Comparing league formats with respect to match unimportance in Belgian football

Dries Goossens · Jeroen Beliën · Frits C.R. Spieksma

Received: date / Accepted: date

Abstract Recently, most clubs in the highest Belgian football league have become convinced that the format of the league should be changed. Moreover, the TV station that broadcasts the league is pleading for a more attractive competition. The clubs, however, have not been able to agree on a new league format, mainly because they have conflicting interests. In this paper, we discuss the current league format, and two other formats that are presently being considered by the Royal Belgian Football Association. We simulate the course of each of these league formats, based on historical match results. The attractiveness of each of the formats is measured by the number of unimportant matches: the less unimportant matches, the more attractive the competition. Furthermore, we provide an overview of how each league format aligns with the expectations and interests of each type of club.

Keywords tournament design · match importance · football · simulation · optimization

1 Introduction

For decades, the first (and highest) division in Belgian football has been organized as a double round robin tournament, i.e. a tournament in which each team plays each other team twice, once at home and once away. The games of the tournament are grouped in rounds, such that each teams plays exactly once per round. Since the season 1977-1978,

Dries Goossens
K.U.Leuven - Faculty of Business and Economics,
Naamsestraat 69, 3000 Leuven, Belgium
E-mail: dries.goossens@econ.kuleuven.be

Jeroen Beliën
H.U.Brussel - Center for Modelling and Simulation,
Stormstraat 2, 1000 Brussels, Belgium
E-mail: jeroen.belien@hubrussel.be

Frits C.R. Spieksma
K.U.Leuven - Faculty of Business and Economics,
Naamsestraat 69, 3000 Leuven, Belgium
E-mail: frits.spieksma@econ.kuleuven.be

18 teams take part in this competition, and only two changes have been implemented since. From the season 1995-1996 onwards, three points are awarded for a win, whereas it used to be two points. Second, since the season 2005-2006, the team that ends one but last in the competition no longer relegates directly, but gets a chance to maintain its place in the first division in a double round robin tournament with three teams from second division. However, during the last years, most clubs in the highest Belgian football league have become convinced that changes in the way the competition is played are needed.

Despite the fact that most teams are convinced of the necessity of changes, there is little agreement on what these changes should be, since arguments and preferences of the teams depend on their (aspired) role in the competition. We can classify the teams involved into four categories. The first category consists of the *traditional top teams*. These teams (RSC Anderlecht, Club Brugge, Racing Genk, and Standard Liège) are the main contenders for the title of league champion nearly every season, and are especially looking for a stronger competition. They reason that by playing more top level games, they will be able to increase their budget and attract and train better players. This should increase their chances in UEFA Cup or Champions League football, where Belgian teams have not been excelling recently. A second category is what we call the *ambitious middle group*. These five teams (Gent, Charleroi, GBA, Lokeren, and Mouscron) have been continuously present in the first division during the last 10 years, and have all been able for a number of seasons in this period to end before at least one traditional top team. These teams are looking for a competition format that preserves or improves their chances for a European ticket, or even a league title. The third group is what we call the *small seven*. These teams are mainly concerned with holding on to their place in the league, and thus are not at all fond of reducing the number of competitors, or increasing the number of teams that relegate to the second division. Finally, a fourth group consists of 12 teams that we call the *second division teams*. These teams are those teams that were playing in or relegated to the second division in the season 2006-2007, but are anxious to promote to the first division. Some of these teams have quite a history in the first division, while others come from lower divisions, and are very ambitious to continue their rise. Apart from the clubs, the TV station that broadcasts the competition is looking for a league format that is as attractive as possible. In other words, the TV station wants to avoid matches that have no importance at all, since these are the games that don't attract viewers. Given the fact that the money from the broadcasting contract is the main source of income for many clubs, the wishes of the TV station carry a considerable weight.

2 Overview of the approach

In order to quantitatively and objectively analyze the different league formats, we developed a unique, integrated approach that combines both simulation and optimization. Simulation is needed to simulate the outcomes of the matches played in a particular league format. Optimization is required both to develop an acceptable match schedule (a so-called calendar) for a selected league format and to determine whether or not a match is classified as unimportant (see further). Analyzing a particular league format involves repeating the following steps for a large number of seasons.

1. Use optimization to develop a suitable match schedule given the teams in the league.
2. For all rounds in the regular competition:
   (a) Use optimization to determine which matches are unimportant.
   (b) Simulate all match outcomes.
3. Simulate the play-offs (if present in the league format).
4. Replace the relegated clubs by the climbers from the second division.

In section 3, we discuss a number of league formats that are currently being considered to be adopted for the highest Belgian football league. Specifics on the way the season schedule is created for each of the formats is provided in section 4. In section 5, we explain how we simulated match results. Our measure of match unimportance is detailed in section 6, and in section 7, we compare the various league formats using this measure of match unimportance. Moreover, we simulate what teams from each category may expect from the various league formats with respect to their chances for the league title, European football, or relegation.

