Forecasting binary outcomes in soccer

Raffaele Mattera¹

Accepted: 26 July 2021
© The Author(s) 2021

Abstract

Several studies deal with the development of advanced statistical methods for predicting football match results. These predictions are then used to construct profitable betting strategies. Even if the most popular bets are based on whether one expects that a team will win, lose, or draw in the next game, nowadays a variety of other outcomes are available for betting purposes. While some of these events are binary in nature (e.g. the red cards occurrence), others can be seen as binary outcomes. In this paper we propose a simple framework, based on score-driven models, able to obtain accurate forecasts for binary outcomes in soccer matches. To show the usefulness of the proposed statistical approach, two experiments to the English Premier League and to the Italian Serie A are provided for predicting red cards occurrence, Under/Over and Goal/No Goal events.

Keywords Sport statistics · Sport analytics · Binary time series · Generalized autoregressive score · Bernoulli distribution

1 Introduction

The development of advanced statistical methods to forecast teams and players' performances (e.g. Adhikari et al. 2020; Galariotis et al. 2018; Yang et al. 2014) or match results (e.g. Diniz et al. 2019; Friesl et al. 2020; Koopman and Lit 2015; Nikolaidis 2015) has recently attracted the interest of many researchers in the field of statistics and econometrics. Since betting on sport events is based on a guess about the future match outcome, the development of statistical models able to improve decision making is of interest to both bettors and odd setters.

In the context of football, the most popular bet is based on whether one expects that a team will win, lose, or draw the next game. Accordingly, many different methodologies have been developed with the aim of constructing profitable betting strategies on the basis of statistical forecasts for the final match outcome.

The result of a football match depends by the number of goals scored and conceded by both the involved teams. Since the number of goals scored (as well as the conceded ones) follows a Poisson distribution, the mainstream statistical approach for forecasting the match

✉ Raffaele Mattera
raffaele.mattera@unina.it

¹ Department of Economics and Statistics, University of Naples "Federico II, Naples, Italy

result is mainly based on Poisson regressions (e.g. Angelini and De Angelis 2017; Koopman and Lit 2015; Maher 1982).

Indeed, Maher (1982) proposed to model the goals scored by two teams by means of two independent Poisson distributions. Therefore, the final match outcome can be predicted by means of the forecasts on the total goals scored by the two opponent teams. The idea of using the model of Maher (1982) for the development of a profitable betting strategy is due to Dixon and Coles (1997) and Dixon and Pope (2004).

Other contributors, as for example Goddard and Asimakopoulous (2004), Goddard (2005), Forrest and Simmons (2000) and Forrest et al. (2005), have adopted a methodology based on the ordered probit or logit regressions to directly model either the outcomes or the probability of future match results.

However, nowadays the sport betting market offers a variety of alternative bets than the classical “win/draw/lose”. For example, it is possible to bet on the event that will be scored more (or less) than a certain amount of goals in the next match, without guessing on the final result. This bet is called over (under). Further, a bettor can bet on the event that both teams will (or not) score at least a goal within the game (Goal/No Goal). Lastly, it is also possible to bet on the occurrence of a red card within the football match.

Exactly as in the case of the “classical bet”, a bettor can be surely interested also in forecasting the probability that these alternative events will occur. Nevertheless, despite their potential relevance, forecasting models for these bets have been poorly studied with some very recent exception (e.g. Wheatcroft 2020).

Moreover, while some of these events (e.g. in the case of Under/Over or of Goal/No Goal) can be predicted by means of forecasts on the number of goals that will be scored by each team, other events are binary outcomes by nature (i.e. occurrence of a red card or of a penalty within the match). Nevertheless, also the aforementioned not binary events as the Goal/No Goal and Under/Over can actually be seen as binary outcomes from the bettor point of view.

Indeed, in these two examples the bettor has only two possible choices for betting: either the event “Goal” will occur or the “No Goal” one. The same also applies to the Under/Over bet.

Therefore, for each team we can construct a binary variable that takes value 1 if in a given match both teams scored at least one goal and 0 otherwise in the case of Goal/No Goal bet. Similarly, we can build a binary variable that takes value of 1 if in a given match at least three goals have been scored and 0 otherwise, as in the case of the Under/Over 2.5 bet.

By considering them as binary bets, in this paper we provide a simple framework to obtain accurate forecasts by implementing appropriate statistical models for binary time series.

While from one side we have the classical *generalized autoregressive moving average* (GARMA) models (Benjamin et al. 2003), that are generalized linear models for modeling binary time series data, more recently (Creal et al. 2013) developed a new dynamic model called *generalized autoregressive score* (GAS), that has already been applied successfully also in the context of football matches forecasting (e.g. Koopman and Lit 2015, 2019). The score-driven models are observation-driven, which means that the likelihood is available in closed form. This allows for a fast estimation process despite the challenges involved with the use of high-dimensional models.

The dynamic characteristic of the class of score-driven models lies on the fact that the time-varying coefficients are updated by means of an autoregressive component and by the score of the conditional observation probability density function.

By assuming that the binary outcomes of football matches are Bernoulli trials, in this paper we show that the GAS model can be successfully applied also for predicting the probability that a binary event will occur in the next match.

To this aim, we provide experimental evidences from two football championships: the English Premier League and the Italian Serie A. For both national leagues, we collect the last 13 seasons of matches and we use the last two to assess the predictive ability of the statistical models under consideration.

Once probability forecasts are obtained, we build a profitable betting strategy that can be summarized as follows: if the probability of a given event is higher (lower) than the average probability for both the teams, there is a suggestion of betting on the outcome 1 (outcome 0).

For example, suppose that we aim to forecast the occurrence of a red card in the next match. If the probability of observing a red card is higher (lower) than the average probability of observing a red card for both teams, the bettor should bet on the event “red card: yes” (“red card: no”). In the case of uncertainty, any bet should be placed. This simple betting strategy allows the bettor to avoid betting on too uncertain matches. This can considerably improve the returns of a betting strategy (e.g. Angelini and De Angelis 2017). Indeed the fact that not always a betting advice is provided allows the bettor to save money and to increase his/her net profit.

We find that, for the event “red card: yes or not”, the proposed betting strategy with GAS probability forecast has an accuracy, computed as the proportion of the correct bets over the total, equal to 100% for the English data and 94% for the Italian ones.

In the case of Goal/No Goal events, instead, the accuracy is on average equal to 93% in both experiments, while for some football teams it is close or equal to 100%.

Lastly, in the case of Under/Over we consider different thresholds. In particular bets on the events Under/Over 1.5, Under/Over 2.5 and Under/Over 3.5 are considered. The difference lies on the considered threshold about the amount of goals scored in the match (one in the first case, two in the second and three in the third one). The betting accuracy is around 90% on average for the English Premier League for all the considered thresholds. In the case of Italian Serie A, instead, the betting accuracy is always higher than 80% (94% in the case of Under/Over 1.5) and is always much better than the considered alternatives.

Therefore, we would recommend the proposed strategy for betting purposes.

