

# Proactive and Reactive Strategies for Football League Timetabling

Xiajie Yi<sup>a,\*</sup>, Dries Goossens<sup>a</sup>, and Fabrice Talla Nobibon<sup>b</sup>

<sup>a</sup>*Department of Business Informatics and Operations Management, Faculty of Economics and Business Administration, Ghent University, Belgium*

<sup>b</sup>*FedEx Express Europe, 1000 Brussels, Belgium*

<sup>\*</sup>*Corresponding author: xiajie.yi@ugent.be*

---

## Abstract

Due to unforeseen events (e.g. bad weather conditions), association football league schedules are not necessarily played as they were announced in the beginning of the season. This paper analyses the impact of uncertainty on the quality of football league schedules by examining fifteen seasons of ten major European football leagues. We describe several quality measures, related to breaks, the fairness of the ranking, and cancelled matches. The empirical study reveals that matches that were rescheduled to another date have a profound impact on the quality of the resulting schedule, indicating that football schedules in Europe deal poorly with uncertainty. Moreover, we present several proactive and reactive approaches in order to mitigate this problem. The former determine where to insert so-called *catch-up rounds* as buffers in the schedule, while the latter reschedule matches to these catch-up rounds when uncertain events occur. We evaluate combinations of proactive and reactive approaches, and provide recommendations to practitioners (e.g. four catch-up rounds usually suffice, and immediate irrevocable rescheduling is not beneficial).

**Keywords:** OR in sports; football schedule; quality measures; uncertainty; proactive and reactive policy

---

## 1. Introduction

Association football (to which we will refer as ‘football’ in this paper) comprises worldwide spectacles like the FIFA World Cup, as well as a plethora of amateur competitions, played simply for the happiness it brings to the players [4]. Each football competition needs a schedule of play that indicates which team is facing which opponent at what time (and where). Most leagues use a round robin schedule, where each team plays against each other team an equal number of times (usually twice). For professional football leagues, a schedule should accommodate numerous wishes and constraints coming from various stakeholders (league, clubs, fans, TV, police, etc.); [13] provide an overview of typical requirements encountered in sport scheduling. Hence, constructing a suitable schedule is quite a challenge. In some cases (e.g. [6]), academics managed to close a contract with a football association, and are providing the official schedule.

Despite hard efforts invested at the start of the season to create a high-quality football schedule, this *initial schedule* is not necessarily fully played as planned. Indeed, several months can span

between the time the initial schedule is announced and the moment that some games are played. During this period, additional information becomes available, e.g. related to weather conditions, the outcome of other competitions (e.g. domestic cup, Champions League, etc.), actors (players and referees), fans' behaviour (e.g. hooliganism, violence), technical problems (e.g. defective stadium lights), or political events (e.g. war, election, protests). This information, which may affect the implementation of the initial schedule, is difficult to estimate and anticipate due to uncertainty related to its occurrence. Nevertheless, it may lead to the pause, postponement, or even cancellation of a match. Games that are postponed or rescheduled induce deviations from the initial schedule, which we call *disruptions* (see a formal definition in Section 2). We refer to the schedule that effectively represents the way the competition was played as the *realized schedule*, which is known only at the end of the season.

As a first contribution, this paper studies the difference between the quality of the initial schedule and that of the realized schedule for fifteen seasons of ten major European football leagues. We introduce quality measures that are based on *breaks*, *fair ranking*, and *failures*. We find that, although unforeseen events rarely interfere with a match, for almost 90% of the seasons, the initial schedule is different from the realized schedule. Furthermore, these disruptions have a profound impact on the quality of the realized schedules.

The main results of our empirical study show the need for our second contribution, namely the development of proactive and reactive scheduling approaches to mitigate the impact of unforeseen events on the quality of the schedule. Proactive scheduling focuses on developing an initial schedule that anticipates the realization of unpredicted events during the season. We do this by inserting so-called *catch-up rounds* (i.e. empty rounds, to which matches can be rescheduled when unexpected events occur) as buffers in the schedule. Reactive scheduling, on the other hand, does not interfere with creating the initial schedule but repairs or re-optimizes the initial schedule when a disruption occurs. Hence, our reactive policies concentrate on rescheduling postponed games to appropriate catch-up rounds. We compare various combined proactive and reactive strategies, and study their effect on the quality of the schedule.

Despite numerous contributions on sport scheduling (see e.g. [10], [15], [16], [19], and [21]), as far as we are aware, the issue of uncertainty (or dealing with it) has not been studied before. However, successful applications of proactive and/or reactive optimization can be found in various domains such as project scheduling [12, 20], berth and quay crane scheduling [8], job shop scheduling [17], inventory systems [18], etc. Although other approaches for dealing with unexpected events (e.g. robust optimization, stochastic programming, see e.g. [9]) are conceivable too, (combined) proactive and reactive optimization has the advantage that (i) it does not require much probabilistic knowledge of the uncertain events, (ii) it looks beyond avoiding some worst-case scenarios and feasibility issues, and (iii) its implementation is relatively easy, making it more realistic for practitioners.

The rest of our paper unfolds as follows. Section 2 sets the stage by introducing basic notions of football scheduling and uncertainty. Section 3 introduces several quality measures to evaluate football schedules. In Section 4, we analyse the schedules of the most prominent European leagues in terms of our quality measures and disruptions. Several proactive policies and reactive policies are proposed in Section 5. Corresponding simulation results are illustrated in Section 6, and we conclude in Section 7.

## 2. Setting the stage

Throughout this paper, we consider only football leagues that have an even number of teams and we denote by  $S = \{1, 2, \dots, 2n\}$  the set of teams. A *game* (or *match*) is an ordered pair of teams, denoted as  $(i, j)$  or  $i - j$ , meaning that team  $i$  plays at home – that is, it uses its own venue

(stadium) for that game – against team  $j$ ; and we say that team  $j$  plays away. A *round* is a set of games, usually played in the same weekend, in which every team plays at most one game. We denote the set of rounds by  $T = \{1, 2, \dots, R\}$ .

An  $l$ -round robin schedule (simply referred to as *schedule* for the remainder of this paper) is a pairing of games to rounds, such that each team plays each other team  $l$  times, and each team plays at most once per round. Most football leagues play a double round robin tournament (2RR), where the teams meet twice (once at home, once away); but triple and quadruple round robin tournaments are also used. A *phased schedule* is such that no team plays against another team for the  $k$ -th time, before it has played at least  $k - 1$  matches against all the other teams (for  $k = 2, \dots, l$ ). Phased schedules can be split into equal parts, such that each part forms a single round robin tournament. Usually, the order of the rounds in these parts is related. Most football competitions use *mirroring*, i.e., the second part of the competition is identical to the first one, except that the home advantage is inverted [7].

A schedule is *compact* if it uses the minimum number of rounds required to schedule all the games; otherwise it is *relaxed*. In a compact schedule with an even number of teams, each team plays exactly one game in each round; in a relaxed round robin schedule, no team plays in each round. In this paper, and in line with common practice in European football schedules [7], we assume compact schedules. We say that a team has a *bye* in a given round if it does not play in that round. The sequence of home matches ('H') and away matches ('A') played by a single team is called its *home-away pattern* (HAP). Given such a HAP, the occurrence of two consecutive home matches, or two consecutive away matches is called a *break*. We consider a sequence of a home (away) game, one or more byes, and another home (away) game also as a break. Teams can have consecutive breaks, causing them to play three or more home (away) games in a row. A series of consecutive away games is called an *away tour*. For team  $i \in S$  and a round  $r \in T$ , we denote by  $h_{i,r}$  (respectively  $a_{i,r}$ ) the number of home (away) games played by that team up to (and including) round  $r$ .

An illustration of a compact, phased, double round robin tournament with six teams, named **A** to **F**, is given in Figure 1(a). The HAPs of this schedule are depicted in Figure 1(b). Note that the HAP of team **A** contains two breaks whereas that of team **F** starts with an away tour of three consecutive away games.

Many of the theoretical results and algorithms in sport scheduling are based on graph theory. As far as we are aware, [1] was the first to use the complete graph  $K_{2n}$  on  $2n$  nodes for constructing single round robin tournaments, with the nodes corresponding with the teams and each edge with a game between the teams of the nodes it connects. A compact schedule can then be seen as a one-factorization of  $K_{2n}$ , i.e., a partitioning into edge-disjoint one-factors  $F_i$  with  $i = 1, \dots, 2n - 1$ . A one-factor is a perfect matching, i.e., a set of edges such that each node in the graph is incident to exactly one of these edges. Each one-factor corresponds to a round in a compact schedule and represents  $n$  matches. One particular one-factorization results in so-called canonical schedules, which are highly popular in sport scheduling [6]. This one-factorization has its one-factors  $F_i$  for  $i = 1, \dots, 2n - 1$  defined as

$$F_i = \{(2n, i)\} \cup \{(i + k, i - k) : k = 1, \dots, n - 1\} \quad (1)$$

where the numbers  $i + k$  and  $i - k$  are expressed as one of the numbers  $1, 2, \dots, 2n - 1 \pmod{2n - 1}$  [1]. For an overview of graph-based models in sports scheduling, we refer to the work by [2].

Games that are postponed or rescheduled in a way that they induce deviations from the initial schedule are labelled *disruptions*. We formally define a disruption as follows:

| round | r1         | r2         | r3         | r4         | r5         | r6         | r7         | r8         | r9         | r10        |
|-------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| date  | 25/11      | 01/12      | 08/12      | 15/12      | 23/12      | 06/01      | 21/01      | 28/01      | 04/02      | 11/02      |
|       | <b>A-B</b> | <b>B-E</b> | <b>A-F</b> | <b>C-B</b> | <b>B-F</b> | <b>E-B</b> | <b>F-A</b> | <b>B-C</b> | <b>F-B</b> | <b>B-A</b> |
|       | <b>C-D</b> | <b>D-A</b> | <b>B-D</b> | <b>E-A</b> | <b>D-E</b> | <b>A-D</b> | <b>D-B</b> | <b>A-E</b> | <b>E-D</b> | <b>D-C</b> |
|       | <b>E-F</b> | <b>C-F</b> | <b>E-C</b> | <b>F-D</b> | <b>A-C</b> | <b>F-C</b> | <b>C-E</b> | <b>D-F</b> | <b>C-A</b> | <b>F-E</b> |

(a) A double round robin schedule for six teams, named **A** to **F**

| round    | r1    | r2    | r3    | r4    | r5    | r6    | r7    | r8    | r9    | r10   |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| date     | 25/11 | 01/12 | 08/12 | 15/12 | 23/12 | 06/01 | 21/01 | 28/01 | 04/02 | 11/02 |
| <b>A</b> | H     | A     | H     | A     | H     | H     | A     | H     | A     | A     |
| <b>B</b> | A     | H     | H     | A     | H     | A     | A     | H     | A     | H     |
| <b>C</b> | H     | H     | A     | H     | A     | A     | H     | A     | H     | A     |
| <b>D</b> | A     | H     | A     | A     | H     | A     | H     | H     | A     | H     |
| <b>E</b> | H     | A     | H     | H     | A     | H     | A     | A     | H     | A     |
| <b>F</b> | A     | A     | A     | H     | A     | H     | H     | A     | H     | H     |

(b) The HAPs of the schedule in Table 1(a)

Figure 1: Illustration of a double round robin schedule with its associated HAPs

**Definition 1.** *Given an initial schedule, if a game  $m$  of round  $r$  was played after at least one game of round  $r + 1$  or played before at least one game of round  $r - 1$  in the realized schedule, we say that there was a disruption. We call the game  $m$  a disrupted game.*

Each disruption will by definition create a difference between the initial and the realized schedule, however, the converse is not necessarily true. Indeed, note that a game which was not played as initially scheduled, but rescheduled (and played) before any game of the next round is played and after all the games of the previous rounds, is not considered as a disruption because the order of the games remains the same. Furthermore, games that are awarded to one or neither team without rescheduling are also not considered as disruptions. Although a game that was not (fully) played at the scheduled date is usually postponed, it is also possible to reschedule a game to an earlier date, provided of course that it is known well beforehand that it will not be possible to play that game as initially scheduled. For example, game between Porto and Marítimo in the 2016-2017 season of the Portuguese league was played earlier than originally scheduled because of Christmas holidays.

