To appear in the *Journal of Applied Statistics*
Vol. 43, No. 7, 2016, 1–21

How to improve a team's position in the FIFA ranking? – a simulation study

Jan Lasek[a,*], Zoltán Szlávik[b], Marek Gagolewski[c] and Sandjai Bhulai[d]

[a] *Interdisciplinary PhD Studies Program, Institute of Computer Science, Polish Academy of Sciences, Jana Kazimierza 5, 01-248 Warsaw;* [b] *IBM Nederland BV, Johan Huizingalaan 765, 1066 VH Amsterdam* [c] *Systems Research Institute, Polish Academy of Sciences, Newelska 6, 01-447 Warsaw, Poland and Faculty of Mathematics and Information Science, Warsaw University of Technology, Koszykowa 75, 00-662 Warsaw, Poland;* [d] *Faculty of Sciences, VU University Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands*

(*Received 00 Month 20XX; accepted 00 Month 20XX*)

In this paper, we study the efficacy of the official ranking for international football teams compiled by FIFA, the body governing football competition around the globe. We present strategies for improving a team's position in the ranking. By combining several statistical techniques we derive an objective function in a decision problem of optimal scheduling of future matches. The presented results display how a team's position can be improved. Along the way, we compare the official procedure to the famous Elo rating system. Although it originates from chess, it has been successfully tailored to ranking football teams as well.

Keywords: association football; FIFA ranking; prediction models; Monte Carlo simulations; optimal schedule; team rankings

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

1. Introduction

In a variety of practical applications ranking of agents or objects is an inherent part of a decision making process and, more generally, decision support systems. In today's era of constant and massive information flow, ranking algorithms are needed to retrieve and evaluate relevant content. They are also useful when any kind of competition between agents arises. Thus, given a particular ranking system, one often poses a question about its efficacy and whether an entity's position in this ranking could be possibly manipulated and/or optimised.

In this paper we focus on application of ranking algorithms in sports. By means of several statistical techniques we define and estimate an objective function for optimal scheduling of games. To this end, we use a statistical model – Poisson regression – for prediction of outcome of future games and employ Monte Carlo simulations to approximate the distribution of appropriate ranks.

*Corresponding author. Email: janek.lasek@gmail.com

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

Individual player and team rankings are an integral part and, in fact, the core of competition. Thanks to them, the involved entities can monitor their progress and interesting match-ups can be arranged. Most importantly, the official ranks are used for scheduling competitions. In association football, it is common that for major competitions and their qualification rounds a draw takes place which determines the schedule of games. In general, higher ranked teams are paired with lower ranked ones. In this way, rankings have crucial impact on the competition. The teams at the top of the ranking are less likely to face other strong opponents. This is advantageous as it helps to avoid potential elimination in early stages of the competition. Designing the optimal policy for increasing a team's rank in the results list is the main focus of our paper. We study the official ranking maintained by FIFA – *Fédération Internationale de Football Association* – the international governing body of association football (soccer), and look for opportunities of exploiting possible chances of advancing in the ranking table. We observe that an improvement of position in a ranking table is a result of not only good overall performance but also an appropriate choice of the opponent. We also discuss efficacy and robustness of the official ranking system.

The paper is structured as follows. In Section 2, we present motivation for our study and discuss related work. In Section 3, we describe in detail how the official FIFA ranking is compiled. In Section 4, we present the algorithm for optimising a team's position in the ranking table and in the consecutive section we present the results of numerical experiments. In Section 6 we discuss the results. In the last section we present our conclusions.

2. Related work

Before we proceed to discuss related work specific to the problem under consideration, let us look at the ranking systems from a broader perspective. Among applications of ranking systems we find, e.g., search engine rankings. In order to attract a user's attention, a given web page needs to match a given query and be ranked high within the possible outcomes. A high ranking often means high profits for a company. Therefore, it is crucial to have a high rank which is often achieved by exploiting a ranking algorithm's properties [24, 31]. Another area in which rankings and, more generally, evaluation metrics are important is the field of scientometrics. For example, a widely recognised tool for evaluation of scientific achievements based on citation analysis is Hirsch's h-index [17]. It might be used in, e.g., awarding research grants and hence is of major importance to scholars. They might be well unfairly manipulated by, e.g., clever self-citation patterns [2, 41]. In general, in the problem of scientific impact assessment, wide range of tools is prone to manipulation. For example, for a certain class of bibliometric impact indices one can arrive at any rankings of evaluated entities, see [13]. The problem of designing manipulation-proof rankings and reputation systems is important in a wide array of on-line peer-to-peer systems, e.g., at stores or gaming services [12]. In this way, the design of an efficient ranking system becomes a constant process of improvement: given a ranking algorithm and some methods of manipulating it, one wants to enhance the ranking method so that it is not vulnerable to manipulation.

Another domain where rankings are widely used is sport. They can serve as, e.g., a measure of the competitive balance between teams or individual competitors [15, 16]. As far as association football is concerned, the official FIFA ranking and its effectiveness is often subject to criticism. For example, Lasek et al. [25] compared several ranking systems and the official ranking according to their predictive capabilities with respect to future games. It was shown that the predictions of several other ranking systems

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

are more accurate than those derived from FIFA ranking points (the points earned by the teams following the ranking procedure). Based on this result a conclusion was drawn that the attribution of rankings points by the official procedure is not effective in a sense that the strength of teams is over or underestimated in many cases. This potentially opens up space for designing strategies of choosing opponents for friendly games so as to maximise the probability of advancing in the ranking table.

Another study that exposes the FIFA ranking's shortcomings was carried out by McHale and Davies [27], who show that the FIFA ranking does not adjust quickly enough to reflect a team's current performance. Wang and Vandebroek [40] describe a simple strategy for low ranked teams to improve their position in the FIFA ranking. The idea is to play many mutual games in which the teams agree to win half of the games. In this way, each team may increase the number of points they gain. The authors also point out several issues about the official ranking, e.g., incorporation of the time dimension in a discontinuous manner and the fact that the method does not account for home team advantage. In various sports (in particular, in association football) it has been observed that the host of the game has some advantage over a visiting team [28–30, 35]. However, in the current FIFA procedure this feature is not taken into account (as opposed to other ratings systems, e.g., Eloratings.net [9] or the official FIFA ranking procedure for women football [10]).

Another topic related to our work is the issue of exploiting competition rules and the manipulation of ranking procedures. Dagaev and Sonin [6] discuss situations in which a team may be better off deliberately losing a match under certain circumstances. The official competition rules do not constitute an incentive for winning in such cases. Santon and Vassilevska-Williams [36] study double-elimination tournaments and the issue of their manipulation. They consider several types of exploiting structure of these tournaments: players throwing matches, a coalition of cooperating players, and agenda fixing. Both papers concern situations which happen in practice at major tournaments like the FIFA World Cup or UEFA Champions League qualifications in association football, or at Olympic Games in badminton. Russell [33] discusses different types of manipulation methods for sport competitions and studies their computational complexity. The author also addresses the problem of how to detect different forms of manipulation.

Given the possible inefficiencies of the official ranking system, that have been pointed out, in this paper we look for opportunities to exploit these flaws. With the use of a better statistical model of teams' strength estimation, we look for opportunities to advance in the ranking table. Before we proceed with the discussion of ranks' optimisation, however, let us recall how the FIFA ranking is calculated.

3. Ranking agents

In this section we present a general framework for ranking agents through pairwise comparisons with the incorporation of time dimension. We model the situation with the use of a graph: this is a common approach to model tournaments (competitions) [1, 33]. With some extra effort with definition of a weighting function for edges we may also add to the model an underlying ranking system. Next, we discuss a specific case of ranking under our consideration – FIFA ranking for national association football teams. For an alternative approach to formalisation of sport rating systems and their comprehensive survey across different disciplines one can refer to [37, 38].

