Winning With Chaos in Association Football: Spatiotemporal Event Distribution Randomness Metric for Team Performance Evaluation

Bandara, I., Shelyag, S., Rajasegarar, S., Dwyer, D., Kim, E-J. & Angelova, M.

DOI 10.1109/ACCESS.2024.3413648
ESSN 2169-3536

Publisher: Institute of Electrical and Electronics Engineers

IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 18 April 2024, accepted 4 June 2024, date of publication 12 June 2024, date of current version 20 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3413648

RESEARCH ARTICLE

Winning With Chaos in Association Football: Spatiotemporal Event Distribution Randomness Metric for Team Performance Evaluation

ISHARA BANDARA[1,2], SERGIY SHELYAG[1,3], SUTHARSHAN RAJASEGARAR[1], DAN DWYER[4], EUN-JIN KIM[2], AND MAIA ANGELOVA[1,5]

[1]School of IT, Deakin University, Burwood, VIC 3125, Australia
[2]Research Centre for Fluid and Complex Systemys, Coventry University, CV1 5RW Coventry, U.K.
[3]College of Science and Engineering, Flinders University, Adelaide, SA 5042, Australia
[4]School of Exercise and Nutrition Sciences, Deakin University, Burwood, VIC 3125, Australia
[5]Aston Digital Futures Institute, Aston University, B4 7ET Birmingham, U.K.

Corresponding author: Ishara Bandara (isharabnd@gmail.com)

This work was supported by the Deakin-Coventry Cotutelle Scholarship.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Coventry University Ethics Approval Committee under Application No. P174511.

ABSTRACT Association football (commonly known as football or soccer) in the modern era places a greater emphasis on collaborating and working together as a team instead of relying solely on individual skills to strategize winning performances. The low-scoring and unpredictable nature of association football makes evaluating team performances challenging. Space creation and space utilization have been discussed in the football world lately. Existing literature evaluates this with on and off-ball runs by players for deceiving defenders to create open spaces. However, the contribution of these team ball movements' enhanced randomness or chaotic nature to winning performances has yet to be explored. This work proposes a novel entropy-based time-series performance evaluation metric, EDRan, for quantifying this enhanced random nature by analyzing the spatial distribution of game events at regular intervals. Additionally, an unexplored cumulative ball possession matrix is used to quantify randomness. The correlation between the match winner and spatial event distribution randomness at regular intervals was analyzed. The significance of the proposed metric was demonstrated using a generalized linear model (GLM), which achieved an average accuracy of 80% for match-winning performance classification. The GLM p-values and coefficients revealed statistically significant relationships between the extracted temporal features and match-winning performances. Findings further revealed dispersed, highly random event distribution by winning teams during the early phases of the game, implying attacking behavior, followed by a compact, cautious playing style toward the end, suggesting that the game's first-half performances are more pivotal. Despite the unpredictability of actual scores in association football, the proposed approach effectively captured the differences in performances between stronger and weaker teams with temporal relationships, highlighting its significance as a time-series metric for performance evaluation.

INDEX TERMS Performance evaluation, soccer, entropy, randomness, spatiotemporal, football, time-series metric.

I. INTRODUCTION

In modern sports, including invasion sports (such as association football, basketball, and hockey), tactics can be the edge between winning and losing a game, as there is little

The associate editor coordinating the review of this manuscript and approving it for publication was Alba Amato.

difference between physicality and skill set at the professional level. Invasion sports are team sports where a team attacks the opposition's territory for scoring while defending their territory from opponents' scoring attempts [1], [2], [3]. Most of the invasion sports are low-scoring by nature. Being considered the most popular sport in the world [4], association football (commonly known as football or soccer) has become a subject of extensive research in recent years due to its very nature and complexity.

Association football is characterized by its low-scoring and unpredictable nature [5], [6] influenced by chance [7]. The chance or the element of randomness is contributed by the goals, which result from unexpected events such as deflections and own goals. Previous literature studies have further revealed that the distribution of goals follows a Poisson distribution [8], [9], [10]. Since Poisson distributions are memory-less and describe independent events, adherence to them implies randomness in goal scoring, and this random component characterizes association football games with more significant uncertainty. Consequently, actual scores may not accurately reflect a team's performance. Thus, performance evaluation can be challenging.

The uncertainty and unpredictability revolving around this game have influenced strategies like "Direct-Play" [11], prioritizing swift ball advancement up the field through long and direct passes. On the other hand, possession play is another tactic where the team's primary focus is maintaining possession with short passes circulating around the field, drawing the opposition defenders to create chances. Although earlier studies have revealed that goal conversion ratio with direct play was better [11], it can be observed that champion teams from most recently concluded top four association football leagues have maintained more significant mean possession percentage and lesser long passes per match than their respective league average values (which implies possession-play) [12]. Due to these contrasting strategies by teams, evaluating performance based on performance evaluation metrics like possession has shown conflicting results. Some studies have identified that possession time may lead to a positive match outcome [13], [14] while some report that there is no significant relationship between ball possession and positive match outcome [15], [16].

Recent works have focused on space creation for performance evaluation. Xavier Hernández (Xavi), current Barcelona manager once stated that "I understand football as a time-space, and it is the use of space" [17]. Space creation and how on and off-ball runs influence space creation has been discussed in literatures [18] and [19]. Considering space and its influence on team performances, "Space-control", "Ball-control", and "Success-Score" have been introduced as key performance indicators (KPI) for evaluating the offensive success of an association football team [20], [21]. They have considered the "space-control" and "ball-control" of events within 30 meters of the opponent's goal, which they define as a "critical area". Nevertheless, some teams may prefer to build up from the back for offensive moves. In the recently concluded 2023 top four association football leagues (English Premier League, La Liga, Serie A, Bundesliga), players who have performed most of the passes are either defenders or defensive midfielders [12]. Literature also reveals that the center-backs contribute most to the team's overall possession play [22].

Furthermore, maintaining possession with no purpose is criticized and the creation of open space is encouraged by Pep Guardiola (the only football manager to win continental treble twice); "It is not about passing for the sake of it. The secret is to overload one side of the pitch so that the opponent must tilt its own defense to cope and so that they leave the other side weak. And when we have done all that, we attack and score from the other side" [23]. The creation of open space is contributed by dynamic team ball movements by ball-carriers and the off-ball runs by the attacking team members, which deceives and draws opposition defenders with them [18].

Adding an element of randomness to these ball movements can make the movements unpredictable for the opposition. In the literature, passing patterns, player to player interactions have been evaluated, and entropy has been used as a measure of uncertainty in these interactions [24], [25], [26] with the game being a complex control system [27], [28], [29], [30], [31]. Neuman et al. have demonstrated that a team's position at the end of the season league table correlates with player-to-player ball passing entropy [24]. Kusmakar et al. have used the same entropy-generating approach to quantify the player-to-player interactions and predict the teams that successfully create goal-scoring opportunities in a given play segment [25]. Further, Berman et al. successfully used passing networks related to entropy to predict player substitutions [26]. However, in association football, players may change their positions during the game based on the game's situation and tactics [32], and thus, player-to-player ball interactions or passing patterns may not accurately quantify randomness in team movements on the field.

Team ball movement refers to the coordinated passing and movement of the ball among the players of a team with the objective of advancing the ball up the field, maintaining possession, and creating goal-scoring opportunities. However, predictable team ball movements and passing patterns can make it relatively easier for the defenders to close down the open spaces and pressure the ball carrier, leading to regaining possession. Adding an element of randomness to these movements can make the team's strategies unpredictable for the opponents. However, the contribution of the randomness in team ball movement for the winning performances and evaluation of its temporal nature is unexplored in the existing literature. Moreover, previous literature has explored player-to-player interactions as a means of quantifying the unpredictability of ball movement, a method that may lack precision due to the dynamic player positions in modern football tactics.

The aim of this work was to quantify the randomness in team movements with spatial event distribution per unit time and evaluate its significance on match-winning performances. Accordingly, the scope of this study was to propose an approach for quantifying temporal randomness in ball movement, independent of player positions, and to evaluate its influence on match-winning performances. Furthermore, temporal trends in team performances were investigated, offering a comprehensive evaluation of team performances over time.

In this work, "randomness" in team ball movement is proposed as a time series metric for team performance evaluation. This paper introduces a novel time-series metric "relative spatial Event Distribution Randomness per unit time" (*EDRan*) to quantify the randomness in team ball movements at regular time intervals relative to the opponent's performances and independent of player-related information and goal locations. This includes the generation of novel region-based cumulative possession matrices and entropy differences per unit time measures.

II. APPROACH

This paper evaluates the significance of randomness in team ball movements for match-winning performances. First, an approach for quantifying randomness and the dynamic nature of team ball movements independent of players and the goal locations was investigated. As a solution, the time-series metric *EDRan* was proposed as a team performance evaluation metric. Initially, the data set was pre-processed. The novel "region-based cumulative possession matrices" were generated at uniform intervals followed by temporal-spatial event distribution randomness quantification. Quantified metric values were analyzed for their correlation and statistical significance with winning performances using generalized linear models. Finally, model results were further analyzed to identify hidden insights into the winner's performances and to evaluate the proposed metrics' ability to evaluate team performances despite the unpredictability of actual scores. Figure 1 shows the framework of this analysis.