The rest of the paper is as follows. In the Sect. 2 the statistical methods for modeling binary time series are presented as well as more details about the proposed score-driven approach based on Bernoulli conditional density. Then, the Sect. 3 is devoted to the data used for the empirical experiments with English Premier League and Italian Serie A data and the betting strategy is also discussed more in detail. Section 4 presents the results of the empirical experiments. At the end are some final remarks are discussed.

2 Forecasting binary time series

As previously discussed, some of the available bets for soccer match are binary (the red cards or penalties occurrence are clear examples) while others can be seen and treated as binary. In general, in many applications to finance or macroeconomics is common to construct binary variables from observed continuous ones (e.g. Christoffersen and Diebold 2006; Harding and Pagan 2011; Hausman et al. 1992; Nyberg 2010; Startz 2008).

Since the team's performances depend on the results realized in the recent matches, assuming the independence of matches played by the same team over time is not reasonable¹. The existence of this persistency effect implies that an autocorrelation structure for event's prob-

¹ Match independence can be a reasonable assumption only for the matches played by different teams within a given match-week, hence by considering a cross-sectional structure of the dataset.

abilities for the same football team exists, thus justifying the usage of a time series approach in forecasting football match outcomes (Angelini and De Angelis 2017).

For modeling binary time series, several authors proposed particular extensions of Generalized Linear Models (e.g. Davis et al. 2003; Benjamin et al. 2003; Moysiadis and Fokianos 2014). One of the most general approach is the so called *generalized autoregressive moving average* (GARMA) model of Benjamin et al. (2003) that can be used to model a variety of time-dependent response variables (e.g. data with Bernoulli, Poisson, negative binomial, or binomial distributions) as well as continuous data. The GARMA model is a direct GLM extension of the usual autoregressive moving average (ARMA).

In the GARMA model we assume that the conditional distribution of each observation y_t , with $t = 1, \dots, T$, belongs to the same exponential family given the previous information set $\mathcal{F}_t$:

$$f(y_t | \mathcal{F}_t) = \exp \{y_t \vartheta_t - b(\vartheta_t) + c_t\}, \quad (1)$$

where the function $b(\cdot)$ defines the specific member of the exponential family of interest and c_t is a sequence of constants. The terms $E(y_t | \mathcal{F}_t)$ and $\text{var}(y_t | \mathcal{F}_t)$ represent the conditional mean and variance of y_t given $\mathcal{F}_t$. We can express a general model for $\eta_t = g(\mu_t)$, without exogenous covariates, as follow (Benjamin et al. 2003):

$$\eta_t = g(\mu_t) = \sum_{j=1}^p \phi_j g(y_{t-j}) + \sum_{j=1}^q \theta_j \{g(y_{t-j}) - g(\mu_{t-j})\} \quad (2)$$

where $g(\cdot)$ is the *link* function while the ϕ_j 's and θ_j 's are the equivalent of the autoregressive and moving average parameters in the standard ARMA. In this case, the parameters $\phi = (\phi_1, \phi_2, \dots, \phi_p)$ and $\theta = (\theta_1, \theta_2, \dots, \theta_q)$ can be estimated by maximum likelihood. For certain functions $g(\cdot)$ it may be necessary to replace y_{t-j} with a transformation y_{t-j}^* in (2) in order to avoid the nonexistence of $g(y_{t-j})$ for certain values of y_{t-j} . Clearly, the kind of transformation y_{t-j}^* , if needed, depends on the particular function $g(\cdot)$.

In the case of binary time series we have that 0 and 1 are the only two possible values that y_t can take. Clearly, the value 1 indicates that some event occurs, and 0 that it does not. A common special case of GARMA for binary time series is the so called *Binary Autoregressive Moving Average* (BARMA) extensively studied by Li (1994) and Wang and Li (2011) that is commonly implemented for macroeconomic applications (e.g. Startz 2008; Kauppi and Saikkonen 2008), that can be written as:

$$\eta_t = g(\mu_t) = \phi_1 g(y_{t-1}) + \theta_1 \{g(y_{t-1}) - g(\mu_{t-1})\} \quad (3)$$

where the function $g(\cdot)$ is the logit to ensure that μ_t , the conditional probability of a success at time point t , takes only values between 0 and 1. As a special case of GARMA, also the parameters ϕ_1 and θ_1 of the BARMA model can be estimated with maximum likelihood but it is common to estimate them also by means of quasi-maximum likelihood (Zeger and Qaqish 1988).

By assuming that the binary outcomes of football matches are Bernoulli trials, an alternative and promising possibility for predicting the conditional probability of success is the *generalized autoregressive score* (GAS) model of Creal et al. (2013), that has been proved to be useful in predicting football matches (Koopman and Lit 2019).

The GAS model can be formalized as follows. Let us consider the teams' matches outcomes over the time ($y_t : t = 1, \dots, T$) with f_t the time-varying parameter vector observed at time t . Let also assume that the available information set at time t , called $\mathcal{F}_t$, it is given by

a collection of some previous realizations of the parameters f_t and the outcomes y_t . Then, define θ a vector of static parameters.

The starting point of the GAS model is the observation density of y_t :

$$y_t \sim p(y_t | f_t, \mathcal{F}_t; \theta), \quad (4)$$

Then, given two integers $0 \leq p, q \leq T - 1$, we can formalize the $GAS(p, q)$ model for the t -th realization f_t of the time-varying parameter vector as:

$$f_t = \omega + \sum_{i=1}^p \mathbf{A}_i s_{t-i} + \sum_{j=1}^q \mathbf{B}_j f_{t-j} \quad (5)$$

where ω is a real vector and the $\mathbf{A}$'s and the $\mathbf{B}$'s are real matrices with an appropriate dimension – all the scalar parameters in $\omega, \mathbf{A}_1, \dots, \mathbf{A}_p, \mathbf{B}_1, \dots, \mathbf{B}_q$ are collected in the vector θ introduced in (4); s_t is a type of score of the conditional distribution defined in (4), and it is a function of the data and the parameters, so that:

$$s_t = S_t \cdot \nabla_t \quad (6)$$

where $S_t = S_t(f_t, \mathcal{F}_t; \theta)$ is a positive definite scaling matrix known at time t and $\nabla_t(y_t, f_t, \mathcal{F}_t; \theta)$ is the score of y_t evaluated with respect to f_t , i.e.:

$$\nabla_t = \frac{\partial \log p(y_t | f_t, \mathcal{F}_t; \theta)}{\partial f_t} \quad (7)$$