3 Reforming the Belgian football league

There are numerous ways to organize a competition, even if we limit ourselves to those formats that are common in football. The World Cup, the Champions League, and the UEFA Cup are currently organized as a two stage event, where the first stage consists of a round robin tournament (single for the World Cup and the UEFA Cup, double for the Champions League), and the second stage is an instant knock-out tournament, with two games in any round but the final for the Champions League and the UEFA Cup, and just one for the World Cup.

More than a decade ago, Griggs and Rosa (1996) published an overview of competition formats that were used then in European national leagues. At that time, the large majority of the football competitions were organized as a double round robin tournament, using so-called canonical schedules (see e.g. De Werra (1980)). Nowadays, most European countries still use a double round robin tournament (e.g. Germany (Bartsch et al, 2006) and Italy (Della Croce and Oliveri, 2006)), although some countries (e.g. The Netherlands) have introduced a play-off stage after the regular competition, in order to decide on relegation or qualification for the European football tournaments.

In some countries, the competition consists of a single round robin tournament. Since in this tournament, each team meets each other team only once, the number of home games and the opponents in these games strongly determines the fairness of the schedule (in a double round robin tournament, this is obviously not an issue). Denmark is one of the few countries that have a triple round robin tournament (i.e. teams meet three times). This format also leads to an uneven distribution of home and away games (see Rasmussen (2008)). In Austria, a quadruple round robin tournament is played with 10 teams, resulting in a schedule with 36 rounds (Bartsch et al, 2006). One of the more peculiar competition formats can be found in Chile, where 20 teams are organized into 4 groups of 5, although they all play once against each of the other 19 teams. After this single round robin tournament, the best two teams of each group advance to a play-off stage, where the actual league champion is decided (see Durán

et al (2007)). In the remainder of this section, we discuss the way in which the Belgian highest football competition is organized currently, followed by two formats that are under consideration by the Royal Belgian Soccer League to be used for the season 2009-2010.

Currently, the highest Belgian football league is played as a straightforward double round robin tournament with 18 teams, spread over 34 rounds. The winner of this league is the champion, and qualifies for the (qualification stage of the) Champions League. The second in the league can also try to qualify for the Champions League, whereas the third in the league goes to the UEFA Cup. The team that ends up last relegates to the second division. The one but last team can however remain in the highest division if it wins a double round robin tournament with three second division clubs. This means that 12 additional games are played after the regular competition, resulting in 318 matches in total (Goossens and Spieksma, 2006).

The first alternative that we consider is loosely based on the competition in The Netherlands, where since the 2005-2006 season, a post-season play-off has been introduced to determine the league champion, relegation, and qualification for the European tournaments (see e.g. Koning (2007)). The format is a double round robin tournament with 16 teams, where the first in the league is the champion, and the last relegates to the second division. Two play-off tournaments decide which teams qualify for European football, and which teams relegate. The European play-offs are played with the teams ranked 2 to 5 in the league: in a first stage with a home and away game, the second ranked team plays the fifth ranked team, while the third ranked team meets the fourth ranked team. The winners of these games play each other twice in order to decide which team can qualify for the Champions League; the losing team goes to the UEFA Cup. The relegation play-offs are played with the teams ranked 14th and 15th, and 6 teams from the second division. These teams play a direct knock-out tournament, with a home and an away game in each stage, until two teams remain. The regular competition, together with the two play-offs results in 260 games per season. Since almost all outcomes in this competition are decided by a play-off stage, we will refer to it as the *play-off league*. Notice that many variants of this format exist, for instance with a play-off deciding on the league title, a play-off in which teams take (part of) the points they collected in the first part of the competition with them, or a setting in which teams that ended up in the middle of the ranking still have a chance to qualify for European football (see e.g. Geril (2008)).

The second league format splits the competition into two parts: an autumn competition and a spring competition. Each of these competitions consists of two series, A and B, of 10 teams each, that play a double round robin tournament. The competition starts with the autumn competition, consisting of 18 rounds. The winner of the A series of this competition qualifies for European football (Champions League qualification). The best five teams of the B series replace the worst five teams in the A series of the subsequently played spring competition, which is again a double round robin tournament for both the A and the B series. The winner of the A series in the spring competition is the league champion, and as such, qualifies for European football (Champions League qualification). The two worst teams in the B series relegate to the second division. The final ticket for European football (UEFA Cup) is awarded to the winner of a play-off tournament, played among the 4 teams that were second in the A

