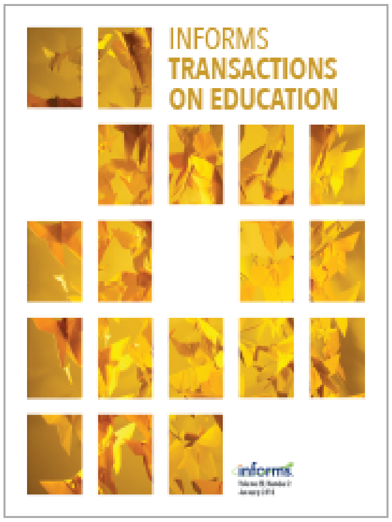


INFORMS Transactions on Education

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Teaching Integer Programming by Scheduling the Belgian Soccer League

Dries Goossens, Jeroen Beliën

INFORMS TRANSACTIONS ON EDUCATION
Articles in Advance, pp. 1–9
ISSN 1532-0545 (online)
http://pubsonline.informs.org/journal/ited

Teaching Integer Programming by Scheduling the Belgian Soccer League

Dries Goossens,[a,b,*] Jeroen Beliën[c]

[a] Department of Business Informatics and Operations Management, Ghent University, 9000 Ghent, Belgium; [b] FlandersMake@UGent – core lab CVAMO, 9000 Ghent, Belgium; [c] Research Centre for Operations Management, KU Leuven, Campus Brussels, 1000 Brussels, Belgium
*Corresponding author

Contact: dries.goossens@ugent.be, https://orcid.org/0000-0003-0224-3412 (DG); jeroen.belien@kuleuven.be, https://orcid.org/0000-0003-2349-7390 (JB)

Received: December 28, 2020
Revised: June 7, 2021; October 15, 2021
Accepted: January 10, 2022
Published Online in Articles in Advance: May 25, 2022

https://doi.org/10.1287/ited.2022.0269

Abstract. This paper presents a didactic approach for teaching integer programming starting from a real-life case on scheduling the Belgian soccer league. We share our experiences as well as didactic resources for two teaching formats. The first format involves hands-on exercises and is more appropriate for small student groups (up to 40 students). The second format is an interactive lecture that focuses more on current research challenges and is better suited for large student groups. During the last decade, both formats have been used repeatedly with consistent, positive feedback from students. The combination of a hands-on, interactive approach that actively involves students with the particularities and experiences of solving a real-life case makes our approach unique and positively affects students' attention.

History: This paper has been accepted for the *INFORMS Transactions on Education* Special Issue on Cases Based on Real-World Projects from the *INFORMS Journal on Applied Analytics*.

Supplemental Material: The supplemental material is available at https://doi.org/10.1287/ited.2022.0269. The spreadsheet exercises file and solution files are available at https://www.informs.org/Publications/Subscribe/Access-Restricted-Materials.

Keywords: integer programming • sports scheduling • spreadsheet optimization • real-life case

1. Introduction

In sports scheduling, the problem is to decide for each match when it is to be played and which team gets the advantage of playing at its own venue. In a schedule, matches are grouped in rounds, which typically correspond to weekends, such that each team plays, at most, one game per round. In this paper, we focus on a case study from the Jupiler Pro League, the highest professional soccer division in Belgium, whose scheduling process is described in a paper by Goossens and Spieksma (2009) in *Interfaces* (now known as the *INFORMS Journal on Applied Analytics*). This league consists of 18 teams, which play a so-called double-round-robin tournament (i.e., a tournament where each team plays against each other team twice). Finding a good schedule can be quite a challenge, as there is a large variety and number of—often conflicting—wishes and requirements from stakeholders (clubs, fans, broadcasters, police, etc.) to be taken into account. The paper describes the transition from a manual scheduling procedure to an integer-programming (IP) approach. First, the authors describe an IP formulation that is based on the well-known assignment model. Later, a more sophisticated two-phased approach is discussed, in which first the home advantage is settled for each team on each round. This serves as input for the second phase, where the opponents are determined. Both phases are again solved using integer programming.

Shortly after the publication by Goossens and Spieksma (2009), the competition design of the Jupiler Pro League changed. The number of teams was reduced to 16, and after these teams finish a regular double-round-robin season, a postseason competition with the best six teams now follows. This postseason competition is again organized as a double-round-robin tournament and termed "play-off 1" (which is, in fact, an incorrect name because the term play-off is normally only used for elimination tournaments). Play-off 1 determines the champion as well as the teams that will represent Belgium in the European tournaments in the next season. A particularity of play-off 1 is that the participating teams start with half of the points they collected in the regular season. Goossens (2018) gives a follow-up on the impact of these changes on the scheduling process, as described by Goossens and Spieksma (2009).

In this paper, we focus on play-off 1. Appealing for its all-decisive nature, yet manageable from a computational point of view because of its limited number of teams, this league has interesting assets for teaching integer programming. This paper describes an interactive lecture that is built around four modeling exercises, where students develop a spreadsheet IP model for scheduling play-off 1. Section 2 discusses the literature on using real-life examples for teaching integer programming. A short introduction to sports scheduling is given in Section 3. The spreadsheet models are discussed in Section 4, followed by a description of our classroom experience in Section 5. We conclude in Section 6.