Since the score depends on the complete density and not only on some moments of y_t , the $GAS(p, q)$ model uses the full density structure for updating f_t . Different choices about the scaling matrix S_t lead to different model specification. Creal et al. (2013) suggested to consider a scale that depends on the variance of the score. Particularly, the authors proposed to scale a matrix S_t to a power $\gamma > 0$ of the inverse of the information matrix of f_t :

$$S_t = \mathcal{I}^{-\gamma} \quad (8)$$

with $\mathcal{I}$:

$$\mathcal{I} = E [\nabla_t \nabla_t'] \quad (9)$$

with ∇_t defined as (7). The parameter γ is fixed by the user and usually takes value in the set $\{0, \frac{1}{2}, 1\}$. When γ is set equal to zero and S_t the identity matrix so $S_t = \mathbf{I}$ there is no scaling. Instead, if $\gamma = 1$ the conditional score ∇_t is premultiplied by the inverse of its covariance matrix $\mathcal{I}^{-1}(f_t)$ otherwise, with $\gamma = \frac{1}{2}$, to its square-root. Therefore, the $GAS(1, 1)$ that can be written as:

$$f_t = \omega + \mathbf{A}_1 s_{t-1} + \mathbf{B}_1 f_{t-1} \quad (10)$$

It is also possible to write the Eq. (10) in the case of $f_t = (f_{1,t}, \dots, f_{K,t})$ by means of:

$$\omega = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_K \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_K \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} b_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & b_K \end{pmatrix}$$

In the case of football matches, the binary outcomes come from a single trial since a match is played just once within the same match-week. Therefore we can assume that these binary observations y_t follow a Bernoulli distribution (Lu 2020):

$$p(y_t | \pi_t) = \pi_t^{y_t} + (1 - \pi_t)^{1-y_t} \quad (11)$$

where $y_t \in \{0, 1\}$ and π_t is the time varying probability. Within the $GAS(1, 1)$ model, following Mesters and Koopman (2012) and Lit and Koopman (2020), we can specify the unobserved time-varying probability π_t as follows:

$$\begin{aligned}\eta_t &= g(\pi_t), \\ \pi_t &= \omega + a\pi_{t-1} + bs_{t-1}\end{aligned}\quad (12)$$

where the link function $g(\cdot)$ is the logit link, such that the values of π_t take value between 0 and 1. The driving force behind the updating equation is, exactly defined as in the Eq. (10), the scaled score innovation s_t where π_t is the only time varying parameter f_t . In this case, the information set consists of lagged variables of π_t . The unknown coefficients, i.e. the constant ω , the autoregressive parameter a and the score updating parameter b can be estimated by maximum likelihood.

3 Data and betting strategy

The empirical experiment has the aim of predicting the match outcomes in the next round of both the considered football competitions, the English Premier League and the Italian Serie A. The considered binary outcomes are the following events: the presence of a red card, if both teams score at least one goal (Goal/No Goal) and if there will be scored not more than or at least a certain amount of goals (Under/Over).

The decision of betting if an event will occur or not is based on some probability forecasts, obtained by a means of the statistical models shown in the previous section. The accuracy of the probability predictions obtained by means of alternative statistical models are compared in terms of the average proportion of correct bets over the total, that we define *betting accuracy*.

For each competition, the total data set consist of the last 13 seasons of football match results. We have partitioned the data set into the in-sample seasons from the 2008–2009 to the 2018–2019, that are used for the parameter estimation step, and the last two seasons 2019–2021 as the out-of-sample ones, that are used to assess the forecasting performance. In doing so we use a recursive forecasting scheme (Koopman and Lit 2019).

After each season, the teams ranked in the last 3 positions are then relegated and new 3 teams are instead promoted into the higher competition (from Championship to Premier League in England and from Serie B to Serie A in Italy). Therefore, to keep the number of teams constant, we consider the teams that participated in the last season. The data used in our empirical study can be found at <http://www.football-data.co.uk>.

The Table 1 shows how dataset for the English Premier League is built.

While one variable (red) is binary, the Goal/No Goal and the Under/Over variables have been transformed in binary events. Indeed, we construct a dummy variable “both teams score” (Goal/No Goal), that takes value of 1 if both teams score at least one goal in the match and 0 otherwise. Similarly, we construct the binary variable “Over: yes” if the sum of the goals scored by both the teams was at least equal to the threshold. This variable assumes value of 0 otherwise. The descriptive statistics for the Premier League are below reported in the Table 2.

Descriptive statistics show that, on average, in the English Premier League the home teams score 1.5 goals while the away ones 1.2. Hence, in a certain measure the home teams are more willing to win a match (probably due to the so called *home bias* effect Goossens et al. 2012; Csató 2020). Moreover, in almost the 50% of the matches both team scores and there are not less than 3 total goals. In the 75% there are at least more than one goal, while only

Table 1 Example of the dataset for the English Premier League

	Home team	Away team	Result		GG
1	Arsenal	West Brom	1	0	0
2	Bolton	Stoke	3	1	1
3	Everton	Blackburn	2	3	1
4	Hull	Fulham	2	1	1
5	Middlesbrough	Tottenham	2	1	1
6	Sunderland	Liverpool	0	1	0
7	West Ham	Wigan	2	1	1
8	Aston Villa	Man City	4	2	1
9	Chelsea	Portsmouth	4	0	0
10	Man United	Newcastle	1	1	1
⋮	⋮	⋮	⋮	⋮	⋮
4930	West Brom	West Ham	1	3	1
4931	Arsenal	Brighton	2	0	0
4932	Aston Villa	Chelsea	2	1	1
4933	Fulham	Newcastle	0	2	0
4934	Leeds	West Brom	3	1	1
4935	Leicester	Tottenham	2	4	1
4936	Liverpool	Crystal Palace	2	0	0
4937	Man City	Everton	5	0	0
4938	Sheffield United	Burnley	1	0	0
4939	West Ham	Southampton	3	0	0
4940	Wolves	Man United	1	2	1

Table 2 English Premier League: descriptive statistics

Statistic	Matches	Mean	SD
Home Goal	4940	1.536	1.310
Away Goal	4940	1.187	1.170
Goal/No Goal	4940	0.508	0.500
Red card: yes/no	4940	0.007	0.085
Under/Over 1.5	4940	0.755	0.430
Under/Over 2.5	4940	0.522	0.500
Under/Over 3.5	4940	0.295	0.456

in the 30% of the matches more than three goals are scored, In the end, red cards occurred only in the 0.7% of the matches within the sample.

Similarly to the Premier League, we also collected the last 13 seasons of the Italian Serie A. The dataset looks like the one in Table 1. The descriptive statistics of the Italian Serie A matches are shown in the Table 3.

The goals statistics are very similar to those of the English Premier League, such that also for the Italian Serie A the home teams are also more willing to win a match. However, Table 3 highlights that the occurrence of a red card is more frequent in the Italian national league than in the English one. The other statistics are instead very similar.

Table 3 Italian Serie A: descriptive statistics

Statistic	Matches	Mean	SD
Home Goal	4940	1.526	1.252
Away Goal	4940	1.190	1.128
Goal/No Goal	4940	0.533	0.499
Red card: yes/no	4940	0.024	0.154
Under/Over 1.5	4940	0.755	0.430
Under/Over 2.5	4940	0.510	0.500
Under/Over 3.5	4940	0.288	0.453

In order to produce betting recommendation about the results of all the matches of the current season for the two considered national league competitions, we take advantage of the statistical model presented in the Sect. 2 of the paper.