series or first in the B series in the autumn or the spring competition. If any of these 4 teams would already have qualified for European football in some other way, it is replaced by the next in line from the subcompetition it qualified from. First, the two B series winners play against each other in a single game with home advantage for the spring winner. Next, the winner meets the team that ended second in the autumn A series, in a single game at the ground of the second from the A series. The winner of that game can challenge the second in the spring A series for the final ticket for European football, again in a single game with home advantage for the second in the spring series. For the next season, the worst five teams of the A series are again replaced by the best five teams of the B series, and two teams promote from the second division, replacing the two last of the spring B series. With this new composition, the autumn competition of the next season can start. This league format has 363 games in total, and is referred to as the *Wijnants league*, named after its inventor Herman Wijnants, chairman of the football club Westerlo VV. With his league format, Wijnants attempted to find a compromise, reconciling the various clubs with conflicting interests. We refer to Hauspie (2007) for more details on the Wijnants league. Also this format has a number of variants, involving e.g. series of 9 or 12 teams instead of 10, or limiting the number of teams that promote from series B to A to three or four.

4 Scheduling the league formats

In this section we describe how the schedules for each of the competition formats were constructed. In Belgium, the wishes of the various stakeholders are collected through the *calendar committee*, which is responsible for creating an acceptable schedule. For reasons of fairness, the committee is convinced that teams should not play more than two consecutive home (away) matches, and that the total number of breaks (i.e. consecutive home (away) games) should be minimal. Furthermore, no team should start or end the competition with a break. As in many countries, in Belgium it is considered fair to schedule the second half of the competition identically to the first half, but with the home advantage inverted (i.e. mirroring). Since most teams prefer not to play against all traditional top teams consecutively, the calendar committee asked that no team would have to play more than twice against a top team in four consecutive games. Furthermore, the calendar committee stated that every team should receive a top team at home at least once in each half of the season. The TV station that broadcasts the games also has a number of wishes regarding the schedule. First of all, they want for each round at least one (and preferably two) of the four top teams to play an away game. The underlying motivation is that a top team's home games are less interesting, since the top team tends to win these games without much effort. Further, they wish to have at most one game between top teams per round, and no top games in the first four rounds. Moreover, the top games should be spread over the season. Apart from these demands, there are still a number of wishes that we don't take into account, since they are valid for only one specific season (e.g. no home game for some team on a given round, because of some other event taking place in the city).

As it turned out that satisfying all demands was simply not possible since they were conflicting, we asked the calendar committee to attach one of five priority levels to each of the wishes. We linked a penalty with each priority level, such that this penalty is

incurred if the corresponding constraint is violated. The goal is then to find a schedule with minimal incurred penalties. Since this scheduling problem is quite a hard nut to crack, we decompose it into two subproblems. In the first phase, each team is assigned a home-away pattern, which indicates when a team plays a home game, and when it plays away. In the second phase, the actual opponents are decided. For more details on the scheduling of the Belgian first division in its current league format, we refer to Goossens and Spieksma (2006).

Although of course, it is not clear yet what the wishes of the calendar committee would be when confronted with a different league format, we think we can safely assume that the demands mentioned above would still be valid. Therefore, we use the same criteria and method to develop schedules for the play-off league and the Wijnants league.

5 Simulating match results

In literature, there are two main approaches to model the outcomes of matches in football. The first approach (see e.g. Dixon and Coles (1997) or Dixon and Pope (2004)) uses bivariate Poisson regression to estimate forecasting models for goals scored and conceded. Forecasts of the match outcome (win-draw-lose) can be derived indirectly, by aggregating the estimated probabilities assigned to appropriate permutations of goals scored and conceded by the two teams. In the second approach (see e.g. Koning (2000), Forrest and Simmons (2000), Audas et al (2002) or Goddard and Asimakopoulos (2004)), the match outcomes are predicted directly by using discrete choice regression models like ordered probit regression. Goddard (2005) presents an extensive study in which the forecasting power of both approaches are compared using a 25-year data set on English league football match outcomes. Although the first approach is favored by most applied statisticians, according to Goddard's study, the differences between the forecasting performance of both approaches appear to be relatively small. Karlis and Ntzoufras (2003) present an alternative approach that combines issues of both the (indirect) goals based models and the (direct) outcomes based models as they try to predict the difference between the goals made by the two teams.

The objective of this study is to evaluate different competition formats rather than predicting match results. Nevertheless, in order to simulate, in a reliable way, all matches in a particular competition format, our analysis heavily relies on the estimated probabilities of match outcomes. Since the number of goals made and conceded are of less importance, we adopted the second approach that directly models the three possible match results (win-lose-draw). These trinomial probability distributions have been estimated for each match starting from 10 years historical data, including all matches of season 1997-1998 till season 2006-2007. If the result of a particular match, let's say team A versus team B, was five times a win for team A, two times a draw, and three times a win for team B, the resulting probability distribution for that match would be: 50% chance team A wins, 20% chance on a draw, and 30% chance team B wins. Remark that we distinguish between the matches A versus B and B versus A, and by doing this, take into account the home advantage.

