TCDformer-based Momentum Transfer Model for Long-term Sports Prediction

Hui Liu[a,1], Xiyuan Huang[a,1], Jiacheng Gu[b,1], Junjie Shi[b], Ning He[b] and Tongtong Feng[c,*]

[a]*School of Urban Rail Transit and Logistics, Beijing Union University, Beijing, 100101, China*
[b]*The Smart City College, Beijing Union University, Beijing, 100101, China*
[c]*The Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China*

ARTICLE INFO

Keywords:
Long-term Prediction
Momentum Encoding
Sports Prediction
TCDformer

ABSTRACT

Accurate sports prediction is a crucial skill for professional coaches, which can assist in developing effective training strategies and scientific competition tactics. Existing sports prediction models use complex mathematical statistical techniques to boost predictability, which are limited by dataset scale and have difficulty handling long-term predictions with potential change points, notably underperforming when predicting point-set-game multi-level matches. To deal with this challenge, this paper proposes *TM²*, a TCDformer-based Momentum Transfer Model for long-term sports prediction, which encompasses a momentum encoding module and a momentum prediction module. The momentum encoding module applies the Maximal Overlap Discrete Wavelet Transform to detect multiscale potential change points through sign variations in wavelet coefficients, and quantifies player momentum by cumulatively aggregating the weighted importance and values of indicators at each time point. The momentum prediction module first decomposes each player's momentum into trend and seasonal components using the local linear scaling approximation (LLSA). The final prediction results are derived from the additive combination of a multilayer perceptron (MLP) for predicting trend components and wavelet attention mechanisms for seasonal components. Comprehensive experimental results show that on the 2023 Wimbledon men's tournament datasets, *TM²* significantly surpasses existing sports prediction models in terms of performance, reducing MSE by 96.26% and MAE by 79.93%, and *TM²* can also be generalized well to other sports event predictions. The source code and datasets are openly available at https://github.com/Liuhui548484/TM2 to support reproducibility and further research.

1. Introduction

Sports hold a pivotal role in modern society, serving not only as a source of entertainment but also as a critical mechanism for enhancing physical and mental health. Iconic events such as the Olympic Games, FIFA World Cup, NBA, and the Championships of Wimbledon have become deeply embedded in global culture and daily life. According to a report from Technavio[2], the global sporting events market size is projected to grow by USD 107.28 billion between 2024 and 2028, with a 22.66% annual growth rate, underscoring the immense economic impact of sporting events on societies worldwide. Iccopr's report[3] highlights a growing focus on physical and mental well-being, with participation in healthy sports activities expanding dramatically. By 2025, an estimated 3.14 billion individuals, or 50.4% of the global population under the age of 60[4], are expected to engage in regular physical exercise. KPMG's report[5] indicates that for every 1,000 sports users, 2.16 coaches are required, making professional sports coaching a rapidly growing and increasingly popular profession.

Accurate sports prediction (Papageorgiou et al., 2024b; Markopoulou et al., 2024) is a crucial skill for professional physical education instructors and coaches because this skill helps them develop effective training strategies and scientific competition tactics (Papageorgiou et al., 2023), and make wise real-time strategy optimization (Sarlis et al., 2023) during games. Specifically, coaches can thoroughly review students' or athletes' sports states and mentality changes, assess their teams' and opponents' strengths and weaknesses (Papageorgiou et al., 2024a), tailor their coaching techniques, and set more realistic and attainable competition objectives (Papageorgiou et al., 2024c) accordingly. Additionally, accurate sports prediction can help prevent injuries (Sarlis et al., 2024a) by adjusting workloads and identifying high-risk situations, contributing to the overall well-being of athletes (Sarlis et al., 2024b).

Existing sports prediction models have made significant progress by leveraging expert experience and machine learning techniques, and can be divided into three categories: 1) Knowledge-driven Models (Panagoulias et al., 2024). These models are constructed based on the experience of

[⋆]This work is supported by the National Natural Science Foundation of China (62272049, 62236006, 62172045)

[*]Corresponding author.

liuhui@buu.edu.cn (Hui Liu); xiyuanhuang@buu.edu.cn (Xiyuan Huang); gujiacheng@buu.edu.cn (Jiacheng Gu); shijunjie@buu.edu.cn (Junjie Shi); fengtongtong@tsinghua.edu.cn (Tongtong Feng); xxthening@buu.edu.cn (Tongtong Feng)

[1]Equal Contribution.

[2]https://www.technavio.com/sample-report/sporting-events-market

[3]https://iccopr.com/wp-content/uploads/2019/03/Sports-Around-the-World-report.pdf

[4]https://datareportal.com/social-media-users/

[5]https://assets.kpmg.com/content/dam/kpmg/cn/pdf/zh/2021/09/olympic-economics-and-sports-industry-outlook.pdf

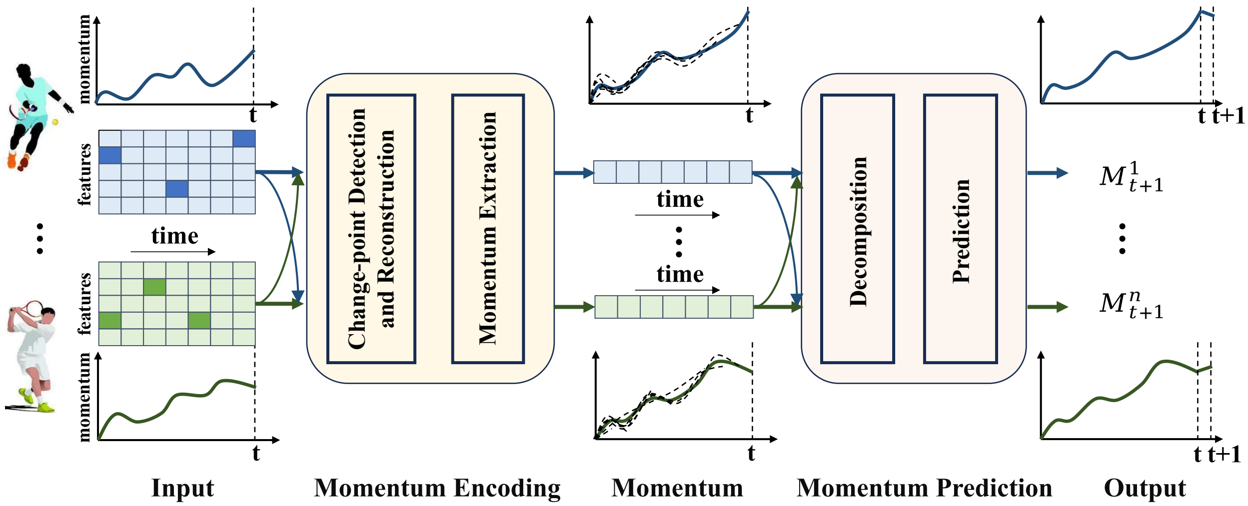

Figure 1: The overview of *TM*^{2}, a TCDformer-based momentum transfer model for long-term sports prediction.

domain experts. They conduct predictions through subjective analysis of key scenarios (such as tennis serving patterns and football corner kick tactics). 2) Data-driven Models (Khan & Ahamad, 2024). These models analyze structured data (such as scores and running distances) and tracking data (player trajectories) by utilizing machine learning techniques (such as LSTM and Transformer). 3) Knowledge-data-driven models (Rodríguez-Rodríguez et al., 2024). These models combine expert knowledge with statistical approaches (such as grey correlation analysis) to balance qualitative and quantitative predictions.

