Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

Data-driven team ranking and match performance analysis in Chinese Football Super League

Yuesen Li[a], Runqing Ma[a], Bruno Gonçalves[b,c,d], Bingnan Gong[e], Yixiong Cui[a,f,*], Yanfei Shen[a,**]

[a] School of Sports Engineering, Beijing Sport University, Beijing, P.R. China
[b] Departamento de Desporto e Saúde, Escola de Ciências e Tecnologia, Universidade de Évora, Évora, Portugal
[c] Comprehensive Health Research Centre (CHRC), Universidade de Évora, Évora, Portugal
[d] Portugal Football School, Portuguese Football Federation, Oeiras, Portugal
[e] Faculty of Physical Activity and Sports Sciences. Universidad Politécnica de Madrid, Madrid, Spain
[f] AI Sports Engineering Lab, School of Sports Engineering, Beijing Sport University, Beijing, P.R. China

ARTICLE INFO

Article history:
Received 17 July 2020
Revised 19 September 2020
Accepted 30 September 2020

Keywords:
Soccer
Match analysis
Ranking
Predictive modeling
Big data

ABSTRACT

Recent years have seen an increasing body of research into the evaluation of the team-level technical-tactical performance in association football using match events data. However, most studies used mono-dimensional approach and modeled the influence of each performance aspects on match result in isolation, which limited the interpretability of the results. The study was aimed to apply a state-of-the-art algorithm to the ranking of team performance and exploitation of key performance features in relation to match outcome based on massive match dataset. Data of all 1200 matches from 2014 to 2018 Chinese Football Super League (CSL) were used. From the original 164 match events, we extracted 22 features that were related to attacking, passing, and defending performance and most. A Linear Support Vector Classifier (LSVC) model was subsequently built with these 22 input features and trained in order to rank the teams by their performance and analyze the features that influence most match outcome (win/not win), with the dataset being divided into a ratio of 4:1 to train and validate the model. The results have shown that the data-driven LSVC model displayed a prediction accuracy of 0.83 and the ranking of teams' match performance and prediction of teams' league standings were highly correlated with their actual ranking. Saves, pass success and shot on target in penalty area were demonstrated as top positive features for winning whereas shots on target during open play, pass and bad shot% were three negative features most influential for the match result. The team ranks of all teams were highly correlated with their real final league rankings. In general, CSL winning teams build their success based on defensive ability and shooting accuracy, and high-ranked teams could always maintain better performance than their counterparts. The team-rank framework could provide a consolidated and complex approach to evaluate the match performance quality of the teams, refining decisions-making during match preparation and player transfer at different periods of the season.

1. Introduction

The performance analysis in association football games can be traced back to the 1950's when Reep and Benjamin manually notated the match events to analyze association football games [45]. Nowadays, with the development of video, data collection and computer science technology, companies like OPTA, Wyscout, TRACAB and Champion can provide detailed and diverse data-sets such as the spatio-temporal information of players [16,20,21] and the technical-tactical events [42]. Thanks to these data-sets, studies that quantify specific aspects of association football performance have sprung up [34].

In team-level, researchers most focused on network metrics to identify and evaluate the players' connection with each other [8,17], rating and ranking methods to provide objective indications of the strength of the teams [29], key performance indicators to model the relationships between match outcomes and match events [32,43] and outcome prediction which is potentially useful to players, team manager and performance analysts [5].

[*] Corresponding author at: School of Sports Engineering, Beijing Sport University, Information Road 48, Haidian District, Beijing 100084, PR China.

[**] Co-corresponding author.

E-mail addresses: cuiyixiong@bsu.edu.cn (Y. Cui), syf@bsu.edu.cn (Y. Shen).

https://doi.org/10.1016/j.chaos.2020.110330

Most of the articles of key performance indicators were using linear models [36], such as discriminant analysis [6,7,28], logistic regression [9,31], multivariate combination of principal-component and cluster analysis [38], Pearson's correlation analysis [47], and generalized mixed linear model [31,36]. These methods are very mature and have standard processes, but their variables are simple, descriptive and isolated [35]. A solution is seeing the problem in a multidimensional view by combining the different technical variables [43], which is the core variable processing method of match outcome prediction models. Among the results of these researches, shots and shots on target were considered as key attacking variables that have positive effects. For variables related to organizing, passes, passes success and crosses were mostly focused by researchers, and the results varied [30,31,36,46]. Tackles, interceptions and clearances were key defensive variables that were studied most. The results varied from different leagues and championships due to the different styles and characteristics both from between- and within-team perspectives [31].

The prediction of sports match outcome has always been deeply concerned by sports experts, fans, and stakeholders due to its unpredictable nature and the existence of sports betting. Machine learning (ML) models for match outcome prediction were first used in 1996, and these models were widely studied since then [5]. In association football area, Reed and colleagues [44] used Artificial Neural Network (ANN) to predict the result of the English Premier League and get an accuracy of 57.9% [44]. Hucalijuk and colleagues [24] predicted the outcomes of the UEFA Champions League by an ANN model with an accuracy of 68% [24]. Odachowski and colleagues [40] compared the difference between a three-class (win, draw, loss) and a binary-class (win, not win) BayesNet model for outcome prediction in various association football leagues. The accuracy was 70.3% for the binary-class model and 46% for the binary-class model [40]. A similar difference between the three-class and binary-class was found in the study by Danisik and colleagues [11]. In 2017, an open-source data-set named the *Open International Soccer Database* was made public in the 2017 Soccer Prediction Challenge [12]. Based on this data-set, researchers have built different three-class models which include XGBoost classification, Hybrid Bayssian Network and kNN etc. and the accuracy was all around 52% no matter how much features they have used [2,10,23].

Because of the difficulty to detect draws [43] and the characteristic of ML (the learning problem tend to be more difficult as the number of classes increase) [5], the above three-class models' accuracy are not ideal. However, it is worth notice that the accuracy of binary-class models is good which makes it possible to use ML methods to rate and rank players and teams. Brooks and colleagues [4] developed a data-driven player ranking model using the predictive model weight [4]. Furthermore, Pappalardo and colleagues [41] developed a player ranking system based on the weights calculated by an LSVC classifier model [41]. These studies have all calculated a property of the performance features – their weights, which can be understand as the importance coefficient of the specific variable. Although the studies on the application of ML methods are still lacking, and the only few studies were player-level, their research paradigms and methods can be applied at the team-level. In addition, the feature weights calculated by the ML models provide the possibility for the application of ML rating and ranking methods in the association football performance analysis.