A. DATA

A publicly available event-log data set [33] was used in this analysis. Although multiple other publicly available association football event-log datasets were evaluated initially (e.g., Wyscout data set [34], [35]), Statbomb open-data dataset was selected because of its detailed information about the event available for evaluation. Statbomb open-data dataset comprises data from numerous top-tier men's and women's club leagues in Europe and international competitions. Considering the physical and strategic differences between men's and women's games [36], top-tier men's competitions and women's competitions were analyzed separately. The men's data set comprised 680 matches from the English Premier League, La Liga, UEFA Champions League, UEFA EURO Cup, and FIFA World Cup. The women's dataset consisted of 445 matches from the FA Women's Super League, National Women's Super League, UEFA Women's Euro, and FIFA Women's World Cup. The Statsbomb dataset contains information on all the events in a match, such as event type, event location, ball possession team during the event, and event duration. The event type considered here refers to a predefined named occurrence during a match, such as passes, free kicks, and tackles. The game's rules govern some events and are identifiable by their event definitions, whereas the referee determines others. The event location is denoted by coordinates on a 120×80 grid, and the direction of attack of the ball possession team is from left to right in the grid. Event duration refers to how long a particular event lasts in seconds.

For this study, games without a winner were removed, as the aim was to evaluate the contribution of randomness in team ball movement to winning performances, and there was a lack of data to evaluate drawn performances. Games decided by extra time or penalties were also excluded due to insufficient data to evaluate such games. Additionally, only events that directly contributed to ball movement and possession were considered (e.g., passes, shots, and carries). Off-ball movements and events were not considered due to dataset limitations. Both the men's and women's datasets were pre-processed by removing games without a winner and games decided by extra time or penalties. The remaining men's dataset consisted of 552 games, with Team 1 winning 308 games and Team 2 winning 244 games (for naming purposes, the two teams are defined as Team 1 and Team 2). The preprocessed women's dataset consisted of 365 games, with Team 1 winning 199 games and Team 2 winning 166 games.

B. REGION-BASED POSSESSION MATRIX

An approach to capture a team's unpredictable ball movement nature with event distribution is analysed in this paper.

In previous literature, three regions in the association football field have been discussed [37], [38]. They are defensive third, mid-field third, and attacking third. Each third was divided into ten equal regions, as shown in Figure 2, hence resulting in a total of thirty regions for the whole association football field. Tianbiao et al. have used a similar approach of dividing the association football field into thirty regions to propose a method for analyzing the passing patterns in association football games using data mining techniques [39]. Unlike their approach, in this work, the field was divided into thirty equal-area regions to quantify the randomness in event distribution fairly across the regions.

Usually, an association football game lasts for ninety minutes plus an additional injury time. This being a time series analysis of team performance, each match duration has been divided into n equal periods (t_i; $1 \leq i \leq n$) and used for analysis, as well as generation of temporal features per match. Considering factors, such as number of training features, over-parameterization, duration of a period, and duration to

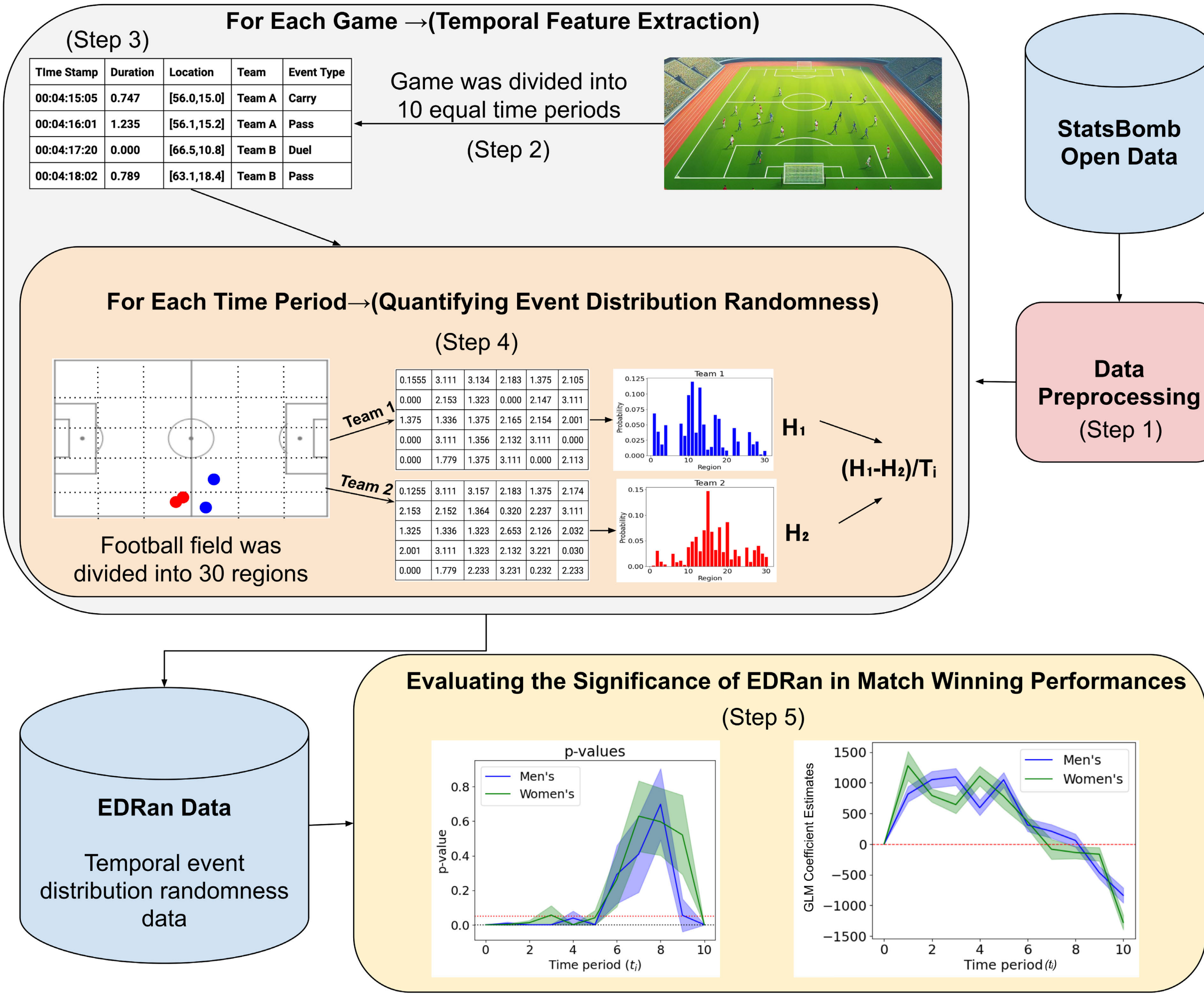


FIGURE 1. The proposed framework. Steps include: (1) Preprocessing data, (2) Dividing each game into 10 equal time periods, (3) Separating events by time period, (4) Generating the "Region-Based Cumulative Possession Matrix" to quantify event distribution randomness (EDRan), and (5) Evaluating the significance of EDRan on match outcomes using a GLM classification model.

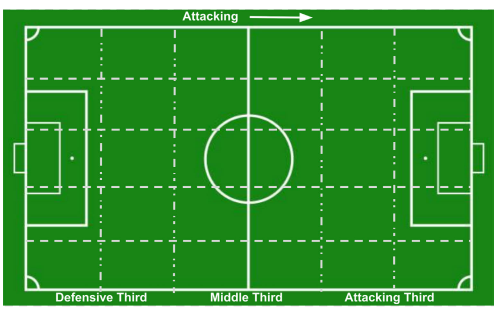


FIGURE 2. Division of the football field into 30 regions.

build up an attack in association football, it has been decided to divide each game into ten equal periods ($n = 10$).

For each time period, t_i ($1 \leq i \leq 10$), region-based cumulative ball possession duration matrices were generated for each team, considering the event location, event duration and the possession team (Figure 3). Only the events that occurred during the corresponding t_i period were considered in the process. In the matrix (5 rows $\times$ 6 columns), each cell indicates a region of the pitch (i.e., $5 \times 6 = 30$ regions in total per matrix). The matrix cell values are computed as follows. First, the events that happened during the considered t_i period were filtered out. For each filtered event, the *updating matrix* was decided by the ball possession team at that event, and the event location determines the updating matrix's *cell position*. Event duration, in seconds, was added to the existing value at the corresponding matrix's cell position. Each matrix was then divided by the total possession duration (in seconds) of the corresponding team, during the considered t_i period, to obtain a probability distribution. A set of probability distributions is then obtained for all the t_i periods in every game that is available in the dataset. Hence, each game will have 20 probability distributions, generated from 20 matrices (10 matrices per team), considering ten time periods (t_i: $1 \leq i \leq 10$). Figure 3 shows the steps involved in the generation of region-based cumulative possession matrices, and how the entropy values per team has been obtained

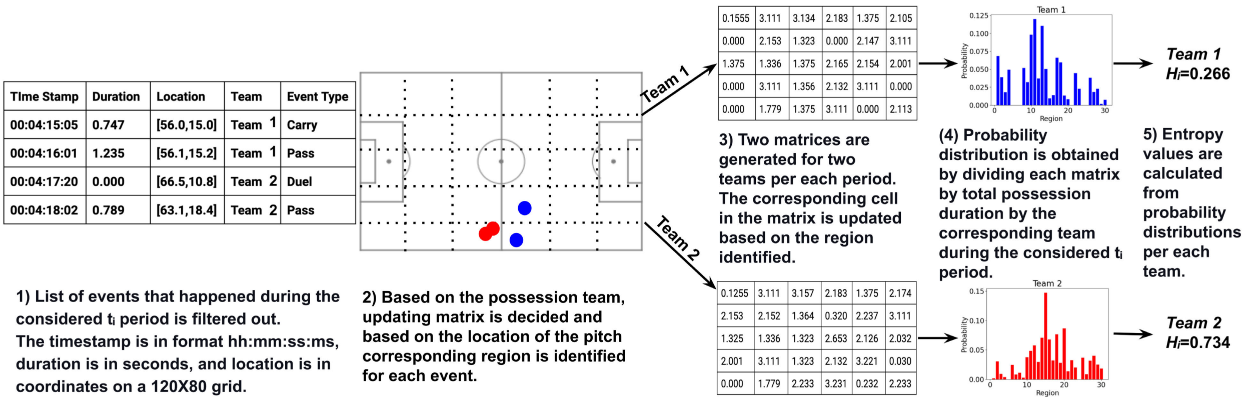


FIGURE 3. Steps involved in generating region-based cumulative possession matrices and computing entropy values for a single t_i period. Steps include filtering out events that happened during the considered time interval, generating cumulative possession matrices, generating probability distributions, and calculating entropy values.

