A simulation model for soccer championships

Koning, Ruud H.; Koolhaas, M.P.; Renes, G.; Ridder, G.

A Simulation Model for Football Championships

Ruud H. Koning* Michael Koolhaas Gusta Renes Geert Ridder

SOM Theme A: Primary processes within firms

Abstract

In this paper we discuss a simulation/probability model that identifies the team that is most likely to win a tournament. The model can also be used to answer other questions like 'which team had a lucky draw?' or 'what is the probability that two teams meet at some moment in the tournament?'. Input to the simulation/probability model are scoring intensities, that are estimated as a weighted average of goals scored. The model has been used in practice to write articles for the popular press, and seems to perform well.

Keywords: Poisson models, football, simulation

1 Introduction

Many people are interested in predicting the outcome of sporting contests. However, one of the reasons that sports attracts so much attention is that the outcome of a contest is not perfectly predictable. Even experts can, after an objective assessment of the relative strength of the contestants, only guess the outcome. In this paper we report our experience with a simulation model for major soccer tournaments, as the World Cup and the European Cup championships. This model uses the outcome of matches that were played in the years before the championship to predict the probability of outcomes of matches during the tournament.

The World Cup for soccer is the sports tournament with the biggest audience. It generates even more attention than the summer Olympics. The continental soccer tournaments as the European Cup, the Copa America and the African Cup are also major events. The simulation model that we have developed has been used to predict the likely outcome of the World and the European Cup, but the methodology can also be used in other tournaments. The World Cup and the European Cup are played every four years

*Corresponding author, Department of Econometrics, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands, email: ruud@soccer-research.com.

by teams that represent the countries that have qualified for that tournament. These teams consist of the best players each country can field; the players themselves play for different club teams, often in countries of which they are not citizens. National teams play fewer matches than club teams. Most teams have to qualify for each tournament (except the organizing country or countries and for the World Cup the winner of the last tournament). The rules for qualification are different for each tournament, but roughly a national team qualifies if it comes first in a full competition of up to six national teams from the same continent. The composition of these qualification groups is determined by a draw, where teams of some countries are seeded and others are not. Beside these qualification matches a national team may play a few friendly or practice matches each year. Because in the years before a major tournament not all teams play each other, it is difficult to assess the strength of the teams that participate in the tournament. Hence we are faced with a problem that is familiar to statisticians: missing data. Even if we knew the strength of each team, we still need a model that translates this into the probability that a certain team wins the tournament. After all, chance plays an important role in football (just as in most other sports).

In this paper we report the experience of the Dutch 'Werkgroep Voetbal& Statistiek'[1], which developed a model that has been used to predict the probability that a national team wins the tournament. The method that we use consists of two parts. First, for all, for all participants in a tournament, scoring intensities (the expected number of goals in a complete match) are estimated. These scoring intensities are match-specific. For example, the expected number of goals scored by The Netherlands against France, is different from the expected number of goals scored by The Netherlands against, say, San Marino. This information is then used as input for a simulation model that computes the probability that each team wins the tournament.

Section 2 discusses data and the estimation method of the scoring intensities. The simulation model is the topic of section 3. Our experiences with the model are discussed in section 4. We end with some conclusions.

2 Data and the Estimation of Scoring Intensities

There are relatively few matches between national teams in football, even though the main tournaments for international teams (World Cup, European Cup, Copa America) are considered to be very important by players and fans. The European football association UEFA has 51 members, so it is not feasible to organize a complete competition to determine which country is the best[2]. Football players are paid by football teams, and they are

[1] The 'Werkgroep Voetbal & Statistiek' consists of the authors of this papers. At the time the model was developed Peter Hopstaken and Ronald van Gelder also participated. More information about this group can be found on www.soccer-research.com.