It is easy to notice that most of the studies above used data-sets from top-level association football leagues or championships, while little has been seen in lower-level leagues. Recent years have seen an increasing body of research into match performance of the Chinese Football Association Super League (CSL)—one of the Asian top and world sub-elite leagues that have a large-scale development in standardization under the globalization of professional football. Although there have been some attempts to describe the technical-tactical and physical demands of CSL games [33,36,48], data-driven evaluation of team performance and league competitiveness is relatively limited. Therefore, the study was aimed to apply a state-of-the-art algorithm to the ranking of CSL teams and exploitation of key performance features in relation to match outcome based on massive match data-set.

2. Methods

2.1. Sample and data source

Chinese Football Association Super League is the highest level of professional association football match in China. There are 240 matches completed by 16 teams in each season (each team played 30 matches in the league). The end-of-season rank is determined by the final points accumulated from each match outcome (3 points for a win, 1 for a draw, 0 for a loss).

Data of all 1200 matches from 2014 to 2018 CSL where 22 teams competed were provided by Champion Technology Co. Ltd. throughout a previously validated computerized notational system Champdas Master System [18].

2.2. Feature selection

A total of 164 match events, technical-tactical performance features related to goal scoring, shooting, passing, organizing, defending and goalkeeping were extracted from the cleaned raw data. As previous research [22,43] revealed that including more features cannot guarantee better model predictions, due to the high unpredictability of association football games, it is therefore determined using the following steps to select most relevant features to match performance in order to proceed to the final training of machine learning (ML) model:

Firstly, a preliminary features selection was done by excluding features related to goals, which reduced the total number of features to 124. Goals only depict team's attacking outcome rather than serves as a performance indicator, so that features related to goals (see Supplementary Table 1 for deleted goal-related features) would produce trivial correlations and provide less insight into the impact of technical features [43].

Afterward, a one-way analysis of variance (ANOVA) was used to further filter the features based on match outcome (win, draw and loss). Those whose differences between three outcomes were not significantly identified by the analysis were screened out ($p > 0.05$).

To avoid the imbalance of absolute match statistics caused by different ball possession time and focus on technical-tactical efficiency, all feature values were adjusted to per 50% of ball possession of the own team (for attacking features) or the opposition team (for defending features) respectively before the analysis according to the following formulas [31]:

$$V_{ajstd} = \left(V_{original}/P_{team}\right) \times \ 50\%$$

$$V_{ajstd} = \left(V_{original}/P_{opposition}\right) \times \ 50\%$$

Where V is value of a feature; p_{team} and $p_{opposition}$ is possession of the own team and the opposite team.

Finally, 22 features were extracted and unified to a same scale by a standardization calculating each feature as its corresponding z-score (x'):

$$x\prime = \frac{x - \bar{x}}{\sigma}$$

Where x is the original value, $\bar{x}$ is the average value of the feature, σ is the Standard deviation of the feature. Such standardization would make the data limited to a certain range and eliminate the impact of singular samples which will increase the training time and may lead the model's failure to converge.

2.3. Classifier

To rank the teams by their performance and analyze the features that most influenced match outcome (win/not win), a performance vector $\boldsymbol{p_T^m} = (x_1, \ldots, x_n)$, $n = 22$, contains values of performance features (x_i) of team T in match m was extracted from the data-set. Combined with the outcome O_T^m (1 for win and 0 for not win) of that match, we solved a classification problem between the team performance vector $\boldsymbol{p_T^m}$ and the match outcome O_T^m.

A Linear Support Vector Classifier (LSVC) was trained to classify the outcome of a match given the teams' performance vectors. The principles are:

Given a set of instance-label pairs (x_i, y_i), with $i = 1,\ldots,l$, $x_i \in R^n$, and $y_i \in \{-1, +1\}$, an LSVC solves an unconstrained optimization problem with a loss function $\xi(w; x_i, y_i)$ [13]:

$$\min_w = \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi(w; x_i, y_i)$$

The loss function in this research is L2-SVM defined as:

$$\xi(w; x_i, y_i) = \max(1 - y_i w^T x_i, 0)^2$$

80% of the data samples are used to train the LSVC model. The cost parameters that had the maximum average Area Under the Receiver Operating Characteristic Curve (AUC) were selected with a 5-fold cross-validation. The model was validated using the remaining 20% of the data.

2.4. The team-rank framework

Fig. 1 shows how the team-rank framework operates. Starting from a data-set contains technical statistics, it consists of three main phases: (a) The performance extraction phase chooses 22 features from the data-set and extracts the performance vector $\boldsymbol{p_T^m}$ and match outcome O_T^m; (b) The learning phase solves a classification problem and learns the weight of each feature; (c) The rating and ranking phase rates the matches base on the feature weights and ranks teams by their season average rating.

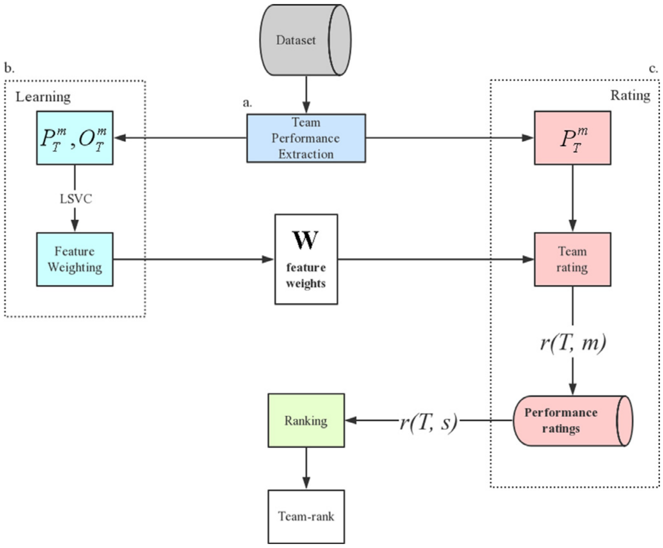


Fig. 1. Schema of the Team-rank framework.

2.4.1. Feature weighting

Different features influence match outcomes at different levels [43]. Therefore, ranking teams on their performance depend on the weight of every single feature of the match and quantifying the specific impact of those features extracted from the data-set on outcomes in CSL matches is an essential step. For each value $\boldsymbol{x}$ in the performance vector $\boldsymbol{p_T^m}$, the LSVC model computes a coefficient $\boldsymbol{w}$ which is used as the weight. Each feature weight models the importance of that feature in the evaluation of the performance quality of any team. The machine learning toolkit Scikit-learn in Python was used to train and obtain the weights.

2.4.2. Rating and ranking

The performance rating of a team T in a single match m is computed as the dot product between the values of the features referring to match m and the feature weights w computed by the LSVC model. Given the performance vector $p_T^m = (x_1, \ldots, x_n)$ and their weights w, the performance rating $r\ (T,\ m)$ of a team T in a match m can be calculated as:

$$r(T, m) = \sum_{i=1}^{n} w_i \times x_i$$

The season performance rating $r\ (T,\ s)$ is the total rating of all the matches for team T, which can be calculated as:

$$r(T, s) = \sum_{i=1}^{n} r(T, m_n)$$

Ranking the $r\ (T,\ s)$ of different teams from high to low, the performance-based team rank could be obtained.