Additionally, we developed probability distributions for the matches of each team against the newly promoted teams from the second division. Note that several different teams contribute to the probability distributions of these matches, since the newly promoted teams tend to differ over the seasons. The advantage of this approach is twofold. First of all, it enables us to simulate the result of a future match between a team and a newly promoted team (which could be a match that has never been played before in the first league). Second, it allows us to construct more reliable probability distributions for those matches that were played only a few times over the past 10 years. Indeed, due to lack of data, the resulting probability distributions for these matches are very unreliable. For such matches, we replaced the team that played the least matches in the first Belgian league over the past 10 years with one of both in that season just promoted teams (a random choice) and considered the probability distributions of the match result between this new team and the other team. Note that the replacing team (the first or the second climber) always played 10 seasons, since there were always two climbers. In case both teams played less than four seasons, both teams were replaced by the first and second climber respectively. The assumption here is that all teams that played less than four seasons in the first division over the past 10 years are equal with respect to their chances to win, draw or loose a particular match and these chances are equal to the chances of a newly promoted team. Our excellent preliminary results on the predicting accuracy of match results (see further) seem to justify this assumption.

In order to test the accuracy of these probability distributions, we used the results of the first half of season 2007-2008, which includes 151 matches, and worked as follows. For each match we predicted the result corresponding with the highest probability in the trinomial distribution. For instance, in the example above, we would predict team A to win. This leaded to a 51.0% accuracy, meaning that the result was correctly predicted for 51.0% of the matches.

The approach described above leaves two important issues out of consideration. First of all, it is reasonable to assume that the results of matches are not independent, and that a team that has won a series of matches (a *hot streak*), is more likely to win the next match as well. Therefore, we repeated the accuracy check described above, but this time we adjusted the three probabilities to take into account the current form of the teams as follows:

$$P(AW_{mod}) = \max\{0; P(AW_{noform}) + (PTSA/12 - PTSB/12) * f_{win}\}$$
$$P(D_{mod}) = \max\{0; P(D_{noform}) + (PTSA/12 - PTSB/12) * f_{draw}\}$$
$$P(AL_{mod}) = P(AL_{noform})$$

$$P(AW_{form}) = \frac{P(AW_{mod})}{P(AW_{mod})+P(D_{mod})+P(AL_{mod})}$$
$$P(D_{form}) = \frac{P(D_{mod})}{P(AW_{mod})+P(D_{mod})+P(AL_{mod})}$$
$$P(AL_{form}) = \frac{P(AL_{mod})}{P(AW_{mod})+P(D_{mod})+P(AL_{mod})}$$

with $P(AW_{noform})$ ($P(D_{noform})$, $P(AL_{noform})$) the probability of team A wins (draws, looses) without taking into consideration its current form; $P(AW_{form})$ ($P(D_{form})$, $P(AL_{form})$) the probability of team A wins (draws, looses) but this time taking into account the current form of the team, $PTSA$ ($PTSB$) the number of points obtained in the last four matches by team A (B) and f_{win} (f_{draw}) a parameter that reflects the

importance of the influence of the current form resulting in a win (draw). The best result was obtained with both f_{win} and f_{draw} equal to 1 leading to an accuracy of 54.3%. With other values for f_{win} and f_{draw} the difference between the accuracy with and without including the form was smaller and sometimes even negative. Given the fact that we fine-tuned the parameters to find the best possible test result, and the fact that in our opinion, the first half of the season 2007-2008 is characterized by some exceptional results, we conclude that the test is somewhat biased in favor of a model that takes into account the form of the moment. Since the differences between the models are nevertheless rather small, we decided to leave the form out of consideration during the simulations.

Second, the current approach considers all played matches over the past 10 years to be equally important. It seems to make sense, however, to give a higher weight to more recent matches when calculating the probability distributions. To take into account this issue, we assigned a weight of 3 to the matches of the three most recent seasons, a weight of 2 to the matches of 4 to 6 seasons ago, and a weight of 1 to the matches of 7 to 10 seasons ago. Surprisingly, using the new distributions, the accuracy of predicting the first half of season 2007-2008 dropped from 51.0% to 44.4% (without form) and from 54.3% to 45.7% (with form). Based on these results, it was decided to keep the original distributions. Table 1 summarizes these findings.

Table 1 Accuracy of predicting the first half of the 2007-2008 first division football season

	Excluding current form	Including current form
Without higher weight to recent years	51.0%	54.3%
With higher weight to recent years	44.4%	45.7%

Note that a fully random prediction would have an accuracy of 33% as one third of the matches would be predicted correctly (in the long run). Compared to the more complex models presented in literature where accuracies between 40% and 50% are reported (see e.g. McHale and Scarf (2006)), our simple approach based solely on historical match results performs remarkably well.