### 3. Quality measures for football schedules

In this section, we present the quality measures that we use to evaluate football schedules. In practice, the quality of such a football schedule is determined by how well it manages to satisfy a number of constraints. For instance, two teams sharing the same home venue must not have simultaneous home games, or team  $A$  may not be able to play a home game against some team  $B$  on a particular date, because insufficient police force is available at that time to guarantee public safety. Usually, each constraint receives a weight relative to its importance, and the quality of the schedule is determined by the weighted sum of satisfied constraints. However, the type of constraints and their importance heavily depend on traditions and geographical conditions, and hence, differ from one league (country) to another. Moreover, the exact constraints and their weights are not

publicly available.

Hence, in this paper, we opt for quality measures that are easy to understand, and that can be applied to any football league. For a given schedule, we say that a team has a break (Section 3.1) if it has two consecutive home games, or two consecutive away games. Fair ranking (Section 3.2) refers to the fact that the number of (home) games played should be balanced over the teams at any point in the season. Failures represent unsuccessful rescheduled matches (Section 3.3), and should obviously be avoided. We also present a number of theoretical results which establish relations among these quality measures. Our motivation to consider these quality measures stems from a survey of European football schedules [7], which shows that in almost all competitions, schedules are used that minimize the number of breaks and balance the home advantage.

### 3.1 Breaks

To reduce travel costs, a team may prefer to have two or more consecutive away games if its stadium is located far from the opponents' venues, and the venues of these opponents are close to each other. Scheduling away tours, however, is quite uncommon in European football, as there is typically a week between consecutive matches and travel distances are relatively small. Moreover, [5] observe that scheduling consecutive home games has a negative impact on attendance. Hence, in most competitions, breaks - and successive breaks in particular - are avoided as much as possible. This observation motivates the use of the *total number of breaks* and the *maximum number of consecutive breaks* (denoted by  $\delta$ ) in a schedule as a measure of the quality of that schedule. The schedule depicted in Figure 1 contains 14 breaks in total and has  $\delta = 2$ .

Ideally, each team has a perfect alternation of home and away games. It is easy to see that only two different patterns without breaks exist (H A H A ... H and A H A H ... A). Moreover, all teams must have different patterns (indeed, two teams with the same pattern can never play against each other), and hence, at most two teams will not have any break. Consequently, any (compact) round robin schedule for  $2n$  teams will have at least  $2n - 2$  breaks; [1] shows that a canonical schedule with exactly  $2n - 2$  breaks can be constructed for any  $n$ . For a double round robin tournament, mirroring a compact single round robin schedule with a minimal number of breaks results in  $6n - 6$  breaks. If  $2n \neq 4$ , this can be achieved without a team having successive breaks [1]. If there is no need for a phased schedule, we can limit the number of breaks to  $2n - 2$  [7].

### 3.2 Fair Ranking

Given the advantage that the home team turns out to have in football [14], it is fair that each team plays half of its games at home, and the other half away. [11] call a schedule balanced if for each team the number of home and away games played at the end of the season differ by at most one. [13] generalize this notion of balanced schedule over all rounds (instead of just the last round), and introduce  $k$ -balancedness.

**Definition 2.** *A schedule is  $k$ -ranking-balanced if and only if*

$$k = \max_{r \in T} k_r := \max_{r \in T} \max_{i \in S} |h_{i,r} - a_{i,r}|. \quad (2)$$

In words, a schedule is  $k$ -ranking-balanced if and only if for each team  $i \in S$  and for each round  $r \in T$ , the difference between the number of home and away games played by team  $i$  up to (and including) round  $r$  is at most  $k$ , and there is a team that achieves this value of  $k$  for at least one round in the schedule. Note that for each schedule we have  $k \geq 1$ . As an illustration, the schedule depicted in Figure 1 has  $k = 3$ .

Goossens and Spieksma [7] point out the importance of having a league table that offers a fair ranking after each round. In this interpretation, ideally, each team will have played the same number of home games after each pair of rounds. They define  $g$ -ranking-balancedness as follows:

**Definition 3.** A schedule is  $g$ -ranking-balanced if and only if

$$g = \max_{r \in T} g_r := \max_{r \in T} \left\{ \max_{i \in S} (h_{i,r}) - \min_{j \in S} (h_{j,r}) \right\}. \quad (3)$$

Literally, a schedule is  $g$ -ranking-balanced if and only if for each pair of teams  $i, j \in S$  after each round  $r \in T$ , the difference between the number of home games played by team  $i$  and the number of home games played by team  $j$  is at most  $g$ , and there are teams that achieve this value of  $g$  at least once in the schedule. We also have  $g \geq 1$  for each schedule. As an illustration, for the schedule depicted in Figure 1 we have  $g = 2$ .

Even more fundamental for a fair ranking than a balanced home advantage over the teams, is the difference between the number of matches played by the teams at any point in the season. In a compact schedule for a competition with an even number of teams, all teams are scheduled to play once per round. However, in case of disruptions, the realized schedule may lose that property, rendering the ranking unfair. Hence, we develop a new measure on the basis of the number of matches played as follows.

**Definition 4.** The ranking difference ( $RD$ ) of a schedule is

$$RD = \sum_{r \in T} RD_r := \sum_{r \in T} \sum_{i \in S} |M_r - m_{i,r}|, \quad (4)$$

where  $M_r$  is the number of matches played by the majority of the teams up to and including round  $r \in T$ , and  $m_{i,r}$  is the number of matches played by team  $i \in S$  up to and including round  $r \in T$ .

In order to have a minimal ranking difference, a schedule should have each team play the same number of games as the majority of the teams after each round; the smaller  $RD$ , the more fair the ranking will be at any point in the season. Note that  $RD \geq 0$  for each schedule; for time-relaxed schedules and tournaments with an odd number of teams,  $RD > 0$ . If a schedule has  $RD = 0$ , we call it a  $RD$ -balanced schedule. Moreover, the ranking difference will always be an even number because a (disrupted) game involves two teams. Figure 2 illustrates the impact on the ranking difference in case the game **A–C** in  $r5$  of the schedule in Figure 1(a) is disrupted and postponed to a catch-up round  $f$  after round 10. Since, as a result, two teams will have played one game less than the majority of the teams for 5 rounds, the ranking difference is now 10.

| round    | r1 | r2 | r3 | r4 | r5       | r6 | r7 | r8 | r9 | r10 | f        |
|----------|----|----|----|----|----------|----|----|----|----|-----|----------|
| <b>A</b> | H  | A  | H  | A  | <b>H</b> | H  | A  | H  | A  | A   | <b>H</b> |
| <b>B</b> | A  | H  | H  | A  | H        | A  | A  | H  | A  | H   |          |
| <b>C</b> | H  | H  | A  | H  | <b>A</b> | A  | H  | A  | H  | A   | <b>A</b> |
| <b>D</b> | A  | H  | A  | A  | H        | A  | H  | H  | A  | H   |          |
| <b>E</b> | H  | A  | H  | H  | A        | H  | A  | A  | H  | A   |          |
| <b>F</b> | A  | A  | A  | H  | A        | H  | H  | A  | H  | H   |          |

Figure 2: Illustration of a HAP set with  $k = 3$ ,  $g = 2$  and  $RD = 10$ .

We point out that the fact that a realized schedule is  $RD$ -balanced does not imply that there were no disrupted games: if a whole round is rescheduled to another date, this will have no impact

on the ranking difference measure. Note that the quality measures  $k$ -ranking-balancedness and  $g$ -ranking-balancedness are determined exclusively by the worst case round and (pair of) team(s); how the other teams fare in the other rounds has no impact. In contrast, each deviation, regardless of the team or the round, contributes to the ranking difference measure.

The following result establishes relations between the quality measures described above (using the notation defined earlier in this section).

**Theorem 1.** *For any given schedule involving  $S$  teams and  $T$  rounds, we have:*

- (a)  $g \leq k$ ;
- (b)  $k \leq 2g - 1$ ;
- (c)  $g = 1$  if and only if  $k = 1$ ;
- (d)  $g = 1$  if and only if there are no breaks on even rounds;
- (e) for schedule with more than two rounds,  $g = 1$  if and only if  $\delta = 1$ ;
- (f)  $\delta \leq 2k - 1$ ;
- (g)  $RD \leq 2k \lfloor |T|/2 \rfloor |S|$ .

**Proof:** (a) Consider a given round  $r \in T$  and  $k_r := \max_{i \in S} |h_{i,r} - a_{i,r}|$ . There exists a team  $i_r \in S$  such that  $k_r := |h_{i_r,r} - a_{i_r,r}|$ . Let us assume, without loss of generality, that  $h_{i_r,r} \geq a_{i_r,r}$  such that  $k_r := h_{i_r,r} - a_{i_r,r}$ . This implies that  $h_{i_r,r} = \max_{j \in S} h_{j,r}$ . Because  $h_{j,r} + a_{j,r} = r$  we infer that  $h_{j,r} \geq a_{i_r,r}$  and hence  $-h_{j,r} \leq -a_{i_r,r}$ , for any team  $j \in S$ . Therefore, we have:

$$\begin{aligned}
 \max_{j \in S} \{-h_{j,r}\} &\leq -a_{i_r,r}, \\
 -\min_{j \in S} h_{j,r} &\leq -a_{i_r,r}, \\
 h_{i_r,r} - \min_{j \in S} h_{j,r} &\leq h_{i_r,r} - a_{i_r,r}, \\
 \max_{i \in S} h_{i,r} - \min_{j \in S} h_{j,r} &\leq h_{i_r,r} - a_{i_r,r}, \\
 g_r &\leq k_r.
 \end{aligned} \tag{5}$$

By taking the maximum over  $r \in T$  in equation (5), we obtain that  $g \leq k$ .