Nevertheless, existing sports prediction models suffer from two key challenges: **1) They are hard to handle long-term predictions with potential change points**. Given the similarity between change points and trends in non-stationary time series data, ignoring the impact of potential change points (specifically, apparent fluctuations driven by external events rather than simple noise) may lead to misleading conclusions. Existing sports prediction models use complex mathematical and statistical techniques to boost predictability, such as CF-LSTM (Xu et al., 2020) and Seq2Event (Simpson et al., 2022), ignoring temporal discontinuities and personnel dynamics across multi-event sequences. Those methods are limited by dataset scale and have difficulty handling long-term predictions with potential change points. **2) They are hard to predict point-set-game multi-level matches**. Sport events are essentially a complex system formed by the real-time coupling of strategic factors (tactical adjustments), psychological factors (stress resistance), and environmental factors (opponent behavior). Existing models are difficult to decouple at multiple scales using statistical analysis, such as ShuttleNet (Wang et al., 2022) and DMA-Nets (Ji et al., 2021), which only focus on each game observation. Those methods lead to difficulties in coordinating short-term fluctuations with long-term trends when predicting multi-level competitions (point-game-set). Ultimately, this results in poor performance in high-risk scenarios, such as comebacks and tiebreaks.

Momentum (Briki, 2017; Morgulev & Avugos, 2023) in sports prediction is defined as the dynamic evolution of competitive advantages driven by real-time performance metrics (e.g., serve accuracy, break-point conversion), psychological states (e.g., confidence, pressure), and contextual factors (e.g., tactical adjustments). It captures both short-term fluctuations (e.g., psychological shifts during rallies) and long-term trends (e.g., physical fatigue across sets) (Liang et al., 2024), offering a comprehensive characterization of athletes' evolving capabilities. At the same time, social relationships between athletes also have an impact on the outcome of sports events (Eckardt & Tamminen, 2023).

Nevertheless, momentum is difficult to directly apply to existing works to solve long-term sports prediction problems. *1) Static intra-point momentum assumption* (Wan et al., 2024). Existing works assume momentum remains constant within discrete time points, such as TFDML (Levy & Lopes, 2021), which relies on fixed momentum representations and encodes point momentum as static values, ignoring intra-point dynamics such as psychological volatility during rallies or mid-game tactical adaptations. *2) Neglected temporal dependencies* (Neumann & Fischer, 2023). Existing works treat sequential actions (e.g., serves across games) as independent events, disrupting autocorrelations between historical and current performance (e.g., fatigue trends or confidence accumulation from consecutive wins). Such as IMO-KF (Qiu et al., 2024) exemplifies this flaw by partitioning temporal sequences into isolated training units. *3) Ignored cross-individual momentum interactions* (Zhou et al., 2024). Existing works analyze individual momentum in isolation, neglecting bidirectional competitive influences, such as DE-TSMCL (Gao et al., 2024) models, which treat momentum as an isolated event-specific property. As demonstrated in Table 2, a dominant opponent (e.g., Novak Djokovic) reduces rivals' psychological factor by 35% on average, but such interactions remain unquantified in current models.

To address these challenges, in this paper, we propose *TM²* (as shown in Figure 1), a TCDformer-based Momentum Transfer Model for long-term sports prediction, which encompasses a momentum encoding module and a momentum prediction module. The encoding module encodes the time series data to identify potential change points, reconstruct new time series features, and calculate the player's real-time momentum curve. The momentum prediction module decomposes each player's momentum into trend and seasonal components and obtains the final momentum prediction results by using a multilayer perception to predict trend components and wavelet attention mechanisms to predict seasonal components. We conduct extensive experiments comparing TM^2 with baseline models (ELO (Neumann & Fischer, 2023), Decision Trees (Kjamilji, 2024), Logistic Regression (Wasi & Abulaish, 2024), Support Vector Machines (Pang et al., 2022), Random Forests (Alfarizi et al., 2022)) and advanced neural network models (Mamba4Cast (Bhethanabhotla et al., 2024), Seq2Event (Simpson et al., 2022)). On the 2023 Wimbledon Men's Championships dataset, TM^2 reduces prediction errors by 96.26% (MSE) and 79.93% (MAE) compared to baseline models, and outperforms advanced models by 6.99% (MSE) and 17.53% (MAE), respectively. On the NBA and beach volleyball datasets, TM^2 outperforms existing models by 12% and 35% in MSE and MAE, respectively.

The rest of this paper is organized as follows. Section 2 reviews related works. Section 3 presents the design details of TM^2. Section 4 presents the implementation details and the evaluation results of TM^2. Section 5 presents the conclusion and future work of this paper.

2. Related Works

2.1. Sports Prediction Models

With the rise of machine learning (ML) and deep learning (DL), sports prediction models have evolved significantly. Zhiqiang Pu et al. (Pu et al., 2024) categorized sports prediction models into knowledge-driven, data-driven, and knowledge-data-driven models. Table 1 shows several excellent models in recent years, evaluated based on their accuracy and efficiency. Those models listed include CF-LSTM (Xu et al., 2020), DNRI (Graber & Schwing, 2020), Seq2Event (Simpson et al., 2022), and ShuttleNet (Wang et al., 2022), each with unique features and applications in sports prediction.

2.1.1. Knowledge-driven Models

Knowledge-driven models traditionally rely on expert knowledge (Panagoulias et al., 2024) to evaluate specific in-game factors and key scenarios, such as serve analysis in tennis or shot selection in basketball. These models focus on expert-defined metrics and situational analysis, which offer valuable insights in specific contexts.

2.1.2. Data-driven Models

Advances in ML and the availability of large datasets have expanded the scope of sports analysis. Data-driven models (Pang et al., 2022) focus on analyzing event data, such as player statistics, and tracking data, such as player movements during matches. Event data typically includes discrete actions like goals, assists, or fouls, while tracking data captures continuous spatial information, such as player positioning and movement trajectories. These models allow for more comprehensive assessments (Khan & Ahamad, 2024), such as evaluating team dynamics or predicting match outcomes based on player performance. However, their reliance on purely data-derived features often prioritizes predictive accuracy over explainability, as complex architectures (e.g., deep neural networks) inherently obscure decision pathways. Therefore, data-driven models lack interpretability.

2.1.3. Knowledge-Data-driven Models

Knowledge-data-driven models (Rodríguez-Rodríguez et al., 2024) combine expert knowledge with ML techniques. For example, such models could integrate tactical expertise with performance data to predict key match events in football. This hybrid approach offers a balance between expert-driven insights and data-driven accuracy (Alfarizi et al., 2022), making it more practical for real-time decision-making in sports. Knowledge-data-driven models fully leverage ML to enhance the predictive power while maintaining interpretability through expert-guided feature selection and model design.

2.2. Time Series Prediction Models

Time series data presents unique challenges due to its dynamic temporal structure and varying patterns across domains. Recent work by Yuxuan Liang et al. (Liang et al., 2024) categorizes time series models into three main types: standard time series models, spatial time series models, and other temporal data models, each with distinct applications and strengths.

2.2.1. Standard Time Series Models

Standard time series models are designed to capture general patterns from large datasets, typically aimed at predicting or classification tasks. These models are often pre-trained on vast amounts of time series data across various domains. For example, Lag-Llama (Rasul et al., 2023) uses a decoder-only transformer architecture, while TimeGPT-1 (Garza & Mergenthaler-Canseco, 2023) adopts an encoder-decoder structure with transformer layers. These models focus on task-specific improvements and are resource-intensive to train from scratch. Other approaches, such as LLM4TS (Chang et al., 2023) and TEMPO (Cao et al., 2023), successfully fine-tune large language models for time series prediction, demonstrating the adaptability of pre-trained models to non-linguistic data.