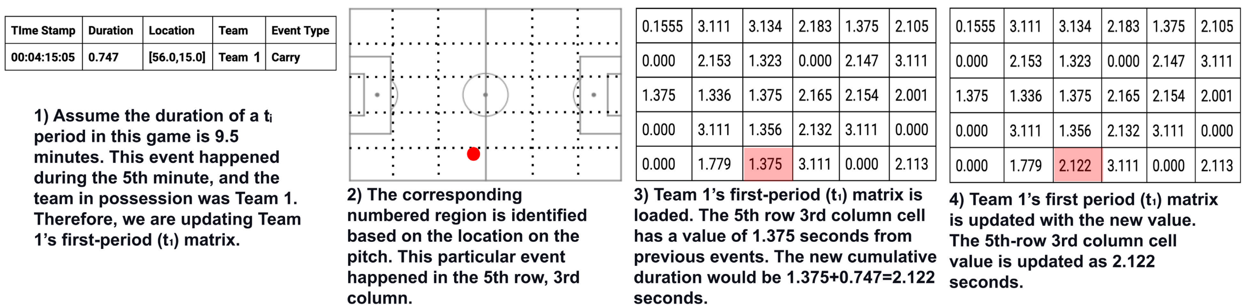


FIGURE 4. Steps of updating the region-based cumulative possession matrix.

(as discussed in section II-C). Figure 4 shows the steps involved in the matrix update process.

In order to analyse the time series nature of randomness in ball movement in a single game, as a case study, consider the UEFA Champions League 2015 final match between Juventus and Barcelona. The event distribution at different time intervals for this match is considered for the below analysis.

This particular game was chosen for the initial analysis based on its importance as a UEFA Champions League final game, where two teams with equal tactical brilliance and strengths compete. The match duration was 98 minutes, including the injury time. The game was divided into ten equal periods of 9.8 minutes. Figure 5 shows the event distribution heat maps of 30 regions obtained for the ten t_i periods. It can be observed that the team Barcelona, which won the match with a score of 3-1, utilized more spatial regions on the field when compared with Juventus during the earlier periods of the game. In the first six t_i periods, Barcelona utilized more spatial regions. However, Juventus has utilized more spatial regions in the subsequent four t_i periods. Moreover, heat maps were relatively uniformly distributed in the early periods with Barcelona, followed by a more uniform distribution at the end of the game with Juventus. Consequently, Barcelona maintained a more randomized spatial distribution of events at the beginning of the game, followed by Juventus towards the end of the game.

C. RELATIVE EVENT DISTRIBUTION RANDOMNESS

Next, an approach to quantify relative randomness in event distribution using obtained probability distributions is proposed.

Shannon entropy, a measure of uncertainty, measures unpredictability or randomness in a given data. Shannon Entropy (H) of a distribution with N number of distinct events is defined as:

$$H = -\sum_{i=1}^{N} P_i \log_2 P_i \tag{1}$$

where P_i is the probability of event i occurrence.

Kullback-Leibler divergence (KL divergence, also known as relative entropy) can be used to quantify the difference between two probability distributions. KL divergence ($D_{KL}(P(x)||Q(x))$ of two probability distributions, $P(x)$ and $Q(x)$, defined on the same sample space X and of a discrete random variable x is:

$$D_{KL}(P(x)|Q(x)) = \sum_{x \in X} P(x) \log_2 \frac{P(x)}{Q(x)} \tag{2}$$

where $P(x) > 0$ and $Q(x) > 0$ for any $x \in X$.

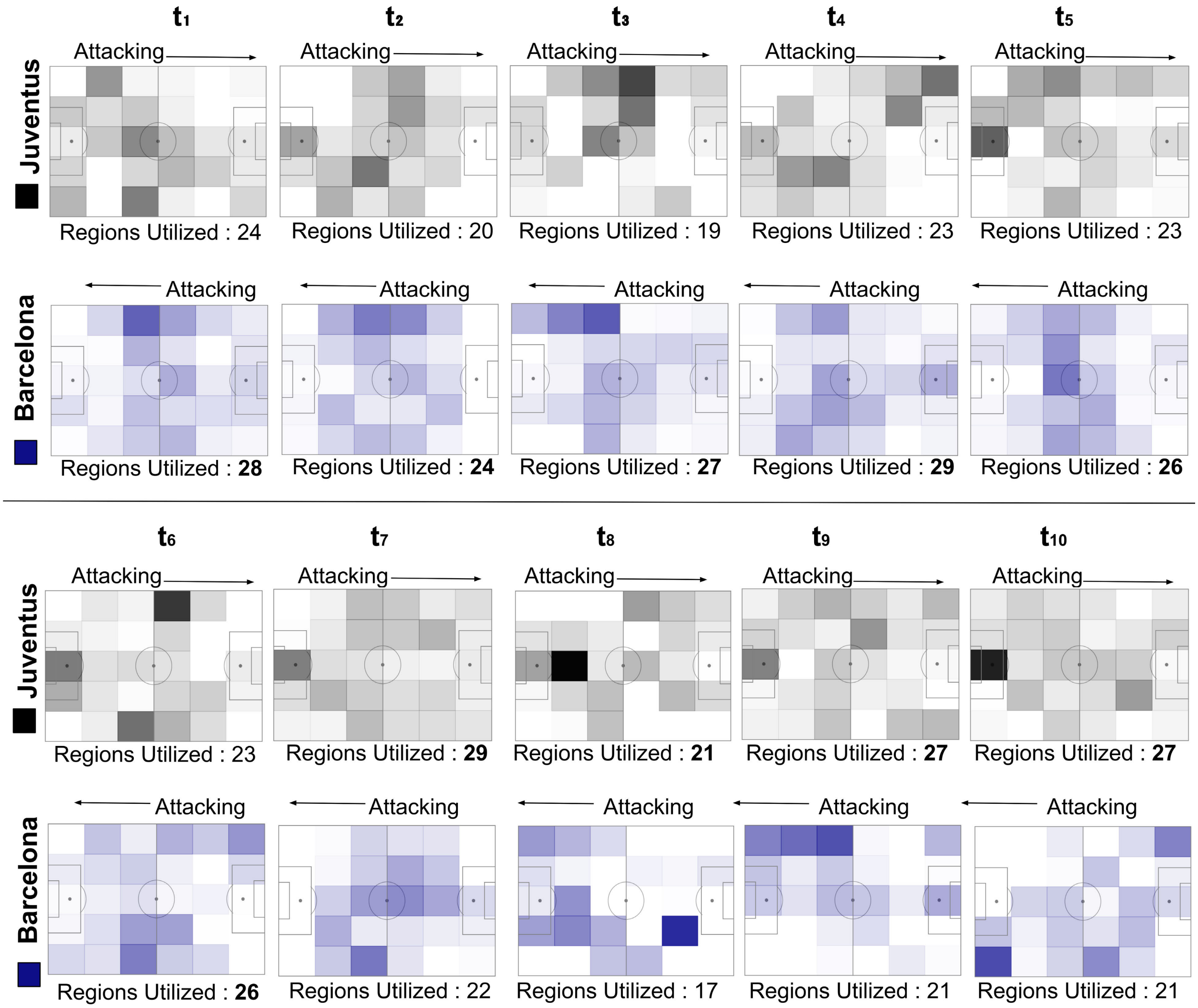


FIGURE 5. **Event distribution heat map of 30 regions for ten t_i periods ($1 \leq i \leq 10$) of the 2015 Champions League final (Juventus vs Barcelona). Color intensity correlates with the total possession duration probability in each region. The number of spatial regions utilized by each team during each t_i period is indicated below the corresponding heatmap (Higher number of regions utilized are indicated with bold font).**

Figure 6 shows the temporal randomness in event distribution, quantified using Shannon entropy, in team performances for each t_i period ($1 \leq i \leq 10$) in games of (1) FIFA World Cup 2018 (64 matches, 32 teams), and (2) UEFA EURO 2020 (51 matches, 24 teams). For this analysis, only games that ended up with a winner in regular playing time were considered. It can be observed that the winner's mean entropy was higher in the early periods. However, the losers' entropy was higher towards the latter.

Inspired from the analysis of figure 6, calculating entropy differences as a relative entropy measure for quantifying relative randomness is proposed. Note the total duration of an association football game may change slightly from game to game with the added injury time. Therefore, a t_i period's duration may vary from game to game. Thus, entropy calculation per unit time is considered here.