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

3.1 *General model for rating of agents*

Rankings are obtained via a proper aggregation of results of pairwise comparisons (e.g., matches) of agents (e.g., players or teams). For each match, a certain number of points is awarded to both teams. This depends on a particular application. We consider a general case of ranking agents, who compete in pairs. To model the situation we use a directed weighted multigraph $G = (A, E)$, with A being the set of vertices and E the multiset of edges. More specifically, let the vertices in the graph represent a finite set of agents $A = \{a_1, a_2, \ldots, a_n\}$ under our consideration. A pair of directed edges $e_{ij}^{(t)}$, $e_{ji}^{(t)}$ between a pair of agents a_i, a_j corresponds to the outcome of pairwise comparison between these agents (e.g., a match) taking place at time t. We adopt the convention that the current timestamp corresponds to $t = 0$. The values $t > 0$ are used to denote how long ago a game was played. A weighting function $w : E \to \mathbb{R}$ for edges defines the number of ranking points obtained by the agents. More precisely, $w(e_{ij}^{(t)})$ and $w(e_{ji}^{(t)})$ denote the number of points awarded to agent a_i and a_j respectively. This weighting function is model specific. In general, its valuation is dependent on time, $w(e_{ij}^{(t_1)}) \neq w(e_{ij}^{(t_2)})$ for $t_1 \neq t_2$. We assume that at a given timestamp t only one comparison between two specific agents a_i, a_j can take place. However, in general multiple comparisons are allowed resulting in a multigraph. Further on, let us denote the set to all outgoing edges from a vertex i with $E(i, \cdot)$. At current time $t = 0$, *a rating* of an agent a_i or the number of her *ranking points*, i.e., a numeric quantity indicator for her performance in pairwise comparisons, is calculated by summing the weights of all outgoing edges from its corresponding vertex in graph G:

$$r_i = \sum_{e_{ij}^{(t)} \in E(i,\cdot)} w(e_{ij}^{(t)}). \tag{1}$$

Next, we sort individual agents' ratings $\{r_i\}_{i=1}^n$ to arrive at a final ranking of competing agents, $\{R_i\}_{i=1}^n$, where the best performing agent is assigned rank of 1. In such an exposition, a variety of problems can be modelled, provided that for a particular problem a weighting function can be defined and that the additive model for ratings derivation in Equation (1) can be assumed. We use this approach to model football competition and ranking of participating teams. It may well be used in modelling general support for political parties, who compete in pairs in a series of debates. In such a setup, the nodes correspond to political parties and weighted edges as the outcomes of debates (in this case the definition of a weighting function is a problem of its own). It could be also used to model a beauty contest, in which the candidates are presented to a jury in pairs. An edge with associated weight corresponds to the number of outcomes toward a particular candidate. The winner of the contest is the person with the highest rating, i.e., the maximal number of votes overall.

3.2 *The official FIFA ranking*

The official FIFA ranking can be seen as an averaging rating system [11, 38]. It is released on an approximately monthly basis. The current procedure of awarding points is applicable since 12 July 2006, following FIFA World Cup Finals in Germany. To calculate current ranking points, for each national team, the results of matches in the last four years are used. During the periods between FIFA's issuing of the rankings, teams compete and gain points. The teams' rankings are recalculated for every ranking release by FIFA.

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

The edge weight function in the FIFA ranking can be decomposed into three factors:

$$w(e_{ij}^{(t)}) = \frac{1}{m_i^{(t)}} \cdot y_t \cdot N_{ij}^{(t)}, \tag{2}$$

where $m_i^{(t)}$ is an averaging constant for a particular season, y_t is a time weighting function and $N_{ij}^{(t)}$ denotes number of points obtained by the team in a particular match. Below, we explain these quantities in detail. The time is expressed in years. For example, $t = 1.5$ means that a match was played one and a half years ago.

The calculation of ranking points awarded for a team a_i for a single game $g = (a_i, a_j, t, k)$, against a team a_j, taking place at time t with the type of the match k (e.g., a friendly or FIFA World Cup Final game) is done according to the following formula:

$$N_{ij}^{(t)} = M_{ij}^{(t)} \cdot I_k \cdot S_{a_2,d}^{(t)} \cdot C_{a_1,a_2}^{(t)}, \tag{3}$$

where the consecutive factors denote: (a) the outcome of the game ($M_{ij}^{(t)}$ points), (b) the importance of the game (I_k), (c) the strength of the opposing team ($S_{a_2,d}^{(t)}$) and the average of confederation strengths ($C_{a_1,a_2}^{(t)}$) of participating teams. Below we explain how these values are obtained.

Teams gain 3 points for a victory, 1 point for a draw and 0 points for a defeat. In a penalty shoot-out, the winning team gains 2 points and the losing team gains 1 point. Based on the outcome of the game, an appropriate value for factor $M_{ij}^{(t)}$ is set as the number of points obtained by team a_i in Equation (3).

Depending on the type of the game k, factor I_k assumes four values as follows: (a) a friendly match – 1.0, (b) FIFA World Cup qualifier or confederation-level qualifier – 2.5, (c) Confederation-level final competition or Confederations Cup – 3.0, (d) FIFA World Cup final competition – 4.0. In this way, we distinguish several types of games and assume that particular kinds of matches (e.g., World Cup games) are more important indicators for a team's strength. Clearly, qualifiers and major competition games have greater impact to a team's total ranking points rather than friendly games. It is reasonable to assume that the results of friendly games should be considered as less important. Often in those games the managers experiment with the team. In particular, usually those matches do not involve the strongest "eleven". However, the way to quantify game importance is not clear. One may argue for or against a particular choice of values for multipliers. It is based on a subjective view on match importance.

The strength of the opposing team (factor $S_{a_2,d}^{(t)}$ in (3)) at time t is calculated as 200 minus the ranking position of that team (from the latest release of the ranking prior to time t). As an exception, the team at the top of the ranking is assigned a maximum value of 200 and the teams ranked at 150th or lower place are assigned the minimum value of 50.

There are six confederations recognised by FIFA that oversee the game in different parts of the world. Each confederation is assigned a number in the range $[0.85, 1]$ that indicates its overall strength. These values are calculated based on the result of the last three World Cup tournaments. For both teams the value of $C_{a_1,a_2}^{(t)}$ in Equation (3) is computed as the mean value of the strength of confederations that the two competing teams belong to [11]. Table 1 presents these values for all six confederations.

An exemplary calculation of the points is presented in Table 2.

There are also further factors constituting edge weighting function in Equation (2).

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

Table 1. Strength of confederations used in ranking points calculation.

Confederation	**Continent**	**After WC 2006**	**After WC 2010**	**After WC 2014**
AFC	Asia and Australia	0.85	0.86	0.85
CAF	Africa	0.85	0.86	0.85
CONCACAF	North and Central America	0.85	0.88	0.85
CONMEBOL	South America	0.98	1	1
OFC	Oceania	0.85	0.85	0.85
UEFA	Europe	1	1	0.99

Table 2. Calculation of the ranking points after the Netherlands 2:1 victory over Brazil in the FIFA World Cup Final competition. On the day the match was played the Netherlands were placed 4th and Brazil 1st.

Team	**M**	**I**	**S**	**C**	**Points**
Netherlands	3	4	200	1	**2400**
Brazil	0		196		**0**

According to the FIFA ranking methodology, four seasons are distinguished corresponding to time intervals $T_s = (s, s+1]$ for $s = 0, 1, 2, 3$. For example, $T_0 = (0, 1]$ corresponds to the last year of play. We also denote with $T_4 = (4, \infty)$ all the other past timestamps. Depending on how long ago a game was played, it receives different weight. Only the results from the last year receive full weight of 1. For $t \in (1, 2]$ we get $y_t = 0.5$, for $t \in (2, 3]$ – $y_t = 0.3$, and for $t \in (3, 4]$ – $y_t = 0.2$. The games older than 4 years are not taken into account at all ($y_t = 0$). We note that the incorporation of time in the model is done in a discontinuous fashion. This can result in sometimes a huge points (and accordingly rank) decrease in accumulated points in consecutive releases of the FIFA ranking. For example, Spain lost 283 ranking points in the release of July 2011 as compared to June since its FIFA 2010 World Cup's winning streak started to be discounted with a smaller weight (however, they were still ranked first).