[2] In this paper we will use the European Cup 2000, also known as EURO2000, as a leading example.

reluctant to let their stars play many international matches. Hence, the European Cup is played in two stages: there is a qualification tournament, and a main tournament. The qualification tournament consists of nine groups, and each group plays a full competition. The winners of each group qualify, and the runners-up have to play a play-off match. Because all countries except the countries that host the tournament (and the reigning champion in case of a World Cup) have to participate in the qualification tournament, there is little room for friendly matches against other strong opponents. Hence, there is sparse evidence that can be used to determine which country has a strong team and which country has a weak team. Of course this problem becomes even more acute when predicting outcomes of a World Cup, as there are hardly any matches between national teams from countries from different continents.

To measure the quality of each country's team, we have first gathered data on international matches, starting in 1960. When this paper was written, the database consists of 8190 matches between national teams. These matches are of three types: friendly matches, qualification matches, and matches played during a tournament. Of course, the number of matches per country varies in the database: it has over 200 matches for the national teams of for example, The Netherlands, Italy, and England, while it has only a few matches of countries that obtained independence recently such as Macedonia. Not all matches are relevant when assessing the strength of a country. It is reasonable to base the estimation of scoring intensities on matches played during the last two or four years only.

In the simulation model of section 3, we assume that the number of goals scored by a team follows a Poisson distribution. The parameter of this distribution is specific to each match, so the parameter of the Poisson distribution that models the numbers of goals by team i against team j is denoted by λ_{ij}^{H}. The superscript indicates that team i plays at home. The same parameter in an away match is denoted by λ_{ij}^{A}.

Figure 1 displays the frequency distribution of the number of goals scored by the home teams during the qualification matches for EURO2000. The average number of goals scored by the home teams is 1.52 (i.e., an average of approximately one goal per hour of playing time), and a Poisson distribution with that parameter has been superimposed in figure 1. Clearly, the fit of the Poisson distribution to the actual distribution is not very good: the actual distribution of number of goals scored has a peak at zero, and has a tail that is fatter than the one of the Poisson distribution. This is not surprising, as the parameter of the Poisson distribution is an average of different scoring intensities. The average number of goals scored in home matches ranges from 0.2 (Andorra) to 7.25 (Spain). Clearly, it is not reasonable to assume that each team has the same scoring intensity. The groups in the qualification consist of both strong and weak teams, and hence it is to be expected that an approximation that assumes identical scoring intensities for each team is not satisfactory.

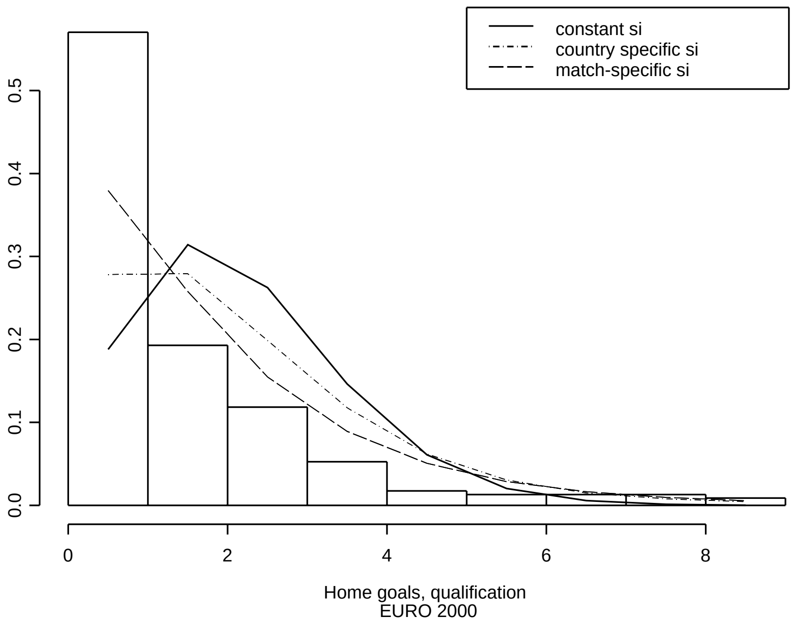

Figure 1: Frequency distribution of goals scored in home matches during qualification for EURO2000 with constant, country specific, and match-specific scoring intensities.