2.5. Model validation

To validate the prediction accuracy of the LSVC model, the final performance ranks for teams in each CSL season were simulated depending on two different outputs of the LSVC model: (1) Team-rank, which is based on the team's performance rating calculated by feature weights, and it represents the overall performance of a team within a single match or a season; (2) Predicted ranking, which is the end-of-season ranking predicted by LSVC model given the actual performance of teams at each match. Since the outcome the LSVC model predicts was binary (win/not win), the score that each outcome would get was set to 3 points for winning and 0.5 points for not winning (the average score of drawing and losing). The rank is separated into 3 parts: teams qualified to the AFC Champions League (top four teams); teams who had a risk to be relegated or were relegated to the second division of the league (bottom four teams); the rest (teams ranked from 5 to 12). The Team-rank in each season and predicted rankings of all teams were tested against with their actual rankings via two metrics: (i) the Pearson's correlation coefficient measuring the relationship between teams' points in the actual ranking and each simulates ranking; (ii) the accuracy of defining the groups of teams (top four, bottom four, all the rest), computed as the ratio of teams in the two performance rankings which resulted to be in their actual ranking group.

3. Results

3.1. Descriptive analysis of selected features

After the feature selection, 22 features were selected. Table 1 shows the average values and the result of one-way ANOVA of each feature (See Supplementary Table 2 for detail results of all 124 features).

Table 1
Differences between winning, drawing and losing teams in game statistics.

Feature name	Win Mean (SD)	Draw	Loss	F	P
Shots	13.33 (4.38)	12.17 (4.22)	11.75 (4.27)	31.43	<0.001
Shots on target	5.54 (2.39)	3.98 (2.19)	3.54 (2.16)	188.95	<0.001
Shot on target in penalty area	3.98 (1.96)	2.63 (1.68)	2.22 (1.64)	233.55	<0.001
Penalty	0.24 (0.46)	0.13 (0.35)	0.12 (0.35)	22.83	<0.001
Bad shot%	0.69 (0.15)	0.78 (0.16)	0.80 (0.16)	115.24	<0.001
Pass	382.23 (61.7)	375.85 (56.80)	386.69 (56.78)	6.3	0.002
Pass success	296.86 (65.73)	286.68 (60.08)	295.28 (58.65)	5.54	0.004
Pass attacking success	63.88 (20.32)	59.96 (21.77)	59.32 (23.21)	10.96	<0.001
Pass forward success%	0.69 (0.07)	0.67 (0.07)	0.66 (0.07)	52.3	<0.001
Possession	0.51 (0.07)	0.5 (0.07)	0.49 (0.07)	5.82	0.003
Cross	14.67 (5.63)	16.69 (6.29)	16.63 (6.23)	30.22	<0.001
Cross success	4.48 (2.40)	4.88 (2.70)	4.52 (2.62)	5.14	0.006
Lost ball	24.92 (6.91)	24.43 (6.98)	25.84 (7.48)	7.78	<0.001
Tackles	17.65 (6.19)	16.46 (5.67)	16.82 (5.37)	8.88	<0.001
Saves	2.34 (1.85)	2.29 (1.88)	2.61 (1.96)	6.7	0.001
Red card	0.05 (0.21)	0.08 (0.28)	0.13 (0.35)	18.2	<0.001
Pen opponent	0.12 (0.35)	0.13 (0.35)	0.24 (0.46)	22.83	<0.001
Interceptions	20.4 (12.07)	19.88 (11.47)	18.61 (10.87)	5.6	0.004
Defensive Foul	14.39 (5.25)	13.64 (5.16)	13.43 (5.11)	8.16	<0.001
Clearances	20.62 (8.15)	19.79 (7.64)	16.99 (6.92)	54.59	<0.001
Shots opponent	11.75 (4.27)	12.17 (4.22)	13.33 (4.38)	31.43	<0.001
Shots on target opponent	3.54 (2.16)	3.98 (2.19)	5.54 (2.39)	188.95	<0.001

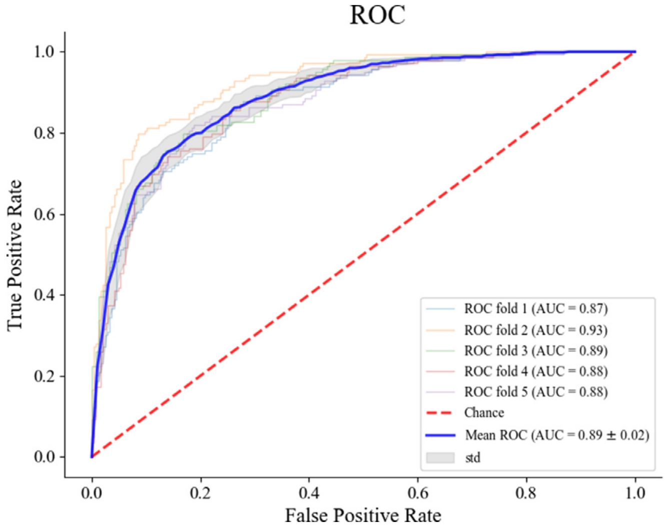


Fig. 2. Mean ROC and ROC of each validate fold Note. Mean AUC is the area under the blue curve (Mean ROC).

3.2. LSVC model

Fig. 2 shows the ROC (Receiver Operating Characteristic Curve) of each fold and the mean ROC. The statistics of the AUC, F1 and prediction accuracy were 0.90, 0.82 and 0.83 respectively after training and validating the LSVC model, which were higher than the predictive result of two baseline classifiers: (a) the classifier that chooses the label at random based on the distribution of win and not win (AUC = 0.5, F1 = 0.49, accuracy = 0.53); (b) the classifier that always predicts the most frequent match outcome not win (AUC = 0.5, F1 = 0.38, accuracy = 0.62).

3.3. Validation of the model

Table 2 shows the a between the team-rank, predicted rankings and the actual ranking of each season. A significant similarity was found between both the simulated rankings (Team-rank and Predicted ranking) and the actual ranking. The correlation between the performance rankings and the actual ranking can reach up to 0.92 in season 2017 combine with high group accuracy: 88% for all teams. On the prediction of league champion, simulated rankings are correct on season 2016, 2017 and 2018 and predicted ranking is much better with only one incorrect on season 2015. The team-rank performs perfectly on predicting the last team of the season without error and the predicted ranking has only two errors.

Table 3 presents the actual ranking, team-rank and predicted ranking from the proposed model for all CSL teams in 2017. On predicting the team groups, team-rank has a high accuracy on the AFC Champions League teams (75%) and a perfect accuracy (100%) of predicting the bottom four teams.