The other component of Equation (2) is responsible for averaging awarded points in a given season (relative to the current timestamp $t = 0$):

$$m_i^{(t)} = \max\left\{5, \left|\{e_{ij}^{(t')} : e_{ij}^{(t')} \in E(i, \cdot) \ \wedge \ t' \in T_{\lfloor t \rfloor}\}\right|\right\}.$$

A game played in a given season T_s, $t \in T_s$ is assigned the same averaging factor as all other games played in that season. If a team played in a given year fewer than 5 games, instead of computing average of points, we divide the total number of points gained by a factor of 5. To put it another way, in case a team played less than 5 games, we add fake losses to its tally so that the total number of games in each year is 5. In such a situation it is always profitable to play at least 5 matches in terms of expected number of ranking points a team can gain. Thus, rarely competing teams are penalised (in 2014, there were 35 national teams that played less than 5 games, e.g., Belize, Faroe Islands, India, Liberia or Nepal). In this way, the FIFA ranking methodology encourages to play a minimal number of games per season.

Once we compute aggregated ranking points for all teams according to Equation (1), we sort these numbers to arrive at the ranking of teams. In case any teams received equal number of ranking points, we solve the ties by assigning the minimum rank.

Summing up, in the official FIFA ranking teams accumulate points by playing matches. The number of points awarded after single match depends on the outcome of the game, its importance, opposition and confederation strength. In each of four most recent years an average of points gained are computed (with correction for teams that played fewer than 5 games). These yearly averages are next added up with weights depending on

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

the time. More recent matches count more towards a team's position in the ranking.

We shall note that when implementing FIFA ranking we encountered some inconsistencies between teams' ranking points and ranks available at the FIFA website and our reproduced results. There are slight differences between our results reported in Section 5 and rankings available at the official website. For example, in a new release of the FIFA ranking usually most recent games (played within last couple of days) are not included. Our implementation of FIFA's methodology includes all the games played prior to a date when new ranking is released. We stick to the description of the algorithm reported in this section.

4. Optimal choice of an opponent

Let us discuss how to choose an opponent for a friendly game so as to maximise the probability of being ranked high. In this part we present subsequent steps in the optimal decision computation process. We define our target and formulate it as an objective function in a decision making problem. We include additional information, e.g., the schedule of matches and the risk levels and describe how the particular components of our analysis are developed. In consecutive subsections we describe the situation from a chosen team's perspective.

4.1 *Basic setup – planning one game ahead*

We start with the definition of a team's target. Let a chosen team $\tilde{a} \in A$ hold a certain position in the most recent ranking release. The target is to advance in the ranking table to at least rank n, with maximal probability. A set of possible actions is identified with the set of all competing teams A. The choice of an action $a \in A$ means that we shall play our friendly game against an opponent a (a decision $a = \tilde{a}$ means that we abstain from playing matches at all).

Let us define a random variable $Z_a^{(t)}$ as the total number of ranking points that we can accumulate in a given forthcoming (in particular, the next) release of the FIFA ranking, issued at time (date) t if we decide to play against an opponent $a \in A$ (according to Equation (1)). This quantity depends on the history of games played over the last four years and on the choice of an opponent to play against a. Since the result of the game is unknown, we use a prediction model for calculating probability of the result [7, 14, 19]. Now, if our goal is to advance in the ranking to at least rank n it means that $Z_a^{(t)}$ shall be not smaller than the nth greatest value of ranking points accumulated by all the ranked teams. That is to say it needs to be not smaller than the nth order statistics when the ranking points by all ranked teams are sorted in decreasing order. Let $P_{(n)}$ denote this quantity. Formally, we should write $P_{(n)}^{(t),a}$ since this random variable is dependent on both our choice of action $a \in A$ and the particular date t of FIFA ranking release under consideration. We suppress superscript t as it should be clear from the context. Moreover, we assume that the outcome of the scheduled game against team a does not influence the distribution of the considered nth order statistic. If considerably large number of games is played, the fixture played by our team has relatively little impact on the distribution of a given order statistic.

The distribution of the random variable $P_{(n)}$ can be computed based on the current state of the FIFA ranking if no games are played. In this setting it is a deterministic quantity. If there are some fixtures until the time point that we are interested in, which is usually the case, it becomes random. If we have (partial) knowledge of the schedule

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

of other games played we shall include it in our computations. Let $\mathcal{S}$ denote a set of scheduled games between some teams from the A set. We obtain the following objective function in a decision problem:

$$d = \operatorname*{argmax}_{a} \mathbb{P}(Z_a^{(t)} \geq P_{(n)} | \mathcal{S}). \tag{4}$$

We may want to introduce a certain level of risk that we can accept. An additional constraint may be added based on the requirement that the probability of being ranked at $m \geq n$ position or lower in the ranking is smaller than a certain predefined level of risk $p \in (0, 1)$ that we are willing to accept:

$$\mathbb{P}(Z_a^{(t)} \leq P_{(n)} | \mathcal{S}) \leq p.$$

In this way, a team may want to secure a spot in the FIFA ranking with sufficiently high probability. Note that the goals of climbing high in the ranking table and – at the same time – accepting lower risk are contradictory. To gain more ranking points, we need to challenge a higher ranked team. This presumably means that the opponent is stronger and the probability of a loss increases.

4.2 *Planning several games ahead*

An extension of the basic set-up is to allow for planning l games ahead. Usually, matches are scheduled up to a few games in advance. On many occasions teams play two friendly games within a week so it is important to be able to schedule more games at once.

Now our decision space is a sequence in the set A^l of length l. We have that $\mathbf{a} = (a_1, a_2, \ldots, a_l) \in A^l$ is the sequence of teams to play against. Our objective function becomes

$$d = \operatorname*{argmax}_{\mathbf{a}} \mathbb{P}(Z_{\mathbf{a}}^{(t)} \leq P_{(n)} | \mathcal{S}). \tag{5}$$

We may extend the framework to risk level p as above.

A natural question that arises is when to play an additional game, i.e., how to set parameter l. There is no clear answer to this question. It depends on the distribution of possible point gains as well as the distribution of order statistic of the rank that we are interested in. However, in certain situations it can be clear that a game against a particular team should not be played. We discuss this issue below.

As described in the previous section, the computation of points is based on summing yearly averages over the last four years of play. Let us assume that over the span of last year a team managed to accumulate an average of μ ranking points In general, this team should not play against an opponent, for which the maximal possible number of points to gain according to Equation (3) is lower than μ. Such a choice would result in pulling the average down. However, this situation may change since the quantity μ varies in time.

As an example, consider the case of Italian national team. On 17/10/2013 Italy occupied rank 8, merely 2 ranking points behind Switzerland at rank 7. By the official decision of FIFA, the ranks for that day were taken into account for seedings for World Cup group stage draw on 06/12/2013. According to the rules, seven top ranked teams were placed in the first pot next to Brazil. These are considered as the strongest teams in the World. Being in the first pot means that in the group stage a team would not be paired with any of these teams. Earlier this year Italy played two friendly games

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

against two lower ranked opponents: Haiti and San Marino. However, since the number of possible points to gain against these opponents was lower than Italy's previous year points average, the team bare a loss in total number of ranking points. If the two games had been avoided, Italy would have overtaken Switzerland in the FIFA ranking release and it would have gained a place in the first pot. Note that Italy, having being paired with Costa Rica, England and Uruguay, was eliminated after the group stage. Finally, we note that such a situation would not happen under the Elo rating system [9, 10]. In case of a win, a team always is awarded some extra points, albeit the gains can be relatively small depending on the strength of an opponent. On the other hand, this rating method does not incorporate time dimension at all. A team's rating remains constant if it does not play any games. In such a situation, we do not have recent evidence of team performance. Under certain circumstances it may also constitute an incentive for a team to abstain from playing a game.