2. Related Literature

Several studies have shown the importance of using real-life examples for motivating students in developing quantitative skills. Popular approaches to bring real-life practice to the classroom are case-based teaching and field-based education. We refer to Drake (2019) for a recent overview on teaching operations research (OR)/management science with cases. The effectiveness of field-based education is described in Gorman (2018a), while Gorman (2018b) presents a recent overview of related literature. Capstone projects, where students work on real-world industry-sponsored projects, are an important manifestation of field-based education. Konrad et al. (2018) provide an interesting study on the impact of capstone projects on student skills development. They turn out to be particularly important for developing skills that go beyond the purely technical capabilities of mastering analytical methodologies and developing models. They train students in defining the problem from a business context; distilling the decision space and objective priorities; collecting and cleaning data; and presenting solutions to stakeholders. Such skills are often neglected in theoretical OR courses. Although direct contact with industrial partners is extremely important, real-life practice can also be introduced in the classroom by less comprehensive and less time-consuming activities as hands-on lectures around real-life applications. Teaching quantitative methods in general can benefit from the students' inherent interest in sports. For instance, Willoughby (2004) presents a well-received one-semester undergraduate course devoted to decision analysis and statistical methodologies in sports. Cochran (2004) also recognizes sports as a context that is more familiar and interesting to many students and can therefore help in improving their understanding and retention. Bickel (2004) demonstrates the use of a baseball example to teach fundamental probabilistic and decision analysis.

An important challenge in teaching IP is to train students in transforming conditional statements of logic into linear relations involving binary decision variables. Stevens and Palocsay (2017) propose a two-step pedagogical approach, in which a conditional requirement is first decomposed into a group of elementary implications, which are subsequently translated into linear constraints. The effectiveness of this approach for instructing business students is demonstrated through an analysis of pretest and posttest results. Many educational papers on teaching IP involve spreadsheet models. Example applications range from simple games or puzzles to functional area business problems and sports, either purely fictitious or (based on) a real-life case. Rasmussen and Weiss (2007) and Weiss and Rasmussen (2007) use a spreadsheet implementation of the Sudoku puzzle to teach IP modeling. Based on a real-life application, Rao and Beliën (2014) present a production-planning case offered in a three-step approach, in which each step introduces a next layer of modeling complexity. Three spreadsheet implementations, as well as their corresponding algebraic formulations, are provided. Pachamanova (2006) presents a distribution-channel selection case that illustrates modeling fixed (entry) costs. Particularly for the education of business students, she emphasizes the importance of the management side, giving priority to making assumptions, modeling, validation, and business decision making, rather than purely focusing on the (spreadsheet) optimization model. Beliën et al. (2013) present spreadsheet optimization models for an online energy-supply game. Compared with industrial applications, these nontraditional applications lie closer to the students' reality, in particular for undergraduate students. As a result, an increase in the students' motivation to develop IP models has been observed. Similar to Rao and Beliën (2014), Beliën et al. (2013) add complexity in a step-by-step approach. We acknowledge the importance of a stepwise approach and adopt a four-step decomposition in this paper.

Sports scheduling is a particularly fruitful didactic application for teaching integer programming, as the problem is intuitive for students to grasp immediately, but, at the same time, the domain is rich enough to illustrate various IP concepts (Trick 2004). Hence, it comes as no surprise that we are not the first to make use of a sports optimization problem in a classroom setting. Chlond (2011) presents a hypothetical, but original, puzzle related to sports league tables that can be solved using an IP formulation. Given a final ranking with summary data (including number of points gained, number of wins, losses, goals for, and goals against) as typically provided in soccer, the challenge of the puzzle is to reconstruct the actual scores for each of the matches. Beliën et al. (2011) present spreadsheet optimization models that can be used to win a fantasy cycling game. The paper by Trick (2004) on teaching integer programming using sports scheduling is most related to our work. It uses a round-robin scheduling problem to illustrate IP modeling, symmetry-breaking constraints, cut generation, and alternative formulations both through a spreadsheet model and algebraic modeling approach. Birge (2004) provides a spreadsheet IP model for a simplified case of developing a schedule for Major League Football involving five teams of the Western Division, where the objective is to minimize the total travel distance. Besides the traveling salesman problem solution methodology, the case illustrates the use of decomposition, column generation, and IP formulation.

3. Sports Scheduling

This section introduces a number of important concepts in sports timetabling. In particular, the (mirrored) double-round-robin tournament format and the clock method are essential for the hands-on soccer league scheduling class (first format of this case; see Section 5). The carry-over effect, home-away patterns, and the phased approach are also touched on in the first format, but are discussed more in-depth in the interactive lecture (second format of this case; see Section 5). For more background, we refer to an extended surveys of the literature and sports-scheduling terminology by Kendall et al. (2010), Rasmussen and Trick (2008), and Ribeiro (2012). A classification of sports-scheduling problems and a repository of problem instances is given by Van Bulck et al. (2020). Goossens and Spieksma (2011) present an overview of schedules used by professional soccer leagues in Europe and discuss their properties.

A schedule for a double-round-robin tournament is often *phased*, in the sense that no team plays any other team for the second time before it has faced all other opponents once. Hence, a phased double-round-robin tournament can be split into two single-round-robin tournaments. If the order of the opponents in each tournament half is identical for each team, we say that the schedule is *mirrored*. Until recently, the vast majority of the tournaments used a mirrored format: It is easier to compute (the second half of the season can simply be copied from the first), and it creates an equal (and maximal) number of rounds between each match A versus B and its counterpart B versus A.