A straightforward extension is to assume that each country has a constant scoring intensity in home matches: $N_{ij} \sim \mathcal{P}(\lambda_i)$. Note that the scoring intensity of team i is independent of the opponent j. λ_i can simply be estimated by the average number of home goals by team i. The resulting frequency distribution over all teams is given in figure 1 by the dotted/dashed line. The peak at 0 is still underpredicted, but much less so than by the solid curve, and the tail fits the actual distribution better.

A further sophistication is the use of match-specific scoring intensities. In this approach, λ_{ij} varies with team j. Team i is expected to score more if team j is a weak team than if team j is a strong team. We propose two such match-specific scoring intensities. First, consider a match between Germany and England. During the period covered by the dataset, England conceded on average 0.5 goals per match. Germany played Finland during the qualification tournament for EURO2000, and won by 2-0. During the sample period, Finland concedes on average 1.625 goals per match. Based on the number of goals scored by Germany against Finland, we expect Germany to score $2 \times \frac{0.5}{1.625} = 0.615$ against England. We expect Germany to score fewer goals against England because their defense is better than the defense from Finland. Of course, Germany played more opponents than only Finland, so we average over all opponents played by Germany.

In the second approach to match-specific scoring intensities, we focus on the offensive capabilities of both teams. Consider again a match between Germany and England. Bulgaria played a 1-1 draw against England. Germany scored on average 2.5 goals per match, and Bulgaria scored on average 0.75 goals per match. Therefore, based on the goals scored against England, one would expect $1 \times \frac{2.5}{0.75} = 3.33$ goals of Germany against England.

These match-specific estimators are just weighted averages. In the first case it is a weighted average of the goals scored by team i, and the weight is the relative quality of the defense of team j. The marginal distribution of goals scored calculated this way is drawn as the dahsed line in figure 1. It fits the observed distribution better than the other two models: it captures the mode at 0 better, and the fit in the tail is also better. The second proposed estimator for match-specific scoring intensities is a weighted average of the goals conceded by team j, with the relative quality of team i's offense as weight. The marginal distribution of goals scored in that case deviates only marginally from the dashed line in figure 1.

The different estimators for λ_{ij} are listed in table 1. In that table, K_i^H denotes the number of home matches played by team i, and N_{ij}^H are the number of goals scored by team i against team j in a home match. $\lambda_{\cdot i}^H$ is the average number of goals conceded by team i in home matches, and $\lambda_{i\cdot}^H$ is the average number of goals scored by team i in home matches. The estimators (2), (4), (6), and (8) take only home advantage into account in the sense that these are based on home results of team i only. The other estimators are based on both home and away results. Note that the summation

(1) $\hat{\lambda}^{(1)}_{ij} = \frac{1}{K^H+K^A} \sum_{i,j} \left(N^H_{ij} + N^A_{ij} \right)$

(2) $\hat{\lambda}^{(2)}_{ij} = \frac{1}{K^H} \sum_{i,j} N^H_{ij}$

(3) $\hat{\lambda}^{(3)}_{ij} = \frac{1}{K^H_i+K^A_i} \sum_k \left(N^H_{ik} + N^A_{ik} \right)$

(4) $\hat{\lambda}^{(4)}_{ij} = \frac{1}{K^H_i} \sum_k N^H_{ik}$

(5) $\hat{\lambda}^{(5)}_{ij} = \frac{1}{K^H_i+K^A_i} \left(\sum_k N^H_{ik} \frac{\lambda^A_{\cdot j}}{\lambda^A_{\cdot k}} + \sum_k N^A_{ik} \frac{\lambda^H_{\cdot j}}{\lambda^H_{\cdot k}} \right)$