While KL divergence is commonly recognized as a relative entropy measure to compare two probability distributions, this work proposes using Shannon entropy difference as an alternative relative entropy measure, for evaluating temporal distributions of winners' and losers' performances. KL divergence distributions are always positive, whereas Shannon entropy differences can yield both positive and negative values, and thus, differences between distributions are easily identifiable. Nevertheless, this work compares the performance of the proposed relative entropy measure, the Shannon entropy difference per unit time (*HD*), with the standard relative entropy measure, the KL divergence per unit time (*KL*), to identify the most appropriate relative entropy per unit time measure for the proposed time-series metric, *EDRan*.

The Shannon entropy difference per unit time (HD_i) is calculated by subtracting the Shannon entropy per unit time of Team 2 for the period t_i ($H2_i/D_i$) from the Shannon entropy

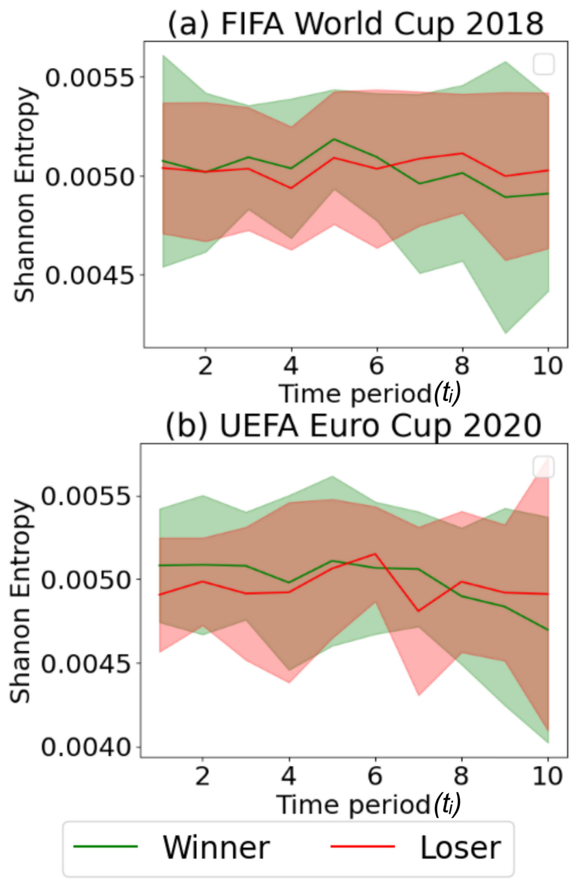


FIGURE 6. Mean Shannon entropy of winners and losers in (a) FIFA World Cup 2018 and (b) UEFA Euro 2020. Green and red shaded areas indicate the standard deviations.

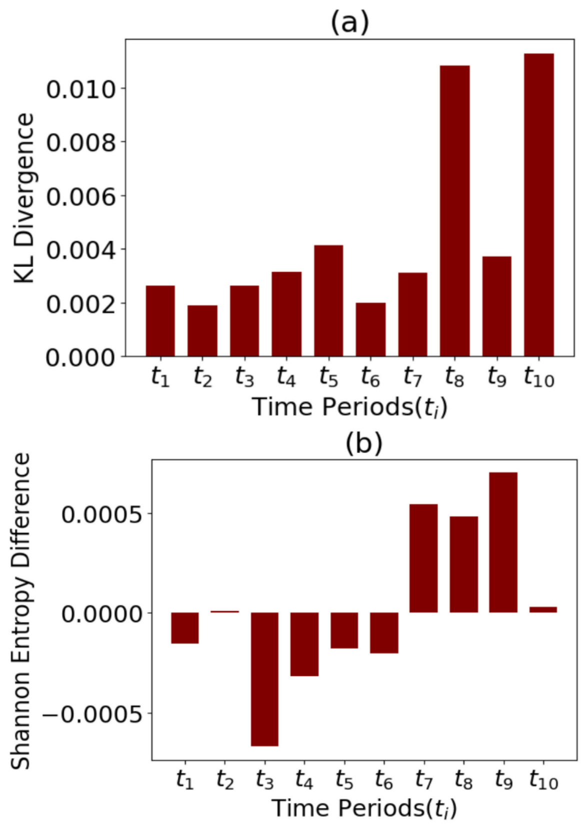


FIGURE 7. Relative entropy measures (a) *KL* (b) *HD* between Juventus and Barcelona of t_i periods in the UEFA Champions League 2015 final.

per unit time of Team 1 for the period t_i $(H1_i/D_i)$.

$$HD_i = \frac{H1_i - H2_i}{D_i}, \tag{3}$$

where D_i is the duration of the time period t_i $(D_i > 0)$, $H1_i$ is the Shannon entropy of Team 1 for the period t_i, and $H2_i$ is the Shannon entropy of Team 2 for the period t_i.

KL divergence of team 1 against team 2 per unit time (KL_i) for period t_i is calculated as follows.

$$KL_i = \frac{D_{KL}(T1_i|T2_i)}{D_i} \tag{4}$$

where D_i is the duration of the time period t_i $(D_i > 0)$, $T1_i$ is the spatial event probability distribution of Team 1 for the period t_i, and $T2_i$ is the spatial event probability distribution of Team 2 for the period t_i,

Figure 7 demonstrates temporal relative entropy measure distributions with (a) *KL* and (b) *HD* in the 2015 Champions League final (Juventus vs. Barcelona). Here, Juventus was considered Team 1, and Barcelona was Team 2. Therefore, the negative entropy difference with *HD* indicates Barcelona's higher spatial event distribution randomness per unit of time in play than that of Juventus during that particular t_i period.

Team Barcelona dominated the game during the first half, creating eight attempted goals with a ball possession of 66%. In contrast, Juventus created only five attempted goals with a ball possession of 34%. However, during the second half, both teams equally performed while creating 9 and 10 attempted goals, respectively. It was observed that the proposed relative entropy measure *HD* demonstrates a more noticeable prominent relationship with the team performance compared to the distribution with *KL*. In *HD* distribution, the winner maintained more entropy per unit time in the first six periods of the game (−ve values) followed by higher entropy per unit time by looser (+ve values) in the subsequent four periods (figure 7), which correlates with the actual match results, where the Barcelona dominated the early parts of the game followed by more offense by Juventus towards the latter. With *KL*, no significant difference can be observed as *KL* is always positive. However, towards the latter, the magnitude increased significantly, suggesting that Team 1 (Juventus) maintained more randomness in event distribution than Team 2 (Barcelona).

D. CORRELATION WITH MATCH RESULTS

Two pre-processed datasets were created for both male and female datasets, derived from the time-series features extracted using the two discussed temporal relative entropy measures (*HD*, *KL*). In order to identify the most appropriate

TABLE 1. Model results.

Gender	*Approach*	*Accuracy*	*F1*	*Precision*	*Recall*
Men's	*KL*	0.7333	0.7491	0.7767	0.7264
	HD	0.7995	0.8189	0.8103	0.8302
Women's	*KL*	0.7296	0.7541	0.7423	0.7755
	HD	0.7660	0.7856	0.7901	0.7870

temporal relative entropy measure for *EDRan*, this work compared the proposed per unit time temporal relative entropy measures (*HD*, *KL*) for their correlation with match-winning performances. Each row of time-series data consisting of ten feature values from the ten intervals (t_i) was retrieved using one of the measures (*HD*, *KL*). Target data was represented with values 0,1, where 0 represents a Team 2 victory result, and 1 represents a Team 1 victory. Both men's datasets comprised 552 rows representing 552 games considered and women's datasets comprised of 365 rows representing 365 games.

The correlation of the two approaches (*HD*, *KL*) with match-winning team performance was evaluated by match-winning performance classification performed using a Generalized Linear Model (GLM). A GLM is a versatile and widely used statistical framework that extends the traditional linear regression model to handle a wide range of data types and data. They enable the evaluation of each input feature's statistical significance with p-values, helping to identify which t_i periods significantly influence match results. Further, from the estimated coefficients of the GLM model, the magnitude and direction of the relationships between predictors and the response variable were identified. In the context of GLMs, a very low p-value (usually below 0.05) suggests that the predictor variable has a statistically significant impact on the response variable (Match winning performance).

A five-fold cross-validation approach was considered for the comprehensive evaluation of results using the GLM model. Fifty rounds of five-fold cross-validation (250 evaluations) were carried out. Based on these 250 evaluation results, the average classification accuracy, p-values, and correlation coefficients were calculated.

III. RESULTS

A. CLASSIFICATION MODEL RESULTS

Average classification accuracy, f1 score, precision, and recall were calculated based on five-fold cross-validation accuracies of 50 attempts ($5 \times 50 = 250$ training and testing combinations). Additionally, average p-values and GLM coefficients were calculated to evaluate the statistical significance and correlation with the match-winning performance. Table 1 presents the match-winning performance classification results with the two temporal relative entropy measures discussed. Figure 8 demonstrates the winning performance classification accuracy distributions with two measures.

HD approach performed better than the *KL* approach with both men's and women's data. *HD* approach achieved an average accuracy of 0.7995 (F1-score of 0.8189) and *KL* approach achieved an average accuracy of 0.7333 (F1-score of 0.7491) with men's dataset. With women's data *HD* approach achieved an average accuracy of 0.7660 (F1-score of 0.7856) and the *KL* approach achieved an average accuracy of 0.7296 (F1-score of 0.7541). Considering the classification results, a significant difference in accuracy was observed. The *HD* approach has performed better in classifying match-winning performance than the *KL* approach.