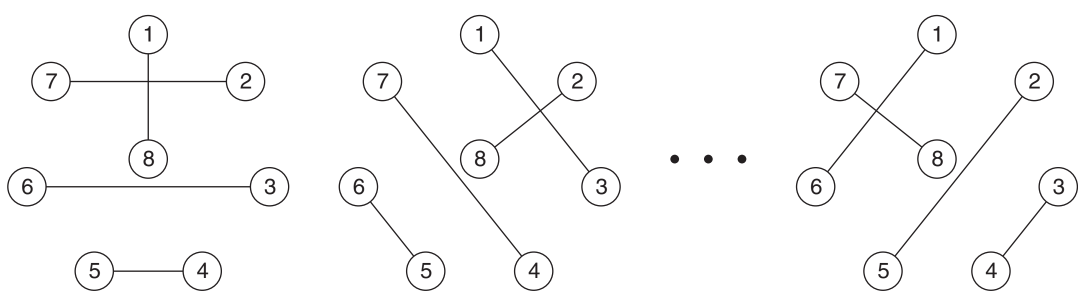
In 1847, Reverend Kirkman, though unconcerned by sports scheduling, published a method that can be used for constructing a schedule for single-round-robin competitions (Kirkman 1847). This method is known as the *clock method* (or *circle method*), and its outcome has been referred to as a *canonical schedule*. A precise description of the circle method for a single round-robin tournament for an even set of teams T is as follows. For each round $r \in \{1, \ldots, |T|-1\}$, we have:

- Team $|T|$ plays against team r,
- Team $[r+k]^+$ plays against team $[r-k]^-$, for each $k \in \{1, 2, \ldots, |T|/2-1\}$,

where $[x]^+ = x$ if $x \leqslant |T|-1$ and $[x]^+ = x - |T| + 1$ otherwise, while $[x]^- = x$ if $x \geqslant 1$ and $[x]^- = x + |T| - 1$ otherwise. Graphically, the clock method is easier to comprehend, as illustrated in Figure 1, where nodes represent teams, and edges correspond to matches. Each round leads to a next round by shifting all edges one stroke clockwise.

It is a common idea in sports that the strength or fitness of an opponent can be impacted by its previous match. For instance, if team A is very strong, its opponent, team B, may be exhausted after playing against team A, or its line-up in the next match may be weakened by suspensions incurred because of (provoked) foul play against team A. Hence, if team C plays against team B in the next round, team C may have a benefit from team A being B's previous opponent. We say that there is a *carry-over effect* from team A to C (Russell 1980). A fair schedule should avoid the situation in which team C frequently plays against the team that played against a strong team (team A) in the previous match. Although there have been attempts to develop schedules that balance the carry-over effects as much as possible (e.g., Anderson 1999), an interesting result is

Figure 1. Representation of the Canonical Schedule for a Competition with Eight Teams; Only Rounds 1, 2, and 7 are Depicted



Notes. In round 2, game 8-2 is obtained after shifting the hand of the clock, which starts at 8 and points at 1 in round 1, one stroke clockwise to 2. For the games that do not involve team 8, both ends of the edge are shifted one stroke clockwise; for example, game 7-2 in round 1 becomes game 1-3 in round 2.

that the canonical schedule actually has the most unbalanced carry-over effects (Lambrechts et al. 2017). In that sense, it is remarkable that this particular schedule is so popular (Goossens and Spieksma 2011).

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*. Nemhauser and Trick (1998) were the first to take advantage of HAPs by splitting the sports-scheduling problem into subproblems. A first subproblem consists of finding a set of HAPs that allows one to construct a schedule. Note that not each set of HAPs is feasible, with each round having an equal number of H's and A's being a simple necessary condition. Another condition is that each pair of HAPs differs, which immediately explains why breaks are unavoidable (for more conditions, see Miyashiro et al. 2003). Next, the opponents are determined for each pattern, taking into account that on any given round, a pattern with an H can only play against a pattern with an A. Finally, the patterns are assigned to actual teams in a third subproblem. Several of the most successful applications in sports scheduling make use of a phased approach, even though the order of the subproblems and the methods used to solve them may differ.

Integer programming plays an important role in solving sport-scheduling problems. In fact, several real-life sport-scheduling applications that have been tackled with integer programming have been reported in the academic literature. We mention soccer leagues in Italy (Della Croce and Oliveri 2006), Chile (Durán et al. 2007), and Ecuador (Recalde et al. 2013); the German basketball league (Westphal 2014); volleyball leagues in Argentina (Bonomo et al. 2012) and Italy (Cocchi et al. 2018); the Canadian football league (Kostuk and Willoughby 2012); and the South American FIFA World Cup qualifiers (Durán et al. 2017) as examples. Integer programming has also been used to solve related problems like referee scheduling (Alarcón et al. 2014), determining whether a team is eliminated for play-offs (Ribeiro and Urrutia 2005), and grouping teams into divisions or series (Toffolo et al. 2019).

4. Hands-On Spreadsheet Exercises

The hands-on part entails developing a schedule for play-off 1 of the highest professional Belgian soccer league. We consider the edition of season 2018–19, which involves the teams Club Bruges FC (CLU), RSC Anderlecht (AND), AA Gent (GEN), Standard Liège (STA), FC Antwerp (ANT), and KRC Genk (GNK). The exercises are incremental: The second, third, and fourth exercise start from the preceding solution and each time add a new modeling challenge. As play-off 1 involves only six teams and the exercises focus on scheduling the first half of the tournament (the second half of the tournament is assumed to be mirrored from the first), the optimization models are limited in terms of the number of decision variables and constraints, making them easily solvable with Microsoft Excel's standard solver.