(b) For a given round  $r \in T$ , either (1)  $h_{i,r} = a_{i,r}$  for all team  $i \in S$  or (2) there exists a team  $j \in S$  such that  $h_{j,r} \neq a_{j,r}$ . Let  $T_ = \{r : h_{i,r} = a_{i,r}, \forall i \in S\}$  and  $T_ \neq = \{r : r \notin T_ =\}$ . Because  $k \geq 1$  we infer that  $k := \max_{r \in T_ \neq} k_r$ . Also, because  $g \geq 1$  we have  $g := \max_{r \in T_ \neq} g_r$ . We now argue that for all  $r \in T_ \neq$  we have  $k_r \leq 2g_r - 1$ . Let  $r \in T_ \neq$  and  $i_r \in S$  such that  $k_r := h_{i_r,r} - a_{i_r,r}$ . Because  $r \in T_ \neq$  we have that  $\max_{j \in S} a_{j,r} \geq \min_{i \in S} h_{i,r} + 1$ . Therefore,

$$\begin{aligned}
 k_r &= h_{i_r,r} - a_{i_r,r} \\
 &\leq h_{i_r,r} - \min_{j \in S} h_{j,r} - 1 + \max_{j \in S} a_{j,r} - a_{i_r,r} \\
 &\leq 2g_r - 1.
 \end{aligned} \tag{6}$$

By taking the maximum over all  $r \in T_ \neq$  in equation (6), we obtain that  $k \leq 2g - 1$ .

(c) The result (a) implies that any schedule with  $k = 1$  has  $g = 1$ . We now consider a schedule with  $g = 1$  and we argue by contradiction that  $k = 1$ . If  $k = 2$  then there exists  $r \in T$  such that  $k_r := \max_{i \in S} |h_{i,r} - a_{i,r}| := 2$ . Without loss of generality, we assume that there exists a team  $i_r$

such that  $h_{i_r,r} \geq a_{i_r,r}$  and  $k_r := h_{i_r,r} - a_{i_r,r} = 2$ . Because  $g = 1$  we have  $h_{j,r} \geq h_{i_r,r} - 1$ , for all  $j \in S$ . Let  $S_r := \{j : h_{j,r} = h_{i_r,r}\}$  be the set of teams that have played the highest number of home games up to round  $r$ . We then have:

$$\begin{aligned} \sum_{j \in S} h_{j,r_k} &= \sum_{j \in S_r} h_{j,r} + \sum_{j \notin S_r} h_{j,r} \\ &= |S_r| h_{i_r,r} + (2n - |S_r|) (h_{i_r,r} - 1) \\ &= |S_r| + 2n (h_{i_r,r} - 1). \end{aligned} \tag{7}$$

On the other hand, because  $g = 1$  and  $h_{j,r} + a_{j,r} = r$  we also have that  $a_{j,r} \geq a_{i_r,r} + 1$  for all  $j \in S$ . Therefore,

$$\begin{aligned} \sum_{j \in S} a_{j,r} &= \sum_{j \in S_r} a_{j,r} + \sum_{j \notin S_r} a_{j,r} \\ &= |S_r| a_{i_r,r} + (2n - |S_r|) (a_{i_r,r} + 1) \\ &= 2n (a_{i_r,r} + 1) - |S_r|. \end{aligned} \tag{8}$$

By subtracting (7) from (8), we obtain:

$$\begin{aligned} \sum_{j \in S} a_{j,r} - \sum_{j \in S} h_{j,r} &= 2n (a_{i_r,r} + 1) - |S_r| - |S_r| - 2n (h_{i_r,r} - 1) \\ &= 2n (a_{i_r,r} - h_{i_r,r}) + 4n - 2|S_r| \\ &= -4n + 4n - 2|S_r| \\ &= -2|S_r| \neq 0, \end{aligned}$$

which implies that the sum of all home games is greater than the sum of all away games. A contradiction! A similar analysis can be developed for  $k \geq 3$ . This completes the proof of case (c).

(d) On the one hand, suppose that we have a schedule with  $g = 1$  and a break on an even round. It is not difficult to see that this will imply that  $k > 1$  and we have a contradiction with (c). On the other hand, if a schedule does not have a break on even rounds then clearly  $g = 1$ .

(e) This result follows from the fact that any schedule with more than two rounds contains a break, which implies that  $\delta \geq 1$  for such schedule. Furthermore, because we consider a schedule with  $g = 1$  we conclude that  $\delta = 1$  for that schedule.

(f) Let us consider a schedule with  $\delta > 0$ ; note that if  $\delta = 0$  then the inequality is trivially true. Let  $i$  be a team with  $\delta$  consecutive breaks and in addition, we assume that  $i$  is the first team (with respect to rounds) to have  $\delta$  consecutive breaks. Without loss of generality, we assume that team  $i$  is playing  $\delta + 1$  home games and we denote by  $r_{LB}$  and  $r_{UB}$  the first (respectively the last) round involving that series of  $\delta + 1$  home games. This means that in round  $r_{LB}-1$  team  $i$  is either playing away or it has not yet started the competition; a similar explanation holds for  $r_{UB}+1$ . We distinguish the following two cases:

(1) On the one hand, if  $LB = 1$  then team  $i$  starts the competition with  $\delta + 1$  home games and the schedule has  $k \geq \delta + 1$ . Clearly, the inequality  $\delta \leq 2k - 1$  is satisfied.

(2) On the other hand, if  $LB > 1$  then team  $i$  has played  $LB - 1$  games before its series of  $\delta + 1$  home games. If  $h_{i, LB-1} \geq a_{i, LB-1}$  then, by focusing on the schedule of team  $i$  up to round  $UB$ , we have that  $k \geq \delta + 1$  and hence  $\delta \leq 2k - 1$ . We now suppose that  $h_{i, LB-1} < a_{i, LB-1}$  and let  $\alpha = a_{i, LB-1} - h_{i, LB-1}$ . Note that  $k \geq \alpha$  and as a consequence, if  $\alpha \geq \frac{\delta+1}{2}$  then we have  $\delta \leq 2k - 1$ . However, if  $\alpha < \frac{\delta+1}{2}$  then by focusing on the schedule of team  $i$  up to round  $UB$ , we infer that



$k \geq \delta + 1 - \alpha$ . Furthermore,

$$\begin{aligned} \alpha < \frac{\delta + 1}{2} &\Rightarrow -\alpha > -\frac{\delta + 1}{2} \\ &\Rightarrow \delta + 1 - \alpha > \delta + 1 - \frac{\delta + 1}{2} = \frac{\delta + 1}{2} \\ &\Rightarrow k \geq \delta + 1 - \alpha > \frac{\delta + 1}{2} \\ &\Rightarrow 2k - 1 \geq \delta. \end{aligned}$$

(g) Consider a given round  $r \in T$  and  $k_r := \max_{i \in S} |h_{i,r} - a_{i,r}|$ . There exists a team  $i_r \in S$  such that  $k_r := |h_{i_r,r} - a_{i_r,r}|$ . Let us assume, without loss of generality, that  $h_{i_r,r} \geq a_{i_r,r}$  such that  $k_r := h_{i_r,r} - a_{i_r,r}$ . This implies that  $h_{i_r,r} = \max_{j \in S} h_{j,r}$  and  $M_r \leq 2h_{i_r,r}$ . Assume that we have a compact schedule, then  $h_{j,r} + a_{j,r} = m_{j,r} = r$ , we infer that  $h_{j,r} \geq a_{i_r,r}$  and  $a_{j,r} \geq a_{i_r,r}$ , for any team  $j \in S$ . Therefore,

$$\begin{aligned} M_r - m_{j,r} &\leq 2h_{i_r,r} - (h_{j,r} + a_{j,r}) \\ &\leq 2h_{i_r,r} - 2a_{i_r,r} \\ &\leq 2k \end{aligned} \tag{9}$$

Since games played by at least half of the teams, i.e.,  $\lceil T/2 \rceil$  teams can reach  $M_r$ , it means at most  $\lfloor T/2 \rfloor$  teams play less than  $M_r$ . Thus,

$$\begin{aligned} RD &= \sum_{r \in S} \sum_{j \in T} (M_r - m_{j,r}) \leq \sum_{r \in S} 2k \lfloor T/2 \rfloor \\ &\leq 2k \lfloor T/2 \rfloor S \end{aligned} \tag{10}$$

□

We mention that the inequalities (b) and (f) are tight. Indeed, we propose two schedules and the corresponding HAP sets in Figures 3 and 4, consisting of three rounds and involving four teams named **A** to **D**. The schedule in Figure 3 has  $k = 3$  and  $g = 2$ , which indicates that the bound provided by (b) is tight; whereas the schedule in Figure 4 has  $k = 1$  and  $\delta = 1$ , which also indicates that the bound provided by (f) is tight.

|            |            |            |   |   |   |
|------------|------------|------------|---|---|---|
| r1         | r2         | r3         | H | H | H |
| <b>A-B</b> | <b>A-D</b> | <b>A-C</b> | A | H | A |
| <b>C-D</b> | <b>B-C</b> | <b>D-B</b> | H | A | A |
|            |            |            | A | A | H |

Figure 3: Illustration of schedule and HAP set with  $k = 3$  and  $g = 2$ .

|            |            |            |   |   |   |
|------------|------------|------------|---|---|---|
| r1         | r2         | r3         | H | A | H |
| <b>A-B</b> | <b>B-A</b> | <b>A-C</b> | A | H | A |
| <b>C-D</b> | <b>D-C</b> | <b>D-B</b> | H | A | A |
|            |            |            | A | H | H |

Figure 4: Illustration of schedule and HAP set with  $k = 1$  and  $\delta = 1$ .

[1] showed that for any  $2n$  teams, it is possible to construct a canonical single round robin schedule with minimal number of breaks and without breaks on even rounds. As a consequence, there exists a schedule for which all the above quality measures are minimal.

### 3.3 Failures

Normally, in a round robin schedule, each team should have played the same number of games at the end of the season. However, not every disrupted match can be rescheduled successfully. Indeed, UEFA regulations prescribe that teams should have at least two rest days between consecutive matches (i.e., a team that plays on Thursday cannot play again until Sunday at the earliest). Furthermore, police may forbid the use of certain dates for rescheduling high-risk matches. Hence, it may happen that none of the remaining dates in the season are suitable for rescheduling a match of a given team, particularly if that team already faces a number of postponed games and/or a busy schedule in the domestic cup or European competitions. We call a match a *failure* if it cannot be played because no suitable date is available on which it can be rescheduled. Ultimately, a failure is cancelled and the victory is awarded to one or neither opponent.

## 4. Empirical Study of European Football League Schedules

We consider ten major European football leagues, comparing the quality of initial schedules with that of the realized schedules. We first give details about the leagues and how the data was collected in Section 4.1; the obtained results are presented in Section 4.2.