6 Measuring match (un)importance

The importance of a game is useful not just for comparing various tournament designs, but also for selecting games to broadcast, assigning referees, or explaining the attendance. Logically, the concept of match importance has been discussed before in a number of papers, and various rather simple measures have been developed. Jennett (1984) measures the importance of a game with respect to winning the league as the inverse of the number of remaining games that still need to be won. When a team can no longer win the league, the measure gets the value 0. Audas et al (2002) call a game important if it is possible for either of the opponents to still win the league (or relegate), if all other teams draw in the rest of their games. The most commonly accepted

measure for match importance however, is what Schilling (1994) calls the conditional importance $S_i(X)_{t,t+k}$ of match $t+k$ for a team i at time t with respect to outcome X, and is defined as follows.

$$S_i(X)_{t,t+k} = p(X_i|W_{i,t+k}, H_t) - p(X_i|L_{i,t+k}, H_t) \tag{1}$$

We use the notation X_i for an outcome X that is achieved by team i. This outcome may be the league championship, but just as well qualification for European football or relegation. The event where team i wins its game scheduled at time $t+k$ is represented by $W_{i,t+k}$; the event of team i losing this game by $L_{i,t+k}$. Finally, H_t represents the history of games that have already been played at time t. Note that this measure does not take into account a draw; extending this definition to include draws is not straightforward.

We propose to measure the attractiveness of a competition format through match unimportance. We define a match as unimportant for a team and with respect to some outcome, if it can have no influence on the outcome for that team. The underlying idea is that a game between two teams that have nothing to gain or to loose is no longer interesting for a TV station to broadcast, and will attract less fans to the stadium. Bojke (2007) confirms that this is the case in the English Premier League. To a lesser extent, this is also true if the game is relevant for only one of the two teams. Of course, we do not claim that unimportant games cannot be a tremendously spectacular, but beforehand, they will typically not attract big crowds. Thus, we suggest that the lower the number of unimportant games in a competition is, the more attractive this competition is. Notice that our definition does not take into account expectations. In other words, when two teams that are not favorites to win the league meet on the first round, this game will never be unimportant, because it is possible at that time that any of these teams becomes the league champion. This can contradict with the perception of the teams or the fans, who may think this game is very unlikely to influence the outcome of the championship, since the chances of any of these teams to win the league would be very slim anyway, regardless of which team wins the game.

In order to know whether a game still matters for a team t with respect to some outcome, we need to know the highest and the lowest ranking that team can still reach at the end of the season, before the game is played. We use the following notation. We define T as the set of teams in the competition, and $G(m)$ as the set of games that are yet to be played, given that there are m rounds remaining. We define the variable w_{ij} to be 1 if i wins its home game against j, and 0 otherwise. Furthermore, we say that l_{ij} is 1 if i loses its home game against j, otherwise $l_{ij} = 0$. The remaining decision variables are p_i, the number of points that a team i has at the end of the season, and r_i, which is 1 if team i is ranked higher than team t at the end of the season. The highest position a team t can possibly reach, given that m rounds remain to be played, and that team t collected a_t points from rounds already played is given by an optimal solution of the following formulation.

minimize

$$1 + \sum_{i \in T \setminus \{t\}} r_i \tag{2}$$

subject to

$$a_i + m + \sum_{j \in T \setminus \{i\}: ij \in G(m)} (2.01 w_{ij} - l_{ij}) + \sum_{j \in T \setminus \{i\}: ji \in G(m)} (2.01 l_{ji} - w_{ji}) = p_i, \forall i \in T \quad (3)$$

$$w_{ij} + l_{ij} \leqslant 1, \forall ij \in G(m) \quad (4)$$

$$p_t \geqslant p_i - B r_i, \forall i \in T \setminus \{t\} \quad (5)$$

$$w_{ij}, l_{ij} \in \{0, 1\}, \forall ij \in G(m) \quad (6)$$

$$r_i \in \{0, 1\}, \forall i \in T \setminus \{t\} \quad (7)$$

The goal function minimizes the number of teams that are ranked before team t, and is scaled with the term 1 to indicate that the highest ranking it can obtain is the first place. The first set of constraints states that the number of points a team i has at the end of the season equals the points this team already has, plus 3.01 points for each win and 1 point for each draw in the games that are yet to be played. Since in case of an equal number of points, the team with that won the highest number of games is to be ranked first, 3.01 points are added instead of 3. Notice that in case of a draw, both w_{ij} and l_{ij} equal 0. The situation where both w_{ij} and l_{ij} equal 1 is not allowed by the second set of constraints. Finally, we need to make sure that a team i will be ranked higher than t, if it obtained more points than t. When the parameter B is chosen equal to the total number of points that can be won in the competition, the final set of constraints will do just this.

With a limited number of changes, the above formulation can be used to determine the lowest ranking that this team t could still end up with. The goal function should be changed to maximize the number of teams that has a higher ranking than team t. Further, constraint set (5) should be replaced by

$$p_t \leqslant p_i + B(1 - r_i), \forall i \in T \setminus \{t\}, \quad (8)$$

such that a team i is only ranked higher than team t, if team i has more points.