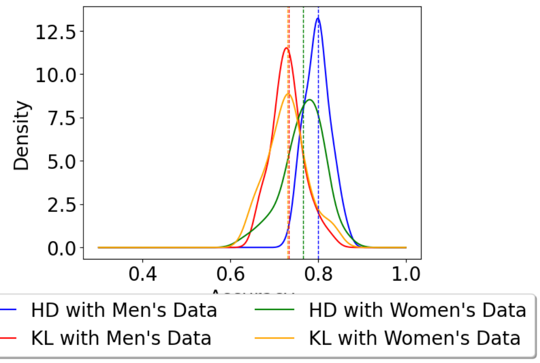


FIGURE 8. Winning performance classification accuracy distribution with *HD* and *KL* measures from 250 evaluations (Mean accuracies are indicated with dotted lines).

Considering these classification model results and significance observed with distributions, for the proposed time-series metric *EDRan*, Shannon Entropy difference (*HD*) was selected as the measure to quantify temporal relative randomness in event distribution from probability distributions obtained using the proposed "Region-based Cumulative Possession Matrix".

Hence, the proposed approach for quantifying *EDRan* includes,

1) Segmenting the game timeline into several time periods
2) For each team in each time period, retrieving probability distribution of randomness in ball-carrier event distribution using the proposed "region-based cumulative possession" matrices
3) Quantification of randomness with *HD* in each time period

B. FEATURE STATISTICAL SIGNIFICANCE

Average p-values and GLM coefficient values were calculated from 250 evaluations with the models trained with *EDRan* data (models trained using features extracted with *HD*). For both men's and women's models trained with the *EDRan* data, models have recorded similar p-values and correlation coefficients. Time periods t_1, t_2, t_3, t_4, t_5, t_{10} achieved very low p-values and relatively higher magnitude coefficients proving their statistical significance. In comparison, t_6, t_7, and t_8 have received higher p-values (> 0.05) and relatively lower magnitude coefficients, implying their statistical insignificance. Table 2 includes the average p-values and feature coefficients with the GLM

TABLE 2. Average p-values and feature coefficients.

Gender	Type	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}
Men's	$p - value$	7.6×10^{-3}	3.8×10^{-4}	5.5×10^{-4}	3.9×10^{-2}	1.0×10^{-3}	2.9×10^{-1}	4.1×10^{-1}	7.0×10^{-1}	5.5×10^{-2}	5.2×10^{-4}
	Coefficient	815.4	1052.2	1097.2	593.7	1049.7	311.6	208.7	59.9	-464.8	-837.0
Women's	$p - value$	2.2×10^{-3}	1.2×10^{-2}	5.6×10^{-2}	7.2×10^{-3}	4.1×10^{-2}	2.6×10^{-1}	6.3×10^{-1}	6.0×10^{-1}	5.2×10^{-1}	2.6×10^{-6}
	Coefficient	1278.4	793.5	643.9	1110.6	780.4	368.11	-83.5	-137.5	-164.8	-1278.0

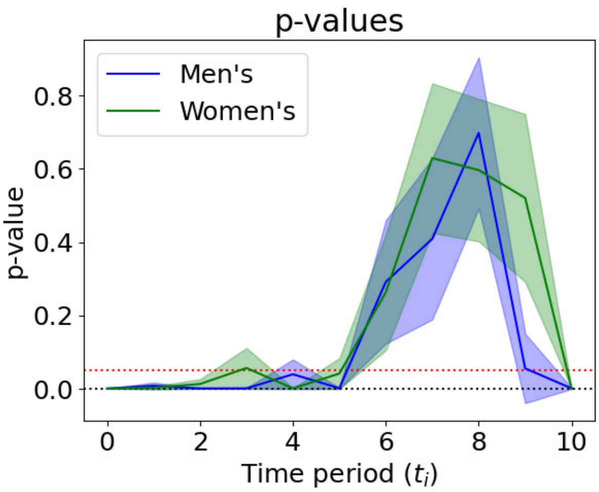


FIGURE 9. p-values for each t_i time period (Threshold for significance (0.05) is indicated with the red dotted line).

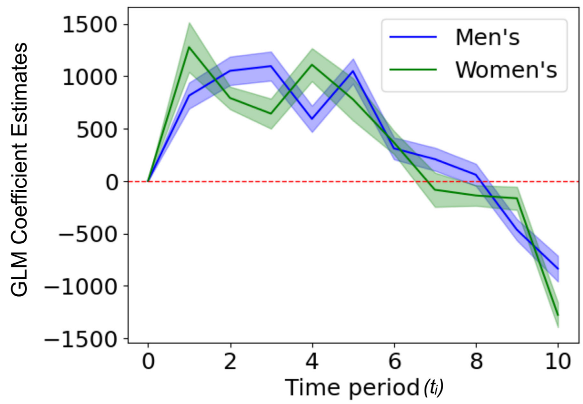


FIGURE 10. GLM coefficient estimates variance with t_i periods.

model while Figure 9 demonstrates the variance of p-values with t_i time periods.

C. TEMPORAL PATTERN ANALYSIS

In addition to the statistical significance of features, direction (positive or negative) and magnitude of the relationship between temporal features and match-winning performance were evaluated with GLM coefficient estimates. GLM coefficient estimates demonstrated a positive direction relationship between *EDRan* metric values and game results for the first eight periods with men's data and six periods with women's data followed by negative coefficient estimates for the subsequent periods (Figure 10). Further, a relatively higher magnitude in GLM coefficient estimates was observed in the first five time periods, followed by a relatively lower magnitude in coefficient estimates in t_6, t_7, and t_8 with both men's and women's data.

Average p-values and coefficient estimates from GLM suggested that first-half performance is more crucial to a match-winning performance. A higher magnitude in GLM coefficient estimates and low p-values for the first five periods suggested that the GLM model has given more importance to the first five periods for match-winning performance classification. This indicated that the first-half performance is more pivotal for a match-winning performance. Periods t_6, t_7, t_8 were observed to be the least important in a game for a match-winning performance based on these results. However, importance improved for t_9 and t_{10}.

Positive values for GLM coefficient estimates in the first eight periods in men's games and the first six periods in women's games indicated that most winners have maintained more spatial event distribution randomness during the game's early phases than their opponents. However, towards the latter part of the game (t_9,t_{10} in men's games and t_7, t_8, t_9, t_{10} in women's games), the negative coefficient estimates between *EDRan* and the match-winning performance indicated that most of the winners played compactly and with lesser randomness towards the latter. It was also observed that the mean time for player substitutions and tactical shifts in the available match data is around 62 minutes and 63 minutes, respectively (Tactical shifts refer to changes in formation and strategies during a game). Over time, it has been observed that the team leading may opt for a cautious defensive strategy towards the end [40]. Teams may opt for substitutions and tactical shifts around this time, where the leading teams may change their strategy to a cautious defensive approach to secure their lead.

Both men's and women's data receiving similar p-value and GLM coefficient estimate patterns may imply that both men's and women's teams utilize similar event distribution tactics where the winning teams opt for a cautious defensive approach towards the end of the game despite the strategical and physical differences between men's and women's games observed in existing literature [36]. However, it was also noted that women's winning teams have opted for a defensive approach earlier (t_7) than in men's games (t_9).

D. COMPARISON WITH EXISTING LITERATURE

A temporal team performance evaluation metric that correlates with match-winning performances has been proposed in this work. To the best of the author's knowledge, no similar work has been done proposing a temporal team performance evaluation metric that correlates with match-winning performances or has been used for the development of match-winner prediction models. However, by using only the proposed metrics' data a match-winner prediction model has been developed in this work to evaluate the proposed metric's correlation with match-winning performances and to identify hidden insights on temporal features. The developed match-winning prediction model has performed with better accuracy compared to match-winner prediction models trained with historical data and multiple performance evaluation metrics. Table 3 presents a comparison between match winner prediction models in recent publications and this work.

Although [42] have used a temporal feature extraction approach from game data, they have used 22 player evaluation metrics and used 396 data points for prediction. Further, they have mainly focused on the development of an accurate match-winner prediction model using a computationally expensive deep learning approach, Gated Recurrent Unit (GRU). Being a black-box model, GRU does not provide insights into temporal features. On the other hand, this work focused mainly on the development of a time-series performance evaluation metric to capture and evaluate team performance. A match-winner prediction model trained with developed performance evaluation metric data achieved similar match-winner prediction accuracy and a better F1 score with a simpler and computationally inexpensive GLM model. Additionally use of GLM has enabled to exploration of hidden insights on temporal features using p-values and correlation coefficients. As both this paper's work and [42] work extract temporal features, this allows these approaches to develop match-winner prediction models to predict match winners at different time stamps. This could be explored in future work.

E. TEAM PERFORMANCE ANALYSIS

After determining the statistical significance and importance of the proposed time-series metric compared to existing literature, it was applied to assess the performance of teams during the 2018 FIFA World Cup.

FIFA World Cup 2018 comprised two stages: The group stage and the knockout stage. Thirty-two teams were divided into eight groups in the group stage and played against other teams in the same group. The first two teams from each group qualified for the knockout stage, while the last two teams were eliminated from the group stage. In the knockout stage, losers were eliminated, while winners advanced to the next knockout round. Round of 16 was the first knockout round, where eight matches were played among 16 teams. Eight teams advanced to the quarter-finals, and four winners from the quarter-finals advanced to the semi-finals. Winners of the two semi-finals played in the grand final, and losers in the semi-finals played for the bronze medal.