Although the team-rank is overall in line with the actual situation, there are still some visible errors: (a) Rating-actual error: According to the team rating, Shandong Luneng is one of the four teams who will participate in the next season's AFC Champions League, but the actual ranking shows that instead of Shandong Luneng, Tianjin Quanjian is an actual Top-4 team; (b) Rating-prediction error: Although the Team-rank and the Predicted ranking are all obtained from the LSVC model, they have different results on the performance rank. Take Shandong Luneng as an example, in Team-rank, its match performance during all season is rated as the 4th place, but in the Predicted ranking, the team is in the 8th place.

The exploration of predicted match ratings from the model showed that the cut-off value that distinguishes the match outcome (win/not win) was 0.100. To further exemplify the finding, Table 4 is built and presents the predicted results and match ratings for Shandong Luneng, Guangzhou R&F and Tianjin Quanjian.

3.4. Feature weights

Fig. 3 shows the feature weights resulting from the LSVC model. The most important positive feature is saves, and the most negative feature is opponent shots on target. Although there are significant differences between the three match outcomes in shots, lost-ball and defensive foul, their feature weights are much smaller than other features, showing only tiny effects on performance rating.

Fig. 4 shows the normalized match performance and match ratings for Beijing Guoan, which is one of a middle-ranked team (end-of-season ranking: 9) in season 2017 and played at home and away against teams of different strengths. An empirical eval-

Table 2
Group accuracy and similarity between simulated rankings & Actual ranking.

	TEAM-RANK							PREDICTED RANKING						
Year	r	AFC	Rest	Bottom	All	Champion	Last	r	AFC	Rest	Bottom	All	Champion	Last
2014	0.85	0.75	0.63	0.50	0.63	×	√	0.78	0.75	0.63	0.50	0.63	×	×
2015	0.88	1	0.75	0.50	0.75	×	√	0.74	0.75	0.63	0.5	0.63	×	×
2016	0.91	0.75	0.75	0.75	0.75	√	√	0.92	0.75	0.75	0.75	0.75	√	√
2017	0.92	0.75	0.88	1	0.88	√	√	0.90	0.75	0.75	0.75	0.75	√	√
2018	0.92	0.75	0.75	0.75	0.75	√	√	0.87	0.75	0.75	0.75	0.75	√	√

Note. r: Pearson's r between Simulated ranking and Actual ranking; AFC: teams qualified for the AFC Asian cup (ranked between 1 and 4); Rest: teams ranked between 5 and 12; bottom: teams ranked between 13 and 16; all: All 16 teams during the season; Champion & Last: whether the simulated ranking predicts the league champion and the last place of the league correctly, where √ stands for yes and × for no.

Table 3
Actual ranking, Team-rank, and Predicted ranking of CSL 2017.

Actual ranking		**Team-rank**		**Predicted ranking**	
GZFC	64	GZFC	0.145	GZFC	60
Shanghai SIPG	58	Shanghai SIPG	0.145	Shanghai SIPG	60
Tianjin Quanjian	54	Hebei CFFC	0.074	Hebei CFFC	47.5
Hebei CFFC	52	Shandong Luneng	0.043	Guangzhou R&F FC	45
Guangzhou R&F FC	52	Guangzhou R&F FC	0.041	Changchun Yatai	45
Shandong Luneng	49	Changchun Yatai	0.040	Beijing Guoan	42.5
Changchun Yatai	44	Beijing Guoan	0.017	Tianjin Quanjian	40
Guizhou Hengfeng	42	Tianjin Quanjian	0.001	Shandong Luneng	37.5
Beijing Guoan	40	Chongqing Lifan	−0.019	Shanghai Shenhua	35
Chongqing Lifan	36	Jiangsu Suning FC	−0.020	Guizhou Hengfeng	35
Shanghai Shenhua	35	Shanghai Shenhua	−0.038	Tianjin Teda	35
Jiangsu Suning FC	32	Guizhou Hengfeng	−0.039	Chongqing Lifan	32.5
Tianjin Teda	31	Henan Jianye	−0.066	Jiangsu Suning FC	30
Henan Jianye	30	Yanbian	−0.068	Henan Jianye	30
Yanbian	22	Tianjin Teda	−0.094	Yanbian	30
Liaoning FC	18	Liaoning FC	−0.163	Liaoning FC	22.5
		$r = 0.93$	$r = 0.87$		

Notes. Actual ranking is the real ranking after the season (3 points for winning, 1 for drawing, 0 for losing); Team-rank is the performance ranking based on the performance rating calculated by the team-rank framework; Predicted ranking is the rank based on the predicted results of the LSVC model (3 points for winning, 0.5 points for not winning).

Table 4
Differences in match ratings of Shandong Luneng, Guangzhou R&F FC and Tianjin Quanjian.

Team	**Predicted outcome**	**Average rating (SD)**	**Predicted number**	**Actual number**
Shandong Luneng	Not win	−0.05 (0.20)	21	17
	Win	0.22 (0.19)	9	13
	All	0.04 (0.20)		
Guangzhou R&F FC	Not win	−0.09 (0.21)	18	15
	Win	0.24 (0.21)	12	15
	All	0.04 (0.21)		
Tianjin Quanjian	Not win	−0.10 (0.19)	20	15
	Win	0.20 (0.18)	10	15
	All	0.001 (0.19)		

uation shows that features positive to match success were overall higher when playing at home than away. Nonetheless, the performance of the team greatly varied when against top ranked team (Guangzhou Evergrande, end-of-season ranking: 1), middle-ranked teams (Guizhou Hengfeng: end-of-season ranking: 8) and low-ranked team (Liaoning FC, end-of season ranking: 16)

4. Discussion

The purpose of this study was to apply a state-of-the-art framework to the ranking of CSL teams and exploitation of key performance features in relation to match outcome (win/not win) based on massive match data-set. The results have shown that the data-driven LSVC model displayed a prediction accuracy of 0.83 and the ranking of teams' match performance and prediction of teams' league standings were highly correlated with their actual ranking. Saves, pass success and shot on target in penalty area were demonstrated as top positive features for winning whereas shots on target during open play, pass and bad shot% were three negative features most influential for the match result.

