrounding that match can be extraordinary. This pressure has the potential to either motivate the player to perform exceptionally well or adversely affect their psychological state, hindering them from getting into the desired mindset.

6. Player's recent form: If a player is in a slump in their recent performances, both their confidence and momentum can be significantly impacted. Adjusting and bouncing back from such a phase is not an easy task in a short period.

7. Tournament format: The rhythm control and energy allocation for players differ between formats like best-of-five sets and best-of-three sets. This variation can also influence their performance.

6 Promotion of the Model

6.1 Rational Analysis of Momentum Indicator M in Women's Matches

According to the literature, we analyzed the scoring performance of world-renowned female tennis players in the context of Breakpoints during the period from 2018 to 2019 under two conditions: 'whether serving' and 'whether leading in game points.' We observed that both being the server and leading in game points can enhance the scoring rate in crucial Breakpoint situations.[8] [9] The identity of the server remains a significant

factor, showing a strong correlation with the momentum indicator M we constructed. Additionally, leading in game points can serve as another factor to assess the magnitude of momentum and be incorporated into our model.

NAME	Breakpoint Win Rate			
	Is Server		Not Server	
	Ahead in score	Behind in score	Ahead in score	Behind in score
Ashleigh Barty	61.11%	60.00%	41.66%	66.66%
Naomi Osaka	67.35%	48.15%	50.00%	26.47%
Simona Halep	68.06%	46.88%	59.09%	45.45%
Aryna Sabalenka	77.78%	62.50%	46.15%	50.00%
Petra Kvitova	76.92%	75.00%	37.50%	36.36%
Serena Williams	75.56%	38.89%	48.27%	37.28%
Madison Keys	63.16%	33.33%	7.69%	37.28%
Elina Svitolina	37.50%	35.71%	25.00%	33.33%
Elise Mertens	75.00%	62.50%	66.66%	41.17%
Angelique Kerber	37.50%	62.50%	70.00%	35.71%
Sloane Stephens	66.67%	14.29%	60.00%	32.00%
Karolina Pliskova	25.00%	37.50%	50.00%	30.00%
Angelique Kerber	55.56%	20.00%	25.00%	20.00%
Garbine Muguruza	50.00%	20.00%	33.33%	60.00%
Caroline Wozniacki	57.89%	53.85%	33.33%	37.50%
Amanda Anisimova	55.56%	28.57%	20.00%	33.33%
Average Skill Level	**64.92%**	**45.77%**	**46.85%**	**38.40%**

Figure 16: Women Breakpoint Win Rate in Different Situations

In fact, women's tennis matches share many factors influencing the course of the game with men's matches, such as the psychological state of players, technical abilities, and external environmental factors, all of which can impact the players' performance. However, the two cannot be entirely equated, and women's matches differ from men's matches in the following aspects:

1. Differences in match duration and format

2. Variances in physical factors leading to variations in ball speed and running distances

3. Distinct preferences for winning through serving or receiving

4. Disparities in the rate of physical recovery during consecutive matches

5. Differences in handling psychological pressure

Due to the differing factors influencing momentum between men's and women's matches, simply transferring parameters obtained from men's matches into women's matches could

potentially result in decreased predictive accuracy. However, employing a Bayesian spatio-temporal model with logit regression analysis can still recalibrate weight coefficients based on data from women's matches. Additionally, normalizing the data can mitigate biases arising from differences in average match duration and running distances between the two types of matches.

6.2 Analysis of the Rationality of Momentum Indicator M in Different Court Conditions

References	Surface	Level	HR / *bpm*	[La] / *mmol L-1*	Volumn of O2 / *ml min-1 kg-1*	% Volumn O2*max*
Murias et al. 2007	*Clay*	*National*	143	1.65	26.33	47.6
	Hard	*National*	135	1.16	27.48	49.5
Girard and Millet 2004	*Clay*	*Club*	181.8	2.36	40.3	80.7
	Hard	*Club*	172.8	3.08	35.9	71.6
Martibn et al. 2011	*Clay*	*National*	154	5.7		
	Hard	*National*	141	3.6		

Figure 17: Geographical Variations

According to the literature, different playing surfaces represent distinct spatial environments. For instance, the oxygen content in the environment of club competitions is relatively high, which may have an impact on the players' stamina. The differences between gravel and hard surfaces also manifest in varying levels of physical activity, reflected in heart rate (HR). Therefore, we conclude that different spatial environments can influence serving, receiving, running distances, and even the athletes' physical fitness.[10]

To mitigate these influences, we can incorporate spatio-temporal effect factors into the model expression. This involves adding factors to characterize random effects in different playing environments. By obtaining fitted values for these factors, the model's generalization capability is expected to be enhanced.