More in details, by means of equations (3) and (12), we are able to forecast the probability that a certain binary event will realize (event “yes”) in the next match or not (event “no”). We call this probability for the i -th team $\hat{\pi}_{i,t+1}$.

Before to compute the forecasts for the $t + 1$ match probability, all the parameters in (3) and (12) are estimated using all data up to time t . Therefore, the first forecasts of all matches at time $t + 1$ (i.e. the first match of the season 2019-2020) are based on the parameter estimates from the data of the previous seasons 2009-2019.

In order to decide on what event to bet, we consider the following simple betting rule:

$$\begin{aligned}
 &\text{if } \hat{\pi}_{i,t+1} \geq \bar{\pi}_i \text{ and } \hat{\pi}_{j,t+1} \geq \bar{\pi}_j \quad \hat{y}_{t+1} = 1 \\
 &\text{if } \hat{\pi}_{i,t+1} \leq \bar{\pi}_i \text{ and } \hat{\pi}_{j,t+1} \leq \bar{\pi}_j \quad \hat{y}_{t+1} = 0 \\
 &\text{if } \hat{\pi}_{i,t+1} \geq \bar{\pi}_i \text{ and } \hat{\pi}_{j,t+1} < \bar{\pi}_j \quad \text{no bet} \\
 &\text{if } \hat{\pi}_{i,t+1} \leq \bar{\pi}_i \text{ and } \hat{\pi}_{j,t+1} > \bar{\pi}_j \quad \text{no bet}
 \end{aligned} \tag{13}$$

In other words, if the probability of an event is greater (lower) than the average probability $\bar{\pi}$ for both the teams, the bettor has to guess on the event “yes” (event “no”). Otherwise, in the case of uncertainty the bet should be avoided.

Betting avoidance on uncertain events is a useful feature of the proposed strategy. Indeed the fact that not always a betting advice is provided allows the bettor to save money and to increase his/her net profit. Indeed, for evaluating the net profit of a betting strategy one should consider the differences between the amount of money earned and the expenses faced by the bettor. Clearly, betting avoidance dramatically reduces bettor’s expenses.

Following the score-driven model literature, the average (unconditional) probability $\bar{\pi}$ for the team i can be computed by means of:

$$\bar{\pi}_i = \frac{\hat{\omega}}{1 - \hat{b}} \tag{14}$$

where $\hat{\omega}$ and $\hat{b}$ are the maximum likelihood parameters estimated from the equation (12) for the i -th football team. The same apply for the j -th team.

For the next round of football matches $t + 2$ and its probability forecasts $\hat{\pi}_{i,t+2}$, we re-estimate both the parameter vector and the average probability after including the football match results of the most recent round in our data set. Hence we have an expanding estimation sample, i.e. we follow a recursive forecasting approach, to ensure that we can utilize as much data we can for estimation phase (Koopman and Lit 2019).

As a benchmark for (13), based on the models (3) and (12) forecasts, we consider a naive betting strategy where the bettor suppose that, looking at the favorite team², the outcome realized at time t will realizes also at time $t + 1$, so no forecasts are used. In other words the naive betting strategy can be written as:

$$\begin{aligned} \text{if } y_{i,t} = 1 \quad \hat{y}_{t+1} &= 1 \\ \text{if } y_{i,t} = 0 \quad \hat{y}_{t+1} &= 0 \end{aligned} \quad (15)$$

The naive strategy (15) is also necessary to assess if statistical models are really useful in predicting binary outcomes. If the naive strategy has a lower accuracy than (13), that is based on statistical forecasts, then we can argue that the models discussed in the Sect. 2 should be used to this aim.

Once a bet indication is obtained by means of either (13) or (15), we can evaluate a loss function to measure the betting accuracy. The betting accuracy is computed ex-post as the proportion of the correct bets over the total:

$$ACC_{i,k} = \frac{\text{Correct Bets}_{i,k}}{\text{Tot. Bets}_{i,k}} \quad (16)$$

for the k -th bet (i.e. red card, Goal/No Goal or Under/Over) for the matches played by the i -th team, given that we make a correct bet if:

$$\{\hat{y}_{t+1} = 1 \text{ and } y_{t+1} = 1\} \text{ or } \{\hat{y}_{t+1} = 0 \text{ and } y_{t+1} = 0\} \quad (17)$$

where $\hat{y}_{t+1}$ is the predicted outcome for the next match $t + 1$ according to (13) and y_{t+1} is the realized one. In the end, we have to highlight that the football matches without any bet recommendation are excluded from the computation of (16).

4 Forecasting accuracy

In this section the performances of the betting strategy are discussed. Since the matches of the two last seasons are used for out-of-sample evaluation, we have a total of $N = 1560$ matches for the out-of-sample assessment for both the national leagues.

4.1 Red card

First of all, we aim to predict the binary outcome “red card: yes/no”. Since the probability of a red card is generally low, most of bookmakers pay a few amount of money for the “red card: yes” bet³. In other words, guessing an extremely rare event generally result in a low payout.

The Table 4 shows the betting accuracy of the proposed strategy for the English Premier League matches. The probability predictions needed to implement the strategy (13) are obtained, as explained in the Sect. 2, by means of the BARMA and the GAS models based on the Bernoulli distribution. Those models are then compared with also the naive approach in (15).

Clearly, the out-of-sample performance is positive, with a very high betting accuracy close to a 100% of average correct bets, for both the statistical models. However, the score-driven

² For sake of simplicity, as in Angelini and De Angelis (2017), we assume that the favorite team of the bettor it is the home one.

³ see the <https://www.betexplorer.com/> website.

approach shows a bit higher accuracy rate, equal to 100%, than the BARMA whose accuracy is equal to 99,8%. The naive approach performs well in this case but it is the worst one among the considered alternatives.

We also study the betting accuracy of the proposed strategy based on probability forecasts with the Italian Serie A data (see Table 5).

In this second empirical experiment we still observe a very high betting accuracy level, much greater than 90%. More in detail, the strategy based on the GAS forecasts reach an accuracy equal to 94%, while the one based on BARMA forecasts ensures an accuracy equal to 97%. Naive strategy is around 95% of accuracy.

In general both the approaches are very effective in predicting football matches outcomes in the case of red cards, even if in the case of Italian Serie A the benefit seems to be lower.

Moreover, the statistical models show variable accuracy regarding the different football teams. In the case of the GAS model, for example, while for some teams the betting accuracy is really high (e.g. Sassuolo, Napoli and Udinese with 100% of correct bets), for some others we get results lower than 90% of accuracy (e.g. Genoa with 84% and Parma with 83%). Indeed, the overperformance of the naive and BARMA-based forecasts is mainly due to the lower variability within the sample of teams.

For example, in the case of Genoa football team, the GAS-based forecasts allow an accuracy equal to 88% (94% in the average), while the BARMA forecasts reach an accuracy greater than 95%. Therefore, it is reasonable to assume that combining both forecasting approaches in this case it would be possible to dramatically increase the betting accuracy.