The above formulations allow us to determine whether a game is important for some team for some outcome or not. For instance, a game is important for the league title, if the highest position this team can still reach is the first, and if the lowest position this team can still drop to is lower than the first. Indeed, if the former was not the case, the team would no longer be able to win the championship, and if only the latter was not the case, the league title could no longer escape them. A similar reasoning can be made for the other relevant outcomes: qualification for European football and relegation.

Note that Ribeiro and Urrutia (2005) developed an integer programming model to determine the minimum number of points any given team has to win in order to be sure it will reach some outcome. A similar model results in the minimum number of points this team has to win in order to have any chance of reaching that outcome. With these numbers, it is also possible to determine whether a match still has any importance or not. Furthermore, we point out that our approach is in line with the Schilling measure (1), since when we find through optimization that a match is unimportant, this means

that $p(X_i|W_{i,t+k}, H_t) = 0$ and thus $S_i(X)_{t,t+k} = 0$. The probabilities in the Schilling measure are however usually determined using a Monte Carlo simulation (see Scarf and Shi (2008), who use the Schilling measure in this way and apply it to the games in the English Premier League). Notice also that matches for which $S_i(X)_{t,t+k} = 0$ according to a Monte Carlo simulation need not be unimportant using our optimization approach.

Although as far as we are aware, no research has been done on the influence the schedule has on the outcome of the competition, it seems reasonable that some influence exists. For sure, the schedule will have its consequences for the (un)importance of a specific game. Indeed, a game that is scheduled in the beginning of the season can never be unimportant, whereas this same game may very well be unimportant when scheduled on the final round. It is however harder to say what the influence of the schedule on the number of unimportant games of the competition could be. The sequence of opponents for some team may also have an influence on the importance of its games. A team that plays all the weak opponents in the beginning of the season will probably win most of these games, and will be in the race for the league title until the last rounds, despite a series of tough opponents that remain to be played. Indeed, we always compare the best case scenario, where all the tough opponents are defeated, with the worst case scenario to determine whether a game still has importance (see further). On the other hand, a team that loses its first games, will have unimportant games (with respect to the league title) earlier in the season, because even the best case scenario will eventually no longer suffice to win the league title. However, when considering the match unimportance in the competition as a whole, it is conceivable that these effects cancel out each other, since a schedule can never offer an easy start for all teams. Finally, notice that because of our choice to simulate the outcome of games independently of the outcome of previous simulated games, the influence of the schedule on the outcome of the competition (and possibly also on match unimportance) is reduced. Indeed, if a series of wins would increase the chance of winning a next game, the effect of, for instance, a good start of the season would be increased.

7 Results

For each of the three league formats discussed in section 3, we have simulated 10 series of five consecutive seasons. We think five seasons is enough to grasp the dynamics of the league composition with respect to relegation. On the other hand, the strength of the teams, as determined by historical results over 10 years may no longer be valid when simulating series of more than five consecutive seasons. A direct climber from the second division is selected randomly from a set including all clubs that are not playing in the first division at that time. A team that needs to play a play-off tournament to claim its place in the first division (as in the current league format, and the play-off league) is chosen with equal probability among the participants of this tournament. The participants of this tournament other than the one(s) coming from the first division are selected with an equal probability among the teams that did not play in the first division that season. The play-offs to determine the teams that qualify for European football are simulated using the probability distributions derived in section 5. When possible we have used the same match results in the various league formats. This technique is known as *common random numbers* (we refer to Kelton et al (2002) for

more information). Notice, however, that not all league formats have the same teams taking part (due to a different number of competitors and relegation rule), and that the number of occurrences of a game can differ as well between the various league formats. Therefore, we cannot point out the league format as the only factor that determines whether a team wins the league, qualifies for European football or relegates, but we have tried to limit the influence of random match results.

We made the following assumptions and simplifications. Presently, each season, four clubs qualify for European football. Two of them go to (the qualification stage of) the Champions League; the other two go to the UEFA Cup. We assume that over the five season period used in our simulation, Belgium will maintain these four tickets for European football. Three tickets are decided through the league, but the fourth team to qualify for European football is the winner of the Belgian Cup. The Belgian Cup is a tournament that is not influenced by the way in which the league is organized, and the Cup Final is typically played after the final round of the league. Although the Cup can have an influence on the distribution of the tickets for European football in the league (e.g. if the Cup finalists end on two of the three first positions in the league, the fourth in the league qualifies for European football), and consequently also on match unimportance, we have not taken this into account. Further, the matches are normally played on Saturdays, but the TV station has the right to move one game to Friday, and two to Sunday. Obviously, this can have consequences for match unimportance, but since there is no clear rule to determine which games will be moved to an earlier or later date, we assume that all games in the same round are played at the same time. Finally, in the Wijnants league, the winner of the A series of the autumn competition qualifies for European football. This means that, from the point of view of qualification for European football, the games in the spring competition are no longer important for this team. On the other hand, if this team would end second in the spring league, its ticket for the play-off for European football goes to the third in the spring league. In order not to complicate the model, we have not taken these particularities into account, and assume that their effects on match unimportance will cancel out each other.