6.3 Rational Analysis of Momentum Indicator M in Table Tennis Matches

According to the literature, there is a strong correlation between the momentum value in table tennis matches and the serving side, serving methods during the serving and receiving phases, as well as the hitting methods during the receiving phase.

	Serve and Receive Phase				
	Position	*Placement*	*3rd-ball attack position*	*Skill and strategy*	*Is change of direction*
$K \times K$	22.046	49.693	12.76	104.244	30.358
p	<0.001***	<0.001***	<0.01**	<0.001***	<0.001***

	Receive-Serve-Attack Phase				
	Position	*Placement*	*4th-ball attack position*	*Skill and strategy*	*Is change of direction*
$K \times K$	14.818	24.492	12.034	63.656	12.957
p	<0.05*	<0.001***	<0.01**	<0.001***	<0.05*

Figure 18: Independence Test for Chi-square Plots

Thus, we observe that, similar to tennis matches, the serving side, serving methods, hitting techniques, time sequence, and other factors continue to influence the momentum indicator M. Therefore, it remains reasonable to use factors such as (Score, Server) to characterize M. From the aforementioned data, it is evident that the winning strategies in table tennis have a stronger correlation with tactics compared to tennis. Hence, there is a greater need to amplify and refine the factor of 'strategy' and incorporate it into the logistic regression expression for M.

7   Sensitivity Analysis

We adjust the weights of $\prod_{i=1}^{n} K_i$ and $\prod_{j=1}^{m} P_j$. Let's assume the weight of $\prod_{i=1}^{n} K_i$ is denoted by α, and the weight of $\prod_{j=1}^{m} P_j$ is $2 - \alpha$.

Then, through sensitivity analysis:

In the generation of accuracy, two methods, Kendall and Pearson, were employed to assess data correlation. These two methods are commonly used in statistical analysis to measure the correlation between two variables, especially for evaluating the rank correlation of two random variables.
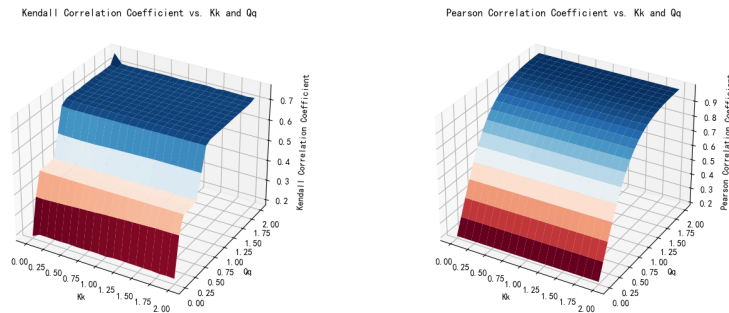

Figure 19: Correlation Coefficient

After k exceeds 0.9, similar effects are observed, and the coefficient does not consider the influence of dDELM on the correlation with winning or losing excessively. It can be considered that Kk has negligible impact on correlation after exceeding 0.9. Choosing 1 is more reasonable, and further adjustment is unnecessary. The chart shows that the influence of Qk is not significant, hence selecting a coefficient of 1 is more appropriate, and no further adjustment is required. This indicates that when k exceeds 0.9, the model is not sensitive to parameter changes and remains relatively stable.

8 Model Evaluation

8.1 Advantage

1. Use the ARIMA time series model to forecast the trend of momentum (M). The differential analysis and shifting methods of the ARIMA time series model are particularly effective in short-term forecasting of trends related to time factors, especially for trends with significant fluctuations.

2. Utilize a logistic regression model under Bayesian spatiotemporal theory to assess the factors influencing M. Compared to using simple classification algorithms that treat volatility as a binary distribution, transforming our model into conditional probability provides some continuity to the dependent variable. Additionally, the model's identified influencing factors generally align with the model's prior assumptions, indicating good predictive performance.

3. Employ cross-validation to evaluate the first two models, thereby improving and optimizing our models for greater stability and generality. Post-validation, our model achieves a prediction accuracy of 93.17% on other datasets. Sensitivity analysis indicates that our model exhibits good generalization capability.