A final comment has to be devoted to the number of teams not included in the samples in both Tables 4 and 5. The main reasons for the Teams' exclusion is due to the lack of observations to generate the probability forecasts. For example, in the case of Benevento and Spezia for the Serie A or Leeds and Sheffield United for the Premier League, have been excluded since they participated only to the current season or in only few occasions.

Overall, the very high accuracy level of the proposed statistical models ensures an enormous advantage from the point of view of a bettor.

4.2 Goal and no goal

Let now consider the model's predictive ability with respect to the event "both teams score at least a goal", also called "Goal/No Goal". At the best of our knowledge, as well as the red card outcome, the "Goal/No Goal" is another poorly explored bet by the sport forecasting literature.

As we mentioned previously, summing the number of goals does not returns a binary variable. However, bets on this event involve a binary choice: if a bettor believe that both teams will score a goal, whatever the final result will be, she bets on "Goal" otherwise she bets on "No Goal".

Hence, we construct a binary variable with the aim of forecasting the probability of the "yes" event defined as $\hat{\pi}_{i,t+1}$.

The betting accuracy for the English Premier League is shown in Table 6. Differently from the red card outcome, in the case of Goal/No Goal the superiority of the score-driven approach is evident. Indeed, the average betting accuracy is equal to 92% while in the case of BARMA forecasts the accuracy is less than 50%. The naive approach performs even poorer with 47% of accuracy. Therefore, in this case the usage of statistical models, especially the score-driven one, is very useful.

Table 4 Red card betting accuracy—English Premier League

	Accuracy GAS (%)	Accuracy GARMA (%)	Accuracy naive (%)	Obs.
Arsenal	100.00	100.00	100.00	494
Aston Villa	100.00	98.33	97.37	380
Brighton			100.00	152
Burnley	100.00	100.00	100.00	266
Chelsea	100.00	100.00	100.00	494
Crystal Palace	100.00	100.00	100.00	304
Everton	100.00	100.00	100.00	494
Fulham	100.00	100.00	100.00	304
Leeds			100.00	38
Leicester	100.00	100.00	100.00	266
Liverpool	100.00	100.00	100.00	494
Man City	100.00	98.33	97.37	494
Man United	100.00	100.00	100.00	494
Newcastle	100.00	100.00	100.00	418
Sheffield United			100.00	76
Southampton	100.00	100.00	100.00	342
Tottenham	100.00	100.00	100.00	494
West Brom	100.00	100.00	100.00	380
West Ham	100.00	100.00	100.00	456
Wolves	100.00	100.00	100.00	228
Average	100.00	99.80	99.71	353

The accuracy is computed as the proportion of correct bets over the total according to the rules (13) and (15). Observations is the number of played matches. The out-of-sample size is given by the matches played in the last two seasons. Blank rows are reported for teams without any bet recommendation

In the case of score-driven forecasts, for some teams the percentage of accurate bets is even equal to 100% (e.g. Arsenal, Manchester United and the West Ham) but for all the teams it is greater than 80% (the only exception is given by the West Brom that has an accuracy of 72%).

On the other side, the results for the Italian Serie A are shown in the Table 7. In this case the superiority of the score-driven approach is even more evident with respect the experiment with Premier League data. Indeed, the average of correct bets is greater than 93%, while for the strategy based on BARMA forecasts it is equal to 40%. The naive strategy in this case performs better than BARMA-based forecasts since the average accuracy is equal to 52%.

For the Italian data, the betting accuracy is always greater than 80% for the GAS-based forecasts while is always lower than 50% for the BARMA ones. Moreover, for some teams the BARMA approach shows also an accuracy level lower than 40% (e.g. Bologna and Lazio with 33% or Sassuolo with 27%) with the best result equal to 50% (in the case of Genoa football team). On the side of the GAS-based forecasts, instead, we also observe many teams with an accuracy levels equal to 100% (e.g. Fiorentina, Inter and Napoli).

In the end, as in the previous experiment, the matches with any betting recommendation have been excluded from the computation of the betting accuracy. In the case of Premier League on the Leeds has been excluded, while in the case of Serie A only the Spezia. Clearly, this result is due to the short time series available.

Table 5 Red card betting accuracy—Italian serie A

	Accuracy GAS (%)	Accuracy GARMA (%)	Accuracy naive (%)	Obs.
Atalanta	100.00	98.31	94.74	456
Benevento			97.37	76
Bologna	96.00	96.61	94.74	456
Cagliari			94.74	456
Crotone	100.00	100.00	100.00	114
Fiorentina	97.96	98.31	97.37	494
Genoa	84.85	94.92	92.11	494
Inter	92.50	93.22	88.16	494
Juventus	93.02	96.61	94.74	494
Lazio	97.87	96.61	94.74	494
Milan	88.10	96.61	94.74	494
Napoli	100.00	100.00	100.00	494
Parma	82.93	98.31	97.37	342
Roma			97.37	494
Sampdoria	90.70	98.31	97.37	456
Sassuolo	100.00	98.31	97.37	304
Spezia			94.74	38
Torino	87.50	100.00	100.00	380
Udinese	100.00	100.00	97.37	494
Verona	95.65	96.61	92.11	228
Average	94.19	97.67	95.83	388

The accuracy is computed as the proportion of correct bets over the total according to the rules (13) and (15). Observations is the number of played matches. The out-of-sample size is given by the matches played in the last two seasons. Blank rows are reported for teams without any bet recommendation

4.3 Under and over

The last considered experiment is with the Under/Over bet for both national leagues. Differently from the previous two cases, the forecasts on Under/Over event has only recently been explored by Wheatcroft (2020) but with a completely different approach.

As in the case of “Goal/No Goal”, the “Under/Over” is not a binary variable. However, we construct three different binary variables that takes value of 1 if there have been scored more than a certain threshold of goals in a match (Over) and 0 otherwise (Under). The selected thresholds of goals are respectively equal to one (Under/Over 1.5), two (Under/Over 2.5) and three (Under/Over 3.5).

Indeed, nowadays the agencies provide the possibility of betting on these thresholds. Clearly, we expect that the predictive power of the proposed strategy for the event Under/Over will be the same regardless a given threshold. For sake of completeness, we report the results for all the three mentioned Under/Over thresholds.