4.1. Performance modeling and team rank framework

Previously, the most commonly used methods in modeling football performance are linear [36]. These methods are very mature and have standard processes, but the variables are simply modeled in isolation [35]. While association football is a multifaceted and complex sport, the performance variables are influenced by the interactions between different technical and tactical outcomes. Therefore, accessing association football performance in a mono-dimensional way might not reveal the non-linear relationship between performance and game outcome, not to mention ranking team performance. ELO is an algorithm that is widely accepted in many fields and it is a standard method to rank players and

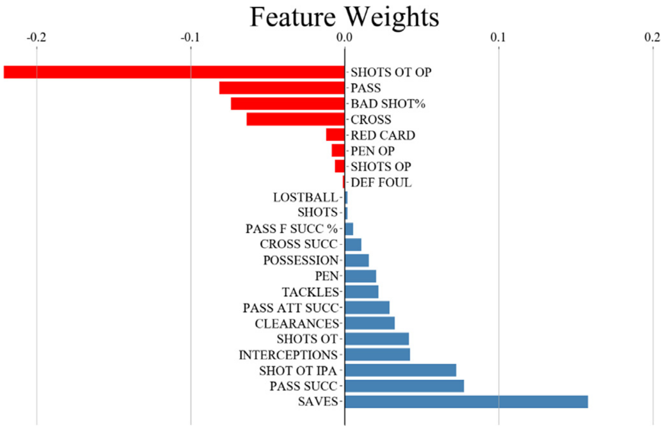


Fig. 3. Weights of each feature (OT = on target, IPA = in penalty area, *F* = Forward, OPA = out penalty area, OP = opponent, PEN = penalty, DEF = defensive).

teams based on their recent performance [25,39]. However, ELO is merely a measurement of the teams' strength based on their recent results, unable to account for their actual match performance. In comparison, the current team-rank model takes into consideration the interaction of different performance in a multidimensional view by extracting performance vectors composed by performance features. Furthermore, the team-rank framework is a rating and ranking framework based on the theory that technical performances can explain part of a team's success [43], which can be utilized in the works of performance analysis.

Although being a machine learning (ML) approach, it should be noticed that our framework is differentiated from traditional ML techniques [5] by the following aspects. Firstly, because of the need to predict the outcome of football games, which have three outcomes (win, draw, loss), most prediction ML models were formulated as a three-class classifier. But along with the increasing of classes, the accuracy tends to reduce. The team-rank model aims to rate and rank the teams by their performance, a high accuracy is needed to simulate and evaluate a team's performance. Therefore, the team-rank model was formulated as a binary-class model (win/not win). Secondly, the ML prediction models aim to predict the match outcomes in the future, so features like player strength and home advantages were included in most of the models. On the contrary, the team-rank model is used in the games that are already finished, it aims to evaluate the teams' absolute performance, regardless of any features outside the game itself.

4.2. Feature importance

For the features related to shots and shots quality, shots on target in the penalty area and shots on target are much more important than shots according to Fig. 3. Furthermore, a bad shot rate has a great negative impact on the match outcome. The results corroborate the previous finding that shots on target essentially affects the probability of winning in CSL [36], which implied that shots accuracy and quality are key performance features in CSL games as in other top leagues [26,37]. Moreover, it is shown that shots on target and passes has strong weights in positive and negative features respectively. Zhou and colleagues also found it in their research about key performance indicators in CSL games which indicates that CSL teams tend to gain a success in a more direct way [48].

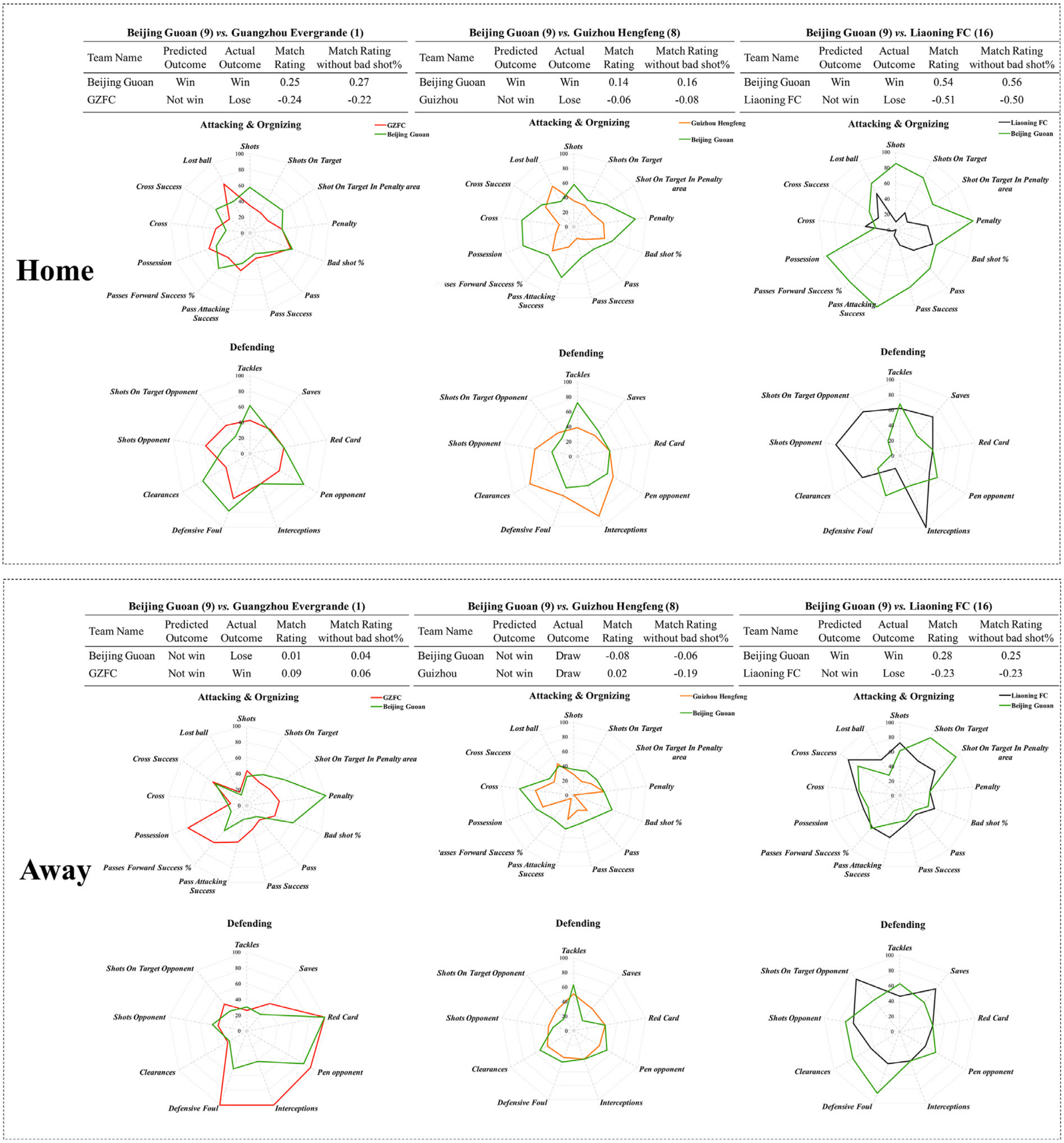
Concerning organization performance, passes success, attacking passes success and ball possession showed positive effects on winning and the passes were shown to be a negative feature. Previous researches indicated that keeping the ball moving continuously and aggressively could lead to a higher percentage of ball possession and more scoring opportunities, which were key performance indicators for successful teams in European leagues and CSL [3,36]. It is worth noting that the weight of pass success is higher than pass attacking success, which may indicate that CSL teams tend to adopt a relatively conservative and stable approach when building up their offense. But according to the research on the 2018 FIFA World Cup [46], better teams tend to make more passes and deliveries into the attacking third regardless their playing style (possession-based or direct-play), implying that world's top teams shall have the ability to make more successful aggressive passes instead of making conservative passes to maintain meaningless ball control. In addition, cross is determined as a negative feature by the LSVC model, which is the same as the results of several previous researches [31,32,36,48]. A proper explanation is that weak teams are less developed and worse prepared in organizing their offense [31], and hence it is probable that low-ranked teams in CSL lack the skills to send the ball into a dangerous position via structured offensive passing or efficient counterattacks. Furthermore, we still noticed that, unlike the total cross number, the LSVC model accounted cross success a positive feature. This may indicate unlike other top association football leagues in Europe where crossing is a forced tactic for most teams [30], it can still be a feasible attacking tactic in CSL top teams.