### 4.1 Data

For the seasons 2002–2017, the following European football leagues are taken into consideration: Belgium, England, France, Germany, Italy, Netherlands, Portugal, Russia, Spain, and Ukraine (constituting the top 10 of the UEFA Country Ranking). In order to be able to make an aggregated analysis, we opted to include only European leagues, since these all abide by the same (UEFA) rules and delegate teams to the same international competitions (Champions League, Europa League). Moreover, they are influenced by a similar climate (e.g. winter in December-February). We gathered information about initial and realized schedules from the following websites:

<http://www.the-sports.org/football-foot-s1-c0-b0.html>

<http://www.rsssf.com>

In some competitions (e.g. Belgium, Netherlands), the regular round robin stage of the competition is followed by a post-competition stage. Typically, post-competitions involve only a subset of the teams in the league, and it is not clear from the beginning of the season which teams will take part in the post-competition stage (this depends on their performance in the regular stage). Consequently, no post-competition schedule can be made at the beginning of the season. Furthermore, post-competitions are often implemented in a direct knock-out format according to a fixed bracket, which leaves few or no scheduling options. For these reasons, we chose to ignore post-competitions in this paper.

In all countries except the Netherlands, the initial schedules are compact. Indeed, in the Netherlands, two extra rounds were used for seasons 2002–2003 and 2003–2004; the 2004–2005 season had one extra round scheduled. The Belgian league had two seasons in which a team was not able to finish the season due to bankruptcy, and hence, these teams' matches were awarded 5-0 to the opponent. Besides, a team withdrew from the Ukrainian league in season 2013–2014 and their results were annulled during the winter break. As explained before, we do not consider these cases as disruptions of the initial schedule.

## 4.2 Results

We analyze the variation of the quality measures between initial and realized schedules for the above-mentioned ten leagues for fifteen consecutive seasons, from 2002 to 2017. Figure 5 indicates various reasons for disruptions, as well as their relative occurrence. The main factors that have affected the implementation of the initial schedules are adverse weather conditions (heavy rains, snow, ...) and conflicts with the schedules of other competitions (cup, Champions League, ...). In some countries, teams were entitled to have their match rescheduled when they missed a number of international players (e.g. because of African Cup of Nations). Political issues can also have a profound influence. For example, in Ukraine, two rounds in season 2014–2015 were postponed for a couple of months for the sake of elections; the Crimea crisis and the conflict with Russia resulted in even more disruptions and even a failure (Chornomorets Odesa vs. Metalist Kharkiv). Occasionally, an entire matchday has been put off in Northern Ireland (2008) and Spain (2011) due to a strike of the referees and the players, respectively.

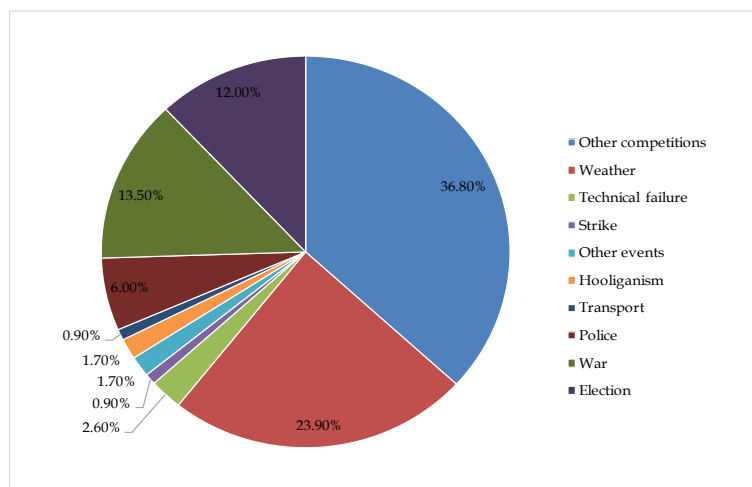


Figure 5: Reasons behind disruptions, with their relative proportion

Figure 6 gives an overview of the percentage of disrupted matches per league (Figure 6(a)), and per month (Figure 6(b)). For 121 seasons out of the 150 seasons that are studied, at least one disruption was recorded. In total, 637 matches out of 46,902 (i.e., about 1.36% of the matches) were rescheduled leading to a disruption. Although having a mild climate is not a guarantee for a low number of disruptions, winter clearly plays a role, as the months December, January, and February record more disruptions than the average month. July and August also have a high disruption percentage, which is due to the large number of teams that are participating in European competitions in that period. June has no disruptions recorded, which is not a surprise as most leagues do not schedule matches in that month. The frequency of disruptions differs considerably between different leagues. For instance, England has over thirteen times as many disruptions than Germany. Unlike most other competitions, the English Premier League has many games played during winter (the Premier League is set to trial a winter break in season 2019/20 for the very first time though). Moreover, there are recurring scheduling conflicts with the FA Cup competition. Indeed, contrary to other competitions, the quarter and semi-finals of the FA Cup coincide with a regular Premier League matchday. Hence, teams who are still active in the FA Cup at that stage see their Premier League match rescheduled.

As shown in Figure 7, leagues usually have at most one disrupted game per round. In fact, two

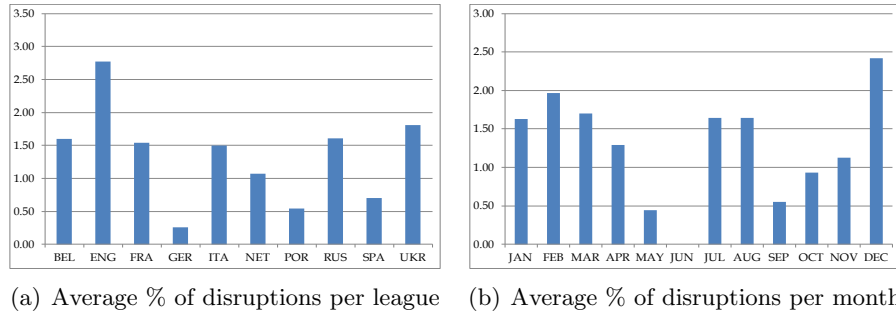


Figure 6: Disruption occurrence per league and per month

thirds of the disruptions are isolated events. Figure 7 also shows the likeliness of having multiple disrupted games in the same round (note however that not all leagues have the same number of matches per round). Important differences between countries can be noted: in Belgium, England, France, the Netherlands and Russia, it happens frequently that multiple games from the same round are disrupted, while this is much less the case in Germany, Portugal or Italy. In the Russian, Spanish, Ukrainian, and Belgian league, it has occurred that an entire round was disrupted.

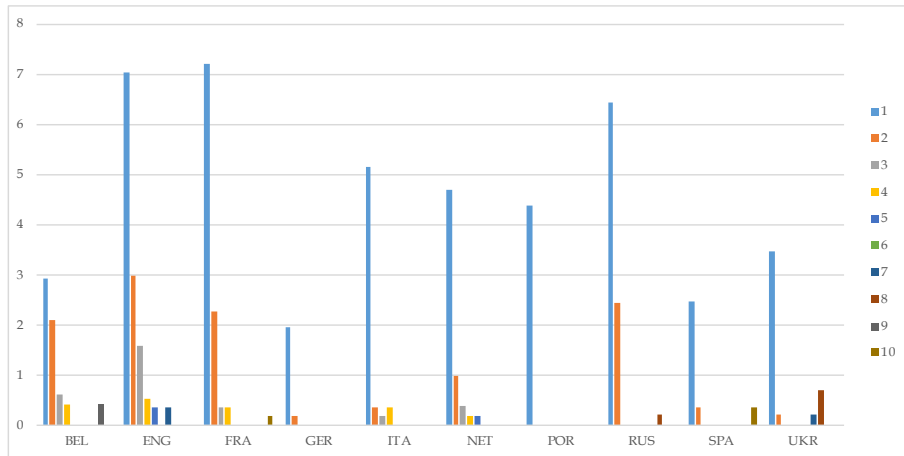


Figure 7: Occurrence (as a %) of rounds with 1, 2, ..., 10 disruption(s) per league

In Table 1, we display the value of each quality measure for each league during the last 15 seasons. Each cell consists of an initial schedule and a realized one, which have four common numbers; the first (respectively the second, the third and the fourth) value represents the number of breaks (respectively  $\delta$ ,  $k$ , and  $g$ ). For realized schedules, these numbers are followed by the value of  $RD$ . The quintuple (90|2|4|4|34), for instance, means that the realized schedule has 90 breaks, there is a team with 2 consecutive breaks, the  $k$ - and  $g$ -ranking-balancedness both equal 4, and the ranking difference is 34. We point out that it is possible that an initial schedule has more rounds than strictly needed, and as such has  $RD > 0$ . In the last fifteen years, the Netherlands had three such initial schedules, namely 2002–2003, 2003–2004, and 2004–2005, resulting in  $RD$  values equal to 30, 26 and 10 respectively. Since all other initial schedules had  $RD = 0$ , we dropped this notation for initial schedules in Table 1.

Out of 150 cells in Table 1, only 28 have the same values on the first four numbers for the