The percentage of correct bets for the English Premier League data is shown in Table 8. The score-driven approach ensures stable level of average accuracy close to 90% regardless the threshold. The BARMA-based forecasts, instead, are competitive only with the scenario of Under/Over 1.5, showing an accuracy equal to 80%. Nevertheless, also in this best case its performances are much lower than the GAS-based forecasts. This result is confirmed also for

Table 6 Goal/No Goal betting accuracy—English Premier League

	Accuracy GAS (%)	Accuracy GARMA (%)	Accuracy naive (%)	Obs.
Arsenal	100.00	47.06	48.68	494
Aston Villa	86.67	45.59	51.32	380
Brighton	98.08	44.12	43.42	152
Burnley	86.67	57.35	46.05	266
Chelsea	86.36	51.47	40.79	494
Crystal Palace	93.48	54.41	43.42	304
Everton	94.44	46.27	46.05	494
Fulham	81.08	53.73	40.79	304
Leeds			40.79	38
Leicester	98.18	48.53	53.95	266
Liverpool	86.11	43.28	60.53	494
Man City	93.88	55.88	48.68	494
Man United	100.00	44.12	51.32	494
Newcastle	96.97	52.24	47.37	418
Sheffield United	97.62	67.92	46.05	76
Southampton	88.64	46.38	40.79	342
Tottenham	100.00	39.71	47.37	494
West Brom	72.22	50.77	47.37	380
West Ham	100.00	45.59	46.05	456
Wolves	95.92	48.53	56.58	228
Average	92.44	49.63	47.81	353

The accuracy is computed as the proportion of correct bets over the total according to the rules (13) and (15). Observations is the number of played matches. The out-of-sample size is given by the matches played in the last two seasons. Blank rows are reported for teams without any bet recommendation

Table 7 Goal/No Goal betting accuracy—Italian Serie A

	Accuracy GAS (%)	Accuracy GARMA (%)	Accuracy naive (%)	Obs.
Atalanta	94.55	35.29	48.68	456
Benevento	100.00	48.98	50.00	76
Bologna	83.67	33.82	52.63	456
Cagliari	88.10	44.12	56.58	456
Crotone	96.00	58.33	61.84	114
Fiorentina	100.00	40.30	57.89	494
Genoa	86.54	40.30	52.63	494
Inter	100.00	37.88	50.00	494
Juventus	97.83	33.82	56.58	494
Lazio	72.00	42.65	46.05	494
Milan	100.00	44.12	59.21	494
Napoli	100.00	42.65	44.74	494
Parma	98.11	41.79	48.68	342
Roma	90.20	38.81	47.37	494
Sampdoria	89.80	40.30	50.00	456
Sassuolo	98.21	27.94	69.74	304
Spezia			44.74	38
Torino	83.72	37.31	52.63	380
Udinese	100.00	49.25	36.84	494
Verona	89.58	37.31	52.63	228
Average	93.07	40.79	52.49	388

The accuracy is computed as the proportion of correct bets over the total according to the rules (13) and (15). Observations is the number of played matches. The out-of-sample size is given by the matches played in the last two seasons. Blank rows are reported for teams without any bet recommendation

the other two bets. Indeed, in the case of Under/Over 2.5 the average accuracy of the BARMA model dramatically reduces to 42%, versus the 89% of the GAS. In the case of Under/Over 3.5 the BARMA has an accuracy equal to 73% versus the 91% of the GAS model. The naive strategy is, instead, competitive only with Under/Over 1.5 event, with an average accuracy of 70% but still remains the worst strategy among the considered alternatives. Therefore, also in this case statistical forecast and the proposed betting strategy are useful. In the other two scenarios, the naive strategy still perform poorly.

More in detail, on the side of GAS forecasts, there are some teams whose percentage is equal to 100% (e.g. Leicester) and much others with an accuracy level higher than 90% (e.g. Sheffield and Tottenham with 97%) but there are some exceptions (e.g. Nwecastle and Wolves with 77%) that pushes down the average accuracy for the English Premier League in the case of Under/Over 1.5. In the case of Under/Over 2.5, instead, the number of teams with maximum accuracy of 100% is even higher (e.g. Everton, Leicester and Tottenham) but some outliers as the West Brom (with a poor 60%) pushes down the average accuracy. In the case of Under/Over 3.5 the percentage of accuracy is higher for most of the considered teams.

On the side of the BARMA-based predictions, instead, the level of accuracy is rarely more than 80% in the case of Under/Over 1.5, while it is dramatically low for all the teams in

the case of Under/Over 2.5. Similarly for the case Under/Over 3.5. In general, the BARMA model is not competitive at all with respect to the GAS one.

The results for the Italian Serie A are shown in the Table 9. Also with the Italian data the superiority of the score-driven approach appears evident. Indeed, in the case of Under/Over 1.5 the GAS shows an average accuracy of 94% versus the 81% of the BARMA and the 70% of the naive strategy. For the Under/Over 2.5 bet, instead, the accuracy of the GAS is extremely high with respect to the alternatives, since it is almost double (86% versus 42% and 53%). The BARMA performs poorer than the naive strategy in this case. In the end, for the Under/Over 3.5 the BARMA accuracy rises and becomes higher than the naive strategy (65% versus 54%), but both of the considered alternatives perform poorer than the GAS model (82%).

Once again statistical models discussed in the Sect. 2 are able to predict binary outcomes and should be used at this aim.

Also in this case a great heterogeneity across football teams can be highlighted. Considering the Under/Over 1.5 bet, the strategy (13) with GAS-based forecasts reaches the 100% of the accuracy for some teams (e.g. Crotone and Sassuolo) and, despite one exception (i.e. Torino with 84%), all the teams have an accuracy higher than 90%. For the BARMA, instead, no team reaches the 90% of the accuracy, that is lower than 80% in the case of the naive betting strategy that does not employ forecasting at all.

The Under/Over 2.5 highlights some outliers in terms of accuracy also for the GAS-based forecasts (e.g. Sampdoria with 68%) but the accuracy is still much higher than the alternatives. The accuracy for most of the teams is, in the case of BARMA forecasts, lower than 50% with some teams (e.g. Juventus) with a value of 33%. Indeed, as we have said, the naive strategy performs better than the BARMA. Therefore, excluding the GAS, one would consider the employment of statistical techniques useless by looking at these results.

In the end, the Under/Over 3.5 case is peculiar in terms of heterogeneity. The GAS model is still the best one with much higher average accuracy, but the heterogeneity is greater than the other two scenarios. For example, there is a huge outlier (i.e. Lazio) with an accuracy equal to 47% that reduces the average around to 82%. The other two strategies instead, do not present such outliers and the percentage of correctly guessed bets is similar across the teams.

As in the previous experiments, Leeds and Spezia are the only teams without any bet recommendation. This fact is due to the length of the time series associated to these two football teams.

In general, the overperformance of the betting strategy based on GAS forecasts is evident for both experiments also in the case of Under/Over event.

5 Conclusions

Forecasting football match results is important since such predictions can be used to construct profitable betting strategies. Even if the most popular bets are based on whether one expects that a team will win, lose, or draw in the next game, nowadays it is possible to bet on a variety of other outcomes. While some of these events are binary in nature, others can be seen as binary outcomes. In this paper we aim to provide a simple framework to obtain accurate forecasts for binary outcomes in soccer matches.