Table 1
Existing advanced sports event prediction models

Author	Modeling Technique(s) and Features	Data Source	contribution	Years
CF-LSTM (Xu et al., 2020)	Integrate the feature information of the pedestrians from the first two time steps into a separate input to the LSTM and focus on the internal features of the dynamic interactions.	ETH Dataset: 750 pedestrians, 2 scenes (ETH, Hotel); UCY Dataset: 786 pedestrians, 3 scenes (ZARA01, ZARA02, UCY). Both datasets were collected from real-world environments.	Performance improved by approximately 60%	2020
DNRI (Graber & Schwing, 2020)	Formulate explicit recovery of system interactions as NRI of latent variables.	The CMU motion capture database	Average relation prediction F1 of 27.1	2020
DMA-Nets (Ji et al., 2021)	End-to-end RNN-based model with a hierarchical dynamic attention layer is introduced that uses two temporal attention mechanisms to enhance the model's ability to represent complex conditional dependencies in real-world datasets, while the temporal prediction layer ensures that predicted citations are monotonically increasing along the temporal dimension.	The United States Patent and Trademark Office (USPTO) and the Microsoft Academic Graph (MAG)	Average accuracy increased by 25%	2021
Seq2Event (Simpson et al., 2022)	A Combined Model of Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRUs) and Transformers.	Matching event data from the WyScout open access dataset	All Seq2Event models outperformed the best baseline; worst Seq2Event (Transformer: 17 heads, hidden size 4096) achieved test loss 0.548.	2022
ShuttleNet (Wang et al., 2022)	Neural Network with two encoder-decoder extractors and a fusion network.	75 high-ranking matches (2018–2021), 31 players (men's and women's singles), manually labeled by domain experts, collected from public sources.	The proposed approach achieves improvements of at least 12.0% (CE) and 3.4% (MSE) over all baselines.	2022

2.2.2. Spatial Time Series Models

Recent advancements in spatio-temporal graph analytics have demonstrated significant potential for enhancing sports prediction and have become the core foundation of spatial time series models. By leveraging GPS-derived player tracking data, researchers have explored graph-based representations to capture dynamic interactions (Antonini et al., 2024; Kim et al., 2022) and movement patterns (Sheridan et al., 2024). For instance, TGNets (Raabe et al., 2023) employs multilayer graph neural networks to predict possession changes in football by modeling player proximity and temporal connections, highlighting the value of spatial-temporal dependencies. Similarly, Jose Gama et al. (Gama et al., 2014) analyze intra-team passing networks using centrality metrics to identify key players during attacking phases, while uPATO (Silva et al., 2017) developed adjacency matrices to assess team performance across match intervals. These approaches underscore the importance of spatial relationships and temporal evolution in team dynamics. Additionally, Solan Malone et al. (Malone et al., 2016) utilized GPS data to quantify positional running profiles in Gaelic football, providing foundational insights into external load metrics.

2.2.3. Other Temporal Data Models

Beyond traditional time series and spatial-temporal data, other domains such as trajectory data and clinical records also involve temporal dynamics. In human mobility prediction, models like AuxMobLCast (Xue et al., 2022) fine-tune large language models to predict movement patterns,

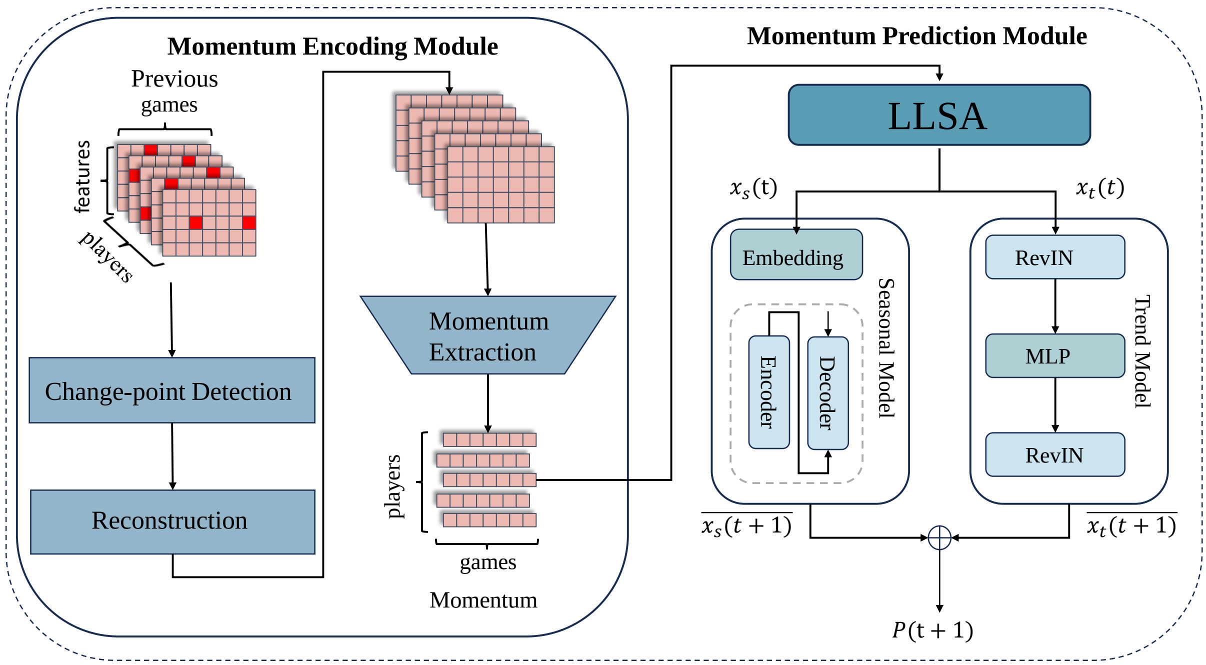

Figure 2: The design details of TM^2 model.

while DiffTraj (Zhu et al., 2023) reconstructs geographic trajectories using diffusion processes. These models illustrate the flexibility of time series methodologies across various industries and their increasing importance in predicting and decision-making processes.

2.3. Momentum prediction Models

Momentum in sports (Lv et al., 2024) encompasses dynamic factors influencing match outcomes, including psychological states, strategic adjustments, and critical turning points. It quantifies the evolving competitive advantage of players or teams through measurable indicators (e.g., scoring streaks, physical performance) and qualitative aspects (e.g., confidence, pressure). For example, Page et al. (Page & Coates, 2017) demonstrate the "win/loss effect" through experimental designs, linking momentum shifts to psychological and strategic factors in ball games; Gauriot et al. (Gauriot & Page, 2018, 2019) analyze psychological momentum in football and strategic momentum in tennis, revealing nuanced dependencies between sequential successes and tactical adaptations; Fitzpatrick et al. (Fitzpatrick et al., 2019) identify key performance characteristics in tennis using statistical methods, while Cui et al. (Cui et al., 2018) and CBRF (Lv et al., 2024) leverage machine learning to profile Grand Slam performances. These works predominantly rely on static metrics (e.g., serve speed, unforced errors) or aggregated match statistics, often employing regression models, decision trees, or ensemble methods like random forests.

Despite advancements, existing momentum prediction models treat momentum as a latent variable derived from discrete events (e.g., points won) rather than a dynamic, time-dependent process. First, they often assume static intra-point momentum, where momentum is represented as fixed values between discrete events (e.g., points or games) (Wan et al., 2024), neglecting intra-rally psychological volatility or mid-game tactical adjustments. Second, temporal dependencies between sequential actions are frequently disregarded (Neumann & Fischer, 2023), e.g., fatigue trends or confidence accumulation from consecutive wins, which isolates training units and disrupts autocorrelation. Third, cross-individual momentum interactions remain unquantified. For example, DE-TSMCL (Gao et al., 2024) treats momentum as an isolated, player-specific property, ignoring bidirectional competitive influences.

To address these limitations, this paper proposes TM^2, a TCDformer-based Momentum Transfer Model for long-term sports prediction, which encompasses a momentum encoding module and a momentum prediction module. TM^2 can identify potential change points through interaction among individuals, adjust the momentum in real time and dynamically, and obtain long-term momentum predictions by separating trend and seasonal components.

3. TM^2 Model

In this section, we present the design details of TM^2 model, which consists of two key modules: the *momentum encoding module* and the *momentum prediction module*. As shown in Figure 2. We also discuss dataset preprocessing and feature selection.