| League | Belgium             | England                        | France                        | Germany                     | Italy                        | Netherlands                      | Portugal                    | Russia                        | Spain                        | Ukraine                     |
|--------|---------------------|--------------------------------|-------------------------------|-----------------------------|------------------------------|----------------------------------|-----------------------------|-------------------------------|------------------------------|-----------------------------|
| 02-03  | initial<br>realized | (126 1 2 2)<br>(140 2 3 3 78)  | (40 1 2 2)<br>(70 2 4 3 144)  | (48 1 2 2)<br>(52 1 2 2 6)  | (58 1 2 2)<br>(76 2 2 2 12)  | (122 2 3 3 30)<br>(144 2 3 3 82) | (48 1 2 2)<br>(54 2 2 2 26) | (52 3 3 3)<br>(68 3 4 4 96)   | (54 1 2 2)<br>(62 2 2 2 16)  | (28 1 2 2)<br>(30 1 2 2 18) |
| 03-04  | initial<br>realized | (94 1 2 2)<br>(118 2 3 3 86)   | (40 1 2 2)<br>(60 2 3 3 48)   | (48 1 2 2)<br>(56 1 2 2 20) | (60 1 2 2)<br>(64 2 2 2 28)  | (140 2 3 2 26)<br>(148 2 3 3 68) | (48 1 2 2)<br>(58 2 2 2 28) | (54 3 4 3)<br>(56 3 4 3 16)   | (54 1 2 2)<br>(58 1 2 2 6)   | (28 1 2 2)<br>(28 1 2 2 0)  |
| 04-05  | initial<br>realized | (130 1 1 1)<br>(160 2 2 2 84)  | (40 1 2 2)<br>(58 2 3 3 54)   | (48 1 2 2)<br>(52 1 2 2 10) | (68 1 2 2)<br>(72 1 2 2 8)   | (110 2 2 2 10)<br>(118 2 3 3 42) | (48 1 2 2)<br>(50 1 2 2 10) | (66 2 2 2)<br>(82 2 3 3 34)   | (54 1 2 2)<br>(62 2 3 2 6)   | (60 2 2 2)<br>(40 2 3 3 66) |
| 05-06  | initial<br>realized | (108 1 2 2)<br>(152 2 3 3 272) | (40 1 2 2)<br>(68 1 3 3 80)   | (48 1 2 2)<br>(54 2 2 2 8)  | (66 1 2 2)<br>(70 1 3 3 4)   | (116 1 3 2)<br>(122 1 3 3 44)    | (48 1 2 2)<br>(48 1 2 2 0)  | (78 3 4 4)<br>(96 3 4 5 104)  | (58 1 2 2 8)<br>(58 1 2 2 8) | (42 2 2 2)<br>(46 2 2 2 16) |
| 06-07  | initial<br>realized | (170 1 1 1)<br>(186 2 2 2 118) | (40 1 2 2)<br>(52 1 2 2 22)   | (48 1 2 2)<br>(48 1 2 2 0)  | (68 1 2 2)<br>(104 2 3 3 14) | (120 1 3 3)<br>(124 2 3 3 12)    | (42 1 2 2)<br>(48 1 3 3 86) | (60 3 6 5)<br>(92 4 7 5 36)   | (54 1 2 2)<br>(56 2 2 2 10)  | (42 2 2 2)<br>(44 1 2 2 18) |
| 07-08  | initial<br>realized | (134 2 2 2)<br>(138 2 3 3 120) | (40 1 2 2)<br>(58 1 2 2 94)   | (48 1 2 2)<br>(56 1 2 2 8)  | (66 1 2 2)<br>(82 2 2 2 60)  | (100 1 3 2)<br>(118 2 4 3 58)    | (42 1 2 2)<br>(42 1 2 2 0)  | (44 3 6 5)<br>(52 3 6 5 48)   | (54 1 2 2)<br>(56 1 2 2 26)  | (42 2 2 2)<br>(46 2 2 2 18) |
| 08-09  | initial<br>realized | (130 1 1 1)<br>(140 2 3 3 100) | (40 1 2 2)<br>(44 1 2 2 4)    | (48 1 2 2)<br>(50 1 2 2 10) | (66 1 2 2)<br>(70 2 2 2 20)  | (116 1 3 2)<br>(114 1 3 3)       | (42 1 2 2)<br>(46 1 2 2 6)  | (32 1 2 2)<br>(56 2 3 3 96)   | (54 1 2 2)<br>(54 1 2 2 0)   | (42 2 2 2)<br>(48 2 2 2 18) |
| 09-10  | initial<br>realized | (172 1 1 1)<br>(204 2 3 3 278) | (40 1 2 2)<br>(102 3 3 3 114) | (48 1 2 2)<br>(48 1 2 2 0)  | (66 1 2 2)<br>(86 2 2 2 56)  | (114 1 3 3)<br>(116 2 3 3 20)    | (42 1 2 2)<br>(48 2 3 2 6)  | (40 2 4 4)<br>(48 2 5 5 12)   | (54 1 2 2)<br>(58 1 2 2 6)   | (50 3 4 3)<br>(78 2 5 4 18) |
| 10-11  | initial<br>realized | (124 3 3 3)<br>(160 2 6 5 162) | (40 1 2 2)<br>(54 1 3 3 28)   | (48 1 2 2)<br>(52 1 3 3 4)  | (64 1 2 2)<br>(80 2 3 2 24)  | (112 1 3 3)<br>(120 2 3 3 16)    | (42 1 2 2)<br>(46 2 3 3 6)  | (32 1 2 2)<br>(50 2 3 2 126)  | (54 1 2 2)<br>(54 1 2 2 0)   | (44 2 2 2)<br>(44 2 2 2 0)  |
| 11-12  | initial<br>realized | (126 1 1 1)<br>(134 2 2 2 84)  | (40 1 2 2)<br>(60 2 3 3 36)   | (48 1 2 2)<br>(50 2 3 2 8)  | (66 1 2 2)<br>(142 3 3 3 74) | (124 2 5 4)<br>(130 2 5 4 20)    | (42 1 2 2)<br>(44 1 2 2 2)  | (50 3 4 4)<br>(56 3 4 4 32)   | (54 1 2 2)<br>(84 2 3 3 16)  | (60 2 2 2)<br>(48 2 2 2 16) |
| 12-13  | initial<br>realized | (130 1 2 2)<br>(150 2 3 2 126) | (40 1 2 2)<br>(48 1 2 2 16)   | (48 1 2 2)<br>(48 1 2 2 0)  | (62 1 2 2)<br>(64 1 2 2 4)   | (106 1 3 2)<br>(106 1 3 2 0)     | (42 1 2 2)<br>(44 1 2 2 8)  | (40 2 4 3)<br>(40 2 4 3 0)    | (54 1 2 2)<br>(64 2 2 2 10)  | (54 2 4 3)<br>(54 2 4 3 0)  |
| 13-14  | initial<br>realized | (126 1 1 1)<br>(162 2 2 2 146) | (40 1 2 2)<br>(54 2 3 2 22)   | (48 1 2 2)<br>(48 1 2 2 4)  | (64 1 2 2)<br>(66 2 2 2 20)  | (116 1 3 2)<br>(122 2 3 2 56)    | (42 1 2 2)<br>(42 1 2 2 0)  | (112 4 6 4)<br>(112 4 6 4 4)  | (54 1 2 2)<br>(58 2 2 2 6)   | (42 2 2 2)<br>(80 2 3 3 32) |
| 14-15  | initial<br>realized | (116 1 2 2)<br>(122 2 3 2 36)  | (48 2 3 2)<br>(62 2 3 2 26)   | (48 1 2 2)<br>(48 1 2 2 0)  | (62 1 2 2)<br>(68 2 3 2 24)  | (116 1 3 2)<br>(118 2 3 2 6)     | (48 1 2 2)<br>(48 1 2 2 0)  | (192 5 6 5)<br>(194 5 6 5 32) | (54 1 2 2)<br>(58 1 3 2 6)   | (52 4 5 4)<br>(60 5 6 5 26) |
| 15-16  | initial<br>realized | (138 1 2 2)<br>(164 2 3 3 112) | (84 3 3 3)<br>(98 3 3 3 28)   | (48 1 2 2)<br>(48 1 2 2 0)  | (64 1 2 2)<br>(66 2 3 3 10)  | (110 1 3 3)<br>(110 1 3 3 0)     | (48 1 2 2)<br>(54 2 3 2 20) | (32 1 2 2)<br>(32 1 2 2 0)    | (54 1 2 2)<br>(58 2 3 3 18)  | (36 2 2 2)<br>(56 3 2 2 0)  |
| 16-17  | initial<br>realized | (154 3 3 3)<br>(168 3 3 3 62)  | (74 1 3 2)<br>(90 2 3 2 80)   | (48 1 2 2)<br>(48 1 2 2 0)  | (64 1 2 2)<br>(80 2 2 2 60)  | (110 1 3 3)<br>(110 1 3 3 0)     | (48 1 2 2)<br>(50 1 2 2 10) | (78 4 6 5)<br>(78 4 6 5 0)    | (54 1 2 2)<br>(64 2 3 3 60)  | (32 2 3 2)<br>(35 2 3 2 12) |

Table 1: Value of the quality measures for the initial and the realized schedules of each league between 2002 and 2017.

initial and the realized schedules, 25 of which correspond to the *RD*-balanced seasons without disruptions. Season 2015–2016 of the Ukrainian league was *RD*-balanced, however, the number of breaks increased with 20 compared to the initial schedule because of a rescheduling of a whole round. Table 1 also shows that some countries consistently produce initial schedules that have the same quality (e.g. France, Germany, Spain), which suggests that these leagues have specific rules with respect to breaks and home advantage, that they explicitly take into account in the scheduling process. In some countries (e.g. Belgium, Portugal, and Ukraine), the number of teams has changed during the period of fifteen years, which may explain some of the differences in the number of breaks. Generally speaking, most leagues use initial schedules that achieve good quality measures (small values). England is the only league that has initial schedules which show perfect balancedness (i.e.,  $k = g = \delta = 1$ ), albeit accompanied by a high number of breaks. The realized schedules in England on the other hand suffer a lot from disruptions, reflected by a very high ranking difference and the loss of perfect balancedness in each realized schedule. The quality of the schedules in the Netherlands appears to be quite poor when looking at the high number of breaks and the unevenly balanced home advantage; on the other hand, it often shows relatively low values for ranking difference.

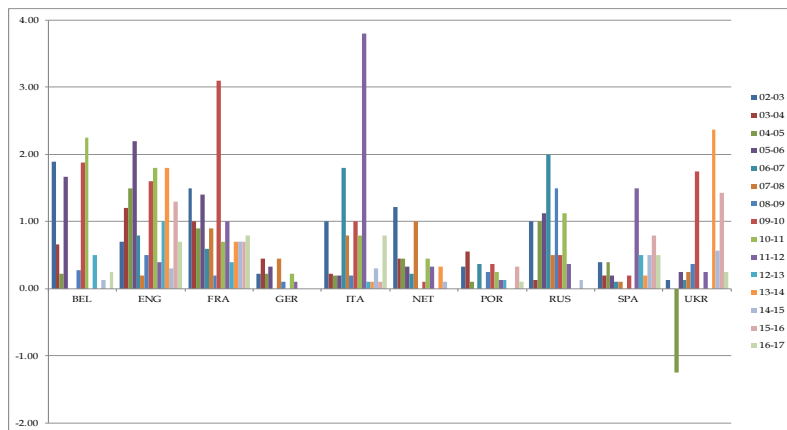


Figure 8: Effect of disruptions on the average number of breaks per team

Figure 8 shows the effect of disruptions on the average number of breaks per team for each season and each league. We observe that the number of breaks is quite sensitive to disruptions. On a few occasions (France 2009–2010, Italy 2011–2012), the total number of breaks even has doubled. As an exception, we mention that for the season 2004–2005 in Ukraine, the number of breaks decreased because of disruptions; indeed, the initial schedule had 60 breaks whereas the realized schedule was completed with 40 breaks (see Table 1). Furthermore, although the initial schedules in England and the Netherlands already have a lot of breaks, disruptions are not handled in a way that minimizes the number of breaks.