4.3 *Finding optimal fixture for a chosen team*

Let us describe individual components used in the computations of the optimal decision. We begin with a discussion on prediction models of future games' outcomes.

4.3.1 Prediction of future games

Being able to simulate schedule outcomes and computing optimal policy in (4) and (5) means that we need a model for prediction of outcomes of scheduled games. In this section we discuss different approaches to this problem.

Match outcome prediction models have been considered by many researchers in the fields of statistics and machine learning. Maher [26] suggests modelling scores by the two teams competing in a match as independent Poisson variables. This is one of the simplest approaches for modelling association football scores and it serves as a basis for more refined models. Maher also suggests bivariate Poisson distribution for both teams' scores, which relaxes the assumption of independence. Bivariate Poisson distribution are studied further by Karlis and Ntzoufras [19]. The authors also suggest diagonal inflated bivariate Poisson models as they find out that the independence assumption in particular causes underestimation of the probabilities of drawn games. Dixon and Coles [7] suggest another modification to the Maher's model by introducing additional parameters to the likelihood function for the low-score outcomes. The Poisson model has been also extended by various authors to account for the fact that a football teams' shape is subject to fluctuations, see, e.g., [5, 22, 32]. By allowing teams' shape parameters to vary over time, the authors adapt their model to a dynamic situation present in association football.

The Poisson approach for modelling football scores takes into account the exact number of goals scored. An alternative way, which focuses solely on the final result of the game – win, draw or loss – is to apply ordered probit or logit regression models. Such methodology was employed by, for example, Koning [21]. However, Goddard [14] in his paper provides empirical comparison on the effectiveness of both approaches. The author concludes that there is virtually no difference between them as they provide similar results. Yet another approach is to model difference in goals scored rather than individual scores of the final outcome. Such studies were conducted by, e.g., Karlis and Ntzoufras [20] or Van Haaren and Van den Broeck [39].

Next to the two statistical approaches for modelling scores, there are various prediction models originating from the field of machine learning. Joseph et al. [18] compare between different techniques and show that a Bayesian network constructed by an expert provides

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

best results among different machine learning algorithms. Incorporation of the expertise in football prediction models is studied further by Constantinou et al. [3, 4]. The authors propose a Bayesian network model that allows for including impact of beliefs on, for example, the perceived current shape of a team, its fatigue or availability of key players.

Based on the popularity of the Poisson model which appears to be well-founded in literature, we employ this very method in our study. Thus, let us describe the basic model introduced by Maher [26]. This is a simple yet effective approach for modelling football scores that is still eagerly used in the area of sport analytics see, e.g., [34]. The main assumption of the model is that the goals scored by a team are treated as independent Poisson variables. Let G_i and G_j be random variables that express the goals scored by a home and away team, denoted with a_i and a_j respectively. We assume that both random variables follow a Poisson distribution with mean λ and μ:

$$\mathbb{P}(G_i = x, G_j = y) = \frac{\lambda^x}{x!} exp(-\lambda) \cdot \frac{\mu^y}{y!} exp(-\mu). \tag{6}$$

A log-linear model for the goal scoring rates is assumed, $\log(\lambda) = c + h + o_i - d_j$ and $\log(\mu) = c + o_j - d_i$, where c is an intercept, o_i, o_j and d_i, d_j stand for offensive and defensive capabilities of teams a_i and a_j, respectively. The parameter h is introduced to capture the advantage of the home team ($h = 0$ if the game is played on aneutral ground). We note that two constraints need to be imposed to identify the model parameters, e.g., that both offensive and defensive teams' capabilities sum to zero, i.e., $\sum_i o_i = 0$ and $\sum_i d_i = 0$. The model parameters may be estimated by constructing an appropriate binary design matrix. In this matrix each row's non-zero entries indicate the intercept, possible home team advantage, offensive and defensive parameters of the two teams involved in a match. By assumption, each match provides two independent samples.

To estimate parameters c, h and $\{o_a, d_a\}$, $a \in A$ we used games played in last four years. This period appears to be long enough to account for a team's shape accurately [25]. Each game in the dataset gives us two observations for goals scored which are used to estimate offensive and defensive strengths of the teams according to the model presented above. Next, the probability of a three-way outcome is computed as follows. If we denote $D = G_i - G_j$ then the probability of a win of team a_i over team a_j and the probability of a draw is computed as $\mathbb{P}(D > 0)$ and $\mathbb{P}(D = 0)$, respectively. To compute these quantities one may use a Skellam distribution [20].

4.3.2 Computing distribution of order statistics

As we noted above, an important part while determining the optimal policy is the knowledge of distribution of $P_{(n)}$, i.e., the nth order statistic, where n is our desired position at a specific time point. Its distribution is dependent on the schedule of games. The longer the horizon between the current timestamp and the date at which we want to compute $P_{(n)}$, the larger the set of games influencing its distribution. In the experiments below, we found that even for relatively short period of time until the next FIFA ranking release, there are many games that can influence $P_{(n)}$. If the distribution of $P_{(n)}$ is influenced by the outcomes of 20 games, exact computation of it requires analysis of $3^{20} \approx 3.5$ billion game outcomes. Therefore in our experiments we decide to simulate match results to approximate distribution of $P_{(n)}$.

To identify the fixtures that can influence our computations we focus on simulating results of only so-called *active* games. To this end, we compute upper and lower bound on $P_{(n)}$. Upper bound can be computed by considering the case when all teams gain maximal number of points for their scheduled games. The lower bound can be computed in an analogous manner. In simulations, we skip games played by a pair of teams for which

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

minimal possible number of points is greater than the upper bound, or their maximal possible points gain is smaller than the lower bound. Although this approximation is rough and may overestimate the upper bound and underestimate the lower bound, it allows us to restrict the set of scheduled games which can influence the distribution of $P_{(n)}$.

4.3.3 Restricting search space to non-dominated choices

When looking for the optimal choice of opponents we note that the search space $|A^l|$ can be large: with 200 teams as possible opponents and 5 games to schedule there are as many as $\binom{200}{5} > 2.5$ billion of possible choices. However, we may greatly restrict computations via elimination of dominated decisions from the set of possible actions A. More precisely, we say that a decision to play against team a_i (weakly) dominates the decision to play against team a_j, written as $a_i \succ a_j$, if

$$a_i \succ a_j \iff N(W_i) \geq N(W_j) \ \wedge \ \mathbb{P}(W_i) \geq \mathbb{P}(W_j) \ \wedge \ \mathbb{P}(L_i) \leq \mathbb{P}(L_j),$$

where $N(W_i)$, $\mathbb{P}(W_i)$ and $\mathbb{P}(L_i)$ denote the maximal number of points that we can obtain when playing against of team a_i (in case of a win; computed according to Equation (3)), probability of winning and losing against team a_i, respectively. In terms of random variables N_i, denoting our point gains depending on the outcome of a match against team $a_i \in A$, this definition expresses the fact that N_i first-order stochastically dominates N_j. Intuitively, this relation defines opportunities that we are looking to exploit. We want to choose opponents for friendly games that are ranked high and against which we are more likely to achieve good result. We refer to [8, 23] for algorithms for computation of non-dominated set of alternatives.

When computing quantities N_i we use the latest FIFA ranking release to retrieve an opponent's position in the ranking. Although these ranks change in time, we assume that we can immediately implement our decision to play against a chosen opponent and assume constant ranks. One can extend this problem by considering ranks of our opponents in a future time point as random variables. From July 2006 until July 2015 the mean and median of absolute change of a team's FIFA ranking position with respect to previous one were 3.2 and 1, respectively. What is more, 90% of changes were within the range of 8 positions.