3.1. Momentum Encoding Module

Given the similarity between change points and trends in non-stationary time series, ignoring the impact of potential change points (specifically, apparent fluctuations driven by external events rather than simple noise) may lead to misleading conclusions. To reduce this risk, TM^2 first applies the Maximal Overlap Discrete Wavelet Transform (MODWT) (Osmani et al., 2024) to the time series data to obtain the wavelet coefficients, representing the differences between moving averages at different scales, as shown in Figure 3. TM^2 then observes the sign changes in the wavelet coefficient vectors on either side of the detected change points to identify the change points, which are key for determining the extent of the change. TM^2 applies a detection rule similar to that of the first jump to detect subsequent jumps. Once the detection of jump positions across all regions is complete, TM^2 performs the inverse MODWT to reconstruct the time series data. TM^2 finally calculates the player's momentum at each time point by the cumulative sum of the importance and value of each indicator at each time point of the player. The specific steps are as follows:

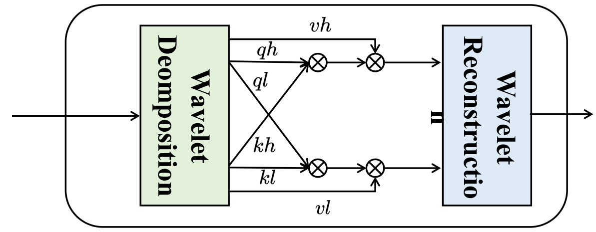

Figure 3: Obtaining potential change points by wavelet attention mechanisms.

$$\begin{cases} n_{\alpha,j} = \sum_{i=l_{max,j}}^{l_j-1} \frac{|W_{j,i+1}-W_{j,t}|}{2} \\ n_{\beta,j} = \sum_{i=l_{j+1}}^{l_{max,j}} \frac{|W_{j,i+1}-W_{j,t}|}{2} \end{cases} \tag{2}$$

Starting with the wavelet coefficient that has the highest absolute value, which signifies the precise change location, the coefficients at the change points are typically represented by the sign changes in the coefficient vectors on both sides. These are denoted as $n_{\alpha,j}$ and $n_{\beta,j}$, but due to their invariance across all wavelet types and transformation orders, they can also be denoted simply as n_α and n_β. To reconstruct the entire jump segment, it is necessary to determine the boundaries of the jump. TM^2 set n_α as the left boundary, α, and n_β as the right boundary, β, defining $\Omega = [\alpha, \beta]$ as the complete jump range. The specific calculation of α and β is as shown in Equation 3 and Equation 4:

$$\alpha = \max\{l \in [1, l-1] \mid \sum_{i=1}^{l-1} \frac{|W_{j,i+1} - W_{j,t}|}{2} \geq n_\alpha\} \tag{3}$$

$$\beta = \min\{l \in [l+1, T] \mid \sum_{i=l+1}^{L} \frac{|W_{j,i-1} - W_{j,t}|}{2} \geq n_\beta\} \tag{4}$$

3.1.1. Step 1: Extraction of Wavelet Coefficients

First, let X represent the time series data with dimensions $T \times D$, where T denotes the length of the series and D represents the number of variables. TM^2 define X as $X = (x_1, \ldots, x_t, \ldots, x_T)^T \in \mathbb{R}^{T\times D}$, where each x_t (for $1 \leq t \leq T$) is a D-dimensional vector representing the values of all D variables at time t, written as $x_t = (x_1^d, \ldots, x_t^d, \ldots, x_T^d)^T \in \mathbb{R}^D$. Similarly, for each variable d (where $1 \leq d \leq D$), the series $x^d = (x_1^d, \ldots, x_t^d, \ldots, x_T^d)^T \in \mathbb{R}^T$ describes its temporal evolution across all T time points.

Next, TM^2 applies the Maximal Overlap Discrete Wavelet Transform (MODWT) (Osmani et al., 2024) to the time series X to obtain the wavelet coefficients. These coefficients represent the differences between moving averages at different scales, as shown in Equation 1:

$$\widetilde{s_j} = \begin{cases} l_j = \arg\max_t(|W_{j,t}|) \\ l_{min}^j = \min\{t \mid t \in \sup(W_j)\} \\ l_{max}^j = \max\{t \mid t \in \sup(W_j)\} \end{cases} \tag{1}$$

Where W_j denotes the set of wavelet coefficients obtained through MODWT, $\arg\max_t$ identifies the value of t that maximizes the given condition, and l_j corresponds to the t value at which $|W_{j,t}|$ reaches its maximum. The notation $\sup(W_j)$ indicates the set of positions in W_j that contain non-zero elements, while l_{min}^j and l_{max}^j represent the positions of the minimum and maximum non-zero elements of t within this set, respectively.

3.1.2. Step 2: Identification and Characterization of Change Points

Change points are identified by observing the sign changes in the wavelet coefficient vectors on either side of the detected change points. These sign changes are key for determining the extent of the change. The changes in sign on either side of the change point in the wavelet coefficient vectors are defined as shown in Equation 2:

3.1.3. Step 3: Detection of the kth Jump at Scale J

To detect subsequent jumps, TM^2 applies a detection rule similar to that of the first jump, with the added step of excluding already-detected jumps and their surrounding regions. The position of the kth jump is defined as:

$$l_k = \arg\max_t(|W_{J,t}| \mid t \notin \bigcup_{1\leq i\leq k} \Omega_i) \tag{5}$$

In this Equation 5, l_k represents the position of the kth jump, Ω_i denotes the ith jump region, and $\cup_{1\leq i\leq k}\Omega_i$ represents the union of all previously identified jump regions. Here, $\arg\max_t$ is used to identify the value of t that maximizes $|W_{J,t}|$, assuming that t does not fall within any of the previously detected jump regions.

Table 2
Pressure values between different players

Player	Carlos Alcaraz	Alexander Zverev	Frances Tiafoe	...	David Goffin	Maximilian Marterer	Novak Djokovic
Alexander Zverev	1	4.78	3.79	...	0.71	3.32	1.27
Carlos Alcaraz	0.21	1	4.85	...	1.72	4.23	2.11
Frances Tiafoe	0.26	0.21	1	...	3.03	3.97	2.72
...	...	...	...	...	...	...	...
David Goffin	1.41	0.58	0.33	...	1	1.25	3.82
Maximilian Marterer	0.31	0.24	0.25	...	0.8	1	3.63
Novak Djokovic	0.79	0.47	0.37	...	0.26	0.28	1

3.1.4. Step 4: Detection at Reduced Scale $j < J$

After identifying the jump regions at scale J as Ω_k, for $1 \leq k \leq K$, the scale is reduced to $J - \Lambda \leq j \leq J$, and $l_{j,k}$ is determined, where $0 \leq \Lambda \leq J$ dictates which scale's details are reconstructed:

$$l_{j,k} = \arg\max_t (|W_{J,t}| \mid t \in \Omega_{j+1,k}) \tag{6}$$

In Equation 6 this ensures consistent jump detection across different scales, with the range covering these jumps at scale j denoted as $\Omega_{j,k} = [\alpha_{j,k}, \beta_{j,k}]$.

3.1.5. Step 5: Signal Reconstruction

Once the detection of jump positions across all regions $\Omega_{j,k}$ is complete, TM^2 perform the inverse MODWT for $1 \leq j \leq J$ to obtain the wavelet coefficients $\widetilde{W}_{j,t}$ that contain the jump information. These are then used to compute the modified $\widetilde{D}_{j,t}$, ultimately leading to the reconstruction of the signal $\overline{X}$.