In terms of defense, saves, interceptions, clearances and tackles are three features that had positive effects on the match outcome while opponent shots on target, red card, shots opponent, and defensive fouls are negative features. Previous researches showed different ideas on tackles, [31] suggested that successful and appropriate tackles could increase the chance of winning while [36] indicated that tackles had trivial effects in the Chinese Football Association Super League. The feature weights showed that tackles had the smallest weight among the three positive defensive features, but it was still a relatively important positive feature. Previous findings [1,31] showed that red card had negative effects because of being send off by a red card is a weakening for a team's strength in terms of goal scoring and match outcome, which corroborates the result of our research. As a negative feature, the weights of opponent shots on target were significantly bigger than opponent shots which are in line with the results of shot related features that it is the quality rather than the number of shots that determines the match outcomes. This indicated that restricting the opponent's shot quality is much more important than decreasing the quantity.

4.3. Simulated rankings

In light of the Rating-actual error, a possible reason would be that Tianjin Quanjian performed worse than Shandong Luneng if we concern solely the values of input features, but the uncertainty of the game a and the inherently unpredictable nature of this sport could determine that they won more games than the latter and there might be other tactical performance features deserved to be considered in the future study [5]. However, it could still be inferred that good performance does not always guarantee the winning of the match, but high-ranked teams could always maintain better performance than their counterparts.

For the Rating-prediction error, the main tasks are: (a) understanding what two simulated rankings represent respectively; (b) analysing why there are differences between the results of two performance rankings. The team-rank is a rank of the CSL teams based on the their performance ratings calculated from the feature weights in the LSVC model, and it represents the specific performance of each team in a single match and in one season. The predicted ranking is based on the game outcome predicted by the



Beijing Guoan (9) *vs.* Guangzhou Evergrande (1)

Team Name	Predicted Outcome	Actual Outcome	Match Rating	Match Rating without bad shot%
Beijing Guoan	Win	Win	0.25	0.27
GZFC	Not win	Lose	-0.24	-0.22

Beijing Guoan (9) *vs.* Guizhou Hengfeng (8)

Team Name	Predicted Outcome	Actual Outcome	Match Rating	Match Rating without bad shot%
Beijing Guoan	Win	Win	0.14	0.16
Guizhou	Not win	Lose	-0.06	-0.08

Beijing Guoan (9) *vs.* Liaoning FC (16)

Team Name	Predicted Outcome	Actual Outcome	Match Rating	Match Rating without bad shot%
Beijing Guoan	Win	Win	0.54	0.56
Liaoning FC	Not win	Lose	-0.51	-0.50

Beijing Guoan (9) *vs.* Guangzhou Evergrande (1)

Team Name	Predicted Outcome	Actual Outcome	Match Rating	Match Rating without bad shot%
Beijing Guoan	Not win	Lose	0.01	0.04
GZFC	Not win	Win	0.09	0.06

Beijing Guoan (9) *vs.* Guizhou Hengfeng (8)

Team Name	Predicted Outcome	Actual Outcome	Match Rating	Match Rating without bad shot%
Beijing Guoan	Not win	Draw	-0.08	-0.06
Guizhou	Not win	Draw	0.02	-0.19

Beijing Guoan (9) *vs.* Liaoning FC (16)

Team Name	Predicted Outcome	Actual Outcome	Match Rating	Match Rating without bad shot%
Beijing Guoan	Win	Win	0.28	0.25
Liaoning FC	Not win	Lose	-0.23	-0.23

Fig. 4. Home and away match performance of Beijing Guoan against opponents of different rankings during 2017 CSL season.

LSVC model, representing the match outcome team's performance would lead to according to.

For example, as it was shown in Table 4, although Shandong Luneng had higher average ratings, Guangzhou R&F FC was predicted to have more wins. In other words, Guangzhou R&F had more matches whose ratings were above the cut-off value (0.100). In addition to these two teams, other teams with the same phenomenon have a similar result in match ratings.

The Rating-prediction error implies that high ranking does not always mean better performance but high-ranked teams could maintain at a performance level that permits higher chance of winning than their rivals. These two errors verified and supplement the previous study that performance can explain a team's success to some extent but it is not absolute, other factors that are either not captured by the technical-tactical data or outside the football game can influence the judgment.

4.4. Individual match performance

The comparison of different matches played by Beijing Guoan verified that the team-rank model is able to detect some common impacts caused by contextual factors in CSL such as home advantage and quality of opposition [33]. During home game, Guoan achieved higher forward pass success rate and ball possession, which leads to more scoring opportunities [27]. Although it had a comparatively higher bad shot rate at both occasions, the same feature for its opponent raised when playing at Guoan's home stadium, implying that teams perform worse on attacking during CSL away game [33].

In addition, the quality of opposition was shown to greatly shape Guoan's match performance and playing style [14,19]. It's worth noticing that Beijing Guoan's bad shot rate was much higher than superior and similar opponent (GZFC and Guizhou Hengfeng), but not the case when facing low-ranked counterpart. This indicates that excluding the impact of low shooting ability might allow Beijing Guoan to perform similar as top-level opponent and superior to the middle-ranked opponent, while causing little influence on their victory when against an inferior team like Liaoning FC. Moreover, the ball possession of Beijing Guoan was higher than low-ranked team, suggestive of the fact that weaker teams are more likely to be forced to play a defensive style, and maintaining the formation closer to the own goal-line, resulting in a lower ball possession [14,15].