Figure 9 shows the effect of disruptions on the average ranking difference per team for each season and each league. It can be seen that the ranking difference in England is at a high level in almost every season. Clearly, during the season, the Premier League ranking is often heavily biased by differences in the number of matches played per team. In contrast, Germany, the Netherlands, Portugal, and Ukraine display a low ranking difference in most seasons (e.g. 6 out of 15 seasons in

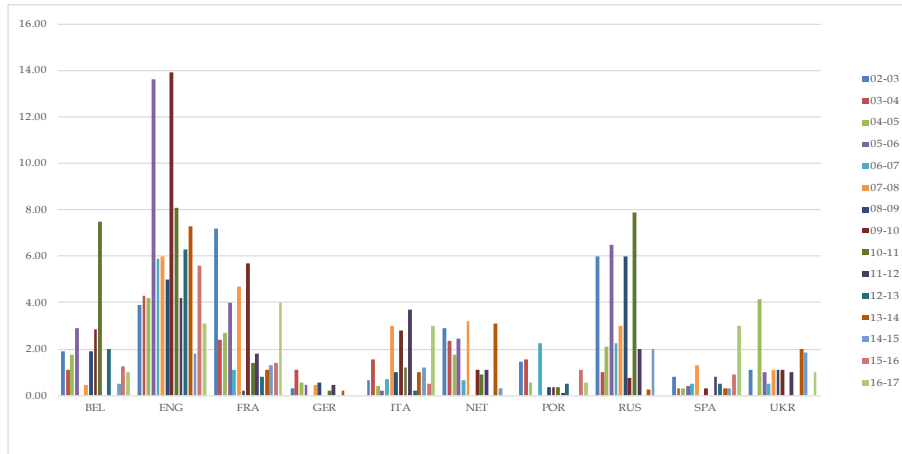


Figure 9: Effect of disruptions on the average value of  $RD$  per team

Germany resulted in  $RD = 0$ ). Belgium, Spain and Russia, sometimes show peaks in  $RD$ , largely corresponding to seasons where disruptions were numerous.

| Differences        |    | BEL | ENG | FRA | GER | ITA | NET | POR | RUS | SPA | UKR |
|--------------------|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| consecutive breaks | -1 | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
|                    | 0  | 11  | 2   | 8   | 13  | 3   | 9   | 10  | 12  | 7   | 11  |
|                    | 1  | 3   | 12  | 6   | 2   | 11  | 6   | 5   | 3   | 8   | 4   |
|                    | 2  | 1   | 0   | 1   | 0   | 1   | 0   | 0   | 0   | 0   | 0   |
| $k$ -balancedness  | 0  | 11  | 1   | 7   | 13  | 9   | 13  | 11  | 9   | 10  | 11  |
|                    | 1  | 3   | 11  | 7   | 2   | 6   | 2   | 4   | 6   | 5   | 4   |
|                    | 2  | 1   | 2   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
|                    | 3  | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| $g$ -balancedness  | 0  | 13  | 3   | 8   | 14  | 11  | 11  | 13  | 10  | 12  | 11  |
|                    | 1  | 2   | 9   | 7   | 1   | 4   | 4   | 2   | 5   | 3   | 4   |
|                    | 2  | 0   | 3   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |

Table 2: Number of seasons with different observations on consecutive breaks,  $k$ -balancedness and  $g$ -balancedness per league

We have also analyzed the sensitivity of the maximum number of consecutive breaks, the  $k$ -balancedness and the  $g$ -balancedness to disruptions, as shown in Table 2. We found that disruptions only have a modest impact on these measures; in some cases these values remain the same despite the occurrence of disruptions. The first measure, i.e., consecutive breaks, has varied in 80 cells in Table 1, with a difference ranging from  $-1$  to  $2$ . There are three cells with the difference of  $2$ , corresponding with Belgium (2009–2010), France (2009–2010) and Italy (2011–2012). Meanwhile, on three occasions, the maximum number of consecutive breaks was reduced because of disruptions: England (2010–2011), and Ukraine (2006–2007 & 2009–2010). The  $k$ -index varies in 54 cells in Table 1, with a difference ranging from  $1$  to  $3$ . Only for England (2010–2011) we have a increase of  $3$ . As for the  $g$ -index, it varies in 44 cells in Table 1, with a difference of  $1$  or  $2$ . Again, only England displays a increase of  $2$ , in seasons 2008–2009, 2009–2010, and 2010–2011. Generally, the impact of disruptions on the  $g$ -index is less profound than the impact on the  $k$ -index, which in turn is smaller than the impact on the maximum number of consecutive breaks. For these three measures, the

difference rarely exceeds one. England experiences a deterioration of the balancedness in all but one season as the competition develops, and as such, never benefits from the high-quality ( $k = g = 1$ ) initial schedules of some seasons. Spain, Germany, and Portugal on the other hand, have the most stable schedules.

Overall, these observations indicate that most leagues use a (re)scheduling policy that is not very effective at minimizing the impact of disruptions on the quality measures that we considered, except with respect to avoiding consecutive breaks. Note however that other constraints (e.g. stadium availability, police requests, ...) may have played a role in this as well.

## 5. Proactive and reactive policies

In the remainder of this paper we try to mitigate the negative effects of disruptions on the quality of football schedules. The quality of a schedule is evaluated on the basis of before-mentioned quality measures: breaks,  $k$ -balancedness, ranking difference and failures. We study several proactive and reactive policies to deal with disrupted matches; typically, a proactive strategy is combined with a reactive strategy (see Section 6).

### 5.1 Proactive policies

Our proactive policies try to anticipate the realization of unforeseen events by inserting catch-up rounds as buffers into the initial schedule. Recall that catch-up rounds are empty rounds to which disrupted matches can be rescheduled. The idea of having such buffers is also mentioned in [3], who develop a time-relaxed schedule for a professional basketball league, where so-called rest weeks can be used a posteriori for any reschedulings that may be required due to unforeseen circumstances. However, the authors consider the position of these rest weeks as given (and the rescheduling process not discussed either). We take four proactive policies into consideration:

- spread catch-up rounds equally over the season (PS)
- spread catch-up rounds equally over the second half of the season (P2)
- position all catch-up rounds near the end of the season (PE)
- spread catch-up rounds over the season according to the distribution of disruptions in Figure 6(b) (PD)

We assume that consecutive catch-up rounds are not allowed (we haven't seen evidence of such practices in our empirical study). Note that each of our proactive policies puts a catch-up round at the end of the season to make sure that disruptions happening in the final round would not automatically lead to a cancellation.

### 5.2 Reactive Policies

Instead of strengthening the initial schedule, reactive policies revise or re-optimize the initial schedule after the unforeseen event has occurred. Hence, our reactive policies concentrate on rescheduling postponed games to appropriate catch-up rounds. More specifically, after each round, it becomes clear which games are disrupted. The reactive policy will determine to which catch-up rounds these disrupted games should be rescheduled, however, taking into account that catch-up rounds – as any other round – cannot have more than one match involving the same team. Note that, in line with common practice in football, we do not allow to make changes to other matches from the



initial schedule. We also assume that a disruption can only be rescheduled to a catch-up round which is not immediately following its original round, since otherwise we would be violating its definition (see Section 2). An exception holds for matches in the last round that cannot be played as planned: they can be rescheduled to the subsequent catch-up round (provided that there is one). Recall that this paper assumes compact schedules, which implies that disrupted matches can only be rescheduled to catch-up rounds (in a time-relaxed schedule, one could reschedule a disrupted match to a regular round, provided that both teams do not play on that round). We consider the following three reactive policies:

- reschedule a disruption to the first available catch-up round (FA)
- reschedule a disruption to the best available catch-up round with respect to the number of breaks (BB)
- reschedule a disruption to the best available catch-up round with respect to  $k$ -balancedness (BK)

The first available (FA) reactive policy reschedules each disrupted game to the earliest possible catch-up round, and as such it is the best one can do to minimize ranking difference. The other two policies are referred to as best break (BB) and best  $k$ -index (BK), and aim at rescheduling disrupted matches with minimal breaks and smallest  $k$  value respectively. We discuss these reactive policies under two settings: (i) each disrupted match should be rescheduled right away and permanently, which is referred to as *fixed* setting; (ii) a disruption can be rescheduled to a catch-up round temporarily and this arrangement can be changed later if more information becomes available. We refer to this as the *flexible* setting.

While the first available policy is easy to implement, we have developed an IP formulation for each of the other two strategies. We use  $T$  to denote the set of regular rounds, and  $C$  for the set of catch-up rounds. The set of disrupted matches is given by  $D$ , where  $D_i \subseteq D$  refers to the set of disrupted matches featuring team  $i \in S$ . Let  $h_d$  ( $a_d$ ) refer to the home (away) team of disrupted match  $d \in D$ . We use  $q_{i,r}$  for the difference between the number of home and away games played by team  $i$  after round  $r \in T \cup C$ , when all disrupted matches are omitted. We define the decision variable  $x_{d,c}$  which is 1 if disrupted match  $d \in D$  is rescheduled on catch-up round  $c \in C$ , and 0 otherwise. Furthermore,  $k_{i,r}$  represents the difference between the number of home and away games played by team  $i$  after round  $r \in T \cup C$  in the repaired schedule. Variable  $y_d$  is 1 if disrupted match  $d \in D$  is a failure, and 0 otherwise. We now present the formulation for the best  $k$ -index policy as an example.

Minimize  $k + |T| \sum_{d \in D} y_d$   
subject to

$$k_{i,r} \leq k \quad \forall i \in S, r \in T \cup C, \quad (11)$$

$$-k_{i,r} \leq k \quad \forall i \in S, r \in T \cup C, \quad (12)$$

$$k_{i,r} = q_{i,r} + \sum_{d:h_d=i} \sum_{c:c \leq r} x_{d,c} - \sum_{d:a_d=i} \sum_{c:c \leq r} x_{d,c} \quad \forall i \in S, r \in T \cup C, \quad (13)$$

$$\sum_{d \in D_i} x_{d,c} \leq 1 \quad \forall i \in S, c \in C, \quad (14)$$

$$\sum_{c \in C} x_{d,c} = 1 + y_d \quad \forall d \in D, \quad (15)$$

$$x_{d,c} \in \{0, 1\} \quad \forall d \in D, c \in C, \quad (16)$$

$$y_d \in \{0, 1\} \quad \forall d \in D. \quad (17)$$

The objective is to reschedule disrupted games such that the  $k$ -index value is minimized, provided that the number of failures is minimal (note that  $|T|$  is an upper bound to  $k$ ). Constraints (11)-(13) compute the  $k$ -index value, while constraints (14) indicate that each team should play at most one match per catch-up round. Constraints (15) enforce that matches that are not rescheduled to a catch-up round count as a failure. The integrality constraints are given in (16) and (17). In the flexible setting, the idea is to solve this formulation after each round  $r$ , with variables  $x_{d,c}$  with  $c \leq r$  fixed to the value they had after solving the model for round  $r - 1$ . In the fixed setting, additionally, each  $x_{d,c}$  which was 1 after solving the model at some round, is fixed to 1 for all following rounds.

## 6. Simulations

In this section, we test the performance of several combinations of a proactive and a reactive policy. We are mainly interested in two things: (i) how does each combined strategy behave with respect to four performance indexes (breaks,  $k$ -balancedness, ranking differences and failures), and (ii) what is the impact of the number of disruptions and the number of catch-up rounds on the quality of the realized schedules.

### 6.1 Experimental design

Motivated by the omnipresence of this competition format in professional European football leagues [7], we consider a mirrored double round robin (2RR) tournament with 20 teams, played using a compact, phased schedule. Moreover, a canonical schedule based on Eq. (1) is used as the initial schedule, determining for each round which team plays against which opponent. Hence, we have 38 rounds in this initial schedule, 54 breaks, a  $k$ -balancedness of 2, and  $RD = 0$ . Disruptions are generated randomly in accordance with the distribution from our empirical study (see Figure 6(b)).