3.1.6. Step 6: Momentum Extraction

In order to obtain the weights, TM^2 first use the analytic hierarchy process (AHP) to evaluate the importance of each indicator and obtain the relative importance of the zth indicator relative to other indicators $\boldsymbol{g}_z$ and the opponent $\overline{\boldsymbol{g}}_z$. According to the calculation results, the athlete's momentum is fitted into the actual change curve, and the fitting curve is intercepted 1 second before and after the corresponding time point. If the time before and after does not meet the requirements, the longest time period that can be intercepted is intercepted, so that the linear momentum corresponding to each time point is obtained. The specific calculation is as follows:

$$\delta_z(t) = d \cdot \boldsymbol{g}_z \cdot \log\left((r_z^t)^k\right) + \frac{1}{t+1} \tag{7}$$

$$\overline{\delta_z}(t) = d \cdot \overline{\boldsymbol{g}_z} \cdot \log\left((r_z^t)^k\right) + \frac{1}{t+1} \tag{8}$$

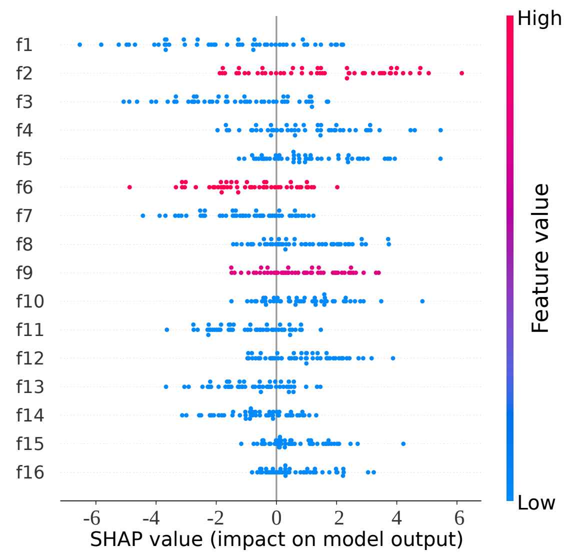

Figure 4: SHAP for momentum weights.

In Equation 7 and Equation 8, where $\delta_z(t)$ represents the weighted value of the zth metric at time t, $\overline{\delta_z}(t)$ represents the impact of the zth metric on the player when the opponent is paired at time t, r^t represents the original value of the metric, d is the limiting factor, reflecting the growth constraint, k is the gap between the same type of data in the metric, and different features may need to square the value or use the original value according to their specific properties.

The weights for each point calculated using the above approach are used to calculate the player's momentum at each time point. By multiplying these weights by each eigenvalue at each time point (denoted as w_1 for player 1 and w_2 for player 2) and summing the products, the player's momentum $M_y(t)$ at a specific moment is determined as shown in Equation 9:

$$M_y(t) = m_{i,j} \cdot \sum_{z=0}^{n} \left(\delta_z(t) \cdot \overline{X}_{i,z}(t) - \overline{\delta_z}(t) \cdot \overline{X}_{j,z}(t)\right) \tag{9}$$

Where $M_y(t)$ represents the momentum of player y at time t, $m_{i,j}$ is the pressure value of each player facing other players (see Table 2), $\overline{X}_{i,z}$ and $\overline{X}_{j,z}$ represent the data of player i and player j, respectively, reconstructed by Local Linear Seasonal Adjustment (LLSA) based on Maximal Overlap Discrete Wavelet Transform (MODWT) (Osmani et al., 2024). n is the total number of features, while $\delta_z(t)$ and $\overline{\delta_z}(t)$ are the influence weights of the zth feature on the player and opponent, respectively. Some athletes' corresponding momentum weight SHAP diagrams are shown in Figure 4.

3.2. Momentum Prediction Module

First, TM^2 decomposes the reconstructed time series data into trend and seasonal components using the Local Linear Scaling Approximation (LLSA) module (Wan et al., 2024). The trend component x_t is calculated using multiple averaging filters of different sizes and integrated into the final trend component through adaptive weighting, while the seasonal component x_s is obtained by subtracting the trend component from the original time series:

$$\begin{cases} x_s = M_y - x_t \\ x_t = \sigma(w(x) * f_2(x)) \end{cases} \tag{10}$$

In Equation 10, $\overline{X}$ represents the data reconstructed via Local Linear Seasonal Adjustment (LLSA) based on Maximal Overlap Discrete Wavelet Transform (MODWT) (Osmani et al., 2024), and x_s and x_t denote the seasonal and trend components, respectively. $\sigma(\cdot)$, $w(x)$, and $f_2(x)$ represent the softmax function, adaptive weights of average filters, and averaging filter, respectively.

For trend prediction, a three-layer MLP is utilized, and to address non-stationarity, RevIN normalization is applied before and after the MLP layers, as shown in Equation 11:

$$\overline{x_t} = \text{RevIN}(\text{MLP}(\text{RevIN}(x_t))) \tag{11}$$

For the seasonal component, wavelet-based attention mechanisms are applied, where attention calculations are performed on the decomposed queries, keys, and values at each scale. The process is detailed in the following equation:

$$Y(q,k,v) = \overline{W}\left(\text{softmax}\left(W(q)\overline{W'(k)}^T\right)W(v)\right) = qk^T v \tag{12}$$

In Equation 12, the final momentum prediction is obtained by summing the output of the next point of trend and seasonal components:

$$P(t+1) = (\overline{x_t}(t+1) + \overline{x_s}(t+1)) \tag{13}$$

In Equation 13, where $P(t_{n+1})$ represents the total momentum of each player at time $t+1$. The continuous nature of match updates ensures that total momentum is refreshed with each time increment and scoring event, maintaining accuracy in the outcome prediction.

3.3. Result Determination Layer

The final match outcome relies on the comparative analysis of total momentum values, which encapsulate the players' capabilities during the match. This approach inherently reflects the players' on-field prowess and status at specific moments. The resultant total momentum also mirrors a player's confidence level, a critical determinant in matches between players of comparable skill.

Given the continuous update of match data, the total momentum of each player is refreshed with each time increment and scoring event. The final momentum comparison between players is derived from the following equation:

$$\eta_{t_{n+1}} = \begin{cases} i & P_t(t_{n+1}) > P_j(t_{n+1}) \\ i \text{ or } j & P_t(t_{n+1}) = P_j(t_{n+1}) \\ j & P_t(t_{n+1}) < P_j(t_{n+1}) \end{cases} \tag{14}$$

In Equation 14, in scenarios where players exhibit identical momentum, historical rankings serve as the tiebreaker. Although these rankings are not directly incorporated into the model's calculations, leveraging them in such instances offers a pragmatic and often accurate resolution approach. While acknowledging that lower-ranked players can occasionally outperform higher-ranked counterparts, the predicted momentum–derived from in-match data–provides a robust and credible basis for determining outcomes.

To conclude, the momentum-based model offers a systematic and dynamic approach for assessing match outcomes, relying on real-time performance metrics. This approach ensures that the final match outcome is reflective of the players' real-time capabilities and their dynamic performance during the match.

3.4. Dataset Preprocessing and Feature Selection

The dataset used in this study is derived from the 2023 Wimbledon men's tournament, containing detailed match data for each player. The dataset initially consisted of 7,285 rows and 49 columns, totaling 356,965 data points. However, many of these data points were not highly relevant to the momentum analysis, necessitating dimensionality reduction. TM^2 applied feature selection techniques to retain 18 key features, as shown in Table 3. These features were chosen based on their relevance to momentum encoding and their ability to capture critical aspects of player performance during matches.

To address potential biases introduced by dimensionality reduction, TM^2 conducted a thorough analysis of the retained features to ensure that they adequately represent the underlying dynamics of the matches. The selected features include metrics such as elapsed time, sets won, games won, and psychological factors, which are crucial for momentum analysis. By focusing on these features, TM^2 aims to minimize the loss of critical information while reducing the computational complexity of the model.

Table 3
Feature selection.

Targets	Explanation
elapsed_time	Time elapsed since the start of the first point to the start of the current point (H: MM: SS)
p_sets	Sets won by player
p_games	Games won by a player in the current set
server	Server of the point
point_victor	Winner of the point
p_ace	Player hit an untouchable winning serve
p_double_fault	Player missed both serves and lost the point
p_break_pt_missed	One player misses a chance to win the match while the other is serving
p_break_pt_won	One player wins while the other is serving.
p_distance_run	Player's distance run during a point (meters)
psychological_factor	The psychological impact of a player's gain or loss during a match

[a] According to the indicators in the table, player 1 and player 2 recorded data separately.

4. Evaluation

4.1. Methodology

Datasets. The experiment utilized the public dataset ETTh1 of the 2023 Wimbledon Championships [6], which was obtained through the International Mathematical Modeling Competition and cross-validated with the official tournament records. The original dataset contains 7,285 rows and 49 columns (356,965 data points), and each player has feature records at key game moments. Through dimensionality reduction, we retained 18 key features related to momentum analysis (as shown in Table 3). In order to verify the generalization of our model, we also selected two open datasets from the Internet, an NBA dataset[7] and a beach volleyball dataset[8]. Those datasets record the details of each game, such as the date of the game, the two sides of the game, the venue of the game, the score, the player's playing time, the scores, rebounds, the assists, and other key data.