The first exercise asks students to model an assignment problem. The students are given a so-called basic match schedule, which is essentially a schedule in which the teams have been replaced by numbers (also referred to as "placeholders"). A basic match schedule has the advantage that it already satisfies the single-round-robin conditions—that is, each number plays exactly once against each other number, and every number plays exactly one game in each round. For six teams, this leads to five rounds and three matches per round. An example of a basic match schedule is given in Table 1. This is, in fact, a canonical schedule (generated with the clock method).

The idea is to assign a unique number $n \in N$ (with N the set of numbers) to each team $t \in T$ (with T the set of teams) to obtain a schedule. So, we have decision variables x_{tn} equal to one if team t is assigned to number n, and zero otherwise. Each possible assignment of a number to a team could lead to a number of wishes that are violated, leading to a certain penalty cost. All such penalty costs c_{tn} are provided. Our objective is to minimize the total penalty cost. This leads to the following assignment problem:

$$\text{Minimize} \quad \sum_{t\in T}\sum_{n\in N} c_{tn}x_{tn}, \tag{1}$$

Subject to

$$\sum_{n\in N} x_{tn} = 1 \qquad \forall t \in T, \tag{2}$$

$$\sum_{t\in T} x_{tn} = 1 \qquad \forall n \in N, \tag{3}$$

$$x_{tn} \in \{0,1\} \qquad \forall n \in N,\ \forall t \in T. \tag{4}$$

Constraints (2) ensure that each team is assigned to exactly one number, while Constraints (3) guarantee that each number is assigned to exactly one team. The spreadsheet exercise file and solution file are provided in the accompanying files soccer_template_1.xlsx and soccer_sol_1.xlsx, respectively, in the supplemental material.

The second exercise introduces a first extension of a constraint that involves more than one team and can

Table 1. Example of a Basic Match Schedule for a Single-Round-Robin Tournament Involving Six Teams

Round 1	Round 2	Round 3	Round 4	Round 5
1-2	2-5	1-6	3-2	2-6
3-4	4-1	2-4	5-1	4-5
5-6	6-3	5-3	6-4	1-3

therefore not be handled using the penalty costs c_{tn}. More specifically, the students are asked to extend the previous model, so that a particular team (Gent) does not play a home game against a particular opponent (Standard) in a specific round (round 2). Given the basic match schedule in Table 1, this leads to the following constraints to be added:

$$x_{GEN,2} + x_{STA,5} \le 1, \quad (5)$$

$$x_{GEN,4} + x_{STA,1} \le 1, \quad (6)$$

$$x_{GEN,6} + x_{STA,3} \le 1. \quad (7)$$

Constraints (5)–(7) ensure that whenever Gent is coupled to a number that plays a home game in round 2, Standard cannot be assigned to the number of the opponent in the corresponding game. The spreadsheet exercise file and solution file are provided in the accompanying files soccer_template_2.xlsx and soccer_sol_2.xlsx, respectively, in the supplemental material.

In the third exercise, students are asked to add a constraint that forbids that two specific teams play a home game in the same round (e.g., because they share the same stadium or the teams are geographically so close that the police cannot guarantee safety if both teams play a home game at the same time). If this requirement holds for the teams Standard and Antwerp, given the basic match schedule in Table 1, one possible way to address this is to add:

$$x_{STA,1} = x_{ANT,6}, \quad (8)$$

$$x_{STA,6} = x_{ANT,1}, \quad (9)$$

$$x_{STA,2} = x_{ANT,3}, \quad (10)$$

$$x_{STA,3} = x_{ANT,2}, \quad (11)$$

$$x_{STA,4} = x_{ANT,5}, \quad (12)$$

$$x_{STA,5} = x_{ANT,4}. \quad (13)$$

This model starts from the observation that in the basic match schedule of Table 1, whenever number 1 plays a home game, the only number that in every round plays an away game is number 6. Thus, if Standard is coupled to number 1, Antwerp must be coupled to number 6 (Constraint (8)) and vice versa (Constraint (9)). The same observation holds for the numbers 2 and 3 (Constraints (10) and (11)) and the numbers 4 and 5 (Constraints (12) and (13)). The spreadsheet assignment and solution of the third exercise are provided in the accompanying files soccer_template_3.xlsx and soccer_sol_3.xlsx, respectively, in the supplemental material.

The additional constraints of exercise 2 and exercise 3 can be generalized as follows:

$$\sum_{t\in T}\sum_{n\in N} a_{tnc}x_{tn} \le b_c \qquad \forall c \in C, \quad (14)$$

with $c \in C$ the set of additional constraints and a_{tnc} and b_c the technological coefficients and right-hand side corresponding to constraint c.