Table 2 Percentage of unimportant games

Outcome	Current league	Play-off league	Wijnants league
League Champion	12.17%	16.15%	28.55%
European football	9.05%	8.45%	4.82%
Relegation	4.65%	8.34%	27.69%
A series	-	-	3.34%
Any outcome	1.51%	1.25%	0.53%

Table 2 shows the percentage of unimportant games for each outcome and each of the league formats. The current league has a low number of unimportant games with respect to the league title and relegation. The Wijnants league gives many chances to all teams to take part in the play-off for European qualification, which results in a very low percentage of unimportant games for this outcome. On the other hand, the number of unimportant games for the league title and relegation are quite high, which can be explained by the fact that teams in the A series of the spring competition cannot

relegate to the second division, and that teams in the B series of the spring already lost their chances on the league title half-way the season. This alone results in almost 25% unimportant games for these outcomes. The Wijnants league however offers an extra objective to play for, namely promotion to (or maintaining a place in) the A series. This explains why the number of games that don't matter for any outcome is very low in this league format. The play-off league scores in between the other leagues for all outcomes.

Table 3 Average number of games against top four teams per season (averaged over the members of each team group

Team group	Current league	Play-off league	Wijnants league
Top four	6.00	7.00	9.79
Ambitious middle	7.87	7.78	9.22
Small seven	5.94	5.40	4.22
Second division	2.59	1.87	2.57

Another consideration is the number of games that a team can expect to play against the top teams per season (see Table 3). Adopting the play-off league would increase the number of games against top teams for the top four teams, but only with one game per season on average. On the one hand, the play-offs for European football offer possibilities for extra games against the top teams, but on the other hand, the fact that there are only 16 teams competing in this league reduces the chances to be in the league. For the ambitious middle teams and the small seven, the result is a status quo, but the second division teams will play less games against top teams than in the current league. The Wijnants league offers a team four games per season against each top team, provided that this team (and the top teams) are in the same series. Thus, the top four and the ambitious middle teams see their number of top games per season increased considerably. Since twice a season, five teams promote to the A series, the decrease in the number of games against top teams remains limited for the small seven and the second division teams.

Table 4 Percentage of presence in simulated first division seasons (averaged for each team group

Team group	Current league	Play-off league	Wijnants league (A+B series)
Top four	100.0%	99.6%	99.5% (89.5%+10.0%)
Ambitious middle	98.4%	90.0%	96.0% (58.6%+37.4%)
Small seven	74.2%	65.4%	76.3% (25.0%+51.3%)
Second division	32.4%	24.6%	49.0% (14.5%+34.5%)

The main concern of the small seven and the second division teams is that the new league formats will reduce their chances of playing in the first division. Therefore, we computed for each team the percentage of the simulated first division seasons in which they participated, and averaged these results for each team group. The results can be found in Table 4 and show that the small seven and the second division teams

indeed play less frequently in the first division with the play-off league, which has two teams less than the current league. However, we notice that the ambitious middle group may also expect to suffer from the reduced number of participants in this league. The Wijnants league has 20 teams, which explains the increased presence for all teams compared with the current league. However, the small seven and the second division teams will mostly play in the (less interesting) B series, although the average top team will also play in the B series for 10% of the time.

Table 5 Distribution of the league champion over the team groups

Team group	Current league	Play-off league	Wijnants league
Top four	98%	98%	94%
Ambitious middle	2%	2%	4%
Small seven	0%	0%	2%
Second division	0%	0%	0%

Table 6 Distribution of qualification for European football over the team groups

Team group	Current league	Play-off league	Wijnants league
Top four	89%	71%	70%
Ambitious middle	6%	24%	24%
Small seven	3%	4%	3%
Second division	2%	1%	3%

Finally, tables 5 and 6 show how the chances for the league title and qualification for European football are distributed over the team groups. The influence of the league formats on the league title is minimal. Only the Wijnants league offers a slightly larger chance to non top teams. However, the consequences for qualification for European football are considerable. The play-offs for European football in both the Wijnants league and the play-off league reduce the chances of the top teams from nearly 90% to about 70%, with especially the ambitious middle group benefiting from this.

In general, the results show that both new leagues can increase the attractiveness, since overall less games will be unimportant, which is the main concern for the TV station. An increased number of games against top teams for the top teams and the ambitious middle is to be expected from the play-off league and especially from the Wijnants league. Both new leagues will increase the chances for European football for the ambitious middle group, at the expense of the top teams. The play-off league however decreases the chances of playing in the first division for the small seven and the second division teams. The opposite is true for the Wijnants league, although in this league those teams would mostly play in the B series. This issue will probably be the largest obstacle on the path towards a change of the league format in Belgian football.