4.5. Practical application

The current framework would provide CSL club managers and coaches a consolidated and complex approach to evaluate the match performance quality of the teams, and compare it with their previous performance and that of their rivals, thus helping technical staffs make better decisions in addition to the ratings of players [42] and during match preparation and player transfer at different periods of the season. Moreover, the findings should be contrasted against other leagues, championships or cups of different levels to reveal the influence of different playing styles on technical-tactical performance values. Finally, physical performance of the teams could be added to the machine learning model to allow comprehensive exploration of key performance features.

4.6. Limitation and feature works

This research only considered the technical-tactical data, other commonly used data types in football analysis like spatio-temporal data had not been used, which is a sequence of samples containing the time-stamp and location of some phenomena which include object trajectories of player and ball movement, and *event* logs that record the location and time of match events [20]. Apart from it, patterns of interaction between players detected by passing network data and physical performance data which contains a player's or a team's running speed, acceleration and distance may also help to further explain team's performance, and finally generate a more diverse and comprehensive rating model [2] Moreover, it should be emphasized that domain knowledge of football need to be meaningfully incorporated in the modeling, rather than just in the result interpretation stage. In fact, contextual factors such as home advantage, weather influence, congested match schedule and previous results are ought to be considered as important features within feature engineering or selection phases. It is possible that integrating them into the current framework would further improve the understanding of how changing contexts condition the importance of performance features and the accuracy of prediction.

5. Conclusion

This work analysed 1200 games from 2014 to 2018 Chinese Football Association Super League and applied a state-of-the-art framework to the ranking of CSL teams and exploitation of key performance features in relation to match outcome based on massive match data-set. The framework solved a classification problem between different game outcomes (win and not win) by a Linear Support Vector Classifier (LSVC) and calculated a weight for each performance feature. The weights showed that shots on target, passes success, saves, interceptions, clearances and tackles are important positive features and opponent shots on target, passes, bad shot rate, crosses and red card are features which have great negative impact. A team rank which expressed the teams' performance was built based on the weights. The errors between simulate rankings and actual ranking are strong evidence that in CSL games, better performance does not mean a winning and high ranking does not always mean a better performance but better teams could maintain a performance that have bigger chance to win than their opponents. Furthermore, the possibility of using Machine Learning methods on the analyzing of association football performance was proved by comparing the feature weights with domain knowledges and former research.

Funding

This work was supported in part by the National Key Research and Development Program of China under grants 2020AAA0103404 and 2018YFC2000600. The corresponding author (Y. Cui) was supported by the China Postdoctoral Science Foundation (2020T130067) and the Fundamental Research Funds for the Central Universities (2019QD033).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.chaos.2020.110330.

CRediT authorship contribution statement

Yuesen Li: Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing. **Runqing Ma:** Investigation, Writing - original draft. **Bruno Gonçalves:** Writing - review & editing. **Bingnan Gong:** Resources, Data curation. **Yixiong Cui:** Supervision, Conceptualization, Writing - original draft, Methodology, Visualization, Funding acquisition. **Yanfei Shen:** Supervision, Methodology, Project administration, Funding acquisition.

References

[1] Bar-Eli M, Tenenbaum G, Geister S. Consequences of players' dismissal in professional soccer: a crisis-related analysis of group-size effects. J Sports Sci 2006;24(10):1083–94. doi:10.1080/02640410500432599.

[2] Berrar D, Lopes P, Dubitzky W. Incorporating domain knowledge in machine learning for soccer outcome prediction. Mach Learn 2019;108(1):97–126. doi:10.1007/s10994-018-5747-8.

[3] Bradley PS, Lago-Peñas C, Rey E, Sampaio J. The influence of situational variables on ball possession in the English Premier League. J Sports Sci 2014;32(20):1867–73. doi:10.1080/02640414.2014.887850.

[4] Brooks J, Kerr M, Guttag J. Developing a data-driven player ranking in soccer using predictive model weights. In: Paper presented at the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016. p. 49–55.

[5] Bunker R, Susnjak T. The application of machine learning techniques for predicting results in team sport: a review. 2019; arXiv preprint arXiv:1912.11762. 10.13140/RG.2.2.22427.62245.

[6] Lago-Peñas C, Lago-Ballesteros J, Rey E. Differences in performance indicators between winning and losing teams in the UEFA champions league. J Hum Kinet 2011;27(2011):135–46.

[7] Castellano J, Casamichana D, Lago C. The Use of match statistics that discriminate between successful and unsuccessful soccer teams. J Hum Kinet 2012;31(1):137–47. doi:10.2478/v10078-012-0015-7.

[8] Clemente FM, Couceiro MS, Martins FML, Mendes RS. Using network metrics in soccer: a macro-analysis. J Hum Kinet 2015;45(1):123–34. doi:10.1515/hukin-2015-0013.

[9] Collet C. The possession game? a comparative analysis of ball retention and team success in European and International Football, 2007-2010. J Sports Sci 2013;31(2):123–36. doi:10.1080/02640414.2012.727455.

[10] Constantinou AC. Dolores: a model that predicts football match outcomes from all over the world. Mach Learn 2019;108(1):49–75. doi:10.1007/s10994-018-5703-7.

[11] Danisik N, Lacko P, Farkas M. Football match prediction using players attributes. In: Paper presented at the IEEE 2018 World Symposium on Digital Intelligence for Systems and Machines (DISA); 2018. p. 201–6.

[12] Dubitzky W, Lopes P, Davis J, Berrar D. The open international soccer database for machine learning. Mach Learn 2019;108(1):9–28. doi:10.1007/s10994-018-5726-0.

[13] Fan RE, Chang KW, Hsieh CJ, Wang XR, Lin CJ. LIBLINEAR: a library for large linear classification. J Mach Learn Res 2008;9(9):1871–4.

[14] Fernandez-Navarro J, Fradua L, Zubillaga A, Mcrobert AP. Influence of contextual variables on styles of play in soccer. Int J Perform Anal Sport 2018;18(3):423–36. doi:10.1080/24748668.2018.1479925.

[15] Gollan S, Bellenger C, Norton K. Contextual factors impact styles of play in the English Premier League. J Sports Sci Med 2020;19(1):78–83.

[16] Gonçalves B, Coutinho D, Exel J, Travassos B, Lago C, Sampaio J. Extracting spatial-temporal features that describe a team match demands when considering the effects of the quality of opposition in elite football. PLoS ONE 2019;14(8):e221368. doi:10.1371/journal.pone.0221368.

[17] Gonçalves B, Coutinho D, Santos S, Lago-Penas C, Jiménez S, Sampaio J. Exploring team passing networks and player movement dynamics in youth association football. PLoS ONE 2017;12(1):e171156. doi:10.1371/journal.pone.0171156.

[18] Gong B, Cui Y, Gai Y, Yi Q, Gomez MA. The validity and reliability of live football match statistics from champdas master match analysis system. Front Psychol 2019;10:1339. doi:10.3389/fpsyg.2019.01339.