In statistics, a common approach for modeling and forecasting binary time series can be found in the *Binary Autoregressive Moving Average* (BARMA) a special case of the class

Table 8 Under/Over betting accuracy—English Premier League

	Accuracy (%)			Accuracy (%)			Accuracy (%)			Obs.
	Under/Over 1.5			Under/Over 2.5			Under/Over 3.5			
	GAS	BARMA	Naive	GAS	BARMA	Naive	GAS	BARMA	Naive	
Arsenal	88.64	78.46	57.89	96.15	45.59	43.42	81.08	73.53	53.95	494
Aston Villa	95.65	79.37	68.42	85.00	44.12	55.26	95.74	75.00	59.21	380
Brighton	94.00	75.81	63.16	98.08	52.94	48.68	93.88	80.88	69.74	152
Burnley	89.13	78.72	67.11	96.08	54.41	53.95	97.56	79.41	67.11	266
Chelsea	94.00	76.92	59.21	86.67	47.06	46.05	98.08	70.59	55.26	494
Crystal Palace	98.15	84.48	65.79	87.80	57.35	43.42	93.88	75.00	59.21	304
Everton	89.36	75.41	60.53	100.00	55.22	44.74	96.08	73.13	63.16	494
Fulham	90.00	79.69	65.79	74.29	56.72	59.21	70.27	77.61	59.21	304
Leeds			68.42			43.42			57.89	38
Leicester	100.00	89.09	82.89	100.00	38.81	51.32	96.08	73.53	53.95	266
Liverpool	91.49	83.87	78.95	85.42	40.30	52.63	91.89	67.16	55.26	494
Man City	92.00	90.48	73.68	84.09	38.24	65.79	66.67	58.82	55.26	494
Man United	95.65	78.69	65.79	90.70	51.47	51.32	92.68	69.12	57.89	494
Newcastle	77.78	77.78	65.79	78.38	53.73	38.16	93.88	71.64	51.32	418
Sheffield United	97.56	77.78	65.79	97.14	62.26	55.26	96.77	86.79	57.89	76
Southampton	92.31	83.08	73.68	91.11	48.53	46.05	86.05	71.01	47.37	342
Tottenham	97.87	84.38	76.32	100.00	41.18	52.63	95.65	66.18	57.89	494
West Brom	81.08	81.48	68.42	60.00	50.77	57.89	94.87	70.77	56.58	380
West Ham	88.00	80.00	68.42	93.88	45.59	50.00	93.18	67.65	51.32	456
Wolves	78.05	73.44	57.89	100.00	61.19	53.95	95.45	79.41	68.42	228
Average	91.09	80.47	67.76	89.73	49.76	50.80	91.04	73.01	57.89	353

The accuracy is computed as the proportion of correct bets over the total according to the rules (13) and (15). Observations is the number of played matches. The out-of-sample size is given by the matches played in the last two seasons. Blank rows are reported for teams without any bet recommendation

Table 9 Under/Over betting accuracy—Italian Serie A

	Accuracy (%)			Accuracy (%)			Accuracy (%)			Obs.
	Under/Over 1.5			Under/Over 2.5			Under/Over 3.5			
	GAS	BARMA	Naive	GAS	BARMA	Naive	GAS	BARMA	Naive	
Atalanta	91.67	89.86	78.95	82.22	36.76	56.58	73.68	49.28	48.68	456
Benevento	97.44	81.13	73.68	96.67	47.73	51.32	89.29	62.26	57.89	76
Bologna	92.98	82.61	71.05	68.75	45.59	52.63	76.32	65.22	44.74	456
Cagliari	97.06	73.53	64.47	85.71	55.88	51.32	80.00	68.12	57.89	456
Crotone	100.00	81.43	72.37	94.59	38.24	56.58	85.71	61.43	48.68	114
Fiorentina	90.91	78.26	60.53	93.55	49.25	50.00	77.78	75.36	60.53	494
Genoa	94.83	81.16	68.42	85.42	34.33	52.63	85.00	66.67	47.37	494
Inter	96.23	85.51	73.68	98.08	36.76	60.53	93.18	65.22	56.58	494
Juventus	93.22	89.86	84.21	93.88	33.82	59.21	80.39	62.32	61.84	494
Lazio	89.09	84.06	73.68	75.00	39.71	51.32	47.06	69.57	48.68	494
Milan	92.73	86.76	73.68	100.00	40.30	53.95	81.40	69.57	48.68	494
Napoli	93.88	86.96	76.32	66.67	47.76	46.05	86.21	68.12	56.58	494
Parma	94.23	81.16	68.42	93.88	35.29	52.63	88.00	59.42	50.00	342
Roma	97.92	86.57	73.68	83.33	35.82	50.00	77.78	56.52	43.42	494
Sampdoria	90.00	81.16	71.05	68.00	39.71	56.58	77.50	71.01	60.53	456
Sassuolo	100.00	84.06	73.68	84.78	35.38	56.58	92.68	60.87	59.21	304
Spezia			71.05			52.63			56.58	38
Torino	84.44	76.81	60.53	75.00	45.59	53.95	71.74	62.32	48.68	380
Udinese	96.23	66.18	52.63	92.16	52.94	52.63	95.83	70.59	61.84	494
Verona	96.00	79.71	71.05	90.91	54.41	43.42	90.32	78.26	73.68	228
Average	94.15	81.93	70.47	85.72	42.38	53.14	81.57	65.37	54.31	387

The accuracy is computed as the proportion of correct bets over the total according to the rules (13) and (15). Observations is the number of played matches. The out-of-sample size is given by the matches played in the last two seasons. Blank rows are reported for teams without any bet recommendation

of generalized linear models especially designed for time series data, called *generalized autoregressive moving average* (GARMA). In a similar fashion, another interesting approach useful in analyzing binary time series is the *generalized autoregressive score* (GAS) model.

The GAS model, to which we refer as *score-driven* approach, updates the time-varying parameters by means of an autoregressive component and the conditional score. Despite it has been already successfully applied in football matches forecasting, by predicting the number of goals scored and conceded by a team, the novelty of this paper lies on its application to the prediction of binary outcomes, still poorly explored by sport statistics literature.

In particular, we aim to forecast the future probability that a binary event will occur or not. On the basis of these probability forecasts, we define a suitable betting strategy comparing the predicted probabilities estimated by either GAS or BARMA (that we use as a benchmark) models with its average. If, for both teams, the probability that an event “yes” (“no”) realizes is greater than its average value, a betting recommendation on the event “yes” (“no”) is provided. Indeed, the proposed strategy provides betting recommendation on either “yes” or “no” events, depending by which one is more likely to occur for both teams. Otherwise, there is no betting advice.

The main idea of the proposed betting strategy is to bet only on matches where the probability estimated on the basis of the statistical model is larger than its average, meaning that this event is more likely to verify than usually does. Hence, a useful feature of the proposed strategy is that not always a betting advice is provided and, hence, too uncertain football matches can be avoided by the bettor that can also save money.

Then, in order to assess if statistical models are really useful in predicting binary outcomes, a naive strategy has been implemented as well.

To show the usefulness of the proposed statistical method, two empirical experiments to the English Premier League and to the Italian Serie A are provided for predicting red cards, Under/Over (with three different thresholds) and Goal/No Goal events.