(6) $\hat{\lambda}^{(6)}_{ij} = \frac{1}{K^H_i} \sum_k N^H_{ik} \frac{\lambda^A_{\cdot j}}{\lambda^A_{\cdot k}}$

(7) $\hat{\lambda}^{(7)}_{ij} = \frac{1}{K^H_j+K^A_j} \left(\sum_k N^H_{kj} \frac{\lambda^H_{i\cdot}}{\lambda^H_{k\cdot}} + \sum_k N^A_{kj} \frac{\lambda^A_{i\cdot}}{\lambda^A_{k\cdot}} \right)$

(8) $\hat{\lambda}^{(8)}_{ij} = \frac{1}{K^A_j} \sum_k N^H_{kj} \frac{\lambda^H_{i\cdot}}{\lambda^H_{k\cdot}}$

Table 1: Estimators for the scoring intensity λ^H_{ij}.

index k may refer to a particular match (in N_{ik}), or to a country (in $\lambda_{\cdot k}$). This difference is important when two countries have played against each other more than once. These matches enter separately in the summations in table 1.

Of course, one can think of variants of the match-specific estimators. Instead of all opponents of a team, one could for example take a weighted average over all teams which have played against both team i and team j. Moreover, the weights $\frac{\lambda^A_{\cdot j}}{\lambda^A_{\cdot k}}$ and $\frac{\lambda^H_{\cdot j}}{\lambda^H_{\cdot k}}$ could be replaced by $\frac{\lambda_{\cdot j}}{\lambda_{\cdot k}}$ so that the weights in $\hat{\lambda}^{(5)}_{ij}$ are based on all matches and are independent of the venue of a particular match.

In order to implement the estimators in a particular case, a few practical issues have to be addressed. First, one needs an appropriate dataset. Matches that have been played too long ago are not very informative on the current quality of a team. When predicting likely results for a particular championship, we have used the matches played during the two years before that championship. Older players tend to leave a national team after a major tournament, and national coaches try to keep a group together during both the qualification matches and the tournament. A second decision is whether or not to include friendly matches in the dataset. On the one hand, friendly matches provide information about the quality of the teams. On the other hand, the best players are not always fielded in a friendly match. A third issue is the problem of outliers. In the estimation of match-specific scoring intensities we need estimates of the marginal scoring intensity $\lambda_{i\cdot}$ and its defense counterpart $\lambda_{\cdot j}$. The first may be overestimated if a team plays very weak teams, and the latter may be underestimated in that case. Also, highly unusual results have to be removed from the database (The Netherlands-Belgium 5-5 in September 1999 is an exceptional result; it does not reflect both the offensive and the defensive skills of both teams). In the end, we aim to estimate scoring intensities for a major tournament, where teams of more or less similar quality meet each other. Finally, we have to decide whether or not to incorporate home advantage. For all teams except the team from the organizing country (or countries, as EURO2000 was organized jointly by Belgium and The Netherlands), there is no home advantage. Simulation results, to be discussed in the next section, are quite insensitive to allowing for home advantage of the organizing country. In most applications we have used $\lambda_{ij}^{(5)}$ as input to the simulation model, this is the average of goals scored by team i, weighted with the relative quality of team j's defense. We choose this estimator because the fit to the observed data as measured by $\sum_{i,j}\left(\lambda_{ij}^{(5)} - n_{ij}\right)^2$ was best among all the estimators of table 1.

3 The Simulation Model

Major soccer tournaments (Euro2000, but also the World Cup tournaments in France (1998) and the US (1994) and the European Championship in England (1996)) consist of two phases. In the first phase, groups of four teams play a half competition. The teams that classify first and second proceed to the second round. In that second round the team that ends first in a group plays the team that ends second in another group. The schedule is such that the teams that classify as first and second in a group can only play each other again in the final[3].