4.3.4 Algorithm for computing optimal decision

In this section we present all steps of computation of one-step optimal decision to attain a given rank. The input of the algorithm is the set of possible opponents A, rank n that we are aiming at, date t of the future FIFA ranking release that we want to advance at and the schedule of upcoming matches $\mathcal{S}$. The steps are given in Algorithm 1.

Algorithm 1 OptimalOneStepDecision(A, n, t, $\mathcal{S}$)

$\bar{A}$ = set of non-dominated decisions from A
$\bar{\mathcal{S}}$ = set of active games from $\mathcal{S}$
Simulate $P_{(n)}$ depending on schedule $\bar{\mathcal{S}}$
for $a \in \bar{A}$ **do**
    $D[a] = \mathbb{P}(Z_a^{(t)} \geq P_{(n)} | \bar{\mathcal{S}})$ ▷ D is an auxiliary array
end for
return $a_{opt} = \mathrm{argmax}_a \, D[a]$

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

The algorithm computes array D of length $|\bar{A}|$ in which under entry $D[a]$ the probability of attaining rank n when playing against team a is stored. The return value of the procedure is the maximal value among all entries in this array. In Line 1 of the algorithm the decision space is narrowed down to non-dominated choices as discussed in Section 4.3.3. The distribution of $P_{(n)}$ is approximated with the use the set of games $\bar{\mathcal{S}}$ in schedule that influence the its distribution as discussed in Section 4.3.2. The Poisson model (Section 4.3.1) is used for prediction of game outcomes in order to determine the set of non-dominated decisions $\bar{A}$, approximation of $P_{(n)}$ and approximation of the calculated probability in Line 5. This probability is computed as:

$$\mathbb{P}(Z_a^{(t)} \geq P_{(n)}|\bar{\mathcal{S}}) = \sum_{r \in \text{Results}(a)} \mathbb{P}(Z_a^{(t)} \geq P_{(n)}|r, \bar{\mathcal{S}}) \cdot \mathbb{P}(r),$$

where the summation runs over possible results of the match. We employ here a simplifying assumption that our choice to play against team a is independent of the distribution of $P_{(n)}$. – thanks to that $P_{(n)}$ needs to be calculated only once (cf., Line 3 in Alg. 1).

The proposed algorithm can be extended for computation of optimal choices when scheduling more than one game. For example, if we want to schedule $l = 2$ games we should modify the loop in Lines 4–6 to iterate over all pairs of choices in set $\bar{A}$.

5. Experimental results

Let us present several case studies that were chosen to illustrate how friendly games can be scheduled so as to maximise the probability of reaching desired targets. In the analysis, we respect the official dates set up by FIFA for playing friendly matches. Playing a match outside these dates does not guarantee that all the players are available to play due to obligations with respect to their clubs.

The FIFA ranking calculation includes many parameters. Most importantly, since the time component is included, we shall fix some dates in our considerations below. Also, in order to maximise chances of winning, all the probabilities of match results are calculated with the assumption that the team under consideration plays at home [29, 30, 35]. This exploits the fact that FIFA ranking methodology does not incorporate home team advantage in its computations as opposed to the Elo rating system [9, 10].

In order to account for scheduled games, we run 100 000 scenarios of possible match outcomes to approximate the distribution $P_{(n)}$ for a given rank n.

5.1 *Ukraine – FIFA 2018 World Cup qualification draw*

As we stated in the introduction, a very important application of ratings is seeding teams for, e.g., World Cup qualification phase as well as group stage draws. On 25/07/2015 a draw took place for the UEFA qualification stage for 2018 FIFA World Cup finals in Russia. During the draw, the teams were seeded according to their ranks in the July 2015 FIFA ranking release. The teams were placed in 6 pots: 9 teams in the first five pots and the last pot consisting of 7 teams. Ukraine was ranked at position 27 in July 2015 FIFA ranking release which placed it at position 19 among countries belonging to UEFA (excluding Russia which does not take part in the qualification stage as the host of the tournament). As a result, Ukraine was ranked just outside top 18 and it was placed in the third pot for the draw. In this subsection we consider how Ukraine could improve its ranking position so as to be ranked among top 18 UEFA teams and be placed among the teams in the second qualifying pot.

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

First of all, on 14/06/2015 Ukraine played European Championships qualification match against Luxembourg at home. This game is automatically included in their schedule. On 09/06 Ukraine played a friendly game away against Georgia, winning 2:1. We want to consider a better possibility of playing a match with respect to the rival. Let us consider the beginning of April as a reference point for analysis. In April 2015 FIFA ranking release Ukraine was ranked at position 33 and position 20 within countries belonging to UEFA. Table 3 presents top 5 decisions with respect of maximising the probability of being ranked among top 18 teams (within UEFA football associations) or higher for a single game (columns of the table are sorted in this order). Numbers in bold indicate that a given decision (given in columns) is optimal when we want to advance at the nth rank (given in rows). At this point, we assume that Ukraine is to play an away game (in line with its actual friendly game played away at Georgia). Ukraine could increase probability of a win by inviting opponents to play at their home ground.

Table 3. One step optimal decision. Next to opponent name its April 2015 FIFA ranking release is provided within brackets, T. and T. stands for Trinidad and Tobago.

Rank	Wales (22)	Haiti (79)	Rwanda (74)	Slovakia (20)	T. and T. (65)
17	**0.2668**	0.1366	0.1240	0.2141	0.1187
18	**0.4384**	0.4112	0.3750	0.3671	0.3602
19	0.5927	**0.6708**	0.6225	0.5272	0.6038
20	0.7338	0.8309	0.7940	0.6860	0.7815

The optimal choice for Ukraine would be to challenge Wales. This yields estimated probability of 0.44 to be ranked among top 18 UEFA teams. Notably, this probability is equal to only 0.17 for the actual scheduled game with Georgia. Clearly, this choice is suboptimal. In fact, among the top 5 opportunities presented in Table 3, wins against Wales or Slovakia would have placed Ukraine in the second pot. According to the estimated probabilities of wins against those teams, Ukraine has higher chances to beat Wales than Slovakia. A win against any other of the opponents – Haiti, Rwanda or Trinidad and Tobago – would not have secured Ukraine a place in the second pot. Ukraine has higher chances to beat any of these teams rather than Wales or Slovakia, however, it would need advantageous results of other games – the teams which compete with Ukraine for a place among top 18 European teams.

5.2 *Hungary – setback and the way up*

As our second example let us consider the situation of Hungary by the end of 2011. In September 2011 FIFA ranking release Hungary soared to position 27 in the ranking. However, a month later Hungary suffered a setback and trailed at rank 36, 9 places down as compared to the previous month. With the perspective of friendly games ahead, we can ask what would be the optimal choice of opponents for the team so as to advance at a given rank in November 2011 FIFA ranking release. Two friendly games by Hungary are presented in the Table 4.

Table 4. Friendly games by the Hungarian national team between 19/10 and 23/11/2011.

Date	Match			Result
11/11/2011	Hungary	–	Liechtenstein	0:0
15/11/2011	Poland	–	Hungary	2:1

Suppose that we are interested in securing one of the places 33-38 with the highest probability. Table 5 presents the choices of the teams which give the highest probability

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

of attaining a given or a higher rank. According to our simulations, if the Hungarian team had wanted to secure their current position with the highest probability, it should have played against Estonia. This decision yields probability of 0.65 of being ranked at 36 position or higher. The same team should have been their opponent if they had wanted to advance at rank 35. If Hungary had wanted to challenge rank 33 or rank 34 it should play against higher rated opponents – Greece and Slovakia, respectively. If the team wanted to make a safe decision to avoid being ranked lower than 37 it should challenge Antigua and Barbuda.

Table 5. One step optimal decision.