In the first part of our simulation 12 disruptions are generated, and four catch-up rounds are available, leading to 42 rounds in total. Table 3 shows where each of the proactive policies positions the catch-up rounds in this case. For the second part, we consider a number of disruptions ranging from 4 to 40, with the number of catch-up rounds varying from 1 to 10.

| Proactive policy | $c_1$ | $c_2$ | $c_3$ | $c_4$ |
|------------------|-------|-------|-------|-------|
| PS               | 11    | 21    | 31    | 42    |
| P2               | 20    | 27    | 34    | 42    |
| PE               | 39    | 40    | 41    | 42    |
| PD               | 20    | 24    | 33    | 42    |

Table 3: Positions of 4 catch-up rounds for each proactive policy

We point out that in general, failures will create fewer breaks than rescheduled matches. Indeed, while rescheduling a match has a double impact on the home-away patterns of the involved teams (around the initial round and the catch-up round to which it is rescheduled), cancelling a match only changes the home-away pattern around initial round. Consequently, in this simulation study, we use breaks per played game ( $PG\_breaks$ ) as an alternative measure:

$$PG\_breaks = \frac{breaks}{total\ number\ of\ games - number\ of\ failures}. \quad (18)$$

## 6.2 Results

### 6.2.1 Comparison of combined proactive-reactive strategies

In this part, we study 4 proactive policies and 3 reactive policies under fixed and flexible settings. This leads to 24 different combinations of proactive and reactive strategies, which all have been tested for a season with 12 disruptions and 4 available catch-up rounds. The results, averaged over 100 season simulations, for each strategy are compared with respect to breaks,  $k$ -balancedness, ranking difference, and failures, in Tables 4, 5, 6, and 7 respectively. The final column in each table indicates whether differences (ordered by breaks| $k$ | $RD$ |failures) between the fixed and the flexible setting are significant. For each proactive policy, the best result per quality measure is bolded; the best values overall for each quality measure are starred.

The results confirm that the FA, BB, and BK reactive policies are the best choices for minimizing the number of the ranking difference, breaks, and the  $k$ -index respectively. With respect to proactive policies, PS performs better in terms of  $RD$  while the worst performance is for the PE policy. This can be explained by the fact that disruptions in the first half of the season can be rescheduled soon under PS, while the opposite is true for PE. On the other hand, PS has a poor performance with respect to avoiding failures, which is the only measure for which PE excels. P2 has superior behaviour on breaks and  $k$ -balancedness. As for the PD strategy, it presents reasonable results on each quality measures: breaks are slightly higher than in P2,  $k$ -balancedness is almost the same as in P2 and smaller than in PS. Moreover,  $RD$  in PD is much lower than in PE, and the number of failures lower than in P2, and only half the number of failures produced by PS. Furthermore, the results show that the number of failures can be reduced significantly when adopting a flexible setting. Except under the proactive policy where all catch-up rounds are pushed towards the end of the season, the flexible setting leads to a significantly different performance than the fixed setting. In general, the flexible setting improves the quality of the results, although in some cases where the reactive policy is not aligned with the quality measure (e.g. the performance with respect to  $RD$  of PS-BK) the opposite is true.

### 6.2.2 Impact of disruptions and available catch-up rounds

In this part, we vary the number of disruptions from 4, 8, 12, ... to 40, and the number of catch-up rounds from 1, 2, ... to 10 (for each combination, 100 simulations are done) in order to study their impact on the quality of the realized schedule. Figures 10(a), 11(a), 12(a), and 13(a) show the results of combined proactive-reactive strategies that perform best with respect to breaks,  $k$ -balancedness,  $RD$ , and failures, respectively, based on the results in Section 6.2.1. Figures 10(b), 11(b), 12(b), and 13(b) show the performance differences between these strategies and the PD proactive strategy coupled with the best reactive strategy for each quality measure. Note that all strategies discussed here assume a flexible setting.

Overall, Figures 10–13 show a rising trend for breaks,  $k$ -balancedness, ranking difference and failures as the number of disruptions increases. Having more catch-up rounds allows to mitigate this effect, although this is much more effective for reducing the number of failures or the ranking difference, than it is for breaks or the  $k$ -index. For the policies depicted in these figures, unless there are more than 20 disruptions per season, the added values of having more than 4 catch-up rounds seems very limited.

Figures 10–13 also give a clearer picture of the performance of a number of strategies in other situations than the one with 4 catch-up rounds and 12 disruptions studied in Section 6.2.1. With respect to breaks (Figure 10), we find that P2-BB and PD-BB behave quite comparably to an increase in number of disruptions, the latter even gains the advantage as the number of catch-up

| Reactive Policy |          | PG_breaks    | $k$          | $RD$          | failures     | significant difference |
|-----------------|----------|--------------|--------------|---------------|--------------|------------------------|
| FA              | fixed    | 0.215        | 3.137        | 181           | 1.252        | Y Y Y Y                |
|                 | flexible | 0.217        | 3.121        | <b>175.2*</b> | 0.996        |                        |
| BB              | fixed    | 0.210        | 3.237        | 233.4         | 1.49         | Y Y Y Y                |
|                 | flexible | <b>0.208</b> | 3.122        | 207.1         | 0.996        |                        |
| BK              | fixed    | 0.215        | 3.137        | 181.3         | 1.254        | Y Y Y Y                |
|                 | flexible | 0.210        | <b>3.084</b> | 236.1         | <b>0.977</b> |                        |

Table 4: Results of the PS proactive policy

| Reactive Policy |          | PG_breaks     | $k$          | $RD$         | failures     | significant difference |
|-----------------|----------|---------------|--------------|--------------|--------------|------------------------|
| FA              | fixed    | 0.215         | 3.169        | 189.8        | 0.63         | Y Y Y Y                |
|                 | flexible | 0.216         | 3.118        | <b>182.7</b> | <b>0.485</b> |                        |
| BB              | fixed    | 0.216         | 3.193        | 194.2        | 0.653        | Y Y Y Y                |
|                 | flexible | <b>0.207*</b> | 3.097        | 209.5        | <b>0.485</b> |                        |
| BK              | fixed    | 0.216         | 3.192        | 195.35       | 0.651        | Y Y Y Y                |
|                 | flexible | 0.212         | <b>3.04*</b> | 244.9        | 0.486        |                        |

Table 5: Results of the P2 proactive policy

| Reactive Policy |          | PG_breaks     | $k$          | $RD$         | failures     | significant difference |
|-----------------|----------|---------------|--------------|--------------|--------------|------------------------|
| FA              | fixed    | 0.2105        | 3.675        | <b>457.9</b> | 0.088        | N N N Y                |
|                 | flexible | 0.2104        | 3.675        | <b>457.9</b> | <b>0.08*</b> |                        |
| BB              | fixed    | 0.2103        | 3.676        | <b>457.9</b> | 0.169        | N N N Y                |
|                 | flexible | <b>0.2102</b> | 3.675        | <b>457.9</b> | <b>0.08*</b> |                        |
| BK              | fixed    | 0.2104        | 3.675        | <b>457.9</b> | 0.163        | N N N Y                |
|                 | flexible | 0.2103        | <b>3.674</b> | <b>457.9</b> | <b>0.08*</b> |                        |

Table 6: Results of the PE proactive policy

| Reactive Policy |          | PG_breaks    | $k$          | $RD$         | failures    | significant difference |
|-----------------|----------|--------------|--------------|--------------|-------------|------------------------|
| FA              | fixed    | 0.218        | 3.169        | 190.2        | 0.539       | Y Y Y Y                |
|                 | flexible | 0.217        | 3.138        | <b>184.3</b> | <b>0.48</b> |                        |
| BB              | fixed    | 0.217        | 3.174        | 193.3        | 0.646       | Y Y Y Y                |
|                 | flexible | <b>0.213</b> | 3.112        | 202.8        | <b>0.48</b> |                        |
| BK              | fixed    | 0.217        | 3.17         | 193.3        | 0.589       | Y Y Y Y                |
|                 | flexible | 0.214        | <b>3.069</b> | 237.6        | <b>0.48</b> |                        |

Table 7: Results of the PD proactive policy

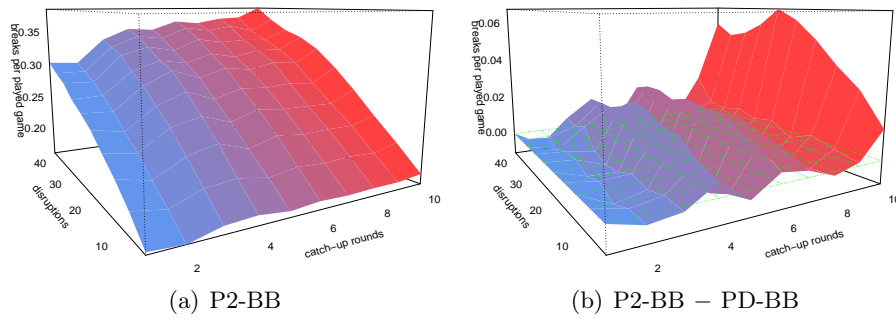


Figure 10: Impact of disruptions and catch-up rounds on breaks

rounds increases. As far as the  $k$ -index is concerned (Figure 11), P2-BK outperforms PD-BK when few catch-up rounds are available or disruptions are rather rare; in the opposite case, the latter is the preferred strategy. On the basis of Figures 12, the ranking difference is similarly handled by PS-FA and by PD-FA. In most cases, the differences on the ranking difference between PS-FA and PD-FA are close to or equal to zero. Comparing PE-FA and PD-FA in terms of avoiding failures, Figure 13 shows PD-FA performing better in situations with a lot of disruptions.

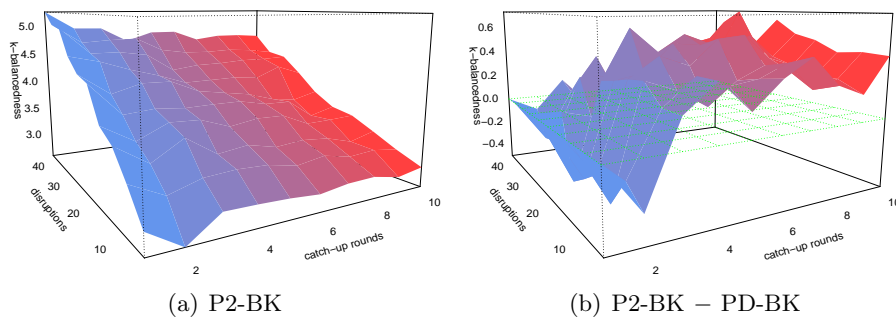


Figure 11: Impact of disruptions and catch-up rounds on  $k$ -balancedness

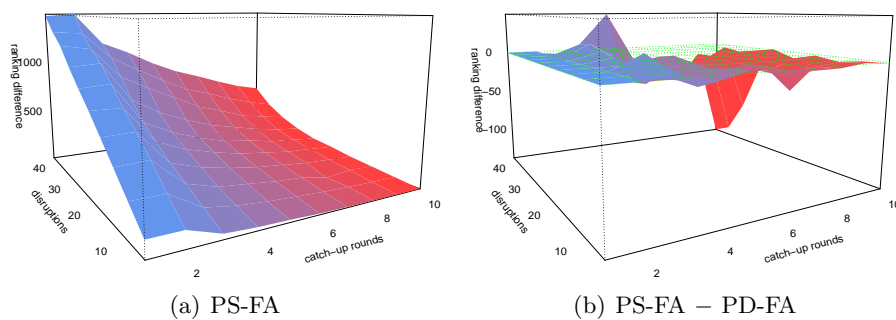


Figure 12: Impact of disruptions and catch-up rounds on ranking difference

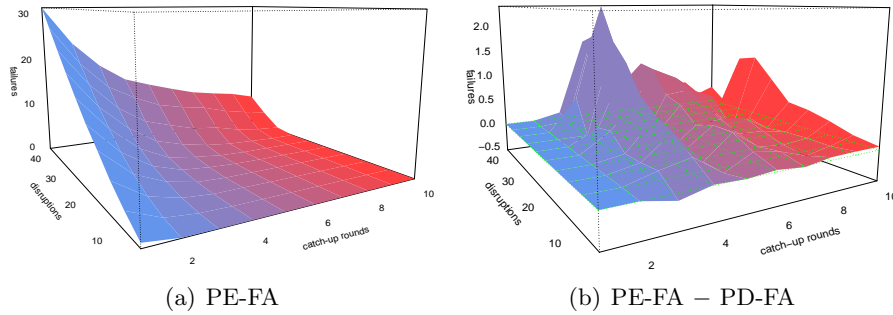


Figure 13: Impact of disruptions and catch-up rounds on failures

### 6.3 Application to real-life football seasons

Since the simulations in Section 6.2 point towards the flexible PD-FA policy with four catch-up rounds as an adequate strategy with respect to several quality measures, we now test the performance of this approach on 11 real-life cases, selected for their high number of rounds with disruptions. Table 8 (third column) shows the total number of games and the number of disruptions for each of these seasons.