The second phase of the tournament is a knock-out tournament. After each match the winning team proceeds to the next round and the losing team is eliminated. A team can win a match in three ways. It may win the match after regular playing time of 90 minutes. If the match is tied after 90 minutes, extra time is played with a maximum of 30 minutes. The team that scores first during this extra time proceeds to the next round, this is called the 'golden goal'-rule. If the match is still tied after the extra time, each team takes penalty kicks until one team wins. Let N_{ij} be the number of goals scored by team i against team j during 90 minutes of play, and let T_{ij} be the waiting time until team i scores during the extra time. The probability that team i proceeds to the next round is then given by

$$\begin{aligned}\Pr(i \text{ beats } j) &= \Pr(N_{ij} > N_{ji}) + \Pr(T_{ij} < T_{ji}, N_{ij} = N_{ji}, T_{ij} < 30) + \\ &\frac{1}{2}\Pr(N_{ij} = N_{ji}, T_{ij} > 30, T_{ji} > 30). \qquad (1)\end{aligned}$$

[3]In the European Cup of 1988, the Soviet-Union won group 2, and the Netherlands ended second in that group. Both teams proceeded to meet again in the final of that tournament.

It is assumed that N_{ij} follows a Poisson distribution with parameter λ_{ij}. The assumption that the number of goals scored in football matches follows a Poisson distribution is made frequently, see for example Maher (1982) and Koning (2000). In these two papers, the scoring intensity is constant throughout the match. It is known, though, that the scoring intensity increases during a match, see Ridder, Cramer, and Hopstaken (1994) and Dixon and Robinson (1997). These inhomogeneous Poisson models require more detailed data than we have available, so we assume that goals are scored according to a homogeneous Poisson process. The scoring intensity during 30 minutes of extra time is $\lambda'_{ij} = \frac{30}{90}\lambda_{ij}$, so T_{ij} follows an exponential distribution with that parameter. The second probability in equation (1) is calculated by:

$$\begin{aligned}
&\Pr(T_{ij} < T_{ji}, N_{ij} = N_{ji}, T_{ij} < 30) = \\
&\quad \Pr(T_{ij} < T_{ji}, T_{ij} < 30 | N_{ij} = N_{ji}) \Pr(N_{ij} = N_{ji} =) \\
&\quad \left(1 - \exp(-\lambda'_{ji}) - \frac{\lambda'_{ji}}{\lambda'_{ij} + \lambda'_{ji}} \left(1 - \exp(-(\lambda'_{ij} + \lambda'_{ji}))\right)\right) \times \\
&\quad\; \Pr(N_{ij} = N_{ji})
\end{aligned} \tag{2}$$

We assume that both teams are equally skilled at taking penalty shots. The outcome of the shoot out is independent of the scoring intensities.

The probability that a team wins the tournament is now determined as follows. The first phase of the tournament consists of the group matches. The result of a group-match between countries i and j is a realization of (N_{ij}, N_{ji}), with $N_{ij} \sim \mathcal{P}(\lambda_{ij})$. Using the simulated results of all matches in a group, we get a ranking and we know which teams proceed to the second phase of the tournament. From now on, we use the probabilities from (1) to calculate the probability that a team proceeds to the next round. Suppose for simplicity that the second phase of the tournament consists of three rounds: a quarter-final, a semi-final, and a final, so that we have a situation as shown in figure 2.