So far, each added constraint has increased the objective value. However, it is not hard to realize that as more such constraints are added, the model may become infeasible. This happens, for instance, if, on top of the previous constraints, we need to enforce that in the first five rounds, when either Club or Anderlecht (or both) play at home, Standard wants to play away. Following the general format of Constraints (14), this could be modeled by adding five constraints, one for each of the five rounds:

$$X_{AND,1} + X_{AND,3} + X_{AND,5} + X_{CLU,1} + X_{CLU,3} + X_{CLU,5} - 2X_{STA,2} - 2X_{STA,4} - 2X_{STA,6} \le 0, \quad (15)$$

$$X_{AND,2} + X_{AND,4} + X_{AND,6} + X_{CLU,2} + X_{CLU,4} + X_{CLU,6} - 2X_{STA,5} - 2X_{STA,1} - 2X_{STA,3} \le 0, \quad (16)$$

$$X_{AND,1} + X_{AND,2} + X_{AND,5} + X_{CLU,1} + X_{CLU,2} + X_{CLU,5} - 2X_{STA,6} - 2X_{STA,4} - 2X_{STA,3} \le 0, \quad (17)$$

$$X_{AND,3} + X_{AND,5} + X_{AND,6} + X_{CLU,3} + X_{CLU,5} + X_{CLU,6} - 2X_{STA,2} - 2X_{STA,1} - 2X_{STA,4} \le 0, \quad (18)$$

$$X_{AND,2} + X_{AND,4} + X_{AND,1} + X_{CLU,2} + X_{CLU,4} + X_{CLU,1} - 2X_{STA,6} - 2X_{STA,5} - 2X_{STA,3} \le 0. \quad (19)$$

A solution to the infeasibility problem is to treat the constraints as soft constraints. We introduce penalty variables y_r for each round r and add them with a penalty cost coefficient to the objective function. The Constraints (15)–(19) become:

$$X_{AND,1} + X_{AND,3} + X_{AND,5} + X_{CLU,1} + X_{CLU,3} + X_{CLU,5} - 2X_{STA,2} - 2X_{STA,4} - 2X_{STA,6} \le 2y_1, \quad (20)$$

$$X_{AND,2} + X_{AND,4} + X_{AND,6} + X_{CLU,2} + X_{CLU,4} + X_{CLU,6} - 2X_{STA,5} - 2X_{STA,1} - 2X_{STA,3} \le 2y_2, \quad (21)$$

$$X_{AND,1} + X_{AND,2} + X_{AND,5} + X_{CLU,1} + X_{CLU,2} + X_{CLU,5} - 2X_{STA,6} - 2X_{STA,4} - 2X_{STA,3} \le 2y_3, \quad (22)$$

$$X_{AND,3} + X_{AND,5} + X_{AND,6} + X_{CLU,3} + X_{CLU,5} + X_{CLU,6} - 2X_{STA,2} - 2X_{STA,1} - 2X_{STA,4} \le 2y_4, \quad (23)$$

$$X_{AND,2} + X_{AND,4} + X_{AND,1} + X_{CLU,2} + X_{CLU,4} + X_{CLU,1} - 2X_{STA,6} - 2X_{STA,5} - 2X_{STA,3} \le 2y_5. \quad (24)$$

In general, we introduce additional binary variables y_c equal to one if constraint c is not satisfied, and zero otherwise. If p_c denotes the penalty for violating constraint c, Objective (1) becomes:

$$\text{Minimize} \quad \sum_{t\in T}\sum_{n\in N} c_{tn}x_{tn} + \sum_{c\in C} p_c y_c. \quad (25)$$

And the additional Constraints (14) are replaced by:

$$\sum_{t\in T}\sum_{n\in N} a_{tnc}x_{tn} \le b_c + My_c \qquad \forall c \in C, \quad (26)$$

with M a big number.

The spreadsheet assignment and solution of the fourth exercise are provided in the accompanying files soccer_template_4.xlsx and soccer_sol_4.xlsx, respectively, in the supplemental material.

Note that in case there are multiple optimal solutions, the Excel solver reports only one of them. These exercises could easily be extended with an assignment of identifying all alternative optima (if any), which entails an iterative approach that adds in each iteration a constraint that cuts away the preceding optimal solution. This constraint ensures that the sum of all x-variables equal to one in the previous solution is smaller or equal than the number of teams minus one.

5. Classroom Experience

The case has been given as a two-hour lecture in two different formats: a hands-on computer class session and an interactive lecture without hands-on exercises, but addressing issues beyond the assignment method. Whereas in both formats, the main teaching objective is to show how optimization methods can be successfully applied to real-life problems, the secondary teaching goals slightly differ between both formats. In the first format, the secondary objective is to give students a hands-on training in developing IP models, while in the second format, the secondary objective is to discuss more advanced methodologies and highlight current research challenges.

In both formats, the case starts with an interactive presentation on the challenges, main objectives, and constraints related to developing a new schedule for the Belgian professional soccer league. The PowerPoint file soccer_scheduling_ITE.pptx (see supplemental material) can be used to guide this lecture and is provided as additional material for interested instructors. A lively discussion starts, in which the students are confronted with two questions: (1) Which stakeholders are involved; and (2) what makes up a good schedule? To direct the discussion, students are asked to come up with specific examples of issues that should be taken into account when building a schedule and which stakeholders are affected. During this discussion, students almost always succeed in bringing up all important constraints and related parties themselves. They are classified by the instructor in four categories, objectives and constraints related to: (1) mayors and police, (2) television, (3) clubs, and (4) the league organizer. We point out that it is typically impossible to satisfy all these wishes, as some of them are clearly conflicting.