[19] Gómez M, Mitrotasios M, Armatas V, Lago-Peñas C. Analysis of playing styles according to team quality and match location in Greek Professional Soccer. Int J Perform Anal Sport 2018;18(6):986–97. doi:10.1080/24748668.2018.1539382.

[20] Gudmundsson J, Horton M. Spatio-temporal analysis of team sports. ACM Comput Surv 2017;50(2):1–34. doi:10.1145/3054132.

[21] Gudmundsson J, Wolle T. Football analysis using spatio-temporal tools. Comput Environ Urban Syst 2014;47:16–27. doi:10.1016/j.compenvurbsys.2013.09.004.

[22] Heuer A, Rubner O. Optimizing the prediction process: from statistical concepts to the case study of soccer. PLoS ONE 2014;9(9):e104647. doi:10.1371/journal.pone.0104647.

[23] Hubáček O, Aourek G, Železný F. Learning to predict soccer results from relational data with gradient boosted trees. Mach Learn 2019;108(1):29–47. doi:10.1007/s10994-018-5704-6.

[24] Hucaljuk J, Rakipović A. Predicting football scores using machine learning techniques. In: Paper presented at the 2011 Proceedings of the 34th International Convention MIPRO; 2011. p. 1623–7.

[25] Hvattum LM, Arntzen H. Using ELO ratings for match result prediction in association football. Int J Forecast 2010;26(3):460–70. doi:10.1016/j.ijforecast.2009.10.002.

[26] Konefal M, Chmura P, Zajac T, Chmura J, Kowalczuk E, Andrzejewski M. Evolution of technical activity in various playing positions, in relation to match outcomes in professional soccer. Biol Sport 2019;36(2):181–9. doi:10.5114/biolsport.2019.83958.

[27] Lago C. The Influence of match location, quality of opposition, and match status on possession strategies in professional association football. J Sports Sci 2009;27(13):1463–9. doi:10.1080/02640410903131681.

[28] Lago-Peñas C, Lago-Ballesteros J, Dellal A, Gómez M. Game-related statistics that discriminated winning, drawing and losing teams from the Spanish Soccer League. J Sports Sci Med 2010;9(2):288–93.

[29] Lasek J, Szlavik Z, Bhulai S. The predictive power of ranking systems in association football. Int J Appl Pattern Recognit 2013;1(1):27–46.

[30] Lepschy H, Wäsche H, Woll A. Success factors in football: an analysis of the German Bundesliga. Int J Perform Anal Sport 2020;20(2):150–64. doi:10.1080/24748668.2020.1726157.

[31] Liu H, Gomez M, Lago-Peñas C, Sampaio J. Match statistics related to winning in the group stage of 2014 Brazil FIFA World Cup. J Sports Sci 2015;33(12):1205–13. doi:10.1080/02640414.2015.1022578.

[32] Liu H, Hopkins WG, Gómez M. Modelling relationships between match events and match outcome in elite football. Eur J Sport Sci 2015;16(5):516–25. doi:10.1080/17461391.2015.1042527.

[33] Liu T, García-De-Alcaraz A, Zhang L, Zhang Y. Exploring home advantage and quality of opposition interactions in the Chinese Football Super League. Int J Perform Anal Sport 2019;19(3):289–301. doi:10.1080/24748668.2019.1600907.

[34] Low B, Coutinho D, Gonçalves B, Rein R, Memmert D, Sampaio J. A systematic review of collective tactical behaviours in football using positional data. Sports Med 2020;50(2):343–85. doi:10.1007/s40279-019-01194-7.

[35] Mackenzie R, Cushion C. Performance analysis in football: a critical review and implications for future research. J Sports Sci 2013;31(6):639–76. doi:10.1080/02640414.2012.746720.

[36] Mao L, Peng Z, Liu H, Gómez M. Identifying keys to win in the Chinese Professional Soccer League. Int J Perform Anal Sport 2016;16(3):935–47. doi:10.1080/24748668.2016.11868940.

[37] Mitrotasios M, González Rodenas J, Armatas V, Aranda R. The creation of goal scoring opportunities in professional soccer. tactical differences between Spanish La Liga, English Premier League, German Bundesliga and Italian Serie A. Int J Perform Anal Sport 2019;19(3):452–65. doi:10.1080/24748668.2019.1618568.

[38] Moura FA, Martins LEB, Cunha SA. Analysis of football game-related statistics using multivariate techniques. J Sports Sci 2014;32(20):1881–7. doi:10.1080/02640414.2013.853130.

[39] Neumann C, Duboscq J, Dubuc C, Ginting A, Irwan AM, Agil M, Engelhardt A. Assessing dominance hierarchies: validation and advantages of progressive evaluation with elo-rating. Anim Behav 2011;82(4):911–21. doi:10.1016/j.anbehav.2011.07.016.

[40] Odachowski K, Grekow J. Using bookmaker odds to predict the final result of football matches. In: *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*. Berlin Heidelberg: Springer; 2012. p. 196–205.

[41] Pappalardo L, Cintia P, Ferragina P, Massucco E, Pedreschi D, Giannotti F. PlayeRank. ACM Trans Intell Syst Technol 2019;10(5):1–27. doi:10.1145/3343172.

[42] Pappalardo L, Cintia P, Rossi A, Massucco E, Ferragina P, Pedreschi D, et al. A public data set of spatio-temporal match events in soccer competitions. Sci Data 2019;6(1):1–15. doi:10.1038/s41597-019-0247-7.

[43] Pappalardo L, Cintia P. Quantifying the relation between performance and success in soccer. Adv Complex Syst 2018;21(03n04):1750014. doi:10.1142/S021952591750014X.

[44] Reed D, O'Donoghue P. Development and application of computer-based prediction methods. Int J Perform Anal Sport 2017;5(3):12–28. doi:10.1080/24748668.2005.11868334.

[45] Reep C, Benjamin B. Skill and chance in association football. J R Stat Soc Ser A 1968;131(4):581. doi:10.2307/2343726.

[46] Yi Q, Gómez MA, Wang L, Huang G, Zhang H, Liu H. Technical and physical match performance of teams in the 2018 FIFA World Cup: effects of two different playing styles. J Sports Sci 2019;37(22):2569–77. doi:10.1080/02640414.2019.1648120.

[47] Yue Z, Broich H, Mester J. Statistical analysis for the soccer matches of the first bundesliga. Int J Sports Sci Coach 2014;9(3):553–60. doi:10.1260/1747-9541.9.3.553.

[48] Zhou C, Zhang S, Lorenzo Calvo A, Cui Y. Chinese Soccer Association Super League, 2012–2017: key performance indicators in balance games. Int J Perform Anal Sport 2018;18(4):645–56. doi:10.1080/24748668.2018.1509254.