The winner of group A is labeled A1, the runner's-up of that group is labeled A2, etc. According to the schedule (which is known before the tournament starts), the winner of group A plays the runner's-up of group D in the quarter final. The winner of that match meets the winner of the match between B1 and C2, etc. Once we have simulated the results from the matches in phase one, we know exactly which teams are A1, A2, etc. Using the probabilities of equation (1) we compute the probability that A1 or D2 and B1 or C2 etc. reaches the semi final, e.g.

$$\Pr(\text{A1 reaches SF}) = \Pr(\text{A1 beats D2}).$$

Next,

$$\begin{aligned}
\Pr(\text{A1 reaches F}) &= \Pr(\text{A1 reaches SF and A1 wins SF}) \\
&= \Pr(\text{A1 reaches SF}) \times \Pr(\text{A1 wins SF}),
\end{aligned}$$

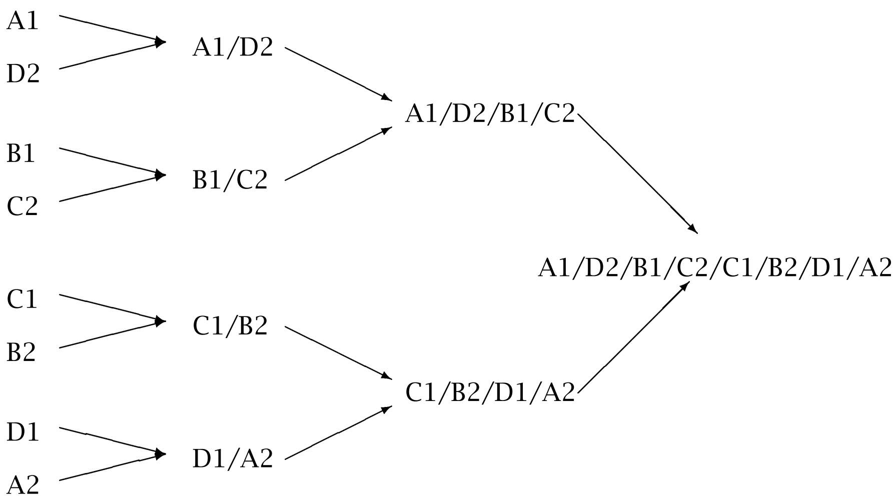

Figure 2: Second phase of the tournament.

by the assumption that outcomes of matches are independent. Now by the law of total probability,

$$\Pr(\text{A1 wins SF}) = \sum_j \Pr(\text{A1 beats} j \text{ and } j \text{ reaches QF})$$

with the summation taken over all possible opponents in the semi-final (in this case, B1 and C2), and again by the independence of outcomes of matches we have that

$$\begin{aligned} &\Pr(\text{A1 beats} j \text{ and } j \text{ reaches QF}) \\ &\quad = \Pr(\text{A1 beats} j) \times \Pr(j \text{ reaches SF}). \end{aligned}$$

Combining these equations we have a recursive relation for the computation of $\Pr(\text{A1 reaches F})$. In the same way we can compute the probability that A1 wins the tournament.

The probability of winning the tournament calculated so far is conditional on the ranking of the first phase. If each group in the first phase consists of K_1 teams, there are K_2 groups, and the best k teams of each group qualify for the second stage, there are in total $(K_1 \cdot (K_1-1) \cdots (K_1-k+1))^{K_2}$ different possible rankings that are used as 'starting values' for the second phase of the tournament[4]. Letting the rankings be indexed as R_l, the unconditional probability that team i wins the tournament is

$$\Pr(i \text{ wins F}) = \sum_l \Pr(i \text{ wins F} | R = R_l) \Pr(R = R_l).$$

[4]In the World Cup tournament 1998 in France, each group consisted of four teams, two of which qualified for the second stage. There were eight groups, so the total number of possible starts of the second phase is $(4 \cdot 3)^8 \approx 4.3 \cdot 10^8$.

The probabilities $\Pr(R = R_l)$ and the different rankings R_l are cumbersome to calculate, so instead we estimate this probability by simulation:

$$\Pr(i \text{ wins F}) = \frac{1}{S}\sum_s \Pr(i \text{ wins F}|R = R_s)$$

where R_s is a ranking obtained by simulation and S is the number of simulations.