Rank	Greece (8)	Slovakia (41)	Estonia (59)	Ant. and Bar. (91)
33	**0.1981**	0.1696	0.1017	0
34	0.2997	**0.3574**	0.318	0.0161
35	0.3442	0.4991	**0.5293**	0.1521
36	0.3963	0.5818	**0.6519**	0.5121
37	0.5135	0.6804	0.7543	**0.8613**
38	0.8767	0.9212	0.9418	0.9896

We also estimated probabilities for optimal decisions when scheduling 2 and 3 games for this period. The optimal decision to secure 36th place or higher is to play against Antigua and Barbuda and Estonia (which yields probability of 0.48) or Antigua and Barbuda, Armenia (rank 46) and Slovakia (with probability of 0.35). On the other hand, if Hungary wanted to attain rank 33 or higher they should have played against Armenia and Switzerland (18) which yields probability of 0.16 of achieving its target, or Armenia, Estonia and Switzerland (with probability equal 0.10).

Playing the schedule of games as shown in Table 4 gives Hungary probability of holding at least rank 36 and 37 with probability 0.15 and 0.63 respectively. From the point of view of Hungary it is profitable to play a single game against Estonia.

5.3 *Poland – friendly schedule at the turn of 2013*

Let us focus our attention on the team of Poland in October 2013. The team played several friendly games which are listed in Table 6.

Table 6. Friendly games by Poland in November 2013–January 2014.

Date	Match			Result
15/11/2013	Poland	–	Slovakia	0:2
19/11/2013	Poland	–	Rep. of Ireland	0:0
18/01/2014	Poland	–	Norway	3:0
20/01/2014	Moldova	–	Poland	0:1

In the FIFA ranking release of 17/10/2013, Poland was ranked 69th with 503 points. We want to choose opponents for the Polish team such that the choice is optimal in the sense discussed in the previous sections. Let us define our target as to advance in the table with highest probability for the next ranking release on 28/11/2013. We focus on scheduling the first two games. Tables 7 and 8 present optimal choices for scheduling $l = 1$ and $l = 2$ games, respectively.

In case of a single choice, if Poland wanted to advance high in the ranking, it should have challenged Switzerland (rank 7) though the probability of advancing is low (it is an optimal choice if we want to tackle any of the positions between 64 and 67). If it wanted to secure its past-current rank (69), it should decided to play against Iceland (rank 46). Choosing the Dominican Republic (78) would allow it not to drop in the ranking lower

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

Table 7. One step optimal decision.

Rank	Switzerland (7)	Slovenia (30)	Iceland (46)	Dom. Rep. (78)
64	**0.0034**	0	0	0
65	**0.0372**	0.0005	0	0
66	**0.1433**	0.0133	0.0002	0
67	**0.2647**	0.1272	0.0082	0
68	0.3199	**0.3389**	0.0935	0.0161
69	0.3279	**0.4748**	0.3043	0.1264
70	0.3282	0.499	**0.5183**	0.4112
71	0.3282	0.5001	0.6139	**0.7296**
72	0.3317	0.5012	0.6298	**0.8909**

than position 71 or position 72, with maximal probability of 0.73 and 0.89 respectively. It can be considered as a decision with the highest safety level.

Let us look at optimal choices when we challenge two teams instead of one. Table 8 presents the results. Playing two games opens possibilities to climb higher in the ranking with higher probability, for example, if Poland have had challenged Iceland and Slovenia. The optimal strategy to secure its place was to play against Dominican Republic and Iceland.

Table 8. Two step optimal strategy of scheduling games. Country codes correspond to ARM – Armenia, CUB – Cuba, DRE – Dominican Republic, ICE – Iceland and SLO – Slovenia.

Rank	ICE (46) SLO (30)	ARM (38) ICE (46)	ARM (38) DRE (78)	DRE (78) ICE (46)	CUB (73) DRE (78)
63	**0.2004**	0.1483	0	0	0
64	**0.2808**	0.2741	0.0055	0.0005	0
65	0.3099	**0.3443**	0.06	0.0096	0
66	0.315	**0.3609**	0.231	0.0714	0.0003
67	0.3154	0.3625	**0.4269**	0.2415	0.0103
68	0.3444	0.3902	**0.5164**	0.454	0.1183
69	0.4376	0.4789	0.5392	**0.572**	0.388
70	0.575	0.609	0.5904	0.6358	**0.6821**
71	0.7158	0.7446	0.6865	0.7345	**0.8598**
72	0.7974	0.8252	0.7534	0.8078	**0.9249**

Let us now reformulate our targets as follows. We would like Poland to advance as high in the ranking as possible for the release on 13/02/2014. Let us assume that now it is allowed to schedule all the four games. For instance, the team may advance at least to position 61 when playing against Algeria (rank 32), Armenia, Dominican Republic and Iceland (41), with the maximal probability of 0.1936. If Poland had wanted to advance to rank 65, they should have played against the same set of opponents. The probability of achieving the target is 0.43. Securing its current position (69) could have been done with relatively high probability of 0.8952 when playing against Antigua and Barbuda (113), Cuba, Dominican Republic and Iceland. We note that the distribution of ranks changed as compared to the case of November 2013 ranking release. This is wider range of games in the schedule are involved in its determination.

5.4 *The case of the hosts of major competitions – Brazil*

A feature that is immanent in FIFA's ranking is that the host(s) of an upcoming major competition usually drops significantly in the ranking table prior to the tournament. This is because the team does not participate in qualification rounds leading up to that tournament. In this way, their ranking points come mainly from friendly games which are assigned lower importance. In this section, we consider the case of Brazil – the host of the 2014 World Cup finals – by the end of 2012.

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

Brazil was steadily failing in the ranking table since 2010. The situation is illustrated in Figure 1, where we can observe a particular drop in July 2012, when the World Cup 2010 results came to be weighted with factor 0.2, down from 0.5. Another disappointment came at the end of 2012, when Brazil trailed in position 19, in the December 2012 FIFA ranking release. Having won the Confederations Cup in June 2013, though, Brazil recovered its earlier better position.

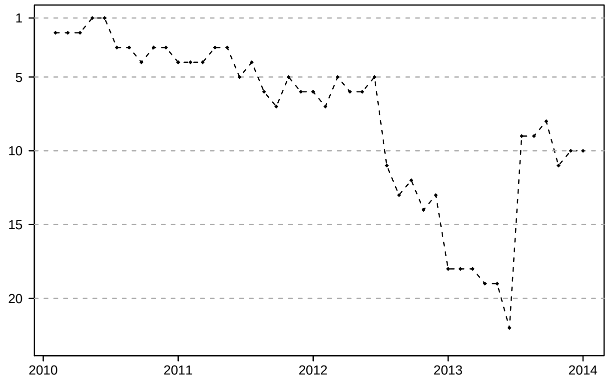

Figure 1. Position of the Brazilian national team over the years 2010–2014

As the problem of sinking in the table appears to be unavoidable, we may pose a question whether it would be possible for Brazil to hold its rank (position 13) that it had in the October 2012 ranking release.[1] Let us look at the friendly games played by Brazil until the December 2012 FIFA ranking release. These are listed in Table 9.

Table 9. Friendly games by Brazil by the end of 2012.

Date	Match			Result
21/11/2012	Argentina	–	Brazil	2:1
14/11/2012	Brazil	–	Colombia	1:1
16/10/2012	Japan	–	Brazil	0:4
11/10/2012	Brazil	–	Iraq	6:0

We may ask the question whether Brazil could have scheduled its games so as to maximise the probability of keeping their current rank, or to progress by 1-2 ranks. Table 10 presents optimal choices for chosen ranks for in one step optimal decision. According to our simulations, playing against Greece (rank 10) is an optimal choice if we want to progress to at least any of the ranks 10–13, challenging Mali (rank 27) is an optimal decision to achieve ranks 14–15, and playing with Haiti (rank 60) could have been a safe decision that would allow the team to secure position 16.