As shown in the empirical analysis, the GAS model outperforms the benchmark approaches represented by the BARMA and the naive strategy. Therefore, a first positive evidence that we get is about the usefulness of forecasting binary outcomes for bettors and odds setters.

More in details, while in the case of red card forecasting the models are competitive each other (the GAS dominates but the advantages with respect the naive strategy is not huge), in the other experiments with Goal/No Goal and Under/Over the betting strategy based on score-driven forecasts results in bets much more accurate. Accuracy is ex-post measured as the percentage of correct bets over the total.

In the end, we have to stress the advantages of a strategy based on betting avoidance in the case of too uncertain events. Indeed, the absence of recommendation, due to probabilistic reasoning, allows to increase the betting accuracy and to reduce the betting costs. Therefore it can be seen as a strength of the proposed approach.

Understanding the determinants of uncorret bets is difficult a priori. However, a future research direction could be the consideration of additional match statistics or performances indicators (i.e. team market value as proxy for the strength) as additional predictors in the GAS model. Perhaps, considering additional covariates could further increase betting accuracy. Nevertheless, we have to highlight that the basic GAS model specification discussed in the paper without covariates already ensures 100% of correct bets also thanks to betting avoidance principle.

Funding Open access funding provided by Università degli Studi di Napoli Federico II within the CRUI-CARE Agreement.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adhikari, A., Majumdar, A., Gupta, G., & Bisi, A. (2020). An innovative super-efficiency data envelopment analysis, semi-variance, and Shannon-entropy-based methodology for player selection: evidence from cricket. *Annals of Operations Research*, 284(1), 1–32.
- Angelini, G., & De Angelis, L. (2017). Parx model for football match predictions. *Journal of Forecasting*, 36(7), 795–807.
- Benjamin, M. A., Rigby, R. A., & Stasinopoulos, D. M. (2003). Generalized autoregressive moving average models. *Journal of the American Statistical Association*, 98(461), 214–223.
- Christoffersen, P. F., & Diebold, F. X. (2006). Financial asset returns, direction-of-change forecasting, and volatility dynamics. *Management Science*, 52(8), 1273–1287.
- Creal, D., Koopman, S. J., & Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5), 777–795.
- Csató, L. (2020). The UEFA champions league seeding is not strategy-proof since the 2015/16 season. *Annals of Operations Research*, 292, 161–169.
- Davis, R. A., Dunsmuir, W. T., & Streett, S. B. (2003). Observation-driven models for Poisson counts. *Biometrika*, 90(4), 777–790.
- Diniz, M. A., Izbicki, R., Lopes, D., & Salazar, L. E. (2019). Comparing probabilistic predictive models applied to football. *Journal of the Operational Research Society*, 70(5), 770–782.
- Dixon, M. J., & Coles, S. G. (1997). Modelling association football scores and inefficiencies in the football betting market. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 46(2), 265–280.
- Dixon, M. J., & Pope, P. F. (2004). The value of statistical forecasts in the UK association football betting market. *International Journal of Forecasting*, 20(4), 697–711.
- Forrest, D., Goddard, J., & Simmons, R. (2005). Odds-setters as forecasters: The case of English football. *International Journal of Forecasting*, 21(3), 551–564.
- Forrest, D., & Simmons, R. (2000). Forecasting sport: The behaviour and performance of football tipsters. *International Journal of Forecasting*, 16(3), 317–331.
- Friesl, M., Libich, J., & Stehlík, P. (2020). Fixing ice Hockey's low scoring flip side? Just flip the sides. *Annals of Operations Research*, 292(1), 27–45.
- Galaritis, E., Germain, C., & Zopounidis, C. (2018). A combined methodology for the concurrent evaluation of the business, financial and sports performance of football clubs: The case of France. *Annals of Operations Research*, 266(1), 589–612.
- Goddard, J. (2005). Regression models for forecasting goals and match results in association football. *International Journal of Forecasting*, 21(2), 331–340.
- Goddard, J., & Asimakopoulos, I. (2004). Forecasting football results and the efficiency of fixed-odds betting. *Journal of Forecasting*, 23(1), 51–66.
- Goossens, D. R., Beliën, J., & Spijksma, F. C. (2012). Comparing league formats with respect to match importance in Belgian football. *Annals of Operations Research*, 194(1), 223–240.
- Harding, D., & Pagan, A. (2011). An econometric analysis of some models for constructed binary time series. *Journal of Business & Economic Statistics*, 29(1), 86–95.
- Hausman, J. A., Lo, A. W., & MacKinlay, A. C. (1992). An ordered probit analysis of transaction stock prices. *Journal of Financial Economics*, 31(3), 319–379.
- Kauppi, H., & Saikkonen, P. (2008). Predicting US recessions with dynamic binary response models. *The Review of Economics and Statistics*, 90(4), 777–791.

- Koopman, S.J., Lit, R. (2015). A dynamic bivariate poisson model for analysing and forecasting match results in the English premier league. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* (pp 167–186).
- Koopman, S. J., & Lit, R. (2019). Forecasting football match results in national league competitions using score-driven time series models. *International Journal of Forecasting*, 35(2), 797–809.
- Li, W. K. (1994). Time series models based on generalized linear models: Some further results. *Biometrics*, 50, 506–511.
- Lit, R., & Koopman, S.J. (2020). Forecasting the 2020 edition of the boat race. *Time Series Lab - Article Series* (pp. 1–12).
- Lu, Y. (2020). A simple parameter-driven binary time series model. *Journal of Forecasting*, 39(2), 187–199.
- Maher, M. J. (1982). Modelling association football scores. *Statistica Neerlandica*, 36(3), 109–118.
- Mesters, G., & Koopman, S. J. (2012). *A forty year assessment of forecasting the boat race*. Tinbergen Institute Discussion Paper: Tech. rep.
- Moysiadis, T., & Fokianos, K. (2014). On binary and categorical time series models with feedback. *Journal of Multivariate Analysis*, 131, 209–228.
- Nikolaïdis, Y. (2015). Building a basketball game strategy through statistical analysis of data. *Annals of Operations Research*, 227(1), 137–159.
- Nyberg, H. (2010). Dynamic probit models and financial variables in recession forecasting. *Journal of Forecasting*, 29(1–2), 215–230.
- Startz, R. (2008). Binomial autoregressive moving average models with an application to us recessions. *Journal of Business & Economic Statistics*, 26(1), 1–8.
- Wang, C., & Li, W. K. (2011). On the autopersistence functions and the autopersistence graphs of binary autoregressive time series. *Journal of Time Series Analysis*, 32(6), 639–646.
- Wheatcroft, E. (2020). A profitable model for predicting the over/under market in football. *International Journal of Forecasting*, 36(3), 916–932.
- Yang, C. H., Lin, H. Y., & Chen, C. P. (2014). Measuring the efficiency of NBA teams: Additive efficiency decomposition in two-stage DEA. *Annals of Operations Research*, 217(1), 565–589.
- Zeger, S. L., & Qaqish, B. (1988). Markov regression models for time series: A quasi-likelihood approach. *Biometrics*, 44, 1019–1031.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.