8 Conclusions and future work

After several months, the debate on the reform of the league in Belgian football is still ongoing. Apart from the different interests of all involved parties, one of the main problems is the lack of a study on what can be expected from a new league format. In this paper, we compared the current league format with two formats under consideration. We have presented tools to evaluate the attractiveness of these league formats by measuring the number of unimportant games that is to be expected. Further, we developed a model based on historical match results to simulate the outcome of each of these formats. This allows us to estimate the number of seasons a team would play in the first division, its number of games against top teams, and its chances for winning the league or qualifying for European football. We hope this paper can contribute to a well-founded choice for a league format and help to overcome the fear for change.

This topic leaves space for quite some future research. It would be very interesting to use our optimization approach to develop a measure for match importance (instead of unimportance), and compare it with other measures, as e.g. the one by Schilling (1994). Further, research could be done to investigate the influence of the schedule on match (un)importance. Also, we would like to find out whether the inclusion of a form parameter when simulating the match results would influence the results. Finally, linking this research to the expected number of spectators would make the financial aspects of adopting a new league format more tangible.

References

Audas R, Dobson S, Goddard J (2002) The impact of managerial change on team performance in professional sports. Journal of Economics and Business 54(6):633–650

Bartsch T, Drexl A, Kroger S (2006) Scheduling the professional soccer leagues of Austria and Germany. Computers and Operations Research 33(7):1907–1937

Bojke C (2007) The impact of post-season play-off systems on the attendance at regular season games. In: Albert, J., Koning, R.H. (eds.) Statistical Thinking in Sports, pp 179–202

De Werra D (1980) Geography, games and graphs. Discrete Applied Mathematics 2(4):327–337

Della Croce F, Oliveri D (2006) Scheduling the Italian Football League: an ILP-based approach. Computers and Operations Research 33(7):1963–1974

Dixon M, Coles S (1997) Modelling association football scores and inefficiencies in the football betting market. Applied Statistics 46:265–280

Dixon M, Pope P (2004) The value of statistical forecasts in the UK association football betting market. International Journal of Forecasting 20:686–700

Durán G, Guajardo M, Miranda J, Sauré D, Souyris S, Weintraub A (2007) Scheduling the Chilean Soccer League by integer programming. Interfaces 37:539–552

Forrest D, Simmons R (2000) Forecasting sport: the behaviour and performance of football tipsters. International Journal of Forecasting 16:317–331

Geril J (2008) Naar een competitie met 16 (in Dutch). De Standaard, January 18th (VUM)

Goddard J (2005) Regression models for forecasting goals and match results in association football. International Journal of Forecasting 21:331–340

Goddard J, Asimakopoulos I (2004) Forecasting football match results and the efficiency of fixed-odds betting. Journal of Forecasting 23:51–66

Goossens D, Spieksma F (2006) Scheduling the Belgian soccer league. In: PATAT' 06: Proceedings of the 6th International Conference on the Practice and Theory of Automated Timetabling, Brno, Czech Republic, pp 420–422

Griggs TS, Rosa A (1996) A tour of European soccer schedules, or testing the popularity of gk_{2n}. Bulletin of the Institute of Combinatorics and its Applications 18:65–68

Hauspie J (2007) Het plan-Wijnants voor de hervorming van de Jupiler League: simpele rekensom (in Dutch). Sport Voetbalmagazine 7(48):24–25

Jennett N (1984) Attendances, uncertainty of outcome and policy in Scottish league football. Scottish Journal of Political Economy 31:176–198

Karlis D, Ntzoufras I (2003) Analysis of sports data by using bivariate Poisson models. Statistician 52:381–393

Kelton WD, Sadowski RP, Sadowski DA (2002) Simulation with Arena. McGraw-Hill, Inc., New York, NY, USA

Koning R (2000) Balance in competition in Dutch soccer. Statistician 49:419–431

Koning R (2007) Post-season play and league design in Dutch soccer. In: Rodriguez, P., Garcia, J., Kesenne, S. (eds.) League Governance, Competition and Professional Sports, pp 183–207

McHale I, Scarf P (2006) Forecasting international soccer match results using bivariate discrete distributions. Working paper, Salford Business School

Rasmussen R (2008) Scheduling a triple round robin tournament for the best Danish soccer league. European Journal of Operational Research 185(2):795–810

Ribeiro C, Urrutia S (2005) An application of integer programming to playoff elimination in football championships. International Transactions in Operational Research 12:375–386

Scarf PA, Shi X (2008) The importance of a match in a tournament. Computers and Operations Research 35(7):2406–2418

Schilling MF (1994) The importance of a game. Mathematics Magazine 67(4):282–288