The results of the group stage matches were analyzed to determine whether there were any differences in performance between the teams that were eliminated from the first round and those that advanced to the knockout stage. Forty-eight matches were played in the group stage. Winning and non-winning performances (losses and draws) of teams that advanced to the knockout stage and eliminated teams were analyzed separately. However, team Geramany's performances were separated from the eliminated group as their performance distributions drastically differed from the rest of the group. Therefore, sixteen advanced teams were considered the qualified group, and fifteen teams, except team Germany, were considered under the eliminated group. Team Germany's performances were analyzed separately.

In both winning and losing performances, it was observed that teams that advanced to the knockout stage maintained more *EDRan* than teams that were eliminated. Teams that were eliminated have maintained lesser *EDRan* than teams that advanced on average, even in their winning performances. Teams that were eliminated failed to maintain *EDRan* above 0 most of the time, even in winning performances, while teams that advanced to the knockout stage managed to maintain *EDRan* above 0 for the majority of the time, even in non-winning performances. This indicated that advanced teams have maintained more random event distribution, utilizing more space and attempting to create more scoring chances on average, even in their non-winning performances (Playing offensively). On the other hand, eliminated teams have maintained a more compact and defensive approach on average (lesser *EDRan*), even in their winning performances. Figure 11 compares the average *EDRan* in winning performances of teams that qualified for knockout stages and teams that were eliminated from the group stage. Figure 12 presents the same comparison with non-winning performances.

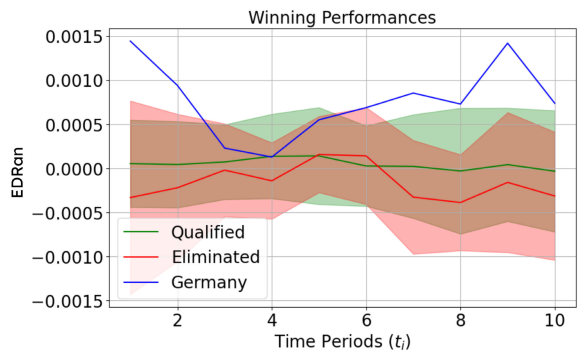


FIGURE 11. Comparison between average *EDRan* (HD) in winning performances of teams that qualified for knockout stages and teams that were eliminated from the group stage (Shaded area indicates the standard deviations).

However, It should be noted that Team Germany's performance differs significantly from that of other eliminated teams. Being the reigning world champion and number one ranked team in the FIFA World Team ranking prior to the World Cup [43], team Germany was one of the favorites at the 2018 FIFA World Cup. Nevertheless, illustrating the unpredictable nature of association football, team Germany placed last in the group, winning only one game and

TABLE 3. Comparison with other match winner prediction models.

Year	*Literature*	*Data*	*Approach*	*No. of features*	*Model*	*Accuracy*	*F1-Score*	*Perf. Eval. Aspect*
2018	Danisik et. al. [41]	Five top tier leagues in Europe (English Premier League, Spanish La Liga, German Bundesliga, Italian Seria A, and French Ligue 1)	Player attributes and results from last five games are considered	139 features (6 player performance evaluation features for all players + match result information from previous games)	LSTM	70.21%	-	To analyze the best players' combinations among the squad
2023	Al Mulla et al [42]	Qatar Stars League	Player performance from player performance evaluation metrics have been collected temporally	396 features (396 features from 22 player performance evaluation metrics collected at six timestamps under three player position categories (attack,defense,central))	GRU	**80.77%**	80.93%	NA
2024	This work	Four top tier leagues in Europe (English Premier League, Spanish La Liga, German Bundesliga, Italian Seria A), FIFA world cup 2018, UEFA Euro 2020	Proposed time-series performance evaluation metric obtained by quantifying event distribution randomness from unexplored region based cumulative possession matrices	10 temporal features (10 temporal features extracted at 10 timestamps from the proposed performance evaluation metric (single metric is used))	GLM	80.00%	**81.89%**	To analyze team performance and tactics (space utilization, event distribution) with proposed time-series performance evaluation metric

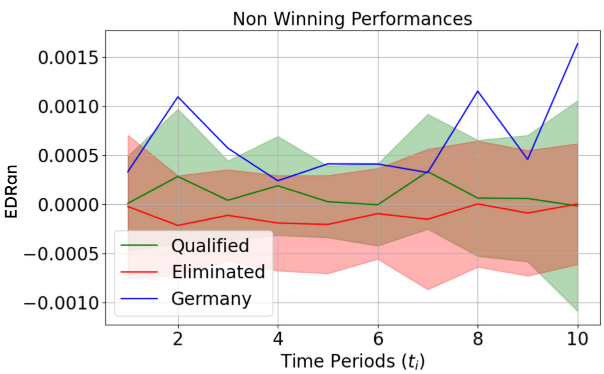


FIGURE 12. Comparison between average *EDRan* (HD) in non-winning performances of teams that qualified for knockout stages and teams that were eliminated from the group stage (Shaded area indicates the standard deviations).

losing two. However, they have still played with better *EDRan* than most other teams in both their winning and non-winning performances. Although they have failed to qualify for the knockout stage based on the real scores, match statistics indicated that they have ranked third with non-penalty expected goals (npxG) per 90 minutes rankings, second in expected goals + assisted goals (xG+xGA) per 90 minutes rankings and ranked sixth in expected goals difference (xGD) per 90 minutes rankings in the FIFA World cup 2018 [44]. Expected goals (xG) is a widely used performance evaluation metric in association football, which is used to quantify the probability of a shot being a goal considering factors like shot location, defender location, and goalkeeper location [45], [46]. Further, xG has been researched for performance evaluation for dealing with randomness in match outcomes [47]. These statistics with xG metrics indicated that Team Germany has created more highly probable goal-scoring opportunities while defending their goal without letting opponents create high-probability goal-scoring opportunities. However, they have failed to utilize those high probable scoring opportunities while their opponents have utilized most of the scoring opportunities, illustrating the unpredictability of association football.

It was noted that, teams which qualified have maintained more mean *EDRan* throughout the match than their opponents during group stage matches. This suggested that teams that qualified have been playing offensively utilizing more space and event distribution randomness in an attempt to create more goal-scoring opportunities. The reason for this may be the fact that in the group stage when teams are tied on points in the points table, goal differences may be important to qualify for the knockout stage. Further, teams that qualify to the knockout stage are often the better side and thus they will try to attack the opposition throughout the match when playing each other. However, a contrasting pattern could be observed with knockout stage matches where the winning team's mean *EDRan* goes below zero while the losing team's mean *EDRan* increases above zero towards the second half of the game (Figure 13). This indicates a compact playing strategy by winning teams while more space is utilized and a spatial event distributed strategy by losing teams towards the end of the game. In knockout stage goal differences or win margins are not considered. Also, all the teams were stronger with equal strengths and ranked number one or two in their group stages. Therefore, teams may consider securing their lead once leading. Therefore, the leading team may opt for compact defensive strategies like "park the bus" towards the end of the game [40], [48]. In order to catch up with their opponents, trailing teams may adopt a more offensive strategy. Similar behavior has been observed in the literature where leading teams are cautious towards the end of the game and more cautious if they are the weaker team [40].

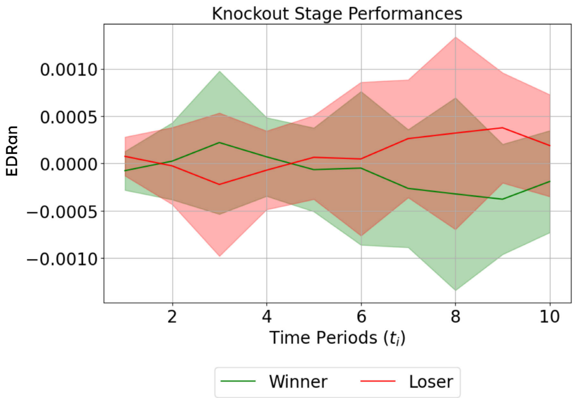


FIGURE 13. Knockout stage performances comparison between winners and losers.

IV. DISCUSSION

In this work, a novel entropy-based time-series metric, *EDRan*, was proposed to quantify the randomness in team ball movement in association football. Higher classification performance compared to existing literature, low GLM p-values, and high GLM coefficient estimates for the classification of match-winning results have demonstrated the significance of the proposed metric with match-winning performances and statistical significance of the temporal features. Additionally, compared approaches in existing literature have mainly focused on developing a match-winner prediction model using computationally expensive deep learning models and numerous features, whereas the main focus of this work was to develop a time-series performance evaluation metric, and thus, it has used computationally inexpensive models and considered only one metric for temporal feature extraction.

Achieving higher classification ac-curacies for match-winning performance classification with both men's and women's data proves the generalizability of the proposed approach. Moreover, the analysis in this work reveals that the first-half performance is more pivotal for the winning performances through p-values and GLM coefficient estimates. Further, it was also revealed that winning teams tend to follow a more compact, less random, and lesser event-distributed (cautious defensive strategy) towards the end of the game.