Evaluation Metrics. We evaluate performance using two categories of metrics. For time series evaluation, we employ Mean Squared Error (MSE) and Mean Absolute Error (MAE). For classification assessment, we include Accuracy, Precision, and F1-score to provide a comprehensive evaluation of prediction accuracy. The F1-score is particularly important for verifying the balance between precision and recall in the model's performance.

Baselines. The baseline models were selected based on three criteria to ensure relevance and comparability. First, we included models widely used in tennis analytics literature, such as the ELO rating system (Neumann & Fischer, 2023), Decision Tree (DT) (Kjamilji, 2024), and Logistic Regression (LR) (Wasi & Abulaish, 2024), which are established benchmarks for general match prediction. Second, we incorporated models with temporal pattern recognition capabilities, including Support Vector Machine (SVM) (Pang et al., 2022) and Random Forest (RF) (Alfarizi et al., 2022), to evaluate their effectiveness in capturing momentum dynamics. Third, we also found some existing models and datasets to verify the versatility and superiority of the model.

[6] https://www.mathmodels.org/Problems/2024/MCM-C/index.html
[7] https://download.csdn.net/download/baidu_38876334/87960136
[8] https://tianchi.aliyun.com/dataset/90450

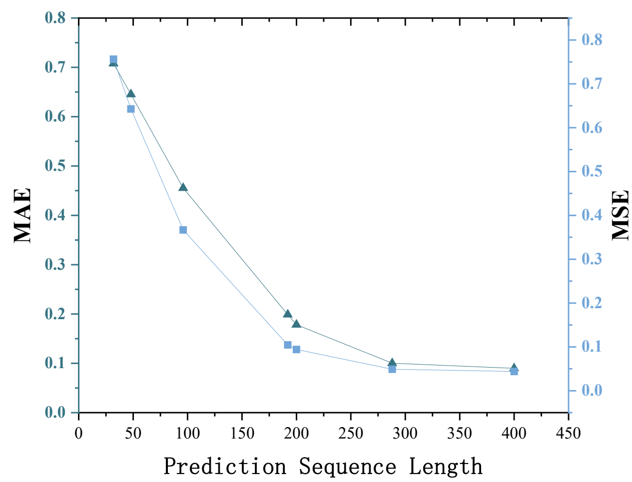

Figure 5: Iterative adjustment curve of TM^2 prediction sequence length.

Training. To prevent data leakage and ensure realistic evaluation, we adopted a strict time-ordered split. The first 80% of chronological matches formed the training set, the next 10% served as the validation set, and the final 10% was used for testing. This approach ensures that the model is trained only on past data and tested on future data, mimicking real-world deployment scenarios. Parameter optimization was conducted exclusively on the training and validation sets, with final evaluation performed only on the test data. Since the data is listed in time, the dataset is partitioned in chronological order.

4.2. Parameter Optimization

For TM^2, we performed systematic hyperparameter tuning, focusing on the prediction window size. We determined the optimal configuration through iterative tuning guided by the minimization of MSE and MAE. To mitigate the impact

Table 4
Model comparison results of TM^2.

Metrics	ELO	DT	LR	SVM	RF	*TM^2*
MAE	0.4945	0.9823	0.9848	0.9885	0.9956	0.0898
MSE	0.4945	1.1316	1.1242	1.1408	1.1420	0.0440
Accuracy	0.5055	0.7178	0.7589	0.7789	0.7562	0.7939
Precision	0.5055	0.6576	0.7629	0.7609	0.7290	0.8021
F1-score	0.6715	0.7675	0.7609	0.7629	0.7735	0.8236

of outliers, the model was trained 10 times for each sequence length, and the results were averaged. After extensive experiments, 400 was finally selected as the optimal setting, and the training process is shown in Figure 5.

4.3. Comparison with Existing Models

We first compare *TM^2* against five baseline models in the ETTh1 datasets. As shown in Table 4, Figure 6, and Figure 7, *TM^2* outperforms the baseline models on all metrics. This is because traditional models such as DT, LR, and SVM rely on specific weighting strategies and perform better when data points are independent of each other. However, in scenarios where consecutive data points influence each other, such as time series data with a large number of change points, these models tend to perform poorly, leading to inductive biases in predictions. The MSE and MAE values of *TM^2* are significantly lower than those of other models. This is because MSE and MAE are more suited for evaluating continuous predictions, which are a key feature of deep learning models. By contrast, Accuracy, Precision, and F1-score, which rely on True Negatives (TN), False Negatives (FN), False Positives (FP), and True Positives (TP), are typically used in classification tasks. Nonetheless, *TM^2* shows a clear advantage in both types of metrics, highlighting its robustness.

We then compare *TM^2* against two advanced neural network models in the ETTh1 datasets: Mamba4Cast (Bhethanabhotla et al., 2024) and Seq2Event (Simpson et al., 2022), which have both achieved excellent predictive accuracy in their respective fields. As shown in Table 5, *TM^2* demonstrates superior performance in terms of MSE and MAE when compared to both Mamba4Cast and Seq2Event, underscoring its development potential and room for further improvement.

To further assess the performance of *TM^2*, we compare *TM^2* against the momentum-based models in the ETTh1 dataset. Since a large number of momentum-based models are strongly coupled with closed datasets or their codes are not open source, only DE-TSMCL (Gao et al., 2024) is compared. As shown in Table 6, DE-TSMCL is better than most baseline models, but *TM^2* is still superior in terms of MSE and MAE compared to DE-TSMCL. This is because DE-TSMCL only focuses on predicting static intra-point momentum for one game and treats momentum as an isolated, player-specific property, ignoring bidirectional competitive influences. Tennis is a two-player sport, the results will be

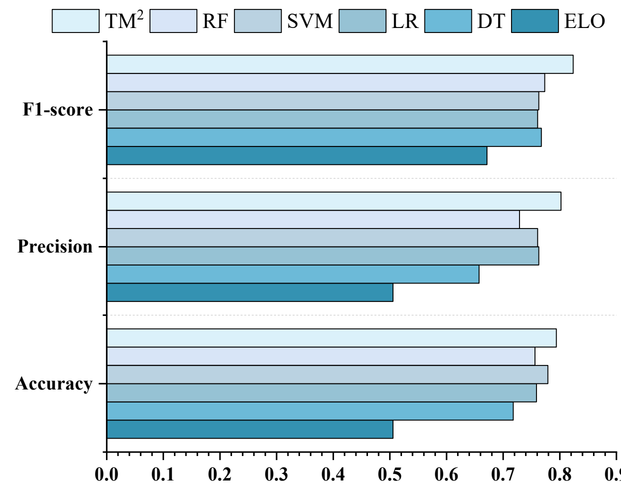

Figure 6: Comparison of TM^2 with existing basic tennis models in terms of F1-score, precision, and accuracy metrics, in the task of predicting tennis match results, with data from the 2023 Wimbledon tournament.

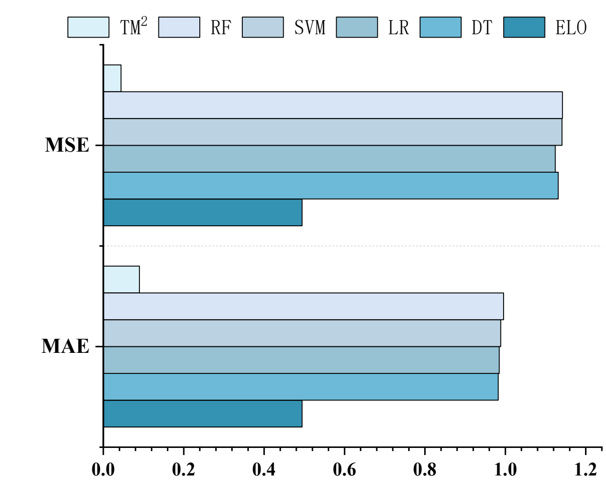

Figure 7: Comparison of TM^2 with existing basic tennis models in terms of MSE and MAE, in the task of predicting tennis match results, with data from the 2023 Wimbledon tournament.

greatly biased when it is ignored. *TM^2* can identify potential change points through interaction among individuals, adjust the momentum in real time and dynamically, and obtain long-term momentum predictions by separating trend and

Table 5
Compare TM^2 with existing models on different datasets.