Next, let us look at the number of games $l = 2, 3, 4$. Let us focus our attention on rank 12, which seems to be an achievable position. According to our simulations, a two step optimal choice would have been for Brazil to play twice against Greece (which yields probability of 0.64 for advancing to at least rank 12) against Greece and Mali (with

[1]According to our calculations, Brazil should have hold rank 13 in October 2012 FIFA ranking release. As we already discussed in Section 3, we identified some inconsistencies in FIFA ranking computations. In our computations we obtained that France should have been ranked behind Brazil at position 14 with 988 ranking points. The source of this mismatch could be the fact that for France three friendly games against Iceland, Estonia and Serbia played between 27/05 and 5/06/2012 appear to not have been included in FIFA ranking points computations.

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

Table 10. One step optimal decision.

Rank	**Greece (10)**	**Mali (27)**	**Haiti (61)**
10	**0.0909**	0.0361	0.0055
11	**0.3117**	0.182	0.0425
12	**0.5928**	0.4626	0.1719
13	**0.7819**	0.7427	0.4175
14	0.8626	**0.898**	0.7004
15	0.9042	**0.9518**	0.8951
16	0.945	0.9747	**0.9763**

probability 0.60). If 3 games are to be played, it is optimal to play two games against Greece and one against Bosnia and Herzegovina (29). This yields a probability of 0.70. A four step optimal decision would have been to play against Algeria (24), Bosnia and Herzegovina, Greece and Mali. Such a decision yields probability of 0.71 of being ranked at 12 position or higher. On the other hand, the schedule of games played by Brazil as shown in Table 9 yields probability of merely 0.17 of advancing at least at rank 12.

6. Discussion

In this section, under consecutive headings, we discuss main points of our study – the choice of number of games to play, risk associated with our decisions and mutual agreements between teams that are necessary for a game to take place.

6.1 *Choosing the number of games to play*

A team can choose to schedule several games to achieve its target. In our experiments we note that the choice of the number of games can have different impact on probability of achieving given rank. For example, in case of Hungary it turned out that the best strategy is to play a single game. On the other hand, as we have seen in the case of Brazil, it is profitable in terms of increasing probability to play more games. Extending the schedule to several games introduces the chance of a loss, which could result in losing ranking points. In case of Brazil, the probability of winning a single game is high enough so that for this team it is better to play more games.

The discussed case of Polish national team shows that playing more games opens more possibilities to advance in the ranking table. Attaining a certain rank may not be possible by taking only one step. Extending schedule to more games opens possibilities to climb the ranking table up with higher probability.

As far as the Ukraine's case is concerned, we note that when Ukraine abstains from playing any friendly match then with estimated probability of 0.50 it can gain a rank among top 18 UEFA teams when we consider scheduled matches of Ukraine's opponents. On the other hand, when Ukraine challenges Wales, which is their one step optimal choice, this probability is equal 0.44. Although this probability is lower, worst–case analysis reveals that only a win against Wales would secure Ukraine spot among top 18 teams. In fact, the results of other games turned out to be not favourable for Ukraine and to qualify for the second pot they needed a win in a friendly game.

6.2 *Risk tolerance*

Teams may have different attitude towards the risk associated with playing a game. If a team wants to secure its rank, then it can play against a lower ranked team and win with high probability as usually it is a weaker opponent. Such a decision, however, limits

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

the possibilities of advancing at a higher rank.

We shall note that in order to climb up the ranking table the team actually needs to *win* its scheduled games. Due to low importance factor of friendly games (multiplier I_k in Equation (3)) even a draw in a match can be treated as a loss (for a draw a team gains three times less points than for a win). Due to this fact, the calculated probabilities are most influenced by the ability of a team to win the game. In this way, playing more games becomes a more risky strategy for attaining certain ranks. As we have seen in example of Hungarian national team, increasing number of games makes attaining certain rank more risky. On the other hand, taking higher risk is necessary for attaining higher ranks.

6.3 *Coalition of teams improving their rank*

Knowing how to exploit the ranking system, teams may also form coalitions in order to inflate their ranking points and advance in the ranking table. We elaborate on the idea, suggested by Wang and Vandebroek [40], below.

Let us assume that two low ranked (below rank 150) teams form a coalition and agree to win and loose half of a large number friendly encounters between them. In limiting case each team will obtain an average of $\frac{3 \cdot 50 + 0}{2} = 75$ points a year (150 points for a win in a friendly game against an opponent ranked lower than at position 150 and 0 points for a loss; we assume that the confederation strength factor in Formula (3) is equal 1). After 4 years they could accumulate $75 \cdot (1 + 0.5 + 0.3 + 0.2) = 150$ ranking points. With that many points they might be ranked higher than the 150th place. We may ask what would be the average position of a team participating in such a coalition. Denoting the rank of a team with n and its ranking points with $P_{(n)}$ we would have the following approximate relation in the long run: $\frac{3 \cdot \max(200-n, 50)+0}{2} \cdot (1+0.5+0.3+0.2) \approx P_{(n)}$. Based on historical data, we may find values of n for which the above equation holds. In mean absolute deviation sense it is most closely met by the values of $n \approx 150$. We conclude that forming such coalitions would be profitable only for low ranked teams which would place them around the 150th position. Higher ranked teams do not have incentive to implement such strategies. Moreover, such alliances are easy to detect when composed of small number of teams.

Let us also discuss how the Elo rating system would behave under such a manipulation. Let us focus our attention on the EloRatings.net system [9]. Similar considerations can be drawn based on the analysis of the Elo rating method employed in the FIFA Women World Rankings [10]. First of all, under the Elo model, in an encounter between two teams, the winning team gains the same amount of points as the other team loses. This implies that, in a coalition, the teams cannot artificially produce points as in the case of the FIFA ranking considered above. Changes in rating points sum up to zero so that there is no added value in a coalition. Secondly, we can expect that the team rating points would converge in a sequence of wins and losses. This is because a higher rated teams out of the two involved in a coalition would gain less ranking points for a winning the match than it can lose to their opponents in absolute terms. Hence, after consecutive matches the teams' ranking points, and in turn their position in the ranking table, shall converge to the same value.

7. Conclusions

Association football is probably the most popular sport in the world as well as a huge industry with millions at stake. Efficacy and robustness of the official ranking system employed by FIFA is of critical importance for fair competition of the involved parties.

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

In this paper we investigated the efficiency of the official FIFA ranking for football teams around the world, and how it could be exploited by scheduling friendly matches in a smart way. This is based on previous research that had established that the official ranking methodology may not award ranking points in the most efficient manner. We defined targets of teams to climb in the ranking, and via our experiments, we showed how opportunities for climbing up the ranking table may be exploited. We quantified opportunities of advancing in the table in terms of associated probabilities. We also discussed other features of the FIFA ranking, i.e., lack of incorporation of the home team advantage and the way it incorporates time into calculation. We discussed what is the optimal choice for a team in terms of the number of games to play and associated risk. We also discussed the robustness of the official procedure to forming coalitions of teams working for improvement of their positions. We compared the FIFA ranking to Elo ranking system, also used in ranking national football teams.

Our proposed methods for improving a team's rank may be employed by national football teams, which may result in better seedings for these teams, and better starting positions for major tournaments. Due to the low importance of the friendly games, it may require patience and a series of good results to climb high in the ranking, however, even a single game can suffice to influence a team's position. In the near future, there are expected changes in the competition rules, e.g. the plans to introduce the Nations League by UEFA (the administrative body for football in Europe) may limit opportunities for playing friendly games and, hence, limit the applicability of optimisation strategies.

In this study we focused on practical aspects of finding optimal decision to play a match. One may consider a situation in which the teams form a market where they offer playing a friendly match. Each team would have its intrinsic targets that would define its strategy of scheduling games. Stability of formed pairs of teams may be investigated. A further study of such scenarios may lead to new results in a game-theoretic framework.