The simulation model of this section mimics the rules of the tournament exactly. Changes in the rules can be incorporated without any difficulty. For example, during the World Cup in 1994, extra time—if any—lasted 30 minutes, independent of the numbers of goals scored during the extra time. The only adjustment needed to allow for that rule is to change (1) to

$$\begin{aligned}\Pr(i \text{ beats } j) &= \Pr(N_{ij} > N_{ji}) \\ &\quad + \Pr(N'_{ij} > N'_{ji}, N_{ij} = N_{ji}) + \frac{1}{2}\Pr(N_{ij} = N_{ji}, N'_{ij} = N'_{ji})\end{aligned}$$

with $N'_{ij} \sim \mathcal{P}(\lambda'_{ij})$.

In the next section we discuss our experiences with this model.

4 Applications of the Simulation Model

Based on the model, we can calculate, for each participating team, the probability that it wins the tournament and hence identify the most likely winner of the tournament. In table 2 we list the three countries with the largest probability of winning the title, the winner, and the losing finalist. For four major tournaments, these predictions have appeared before each tournament in the popular press (Van Gelder, Hopstaken, Koning, Koolhaas, Ridder, and Renes (1994), Werkgroep Voetbal & Statistiek (1996), Werkgroep Voetbal & Statistiek (1998)).

In the interpretation of the probabilities in table 2, one should note that 24 (USA '94) and 32 countries (France '98) participated in the World Cups, and 16 countries participated in the European Cups. If the winner would have been picked at random, the winning probabilities would have been 0.04, 0.03, and 0.06 respectively. In the World Cup of 1994 Brazil was three times as likely to win as by chance alone. The probabilities of winning the tournament are small for each the favorites are small. This is to be expected, because by the very selection of teams that participate in a World Cup or European Cup, one does not expect great differences in playing abilities.

The model seems to do a reasonable job of indicating the countries that are likely to win the title, except for the European Cup in 1996. The World Cup of 1994 and the European Cup of 2000 were won by the countries that were most likely to do so. Also, five out of eight quarter finalists were predicted correctly for EURO2000, and similar results hold for other tournaments.

Tournament	Winning probabilities	Winning finalist	Losing finalist
World Cup '94	Brazil (0.12) Spain (0.11) Italy (0.08)	Brazil	Italy
European Cup '96	Spain (0.17) France (0.15) Croatia (0.11)	Germany	Czech Republic
World Cup '98	England (0.19) Brazil (0.18) France (0.10)	France	Brazil
European Cup '00	France (0.18) Spain (0.12) Portugal (0.11)	France	Italy

Table 2: Winning probabilities per tournament.

During the tournament it is simple to update the winning probabilities, by replacing simulated results and probabilities of the second phase by realized results. The scoring intensities can also be re-estimated after more matches become available. Moreover, counterfactual questions like 'what would have been the probability that Italy would have won EURO2000 had Portugal been the opponent in the final' can be answered easily. Also, intermediate results like the probability of reaching the quarter final, or the probability that any two teams meet at some moment during the tournament can be derived from the model. Using that probability, one can calculate that the probability that Spain would win the tournament conditionally on reaching the second phase is 0.16. The same probability for Portugal is 0.17, which indicates that Spain is the second most likely country to win the tournament because they have a relatively high probability of qualifying for the second phase.

In another experiment, we used the scoring intensities of section 2 to simulate a full competition between all the teams that participated in EURO 2000. A full competition is considered by most people to be the fairest way of determining the champion. Of course, a full competition is not feasible in practice because of the number of matches it would take. Table 3 gives the results for EURO2000. Note that France is twice as likely as Sweden to win the tournament, but this ratio is 31 if a full competition would be played. Note also how the winning probability of Spain drops: it is the second most likely winner of the tournament, but it is not expected to end high in a full competition. The reason is that the draw of the first phase of

Country	Winning probability	Full competition	Actual result
France	0.18	0.31	winning finalist
Spain	0.12	0.05	quarter-finalist
Portugal	0.11	0.27	semi-finalist
England	0.10	0.04	group
Sweden	0.085	0.01	group
Italy	0.079	0.09	losing finalist