Dataset \ Approaches	Mamba4Cast		Seq2Event		TM^2	
	MSE	**MAE**	**MSE**	**MAE**	**MSE**	**MAE**
NBA	0.1248	0.3542	0.1790	0.4580	0.1092	0.3116
Beach Volleyball	0.3805	0.8353	0.4019	0.9255	0.2651	0.5717
ETTh1	0.0989	0.2299	0.1832	0.4619	0.0440	0.0898

Table 6
Comparison TM^2 with momentum-based models.

Metrics	DE-TSMCL	TM^2
MAE	0.4808	0.0898
MSE	0.3858	0.0440

seasonal components. TM^2 can effectively solve the limitations of DE-TSMCL in long-term sport predictions.

4.4. Generalization

To validate the scalability potential of the proposed model architecture across different sports domains, we conducted cross-domain experiments using publicly available NBA basketball and beach volleyball datasets. Experimental results show that TM^2 maintains satisfactory prediction performance despite a 41.25% reduction in MAE (mean absolute error) and a 42.55% reduction in MSE (mean squared error) compared to tennis match prediction, as shown in Table 5. We hypothesize that this performance difference may be attributed to the inherent characteristics of team sports: (1) the error propagation mechanism in collective games, where individual performance differences within a team can cumulatively affect the results; and (2) the limitations of applying the overall momentum metric to team-based scenarios, which proves to be less effective than individual sports prediction. This comparative analysis highlights the adaptability of the architecture while revealing the prediction challenges of specific sports.

4.5. Limitations

Although TM^2 achieves state-of-the-art performance across most metrics, its superior MSE, MAE, and accuracy over Mamba4Cast (Bhethanabhotla et al., 2024) and Seq2Event (Simpson et al., 2022) on diverse datasets can be attributed to its ability to detect dynamic change points in temporal sequences. Our analysis reveals that Mamba4Cast and Seq2Event struggle to distinguish such dynamic shifts, leading to suboptimal performance. Notably, TM^2 exhibits diminished results on the Beach Volleyball dataset. We hypothesize that this stems from the collective nature of team sports, where quantifying "momentum" becomes inherently challenging. In beach volleyball, for instance, an individual player's exceptional skill may theoretically elevate team momentum, yet practical outcomes remain constrained by the sport's inherent interdependence—no single player disproportionately influences match results. Conversely, TM^2 excels on the NBA dataset, as basketball inherently allows star players with significantly above-average skills to exert decisive impacts on game outcomes, aligning with the model's momentum quantification mechanism.

TM^2 remains highly sensitive to data quality. Missing values or noise (e.g., incomplete player statistics or sensor errors) may degrade performance. Additionally, with only 7,285 training samples, deeper architectures risk overfitting despite our use of regularization and iterative cross-validation. Future work should investigate data augmentation or self-supervised pretraining to address data scarcity.

5. Conclusions and Future Work

This study addresses the core limitations of existing sports prediction models in dynamically modeling competitive interactions by proposing the TM^2 approach. By integrating the Local Linear Scaling Approximation (LLSA) module, bidirectional LSTM architecture, and momentum interaction pressure matrix, TM^2 achieves multi-scale dynamic coupling of tactical adjustments, psychological fluctuations, and environmental factors. Experimental results on datasets from Wimbledon and the NBA demonstrate significant reductions in prediction errors (96.26% in MSE and 79.93% in MAE), particularly excelling in high-pressure scenarios such as tiebreaks and comebacks. These findings validate the effectiveness of continuous dynamic modeling in synergizing short-term fluctuations with long-term trends in sports competitions, offering theoretical support for real-time coaching decisions, training optimization, and injury prevention. The cross-sport validation (from tennis to basketball) further highlights TM^2's adaptability in interaction-intensive scenarios, though limitations remain in modeling team-based collective sports.

Future research should expand TM^2's theoretical and practical boundaries across multiple dimensions. First, extending the approach to team sports (e.g., football, hockey) requires addressing the challenges of multi-agent interactions and cooperative strategies to capture complex group dynamics. Second, incorporating multimodal data such as physiological signals (e.g., heart rate variability, muscle fatigue) and environmental variables (e.g., crowd noise, weather conditions) could refine real-time analysis of psychological states and tactical execution. Third, exploring temporal resolution tuning (e.g., point-level vs. rally-level granularity) would enhance the model's adaptability to sports with heterogeneous game structures (e.g., volleyball and basketball), enabling dynamic adjustments to short-term fluctuations (e.g., momentum shifts within rallies) and long-term trends (e.g., fatigue accumulation across sets). These advancements would propel sports prediction from static analytics to dynamic ecosystem modeling, fostering interdisciplinary innovation in competitive science and decision intelligence.

References

Alfarizi, M. G., Tajiani, B., Vatn, J., & Yin, S. (2022). Optimized random forest model for remaining useful life prediction of experimental bearings. *IEEE Transactions on Industrial Informatics*, *19*, 7771–7779.

Antonini, V., Mileo, A., & Roantree, M. (2024). Engineering features from raw sensor data to analyze player movements during competition. *Sensors*, *24*, 1308.

Bhethanabhotla, S. K., Swelam, O., Siems, J., Salinas, D., & Hutter, F. (2024). Mamba4cast: Efficient zero-shot time series forecasting with state space models. *arXiv preprint arXiv:2410.09385*, .

Briki, W. (2017). Rethinking the relationship between momentum and sport performance: Toward an integrative perspective. *Psychology of Sport and Exercise*, *30*, 38–44.

Cao, D., Jia, F., Arik, S. O., Pfister, T., Zheng, Y., Ye, W., & Liu, Y. (2023). Tempo: Prompt-based generative pre-trained transformer for time series forecasting. *arXiv preprint arXiv:2310.04948*, .

Chang, C., Peng, W.-C., & Chen, T.-F. (2023). Llm4ts: Two-stage fine-tuning for time-series forecasting with pre-trained llms. *arXiv preprint arXiv:2308.08469*, .

Cui, Y., Gomez, M.-A., Goncalves, B., & Sampaio, J. (2018). Performance profiles of professional female tennis players in grand slams. *PLoS ONE*, *13*, e0200591.

Eckardt, V. C., & Tamminen, K. A. (2023). A scoping review on interpersonal coping in sports. *International Review of Sport and Exercise Psychology*, (pp. 1–27).

Fitzpatrick, A., Stone, J. A., Choppin, S., & Kelley, J. (2019). A simple new method for identifying performance characteristics associated with success in elite tennis. *International Journal of Sports Science & Coaching*, *14*, 43–50.

Gama, J., Passos, P., Davids, K., Relvas, H., Ribeiro, J., Vaz, V., & Dias, G. (2014). Network analysis and intra-team activity in attacking phases of professional football. *International Journal of Performance Analysis in Sport*, *14*, 692–708.

Gao, H., Ren, Q., & Li, J. (2024). Distillation enhanced time series forecasting network with momentum contrastive learning. *Information Sciences*, *675*, 120712.

Garza, A., & Mergenthaler-Canseco, M. (2023). Timegpt-1. *arXiv preprint arXiv:2310.03589*, .

Gauriot, R., & Page, L. (2018). Psychological momentum in contests: The case of scoring before half-time in football. *Journal of Economic Behavior & Organization*, *149*, 137–168.

Gauriot, R., & Page, L. (2019). Does success breed success? a quasi-experiment on strategic momentum in dynamic contests. *The Economic Journal*, *129*, 3107–3136.

Graber, C., & Schwing, A. G. (2020). Dynamic neural relational inference. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 8513–8522).