Acknowledgement. Jan Lasek would like to acknowledge the support by the European Union from resources of the European Social Fund, Project PO KL "Information technologies: Research and their interdisciplinary applications", agreement UDA-POKL.04.01.01-00-051/10-00 via the Interdisciplinary PhD Studies Program.

References

[1] A. Altman, *The Axiomatic Approach to Ranking Systems*, Ph.D. thesis, Israel Institute of Technology, Israel, 2007.

[2] C. Bartneck and S. Kokkelmans, *Detecting h-index manipulation through self-citation analysis*, Scientometrics 87 (2011), pp. 85–98.

[3] A.C. Constantinou, *Bayesian networks for prediction, risk assessment and decision making in an inefficient association football gambling market*, Ph.D. thesis, Queen Mary, University of London, United Kingdom, 2012.

[4] A.C. Constantinou, N.E. Fenton, and M. Neil, *pi-football: A Bayesian network model for forecasting association football match outcomes*, Knowledge-Based Systems (2012), pp. 322–339.

[5] M. Crowder, M. Dixon, A. Ledford, and M. Robinson, *Dynamic modelling and prediction of english football league matches for betting*, Journal of the Royal Statistical Society, Series D (The Statistician) 51 (2002), pp. 157–168.

[6] D. Dagaev and K. Sonin, *Winning by losing: Incentive incompatibility in multiple qualifiers* (2013), available at SSRN: http://dx.doi.org/10.2139/ssrn.2225463, last access date: 16 January 2015.

[7] M.J. Dixon and S.G. Coles, *Modelling asssociation football scores and inefficiencies in the football betting market*, Applied Statistics 46 (1997), pp. 265–280.

[8] J. Du, Z. Cai, and Y. Chen, *A Sorting Based Algorithm for Finding a Non-dominated Set in Multi-objective Optimization*, in *Proceedings of the Third International Conference on Natural Computation - Volume 04*, IEEE Computer Society, Washington, DC, USA, 2007, pp. 436–440.

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

[9] EloRatings.net, *The World Football Elo Rating System* (2015), `http://www.eloratings.net/system.html`, Last access date: 16 January 2015.

[10] FIFA.com, *FIFA/Coca-Cola Women's World Ranking Procedure* (2015), `http://www.fifa.com/mm/document/fifafacts/r&a-wwr/52/00/99/fs-590_06e_wwr-new.pdf`, Last access date: 16 January 2015.

[11] FIFA.com, *FIFA/Coca-Cola World Ranking Procedure* (2015), `http://www.fifa.com/mm/document/fifafacts/r&a-wr/52/00/97/fs-590_10e_wrpoints.pdf`, Last access date: 16 January 2015.

[12] E. Friedman, P. Resnick, and R. Sami, *Manipulation–Resistant Reputation Systems*, in *Algorithmic Game Theory*, Cambridge University Press, 2007.

[13] M. Gagolewski, *Scientific impact assessment cannot be fair*, Journal of Informetrics 7 (2013), pp. 792–802.

[14] J. Goddard, *Regression models for forecasting goals and match results in association football*, International Journal of Forecasting 21 (2005), pp. 331–340.

[15] F. Halicioglu, *Research on the prediction of the likely winners of the Euro 2008 football tournament*, Journal of Quantitative Analysis in Sports 5 (2009), pp. 1–14.

[16] F. Halicioglu, *An empirical study of relationship between FIFA world ranking and domestic football competition level: the case of Turkey*, in *Proceedings of International Congress on Sports Economics and Management*, Ege University, Turkey, 2012, pp. 44–58.

[17] J.E. Hirsch, *An index to quantify individual's scientific research*, in *Proceedings of the National Academy of Sciences*, Vol. 102, 2005, pp. 16569–16572.

[18] A. Joseph, N.E. Fenton, and M. Neil, *Predicting football results using bayesian nets and other machine learning techniques*, Knowledge-Based Systems 19 (2006), pp. 544–553.

[19] D. Karlis and I. Ntzoufras, *Analysis of sports data by using bivariate Poisson models*, The Statistician 52 (2003), pp. 381–393.

[20] D. Karlis and I. Ntzoufras, *Bayesian modelling of football outcomes: using the Skellam's distribution for the goal difference*, IMA Journal of Management Mathematics 20 (2009), pp. 133–145.

[21] R.H. Koning, *Balance in competition in Dutch soccer*, The Statistician 49 (2000), pp. 419–431.

[22] S.J. Koopman and R. Lit, *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) 178 (2015), pp. 167–186.

[23] H.T. Kung, F. Luccio, and F.P. Preparata, *On finding the maxima of a set of vectors*, Journal of the Association for Computing Machinery 22 (1975), pp. 469–476.

[24] A.N. Langville and C.D. Meyer, *Google's PageRank and Beyond: The Science of Search Engine Rankings*, Princeton University Press, 2006.

[25] J. Lasek, Z. Szlavik, and S. Bhulai, *The predictive power of ranking systems in association football*, International Journal of Applied Pattern Recognition 1 (2013), pp. 27–46.

[26] M.J. Maher, *Modelling association football scores*, Statistica Neerlandica 36 (1982), pp. 109–118.

[27] I. McHale and S. Davies, *Statistical analysis of the effectiveness of the FIFA World Rankings*, in *Statistical Thinking in Sports*, J. Albert and R. Koning, eds., Chapman & Hall/CRC, Boca Raton, Florida, 2007, pp. 77–90.

[28] A.M. Nevill and R.L. Holder, *Home advantage in sport*, Sports Medicine 28 (1999), pp. 221–236.

[29] R. Pollard, *Home advantage in football: A current review of an unsolved puzzle*, The Open Sports Sciences Journal 1 (2008), pp. 12–14.

[30] R. Pollard, C.D. da Silva, and C.M. Nisio, *Home advantage in football in Brazil: differences between teams and the effects of distance traveled*, The Brazilian Journal of Soccer Science 1 (2008), pp. 3–10.

[31] A. Rajaraman and J.D. Ullman, *Link analysis*, in *Mining of Massive Datasets*, Cambridge University Press, 2012.

[32] H. Rue and O. Salvesen, *Prediction and retrospective analysis of soccer matches in a league*, Journal of the Royal Statistical Society: Series C (The Statistician) 49 (2000), pp. 399–418.

[33] T.C. Russell, *A computational study of problems in sports*, Ph.D. thesis, University of Waterloo, Canada, 2010.

[34] P.A. Scarf, M. Yusof, and M. Bilbao, *A numerical study of designs for sporting contests*, European Journal of Operational Research 198 (2009), pp. 190–198.

[35] A. Seckin and R. Pollard, *Home advantage in Turkish professional soccer*, Perceptual and Motor Skills 107 (2008), pp. 51–54.

[36] I. Stanton and V. Vassilevska Williams, *The structure, efficacy and manipulation of double-elimination tournaments*, Journal of Quantitative Analysis in Sports 9 (2013), pp. 319–335.

[37] R.T. Stefani, *Survey of the major world sports rating systems*, Journal of Applied Statistics 24 (1997), pp. 635–646.

Please cite this paper as: Lasek J., Szlavik Z., Gagolewski M., Bhulai S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 2016, pp. 1349–1368, doi:10.1080/02664763.2015.1100593.

[38] R.T. Stefani, *The methodology of officially recognized international sports rating systems*, Journal of Applied Statistics 7 (2011), pp. 1–22.

[39] J. Van Haaren and G. Van den Broeck, *Relational learning for football-related predictions*, in *Latest Advances in Inductive Logic Programming, International Conference on Inductive Logic Programming*, 2011.

[40] C. Wang and M. Vandebroek, *A model based ranking system for soccer teams*, Tech. Rep., KU Leuven - Faculty of Economics and Business, 2013.

[41] L.A. Zhivotovsky and K.V. Krutowsky, *Self-citation can inflate h-index*, Scientometrics 77 (2008), pp. 373–375.