A first take-away is to show students how complex the problem is in terms of (a) the total number of possible solutions and (b) the difficulty to even find a solution that satisfies the basic round-robin requirements. With respect to (a), a small quiz is done, in which the students need to select the highest number among alternatives like the number of stars in the universe, the number of living insects, the number of people ever born on planet Earth, etc., and the number of feasible schedules for a double-round-robin tournament involving 14 teams. That the latter is by far the biggest number is a real eye-opener to many students. With respect to (b), a small example is given involving only six teams, for which a single-round-robin schedule is built from scratch. Already in round 4 of this schedule, the manual procedure gets stuck into an infeasibility with respect to the constraint that every team plays exactly once against each other team, which can only be fixed by altering earlier rounds. At this point, the clock method (see Section 3) is introduced, and it is shown how this method succeeds in developing a feasible round-robin schedule quickly.

Next, we outline the approach of solving this problem by means of an IP formulation. We explain that even though it sounds very reasonable, we faced a number of issues in real life to convince the league owners to adopt this modeling approach. These involve the loss of insight in the scheduling process when moving from a manual scheduling process to a "black box" and having to explicitly list the requirements and their importance instead of sticking to an "I'll tell you whether or not I like the schedule once you propose one" attitude. Determining the relative importance of the requirements was a difficult exercise for the league owners. At first, their reaction was that simply giving a maximal weight to each constraint would result in the best schedule. At the same time, it became clear that, in the past, some teams had been considered "more equal than others." We believe that these issues are common to many settings where operations research replaces legacy decision processes.

From this point on, the formats differ. The first format continues with the four Excel IP modeling exercises, where the students are given time to tackle them on their own. Feedback is provided during class, and the solutions are gradually presented for each exercise (because each exercise builds on the following). In the second format, the different modeling examples of the assignment method are presented by the teacher, without the implementation in Excel. An interactive discussion starts on the limitations of the assignment method. Issues like the carry-over effect, the fact that there are numerous other basic match schedules, and the consequences in terms of solution quality of mirroring the first half of the season to obtain the second half schedule are raised. More advanced solution methodologies as the two-phase decomposition approach starting from a set of home-away patterns, as well as local search approaches, are subsequently discussed. We refer to Goossens and Spieksma (2009) for a more detailed discussion of these issues and the IP models for the two-phase decomposition approach.

In both formats, the case is concluded with a brief discussion on the reception of the solution. Besides the number of violated constraints (and the total penalty cost), which are clearly much lower than the manual approach could obtain, we show the impact of the improved schedule in terms of reduced cost of policing, increased stadium attendance, and revenue from the TV broadcasting contract. Obviously, the schedule is not the only factor in this story (the numbers are, for instance, also determined by the fanbase of the teams that promote to or are relegated from the first division), but the difference before and after the adoption of the IP modeling approach is unmistakable. Nevertheless, we point out that coaches tend to look at the schedule from the point of view of their team's performance: A good schedule is one where their team wins many games. In that respect, no schedule can be acceptable for all teams, and given that coaches typically get more media attention than other (more important) stakeholders, a negative perception of the schedule with the general public remains a possibility.

We have used the first teaching format (hands-on computer class) nine times as part of a course that focuses on real-life applications of operations research for groups of 20–40 graduate business students who already completed a basic course on operations research in their undergraduate studies. Furthermore, this format has been brought eight times in an introductory course on operations research for postgraduate business students with limited background in mathematics (in this case, the focus was on the first exercise; the other exercises were left for the stronger students to explore on their own). The second format (interactive lecture) was brought for an audience of 120–300 graduate business students and was used in total six times. For both formats, we regularly received positive (informal) feedback from the students. After having given the interactive lecture (second format) in March 2021 as an online class (due to the COVID-19 situation at the time), we distributed a questionnaire that has been filled out by 112 students. This survey consists of eight questions with answers in the format of a Likert scale. The questions and the responses are listed in Table 2. The results confirm our feeling that the soccer league scheduling case is well appreciated by our students. Students see closing the gap between theory and practice as the strongest asset of the case. More than 90% of the respondents either agree or strongly agree with question 2 (on the use of optimization modeling for solving real-life problems), question 5 (on the use of optimization to reconcile conflicting interests in real-life settings), and question 6 (on the surplus value of the case, as compared with typical textbook exercises, to learn how to successfully apply optimization

Table 2. Questions and Results of a Survey Filled Out by 112 Students After Having Followed the Interactive Lecture (Second Format)

Question	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
1. The soccer league scheduling case has increased my interest in optimization modeling.	22 (19.6%)	68 (60.7%)	19 (17.0%)	2 (1.8%)	1 (0.9%)
2. The soccer league scheduling case has increased my insight in the use of optimization modeling for solving real-life problems.	37 (33.0%)	60 (53.6%)	14 (12.5%)	1 (0.9%)	0 (0.0%)
3. The soccer league scheduling case has made me aware of the vast number of possible solutions an integer programming problem may have.	48 (42.9%)	52 (46.4%)	10 (8.9%)	1 (0.9%)	0 (0.0%)
4. The soccer league scheduling case has improved my optimization modeling skills.	3 (2.7%)	21 (18.8%)	65 (58.0%)	21 (18.8%)	2 (1.8%)
5. The soccer leagues scheduling case has shown me how optimization can be used to reconcile conflicting interests of various stakeholders in real-life problems.	36 (32.1%)	64 (57.1%)	12 (10.7%)	0 (0.0%)	0 (0.0%)
6. The soccer league scheduling case has a surplus value compared with typical textbook examples with respect to learning how to successfully apply optimization modeling in practice.	39 (34.8%)	53 (47.3%)	17 (15.2%)	3 (2.7%)	0 0.0%)
7. In the soccer league scheduling case, progressively adding complexities to the basic scheduling problem improved my understanding of the modeling process.	32 (28.6%)	57 (50.9%)	19 (17.0%)	3 (2.7%)	0 (0.0%)
8. The soccer league scheduling case has increased my interest in a job (academic, consultancy, business) within the field of operations research.	7 (6.3%)	33 (29.5%)	58 (51.8%)	13 (11.6%)	1 (0.9%)

Note. Questions 3 and 7 have only been answered by 111 students.

modeling in practice). The responses suggest that the case is less successful in actually improving the students' optimization-modeling skills (question 4). This is probably due to the second format (interactive lecture) for which the survey was taken. This format focuses on the real-life aspect and interactive discussion at the cost of hands-on modeling exercises (first format).