However, contrasting patterns were observed in group-stage performances in the FIFA World Cup, where goal differences matter. Stronger teams have maintained higher spatial event distribution randomness throughout the game (offensive strategy). Yet, with knockout performances, the winning teams followed a strategy with compact and lesser random ball movement toward the end of the game. It was also noted that in knockout performances, winning teams opted for a compact strategy earlier than what was observed with other games in general. The proposed time-series metric, EDRan further revealed that the team Germany which was knocked out from the first round has performed with highly random event distribution on average than most of the teams, which agrees with rankings based on widely recognized performance evaluation metrics, like xG+xGA and npxG. However, xG+xGA and npxG metrics are not time-series metrics. Thus, they do not provide the temporal time-series stochastic nature of the performance. These results imply the proposed approach's ability to recognize better performances in a temporal time-series manner, even when the team ended up losing, demonstrating its suitability as a time-series performance evaluation metric.

It is important to acknowledge that this study utilized a smaller dataset consisting of over 1232 games. Future research could investigate this topic using a larger dataset and more computationally intensive models. Moreover, due to the dataset's size, only ten t_i periods were included. The exploration of a greater number of periods could be considered with a larger dataset. The proposed approach allows to prediction of match winners at certain timestamps of the game which could be explored in the future. Furthermore, the proposed approach assigns equal importance to event distribution across all regions of the field. However, the impact of event distribution randomness may vary between different areas. This aspect could be investigated in future work. It is important to note that this research was specifically focused on association football games, but it is worth mentioning that the same methodology can be applied to other similar invasion sports, such as hockey, basketball, and handball. Additionally, this dataset covers matches played after 2004, and game-play strategies can evolve and change over time. As a result, the developed model is likely to perform best under similar conditions to those found in the analyzed dataset. Yet, it may also serve as a valuable tool for investigating how strategies in association football evolve and how randomness in ball movement plays a part in evolving association football strategies.

V. CONCLUSION

With the high classification performance obtained with classifying match winners, it is fair to conclude that maintaining high randomness in team ball movement is significantly important for victories. The proposed time-series metric based on randomness in team ball movement has demonstrated its ability to evaluate the team's performance despite the unpredictability associated with actual game scores. Generalizability of the proposed approach and obtaining similar insights with both men's and women's games conclude that both men's and women's teams tend to play with similar approaches although tactical and physical differences between games of the two sexes have been identified in the literature [36]. It was also observed that winning teams have opted for a compact and cautious behavior towards the end of the game which agrees with existing work [40].

Several key insights were revealed with observed temporal patterns. Generally, winners often maintain more spatial event distribution randomness than their opponents in the early phases of the game. Therefore, teams should always try to maintain a highly distributed random ball movement, eventually creating more open spaces and opportunities, as unpredictability in ball movement will make it difficult for the defense to predict the movements. Results further conclude that the first-half performance is vital in taking control of the game for match-winning performances. Therefore, fielding the best eleven at the start can be more beneficial than saving them for later substitutions. Contrarily, winners may play with less random and compact defensive strategies towards the latter to safeguard their lead. Although this nature can be observed in general, a different pattern was observed during the group stages where the goal differences matter, winning teams tend to play offensively with random and unpredictable ball movement throughout the match.

Capturing the differences in performance between better-performing and weaker teams despite the unpredictability of actual scores showcased the significance of the proposed approach as a performance evaluation metric. Teams can rely on the proposed metric to analyze the team performance in the long run, as association football is always affected by random components or chance. With the revealed insights from this work, teams are recommended to implement enhanced randomness in ball movement in offensive play for winning performances.

ACKNOWLEDGMENT

The authors express their gratitude to Statsbomb for sharing a publicly accessible dataset of association football statistics derived from actual match data.

REFERENCES

[1] L. Lamas, J. Barrera, G. Otranto, and C. Ugrinowitsch, "Invasion team sports: Strategy and match modeling," *Int. J. Perform. Anal. Sport*, vol. 14, no. 1, pp. 307–329, Apr. 2014.

[2] M. Cavus and P. Biecek, "Explainable expected goal models for performance analysis in football analytics," 2022, *arXiv:2206.07212*.

[3] K. A. Lemmink and W. Frencken, "Tactical performance analysis in invasion games: Perspectives from a dynamic systems approach with examples from soccer," in *Routledge Handbook of Sports Performance Analysis*. Evanston, IL, USA: Routledge, 2013, pp. 89–100.

[4] J. Dvorak, A. Junge, T. Graf-Baumann, and L. Peterson, "Football is the most popular sport worldwide," *Amer. J. Sports Med.*, vol. 32, pp. 3S–4S, Jan. 2004.

[5] C. McCarthy, P. Tampakis, M. Chiarandini, M. Randers, S. Jänicke, and A. Zimek, "Analyzing passing sequences for the prediction of goal-scoring opportunities," in *Machine Learning and Data Mining for Sports Analytics*, U. Brefeld, J. Davis, J. Van Haaren, and A. Zimmermann, Eds. Cham, Switzerland: Springer, 2023, pp. 27–40.

[6] F. Wunderlich, A. Seck, and D. Memmert, "Skill or luck? Analysing random influences on goal scoring in football," in *Proc. 13th World Congr. Perform. Anal. Sport, 13th Int. Symp. Comput. Sci. Sport*, Jul. 2023, pp. 126–129.

[7] C. Reep and B. Benjamin, "Skill and chance in association football," *J. Roy. Stat. Soc. Ser. A, General*, vol. 131, no. 4, pp. 581–585, 1968.

[8] A. Heuer, C. Müller, and O. Rubner, "Soccer: Is scoring goals a predictable Poissonian process?" *EPL (Europhysics Letters)*, vol. 89, no. 3, p. 38007, Feb. 2010.

[9] A. Heuer and O. Rubner, "How does the past of a soccer match influence its future? Concepts and statistical analysis," *PLoS ONE*, vol. 7, no. 11, Nov. 2012, Art. no. e47678.

[10] G. M. Zebari, S. Zeebaree, M. M. Sadeeq, and R. Zebari, "Predicting football outcomes by using Poisson model: Applied to Spanish primera división," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 2, pp. 133–140, Nov. 2021.

[11] M. Hughes and I. Franks, "Analysis of passing sequences, shots and goals in soccer," *J. Sports Sci.*, vol. 23, no. 5, pp. 509–514, May 2005, doi: 10.1080/02640410410001716779.

[12] (2023). *Football Stats | Player Comparisons | Team Analyses*. [Online]. Available: https://xvalue.ai/stats

[13] C. Peñas, J. L. Ballesteros, A. Dellal, and M. G. López, "Game-related statistics that discriminated winning, drawing and losing teams from the Spanish soccer league," *J. Sports Sci. Med.*, vol. 9, pp. 288–293, Jun. 2010.

[14] K. Goral, "Passing success percentages and ball possession rates of successful teams in 2014 FIFA World Cup," *Int. J. Sci. Culture Sport*, vol. 3, pp. 86–95, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:36687428

[15] R. Aquino, J. C. Machado, F. Manuel Clemente, G. M. Praça, L. G. C. Gonçalves, B. Melli-Neto, J. V. S. Ferrari, L. H. P. Vieira, E. F. Puggina, and C. Carling, "Comparisons of ball possession, match running performance, player prominence and team network properties according to match outcome and playing formation during the 2018 FIFA World Cup," *Int. J. Perform. Anal. Sport*, vol. 19, no. 6, pp. 1026–1037, Nov. 2019.

[16] A. Kubayi and A. Toriola, "The influence of situational variables on ball possession in the south African premier soccer league," *J. Human Kinetics*, vol. 66, no. 1, pp. 175–181, Mar. 2019.

[17] *'Xavi—I Understand Football as a Time-Space' INTERVIEW*. Accessed: Nov. 21, 2023. [Online]. Available: https://www.youtube.com/watch?v=q61aaEnpYMg

[18] J. Fernández and L. Bornn, "Wide open spaces: A statistical technique for measuring space creation in professional soccer," in *Proc. MIT Sloan Sports Anal. Conf.*, Mar. 2018.

[19] F. Martens, U. Dick, and U. Brefeld, "Space and control in soccer," *Frontiers Sports Act. Living*, vol. 3, 2021, doi: 10.3389/fspor.2021.676179.

[20] J. Perl and D. Memmert, "A pilot study on offensive success in soccer based on space and ball control—Key performance indicators and key to understand game dynamics," *Int. J. Comput. Sci. Sport*, vol. 16, no. 1, pp. 65–75, Jul. 2017.

[21] D. Brinkjans, D. Memmert, J. Imkamp, and J. Perl, "Success-score in professional soccer—Validation of a dynamic key performance indicator combining space control and ball control within goalscoring opportunities," *Int. J. Comput. Sci. Sport*, vol. 21, no. 2, pp. 32–42, Dec. 2022.

[22] P. Kundzelevicius, "Investigating the tactical profile of defensive players in elite football. A network theory analysis," M.S. thesis, Lietuvos Sporto Universitetas, Kaunas, Lithuania, 2023.

[23] S. Swaby. *Pep Guardiola Calls Tiki-Taka 'Rubbish,' Admits That He Loathes All the Passing*. Accessed: Nov. 21, 2023. [Online]. Available: https://bleacherreport.com/articles/2234051-pep-guardiola-calls-tiki-taka-rubbish-admits-that-he-loathes-all-the-passing

[24] Y. Neuman, N. Israeli, D. Vilenchik, and Y. Cohen, "The adaptive behavior of a soccer team: An entropy-based analysis," *Entropy*, vol. 20, no. 10, p. 758, Oct. 2018. [Online]. Available: https://www.mdpi.com/1099-4300/20/10/758

[25] S. Kusmakar, S. Shelyag, Y. Zhu, D. Dwyer, P. Gastin, and M. Angelova, "Machine learning enabled team performance analysis in the dynamical environment of soccer," *IEEE Access*, vol. 8, pp. 90266–90279, 2020.