Ji, T., Self, N., Fu, K., Chen, Z., Ramakrishnan, N., & Lu, C.-T. (2021). Dynamic multi-context attention networks for citation forecasting of scientific publications. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 7953–7960). volume 35.

Khan, N. J., & Ahamad, G. (2024). Fuzzy optimisation based cricket talent identification. *Expert Systems with Applications*, *237*, 121573.

Kim, J., Kim, H., Lee, J., Lee, J., Yoon, J., & Ko, S. K. (2022). A deep learning approach for fatigue prediction in sports using GPS data and rate of perceived exertion. *IEEE Access*, *10*, 103056–103064.

Kjamilji, A. (2024). Privacy-preserving zero-sum-path evaluation of decision tress in postquantum industrial iot. *IEEE Transactions on Industrial Informatics*, .

Levy, B. P., & Lopes, H. F. (2021). Trend-following strategies via dynamic momentum learning. *arXiv preprint arXiv:2106.08420*, .

Liang, Y., Wen, H., Nie, Y., Jiang, Y., Jin, M., Song, D., Pan, S., & Wen, Q. (2024). Foundation models for time series analysis: A tutorial and survey. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 6555–6565).

Lv, X., Gu, D., Liu, X., Dong, J., & Li, Y. (2024). Momentum prediction models of tennis match based on catboost regression and random forest algorithms. *Scientific Reports*, *14*, 18834.

Malone, S., Solan, B., Collins, K. D., & Doran, D. A. (2016). Positional match running performance in elite gaelic football. *The Journal of Strength and Conditioning Research*, *30*, 2292–2298.

Markopoulou, C., Papageorgiou, G., & Tjortjis, C. (2024). Diverse machine learning for forecasting goal-scoring likelihood in elite football leagues. *Machine learning and knowledge extraction*, *6*, 1762–1781.

Morgulev, E., & Avugos, S. (2023). Beyond heuristics, biases and misperceptions: the biological foundations of momentum (hot hand). *International Review of Sport and Exercise Psychology*, *16*, 155–175.

Neumann, C., & Fischer, J. (2023). Extending bayesian elo-rating to quantify the steepness of dominance hierarchies. *Methods in Ecology and Evolution*, *14*, 669–682.

Osmani, S. A., Jun, C., Baik, J., Lee, J., & Narimani, R. (2024). Wavelet-based precipitation preprocessing for improved drought forecasting: A machine learning approach using tunable q-factor wavelet transform and maximal overlap discrete wavelet transform. *Expert Systems with Applications*, *257*, 124962.

Page, L., & Coates, J. (2017). Winner and loser effects in human competitions: Evidence from equally matched tennis players. *Evolution and Human Behavior*, *38*, 530–535.

Panagoulias, D. P., Virvou, M., & Tsihrintzis, G. A. (2024). A novel framework for artificial intelligence explainability via the technology acceptance model and rapid estimate of adult literacy in medicine using machine learning. *Expert Systems with Applications*, *248*, 123375.

Pang, J., Pu, X., & Li, C. (2022). A hybrid algorithm incorporating vector quantization and one-class support vector machine for industrial anomaly detection. *IEEE Transactions on Industrial Informatics*, *18*, 8786–8796.

Papageorgiou, E., Krommidas, C., Digelidis, N., Moustakas, L., & Papaioannou, A. G. (2023). An online pete course on intercultural education for pre-service physical education teachers: A non-randomized controlled trial. *Teaching and Teacher Education*, *121*, 103920.

Papageorgiou, G., Sarlis, V., & Tjortjis, C. (2024a). Evaluating the effectiveness of machine learning models for performance forecasting in basketball: a comparative study. *Knowledge and Information Systems*, *66*, 4333–4375.

Papageorgiou, G., Sarlis, V., & Tjortjis, C. (2024b). An innovative method for accurate nba player performance forecasting and line-up optimization in daily fantasy sports. *International Journal of Data Science and Analytics*, (pp. 1–24).

Papageorgiou, G., Sarlis, V., & Tjortjis, C. (2024c). Unsupervised learning in nba injury recovery: Advanced data mining to decode recovery durations and economic impacts. *Information*, *15*, 61.

Pu, Z., Pan, Y., Wang, S., Liu, B., Chen, M., Ma, H., & Cui, Y. (2024). Orientation and decision-making for soccer based on sports analytics and ai: A systematic review. *IEEE/CAA Journal of Automatica Sinica*, *11*, 37–57.

Qiu, M., Zhang, S., Yi, Q., Zhou, C., & Zhang, M. (2024). The influence of "momentum" on the game outcome while controlling for game types in basketball. *Frontiers in Psychology*, *15*, 1412840.

Raabe, D., Nabben, R., & Memmert, D. (2023). Graph representations for the analysis of multi-agent spatiotemporal sports data. *Applied Intelligence*, *53*, 3783–3803.

Rasul, K., Ashok, A., Williams, A. R., Khorasani, A., Adamopoulos, G., Bhagwatkar, R., Biloš, M., Ghonia, H., Hassen, N. V., Schneider, A. et al. (2023). Lag-llama: Towards foundation models for time series forecasting. *arXiv preprint arXiv:2310.08278*, .

Rodríguez-Rodríguez, I., Campo-Valera, M., Rodríguez, J.-V., & Woo, W. L. (2024). Iomt innovations in diabetes management: Predictive models using wearable data. *Expert Systems with Applications*, *238*, 121994.

Sarlis, V., Papageorgiou, G., & Tjortjis, C. (2023). Sports analytics and text mining nba data to assess recovery from injuries and their economic impact. *Computers*, *12*, 261.

Sarlis, V., Papageorgiou, G., & Tjortjis, C. (2024a). Injury patterns and impact on performance in the nba league using sports analytics. *Computation*, *12*, 36.

Sarlis, V., Papageorgiou, G., & Tjortjis, C. (2024b). Leveraging sports analytics and association rule mining to uncover recovery and economic impacts in nba basketball. *Data*, *9*, 83.

Sheridan, D., Brady, A. J., Nie, D., & Roantree, M. (2024). Predictive analysis of ratings of perceived exertion in elite gaelic football. *Biology of Sport*, *41*, 61–68.

Silva, F. G., Gomes, A. J., Nguyen, Q. T., Martins, F. M., & Clemente, F. M. (2017). A new tool for network analysis on team sports: The ultimate performance analysis tool. In *2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC)* (pp. 439–445). IEEE.

Simpson, I., Beal, R. J., Locke, D., & Norman, T. J. (2022). Seq2event: learning the language of soccer using transformer-based match event prediction. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 3898–3908).

Wan, J., Xia, N., Yin, Y., Pan, X., Hu, J., & Yi, J. (2024). Tcdformer: A transformer framework for non-stationary time series forecasting based on trend and change-point detection. *Neural Networks*, *173*, 106196.

Wang, W.-Y., Shuai, H.-H., Chang, K.-S., & Peng, W.-C. (2022). Shuttlenet: Position-aware fusion of rally progress and player styles for stroke forecasting in badminton. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 4219–4227). volume 36.

Wasi, N. A., & Abulaish, M. (2024). Skeds—an external knowledge supported logistic regression approach for document-level sentiment classification. *Expert Systems with Applications*, *238*, 121987.

Xu, Y., Yang, J., & Du, S. (2020). Cf-lstm: Cascaded feature-based long short-term networks for predicting pedestrian trajectory. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 12541–12548). volume 34.

Xue, H., Voutharoja, B. P., & Salim, F. D. (2022). Leveraging language foundation models for human mobility forecasting. In *Proceedings of the 30th International Conference on Advances in Geographic Information Systems* (pp. 1–9).

Zhou, Z., Zhou, X., Chen, Y., & Qi, H. (2024). Evolution of online public opinions on major accidents: Implications for post-accident response based on social media network. *Expert Systems with Applications*, *235*, 121307.

Zhu, Y., Ye, Y., Zhang, S., Zhao, X., & Yu, J. (2023). Difftraj: Generating gps trajectory with diffusion probabilistic model. *Advances in Neural Information Processing Systems*, *36*, 65168–65188.