4. The momentum expression we propose encompasses the Score, Strength, and Server (3 S)—quantities with clear intuitive relevance to the outcome. We also creatively introduce a fatigue index based on running distance to better establish the momentum expression. The formulation is comprehensive and logically structured.

5. Our preprocessing of raw data involves non-rigid quantization (normalization) and the substitution of dummy variables. The former ensures the generalization capability of the model under different scales (e.g., women's events), while the latter better captures the independent effects of different categories under a single variable.

8.2 Disavantage

1. The model's generalization can have a broader scope by refining the interaction between spatial factors (such as information about both teams and the venue) and temporal factors.

2. The model exhibits a lag effect when predicting certain data and is less sensitive to data with low volatility.

Memorandum

To: Coaches
From: Team 888888
Date: February 5th, 2024
Subject: Enhancing Athlete Performance in Matches through Momentum

In modern professional tennis competitions, players have gained significant experience in their pre-match preparations and tactical planning. Through systematic physical training and cleverly designed game strategies, they ensure that they can perform at their best on the court. However, with the continuous improvement of athletic levels and increasingly fierce competition, relying solely on traditional preparation methods is no longer sufficient. Even the most outstanding professional players may experience fluctuations in their performance and could be overturned in seemingly advantageous situations. We introduce a new metric, momentum, which can comprehensively integrate various aspects of a player's performance, adapt effectively, and accurately predict the next game's fluctuations. This, in turn, allows for the analysis of an athlete's matches and the formulation of targeted recommendations to enhance their performance in games.

After establishing and improving our model and discussing the relationship between influencing factors and outcomes, we have determined that there is a strong correlation of up to 90% between the first-order difference in momentum between two players and the outcome of the match, with a positive value indicating a higher likelihood for the player to score in the next point. Additionally, momentum can effectively consolidate various performance metrics used to evaluate a player's on-court performance and process them to determine the impact of different factors on a player's performance. This helps coaches provide tailored guidance to the athletes.

Based on the data from the 2023 Wimbledon Men's Tennis Tournament, we have found that momentum is primarily correlated with a player's confidence. When a player is on a winning streak or in an advantageous serving situation, their momentum tends to increase significantly. Conversely, when a player experiences consecutive losses or makes serving errors, their momentum tends to decrease significantly. In addition to this, objective factors such as the level of physical fatigue and the characteristics of the playing surface can also influence a player's condition and momentum.

References

[1] Trafimow, D. (2017) Why It Is Problematic to Calculate Probabilities of Findings Given Range Null Hypotheses. Open Journal of Statistics, 7, 483-499. `https://doi.org/10.4236/ojs.2017.73034`.

[2] Braidwood, J. (2023), Novak Djokovic has created a unique rival – is Wimbledon defeat the beginning of the end, The Independent.

[3] Headrick, T. (2016) A Note on the Relationship between the Pearson Product-Moment and the Spearman Rank-Based Coefficients of Correlation. `http://dx.doi.org/10.4236/ojs.2016.66082`.

[4] BOX G E P,JENKINS G M. Time series analysis forecasting and control(Rev ed) [J] . Journal of Time,1976, 31(4) :238-242

[5] Ahn, H. (2014) Effect Modeling of Count Data Using Logistic Regression with Qualitative Predictors. Engineering, 6, 758-772. `http://dx.doi.org/10.4236/eng.2014.612074`.

[6] Besag J, Mollié Y A. Bayesian image restoration, with two applications in spatial statistics[J]. Annals of the Institute of Statistical Mathematics, 1991, 43: 1-20. [48] Bernardinelli L

[7] Clayton D, Pascutto C, et al. Bayesian analysis of space-time variation in disease risk[J]. Statistics Medicine, 1995, 14(21-22): 2433-43

[8] Heinz Kleinoder. The ITF coaching&sport science review. The year issue 24.July2001

[9] Nick Bollettieri. Keep score in mind-play your game accordingl y. Tennis magazine,2007

[10] Barnett, Tristan and Pollard, Graham (2007) How the tennis court surface affects player performance and injuries.* Medicine and Science in Tennis, 12 (1). pp. 34-37. ISSN 1567-2352 `https://www.independent.co.uk/sport/tennis/novak-djokovic-wimbledon-final-carlos-alcaraz-b2376600.html`