At the end of the survey, students could share some additional feedback in an open format. As an illustration, we list some students' quotes that we received:

- "Constructing the session with the underlying story, step by step discoveries, as well as the results of each trial made the approach very clear and attractive! The presentation has really enhanced my interest in the various applications of optimization problems and in research. Thank you very much for this incredible opportunity!"
- "I really found this case very interesting. The professor has made me more curious about real-life optimization problems. Thank you for that!"
- "Even though it was an evening lecture, it was really better than an average lecture. The professor got my attention although I am not really interested in football nor optimization."
- "Choosing another sport could have been more gender-neutral."

Also from this open feedback, we conclude that many students value the soccer league scheduling case. The last comment, however, raises the important issue of this case's potential to reach out to female students or minority groups. Let us start with saying that in Belgium, there is also a (nonprofessional) women's soccer league. The reason why this league was not studied in this case study is that many of the constraints that make the case challenging from a modeling (or computational) point of view do not apply to the women's league. Indeed, matches attract far fewer spectators (none of them hooligans); hence, the police do not forbid certain games at particular moments. Furthermore, apart from a handful of games since season 2020–21, the league is not broadcast on TV. Consequently, there are no wishes related to a favorable broadcasting scheme to be taken into account. In this respect, the women's highest soccer league is comparable to nonprofessional men's soccer leagues in Belgium, for which building a league schedule is far less complex. Also, given that there is much less money involved, fairness issues (e.g., avoiding carry-over and breaks) tend to receive much less attention. All in all, the only aspect that remains is developing a double-round-robin schedule that satisfies the basic constraints (every team meets each team exactly twice), which is fairly easy (any assignment of teams to numbers in the basic match schedule will do). The same observations hold for most parasports (e.g., the Belgium wheelchair basketball league) and nonprofessional sports. To decrease the observed "distance" with this case for students from minority groups, we added an extra slide to the PowerPoint. We advise using this slide to:

- Make students aware that the methodology of this case does not merely apply to men's leagues, but can be extended to all sports leagues that attract many spectators, of which the games are broadcast on TV, and that are organized using a round-robin format (e.g., women's soccer in the United States or wheelchair basketball in Germany).
- Initiate a discussion on which stakeholders or constraints are less important for sports leagues that attract fewer spectators and which aspects still remain in the problem.

6. Conclusion

This paper has presented our 10-year experience with two teaching formats of an integer-programming class around a real-life sports-scheduling problem. In both formats, students are actively involved, either through an interactive class discussion or through several hands-on modeling exercises, or a combination of both. From both the informal, positive student feedback we consistently receive for these classes and the formal results of a survey, we strongly believe that the combination of active involvement and the real-life application positively affects students' attention and interest in optimization modeling. Our hands-on exercises are incremental and do not require students to solve a large, complex scheduling problem from scratch. Nevertheless, by solving these (small) modeling exercises, meaningful results arise that could be presented to the league owners. Although there is room for discussion on issues that complicate the adoption of these models in practice, both teaching formats are mainly instructor-driven. Developing a case for students to solve (either individually or in group) with no or minimal support from the instructor would better reflect the real-life experience of optimization professionals. For instance, our class does not teach students how to cope with missing or incorrect data or translating fuzzy wishes into specific constraints and objectives. Developing such a comprehensive case around this real-life problem is an interesting avenue for future work.

References

Alarcón F, Durán G, Guajardo M (2014) Referee assignment in the Chilean football league using integer programming and patterns. *Internat. Trans. Oper. Res.* 21:415–438.

Anderson I (1999) Balancing carry-over effects in tournaments. Holroyd F, Quinn K, Rowley C, Webb B, eds. *Combinatorial Designs and Their Applications* (CRC Press, Boca Raton, FL), 1–16.

Beliën J, Goossens D, Van Reeth D, De Boeck L (2011) Using mixed-integer programming to win a cycling game. *INFORMS Trans. Ed.* 11(3):93–99.

Beliën J, Colpaert J, De Boeck L, Eyckmans J, Leirens W (2013) Teaching integer programming starting from an energy supply game. *INFORMS Trans. Ed.* 13(3):129–137.

Bickel JE (2004) Teaching decision making with baseball examples. *INFORMS Trans. Ed.* 5(1):2–9.

Birge JR (2004) Scheduling a professional sports league in Microsoft® Excel: Showing students the value of good modeling and solution techniques. *INFORMS Trans. Ed.* 5(1):56–66.

Bonomo F, Cardemil A, Durán G, Marenco J, Sabán D (2012) An application of the traveling tournament problem: The Argentine volleyball league. *Interfaces* 42(3):245–259.

Chlond MJ (2011) Puzzle-integer programming and league table puzzles. *INFORMS Trans. Ed.* 11(3):138–140.