| League | year   | matches | initial         | realized         | simulated       |
|--------|--------|---------|-----------------|------------------|-----------------|
| ENG    | 03-04  | 380/8   | (94,2)          | (118,3,86,0,8)   | (110,3,126,0,4) |
|        | 04-05  | 380/9   | (130,1)         | (160,2,84,0,6)   | (138,2,110,0,4) |
|        | 09-10  | 380/22  | (172,1)         | (204,3,318,0,11) | (181,2,330,0,4) |
|        | 10-11  | 380/15  | (124,3)         | (160,6,318,0,9)  | (162,6,96,0,4)  |
|        | 12-13  | 380/11  | (130,2)         | (150,3,126,0,6)  | (140,3,142,0,4) |
| 15-16  | 380/10 | (138,2) | (164,2,112,0,3) | (150,2,46,0,3)   |                 |
| RUS    | 06-07  | 240/10  | (60,6)          | (92,7,36,0,5)    | (78,6,54,0,4)   |
| ITA    | 11-12  | 380/20  | (66,2)          | (142,3,74,0,6)   | (94,3,204,0,4)  |
| SPA    | 11-12  | 380/22  | (54,2)          | (84,3,16,0,4)    | (80,3,32,0,4)   |
| UKR    | 04-05  | 240/13  | (60,2)          | (40,3,66,0,5)    | (74,4,24,0,4)   |
|        | 14-15  | 182/8   | (52,3)          | (60,6,26,0,2)    | (82,5,20,0,2)   |
| BEL    | 10-11  | 240/14  | (42,2)          | (72,3,120,0,6)   | (88,4,124,1,4)  |

Table 8: Application of the PD-FA policy with 4 catch-up rounds to real-life seasons. For initial schedules, only the number of breaks followed by  $k$ -balancedness is reported; realized and simulated schedules feature the number of breaks,  $k$ -balancedness, ranking difference, failures and the number of used catch-up rounds, in that order.

Table 8 shows that, although in reality many more catch-up rounds were used, our approach handles 10 out of 11 leagues successfully with only four catch-up rounds (Belgium 10-11 being a notable exception with one failure). Our (simulated) realized schedules for Russia (06-07) and England (15-16) outperforms the real-life solutions in terms of all quality measures. In general, our approach needs fewer breaks than the solution developed in reality, although its resulting ranking difference is often larger. The latter is at least to some extent due to the fact that more catch-up rounds are used in real life, which allows rescheduling disruptions sooner.

## 6.4 Recommendations to Practitioners

Our simulations involve a realistic scenario based on professional football in Europe (Section 6.2.1), as well as scenarios with other levels of disruptions and catch-up rounds than common in European football leagues (Section 6.2.2), and have been validated on real-life cases. Hence, the following recommendations are mainly focussed on practitioners from football, but should also have some value for other sports.

In general, policies work better using the flexible than the fixed setting. Hence, we would advise practitioners to wait as long as possible before announcing the round to which a disrupted match is to be rescheduled. Indeed, as more information becomes available, disruptions can be handled more adequately. On the other hand, teams, fans, and TV companies may not be happy with a late decision on a new date for a disrupted game. Hence, practitioners face a trade-off between waiting for useful new information and informing all involved stakeholders well in advance.

Next, practitioners should develop a clear idea about the quality measure they care about most. For instance, rescheduling a disrupted game to the earliest available round is a good idea if one wants to minimize the ranking difference. The preferred quality measure also has implications on which proactive strategy will work best. If failures are to be avoided at all cost, it is best to move the catch-up rounds as much to the end as possible. In contrast, catch-up rounds should be balanced over the entire season to have the best effect if having a representative ranking at each point in the season is paramount. Provided that all quality measures are relevant, balancing the catch-up rounds over the second half the season or according to a disruption distribution is advised.

Finally, our simulations show that four or five catch-up rounds should suffice to handle up to 20 disruptions. Extra catch-up rounds only show added value in cases where even a single failure is highly problematic or when considerably more than 20 disruptions are to be expected. In general, having more than 5 catch-up rounds only marginally improves the results with respect to breaks, the  $k$ -index and ranking difference, which probably would not outweigh the cost of reserving venues and timeslots for catch-up matches.

## 7. Conclusions

This paper studies the impact of uncertainty on the quality of football schedules of leading European leagues, and proposes several proactive and reactive strategies for dealing with disruptions.

Essentially, our empirical study indicates (i) that unforeseen events disrupt about 1.6% of the matches and (ii) that even a small number of disruptions substantially impacts the quality of the realized schedule. Indeed, in almost all cases disruptions (drastically) increased the number of breaks, and in half of the cases, the ranking balancedness worsened as well. Furthermore, disruptions often contribute to a high ranking difference. We conclude that, although there are large differences among the leagues, European football schedules are in general not robust with respect to disruptions.

Thus, four proactive and three reactive policies are proposed for anticipating and handling disruptions. According to our simulations, different combinations of proactive-reactive policies perform best on different quality measures. Overall, positioning catch-up rounds over the season according to the distribution of disruptions is a proactive strategy that allows to perform reasonably well on all quality measures. The advised reactive strategy depends heavily on the preferred quality measure. Furthermore, if 5 catch-up rounds are available, the impact of additional catch-up rounds drastically decreases.

Although this paper has focussed on football, our strategies, methodology, and quality measures can also be applied to other sports that play according to a compact round robin tournament.

Furthermore, a promising area of research would be to apply proactive and reactive strategies to time-relaxed round robin schedules, which are common e.g. in amateur sports.

## Acknowledgments

PhD research of Xiajie Yi is supported by the Chinese Scholarship Council (CSC, China, CSC201606890038) and by the Special Research Fund of Ghent University (Bijzonder Onderzoeksfonds, Belgium).

## References

- [1] D. De Werra. Scheduling in sports. In P. Hansen, editor, *Studies on Graphs and Discrete Programming*, volume 11 of *Annals of Discrete Mathematics*, pages 381–395. North-Holland, Amsterdam, 1981.
- [2] A. Drexler and S. Knust. Sports league scheduling: Graph- and resource-based models. *Omega*, 35:465–471, 2007.
- [3] Guillermo Durán, Santiago Durán, Javier Marenco, Federico Mascialino, and Pablo A Rey. Scheduling Argentina’s professional basketball leagues: A variation on the travelling tournament problem. *European Journal of Operational Research*, 275(3):1126–1138, 2019.
- [4] D. Forrest and I. McHale. Subjective well-being and engagement in sport: evidence from England. In P. Rodríguez, S. Késenne, and B. Humphreys, editors, *The Economics Of Sport, Health And Happiness*, pages 184–199. Edward Elgar Publishing, 2011.
- [5] D. Forrest and R. Simmons. New issues in attendance demand: The case of the English football league. *Journal of Sports Economics*, 7(3):247–266, 2006.
- [6] D. Goossens and F.C.R. Spijksma. Scheduling the Belgian soccer league. *Interfaces*, 39(2):109–118, 2009.
- [7] D Goossens and F.C.R Spijksma. Soccer schedules in Europe: an overview. *Journal of Scheduling*, 15(5):641–651, 2012.
- [8] Xiao Le Han, Zhi Qiang Lu, and Li Feng Xi. A proactive approach for simultaneous berth and quay crane scheduling problem with stochastic arrival and handling time. *European Journal of Operational Research*, 207:1327–1340, 2010.
- [9] W. Herroelen and R. Leus. Project scheduling under uncertainty: Survey and research potentials. *European Journal of Operational Research*, 165:289–306, 2005.
- [10] G. Kendall, S. Knust, C.C. Ribeiro, and S. Urrutia. Scheduling in sports: An annotated bibliography. *Computers and Operations Research*, 37:1–19, 2010.
- [11] S. Knust and M. von Thaden. Balanced home-away assignments. *Discrete Optimization*, 3:354–365, 2006.
- [12] Olivier Lambrechts, Erik Demeulemeester, and Willy Herroelen. Proactive and reactive strategies for resource-constrained project scheduling with uncertain resource availabilities. *Journal of Scheduling*, 11(2):121–136, 2008.



- [13] K. Nurmi, D. Goossens, T. Bartsch, F. Bonomo, D. Briskorn, G. Duran, J. Kyngäs, C.C. Ribeiro, F. Spieksma, S. Urrutia, and R. Wolf-Yadlin. A framework for a highly constrained sports scheduling problem. *IAENG Transactions on engineering technologies*, 5:14–28, 2010.
- [14] R. Pollard. Home advantage in soccer: Variations in its magnitude and a literature review of the inter-related factors associated with its existence. *Journal of Sport Behavior*, 29(2): 169–189, 2006.
- [15] Rasmus V Rasmussen and Michael A Trick. Round robin scheduling—a survey. *European Journal of Operational Research*, 188(3):617–636, 2008.
- [16] Celso C Ribeiro. Sports scheduling: Problems and applications. *International Transactions in Operational Research*, 19(1-2):201–226, 2012.
- [17] I. Sabuncuoglu and M. Bayiz. Analysis of reactive scheduling problems in a job shop environment. *European Journal of Operational Research*, 126:567–586, 2000.
- [18] A. Seidscher and S. Minner. A Semi-Markov decision problem for proactive and reactive transshipments between multiple warehouses. *European Journal of Operational Research*, 230: 42–52, 2013.
- [19] David Van Bulck, Dries Goossens, Jörn Schönberger, and Mario Guajardo. Robinx: a three-field classification and unified data format for round-robin sports timetabling. *European Journal of Operational Research*, 280(2):568–580, 2020.
- [20] S. Van de Vonder, E. Demeulemeester, and W. Herroelen. Proactive heuristic procedures for robust project scheduling: An experimental analysis. *European Journal of Operational Research*, 189:723–733, 2008.
- [21] MB Wright. 50 years of or in sport. *Journal of the Operational Research Society*, 60(sup1): S161–S168, 2009.