[26] Y. Berman, S. Mistry, J. Mathew, and A. Krishna, "Temporal match analysis and recommending substitutions in live soccer games," in *Proc. IEEE Int. Conf. Web Services (ICWS)*, Jul. 2022, pp. 397–404.

[27] J. H. Fewell, D. Armbruster, J. Ingraham, A. Petersen, and J. S. Waters, "Basketball teams as strategic networks," *PLoS ONE*, vol. 7, no. 11, Nov. 2012, Art. no. e47445.

[28] L. Vilar, D. Araújo, K. Davids, and Y. Bar-Yam, "Science of winning soccer: Emergent pattern-forming dynamics in association football," *J. Syst. Sci. Complex.*, vol. 26, no. 1, pp. 73–84, Feb. 2013.

[29] M. S. Couceiro, F. M. Clemente, G. Dias, P. Mendes, F. M. Martins, and R. S. Mendes, "On an entropy-based performance analysis in sports," in *Proc. 1st Int. Electron. Conf. Entropy Appl.* Basel, Switzerland: Springer, Nov. 2014, doi: 10.3390/ecea-1-a008.

[30] G. Cheng, Z. Zhang, M. Kyebambe, and N. Kimbugwe, "Predicting the outcome of NBA playoffs based on the maximum entropy principle," *Entropy*, vol. 18, no. 12, p. 450, Dec. 2016. [Online]. Available: https://www.mdpi.com/1099-4300/18/12/450

[31] P. Silva, R. Duarte, P. Esteves, B. Travassos, and L. Vilar, "Application of entropy measures to analysis of performance in team sports," *Int. J. Perform. Anal. Sport*, vol. 16, no. 2, pp. 753–768, Aug. 2016.

[32] R. Das. (2020). *The Future of Football: Why Player Positions May Not Matter Anymore*. [Online]. Available: https://www.sportskeeda.com/football/the-future-of-football-why-player-positions-may-not-matter-anymore

[33] StatsBomb. (Oct. 2022). *Statsbomb Open Data*. [Online]. Available: https://github.com/statsbomb/open-data.git

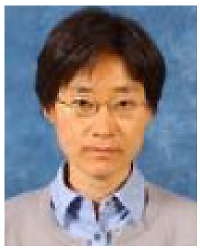
[34] L. Pappalardo and E. Massucco, Feb. 2019, "Soccer match event dataset," *Figshare*, doi: 10.6084/m9.figshare.c.4415000.v5.
[35] L. Pappalardo, P. Cintia, A. Rossi, E. Massucco, P. Ferragina, D. Pedreschi, and F. Giannotti, "A public data set of spatio-temporal match events in soccer competitions," *Sci. Data*, vol. 6, Oct. 2019, doi: 10.1038/s41597-019-0247-7.
[36] L. Pappalardo, A. Rossi, M. Natilli, and P. Cintia, "Explaining the difference between men's and women's football," *PLoS ONE*, vol. 16, no. 8, Aug. 2021, Art. no. e0255407, doi: 10.1371/journal.pone.0255407.
[37] C. Lago, "The influence of match location, quality of opposition, and match status on possession strategies in professional association football," *J. Sports Sci.*, vol. 27, no. 13, pp. 1463–1469, Nov. 2009.
[38] J. P. Guimarães, M. Rochael, A. Andrade, S. Glória, and G. Praça, "How reaching the pitch's final third is related to scoring opportunities in soccer?" *Retos*, vol. 43, p. 171, Jul. 2021.
[39] L. Tianbiao and H. Andreas, "Apriori-based diagnostical analysis of passings in the football game," in *Proc. IEEE Int. Conf. Big Data Anal. (ICBDA)*, Mar. 2016, pp. 1–4.
[40] T. Guan, J. Cao, and T. Swartz. (2021). *Should You Park the Bus?* [Online]. Available: https://www.sfu.ca/~tswartz/papers/bus.pdf
[41] N. Danisik, P. Lacko, and M. Farkas, "Football match prediction using players attributes," in *Proc. World Symp. Digit. Intell. for Syst. Mach. (DISA)*, Aug. 2018, pp. 201–206.
[42] J. AlMulla, M. T. Islam, H. R. H. Al-Absi, and T. Alam, "SoccerNet: A gated recurrent unit-based model to predict soccer match winners," *PLoS ONE*, vol. 18, no. 8, Aug. 2023, Art. no. e0288933, doi: 10.1371/journal.pone.0288933.
[43] G. Sanwal. *2018 FIFA Team Rankings: Germany Leads Before the World Cup*. Accessed: Nov. 21, 2023. [Online]. Available: https://www.fifaworldcupnews.com/germany-leads-fifa-rankings-world-cup/
[44] *2018 FIFA World Cup Player Stats*. Accessed: Nov. 21, 2023. [Online]. Available: https://fbref.com/en/comps/1/2018/stats/2018-World-Cup-Stats
[45] H. Eggels, M. Pechenizkiy, R. Almeida, R. Van Elk, and L. V. Agt, "Expected goals in soccer: Explaining match results using predictive analytics," M.S. thesis, Dept. Math. Comput. Sci., Eindhoven Univ. Technol., The Netherlands, Oct. 2016. [Online]. Available: https://research.tue.nl/en/studentTheses/expected-goals-in-soccer
[46] J. Mead, A. O'Hare, and P. Mcmenemy, "Expected goals in football: Improving model performance and demonstrating value," *PLoS ONE*, vol. 18, no. 4, Apr. 2023, Art. no. e0282295.
[47] M. Brechot and R. Flepp, "Dealing with randomness in match outcomes: How to rethink performance evaluation in European club football using expected goals," *J. Sports Econ.*, vol. 21, no. 4, pp. 335–362, May 2020.
[48] T. Guan, J. Cao, and T. Swartz, "Parking the bus," *J. Quant. Anal. Sports, De Gruyter*, vol. 19, no. 4, pp. 263–272, Dec. 2023.

ISHARA BANDARA received the master's degree in big data analytics from Robert Gordon University, U.K., in 2021. He is currently pursuing the joint Ph.D. degree with Deakin University, Australia, and Coventry University, U.K. Prior to pursuing the Ph.D. degree, he was a Software Engineer for six years. His research interests include machine learning, sports analytics, time-series data, and pattern recognition.

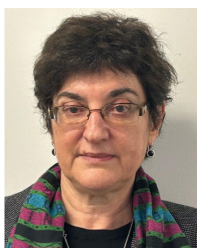
SERGIY SHELYAG received the Ph.D. degree from the University of Göttingen, Germany, in 2004. He was a Research Fellow with the Department of Applied Mathematics, The University of Sheffield, U.K., and the Department of Mathematics and Physics, Queen's University, Belfast, U.K. He was an ARC Future Fellow with the Department of Mathematical Sciences, Monash University, Australia; a Senior Lecturer of applied mathematics with the University of Northumbria, Newcastle, U.K., and a Senior Lecturer of data science with Deakin University, Australia. He is currently an Associate Professor of applied mathematics with the College of Science and Engineering, Flinders University, Australia. His current research interests include mathematical and computational modeling of complex physical processes, data analytics, machine learning, and dynamical systems.

SUTHARSHAN RAJASEGARAR received the Ph.D. degree from The University of Melbourne, Australia. He is currently an Associate Professor with the School of Information Technology, Deakin University, Geelong, Australia. His current research interests include anomaly/outlier detection, distributed machine learning, pattern recognition, computer vision, sports analytics, the IoT, and quantum computing.

DAN DWYER received the Ph.D. degree from Griffith University, Queensland. He was a Research Active-Academic with the University of Tasmania and the University of Newcastle. He was a Sports Scientist with the Victorian Institute of Sport, Melbourne. He is currently a Senior Lecturer of applied sport science with Deakin University. His research interests include measurement, analytics, and prediction in sports. He is a member of the Centre for Sport Research.



EUN-JIN KIM received the B.Sc. degree in physics from Yonsei University, Seoul, South Korea, and the Ph.D. degree in physics from the University of Chicago, USA. She held postdoctoral positions at the University of Leeds, U.K.; the University of Exeter, U.K.; the National Center for Atmospheric Research, Boulder, USA; and the University of California at San Diego, USA. She was an Associate Professor with The University of Sheffield, U.K. She is currently a Professor of physics and applied mathematics with Coventry University, U.K., and a Visiting Scientist with Seoul National University, South Korea. She has a track record of multidisciplinary research in complex systems and non-equilibrium statistical mechanics, in particular information geometry. She has published over 150 peer-reviewed research articles. She was awarded the Leverhulme Trust Research Fellowship.



MAIA ANGELOVA received the Ph.D. degree in theoretical and mathematical physics from the Faculty of Physics, Sofia University, Bulgaria. Since November 2023, she has been a Professor and the Lead of artificial intelligence with the Health Program, Aston Digital Futures Institute, Aston University, Birmingham, U.K. From January 2017 to November 2023, she was a Professor of data analytics and machine learning and the Director of the Data Analytics Laboratory, Deakin University, Melbourne, Australia. From January 1997 to December 2016, she was a Professor of mathematical physics and the Head of the Mathematical and Intelligent Modeling Laboratories, Northumbria University, Newcastle upon Tyne. From October 1991 to December 1996, she was a Lecturer of physics with the Somerville College, Oxford University, UK. Her research interests include data-driven modeling, analysis, and decision-making of time series arising from problems in health, sport, and defense.