Cocchi G, Galligari A, Nicolino FP, Piccialli V, Schoen F, Sciandrone M (2018) Scheduling the Italian national volleyball tournament. *Interfaces* 48(3):271–284.

Cochran JJ (2004) Introduction to the special issue: SpORts in the OR classroom. *INFORMS Trans. Ed.* 5(1):1.

Della Croce D, Oliveri D (2006) Scheduling the Italian Football League: An ILP-based approach. *Comput. Oper. Res.* 33:1963–1974.

Drake MJ (2019) Teaching OR/MS with cases: A review and new suggestions. *INFORMS Trans. Ed.* 19(2):57–66.

Durán G, Guajardo M, Sauré D (2017) Scheduling the South American qualifiers to the 2018 FIFA World Cup by integer programming. *Eur. J. Oper. Res.* 262(3):1109–1115.

Durán G, Guajardo M, Miranda J, Sauré D, Souyris S, Weintraub A, Wolf R (2007) Scheduling the Chilean soccer league by integer programming. *Interfaces* 37(6):539–552.

Goossens D (2018) Optimization in sports league scheduling: Experiences from the Belgian Pro League soccer. Parlier GH, Liberatore F, Demange M, eds. *Operations Research and Enterprise Systems* (Springer, Cham, Switzerland), 3–19.

Goossens D, Spieksma FCR (2009) Scheduling the Belgian soccer league. *Interfaces* 39(2):109–118.

Goossens D, Spieksma FCR (2011) Soccer schedules in Europe: An overview. *J. Scheduling* 15:641–651.

Gorman MF (2018a) An introduction to the special issue on field-based education. *INFORMS Trans. Ed.* 18(3):143–144.

Gorman MF (2018b) A survey of research in field-based education: A summary of process, best practices, and lessons learned. *INFORMS Trans. Ed.* 18(3):145–161.

Kendall G, Knust S, Ribeiro CC, Urrutia S (2010) Scheduling in sports: An annotated bibliography. *Comput. Oper. Res.* 37:1–19.

Kirkman T (1847) On a problem in combinatorics. *Cambridge Dublin Math. J.* 2:191–204.

Konrad R, Hall-Phillips A, Vila-Parrish AR (2018) Are our students prepared? The impact of capstone design pedagogical approaches on student skill development during industry-sponsored fieldwork. *INFORMS Trans. Ed.* 18(3):183–193.

Kostuk KJ, Willoughby KA (2012) A decision support system for scheduling the Canadian Football League. *Interfaces* 42(3):286–295.

Lambrechts E, Ficker AMC, Goossens DR, Spieksma FCR (2017) Round-robin tournaments generated by the circle method have maximum carry-over. *Math. Program.* 172(1-2):277–302.

Miyashiro R, Iwasaki H, Matsui T (2003) Characterizing feasible pattern sets with a minimum number of breaks. Burke E, De Causmaecker P, eds. *Practice and Theory of Automated Timetabling IV* (Springer, Berlin Heidelberg), 78–99.

Nemhauser GL, Trick MA (1998) Scheduling a major college basketball conference. *Oper. Res.* 46(1):1–8.

Pachamanova D (2006) Introducing integer modeling with Excel solver. *INFORMS Trans. Ed.* 7(1):88–98.

Rao BM, Beliën J (2014) Case article-production scheduling at Falcon Die Casting: A comprehensive example on the application of linear programming and its extensions. *INFORMS Trans. Ed.* 15(1):150–153.

Rasmussen RV, Trick MA (2008) Round robin scheduling—a survey. *Eur. J. Oper. Res.* 188(3):617–636.

Rasmussen RA, Weiss HJ (2007) Advanced lessons on the craft of optimization modeling based on modeling Sudoku in Excel. *INFORMS Trans. Ed.* 7(3):228–237.

Recalde D, Torres R, Vaca P (2013) Scheduling the professional Ecuadorian football league by integer programming. *Comput. Oper. Res.* 40:2478–2484.

Ribeiro CC (2012) Sports scheduling: Problems and applications. *Internat. Trans. Oper. Res.* 19:201–226.

Ribeiro CC, Urrutia S (2005) An application of integer programming to playoff elimination in football championships. *Internat. Trans. Oper. Res.* 12:375–386.

Russell KG (1980) Balancing carry-over effects in round robin tournaments. *Biometrika* 67:127–131.

Stevens SP, Palocsay SW (2017) Teaching use of binary variables in integer linear programs: Formulating logical conditions. *INFORMS Trans. Ed.* 18(1):28–36.

Toffolo TAM, Christiaens J, Spieksma FCR, Vanden Berghe G (2019) The sport teams grouping problem. *Ann. Oper. Res.* 275:223–243.

Trick MA (2004) Using sports scheduling to teach integer programming. *INFORMS Trans. Ed.* 5(1):10–17.

Van Bulck D, Goossens D, Schönberger J, Guajardo M (2020) RobinX: A three-field classification and unified data format for round-robin sports timetabling. *Eur. J. Oper. Res.* 280(2):568–580.

Weiss HJ, Rasmussen RA (2007) Lessons from modeling Sudoku in Excel. *INFORMS Trans. Ed.* 7(2):178–184.

Westphal S (2014) Scheduling the German basketball league. *Interfaces* 44(5):498–508.

Willoughby KA (2004) The science of sports: Combining quantitative analysis and sports applications in an undergraduate course. *INFORMS Trans. Ed.* 5(1):88–99.