Table 3: Full competition simulation results for EURO2000.

the tournament was favorable to Spain. The probability that it proceeds to the second round is 0.75, which is very high compared to the same probability of other countries. After reaching the second phase when only eight teams are left in the tournament, the winning probability of Spain is 0.10, lower than the same probability of Italy, which is 0.17. Note that this probability of Spain differs from the conditional probability mentioned earlier (0.16). The reason is that the conditional probability of winning the tournament given qualification for the second phase is an average over all possible schedules of that second phase. The winning probability of Spain of 0.10 is for the *realized* second phase, when all quarter finalists are known. Of course, one can estimate scoring intensities and simulate a full competition at any moment, not only when a tournament is played. The methodology of this paper can thus be used to determine a ranking of national football teams[5]. An ranking of national teams is published by the FIFA (the world soccer association), but that ranking is based on a complex set of rules and it lacks the simple interpretation of being the result of a simulated full comeptition. A ranking obtained by simulating a full competition between all teams, could also be used to seed the best teams in the first stage of a tournament.

Another possible experiment with the model is to determine the effect of the selection of groups in the first phase. The effect of the selection can be calculated by first calculating the average winning probability (where the average is taken over all possible draws), and by comparing these probabilities with the winning probabilities that correspond to the realized draw. This is another way to identify teams with a 'lucky draw'.

5 Conclusions

This paper discusses a simulation model for football tournaments. The model is based on match-specific Poisson parameters. The parameters, to-

[5]The problem of ranking teams when not all teams play against each other is discussed extensively in the literature, with most papers dealing with ranking in American Football. See for instance Stern (1992) and Wilson (1995).

gether with the assumption of Poisson-distributed number of goals, provide a reasonable fit to the observed number of goals in a match. The model is partly a simulation, partly a probability model. The probabilities of proceeding to another round of the tournament are exact in the sense that extra time, the golden goal rule, and a penalty shoot-out are allowed for. The simulation part of the model is used to determine rankings in groups in the first phase of the tournament; in the knock-out phase of the tournament the probabilities can be calculated using a closed-form expression. Our experience with the model is positive: it does a good job of indicating favorites, and it can be used to answer many relevant questions about the tournament. One main advantage of the model is that it mimics the tournament precisely so that the effect of any changes in scoring intensities or probabilities are calculated in a consistent manner.

The methodology of this paper can be extended to other tournaments, like the UEFA Champions League. Moreover, it can be used to determine a ranking of any given set of countries at any moment in time, also when no major international tournaments are played.

References

Dixon, M.J. and M.E. Robinson (1997). A birth process for association football matches. Manuscript, Lancaster University.

Koning, R.H. (2000). An econometric evaluation of the firing of a coach on team performance. Research Report 00F40, SOM.

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

Ridder, G., J.S. Cramer, and P. Hopstaken (1994). Down to ten: Estimating the effect of a red card in soccer. *Journal of the American Statistical Association 89*, 1124–1127.

Stern, H. (1992). Who's number ones?-Rating football teams. *Proceedings of the Section on Statistics in Sport, American Statistical Association*, 1–6.

Van Gelder, R., P. Hopstaken, R.H. Koning, M.P. Koolhaas, G. Ridder, and G. Renes (15 June 1994). Toch bestaat de mogelijkheid dat Zuid-Korea wint. Trouw. (In the end, it is possible that South-Korea wins.).

Werkgroep Voetbal & Statistiek (5 June 1996). Gokken op goud. Trouw. (Gambling on gold.).

Werkgroep Voetbal & Statistiek (6 June 1998). De statistiek spreekt: Engeland is favoriet. Het Parool. (Statistics speaks out: England is the favorite.).

Wilson, R.L. (1995). Ranking college football teams: A neural network approach. *Interfaces 25*, 44–59.