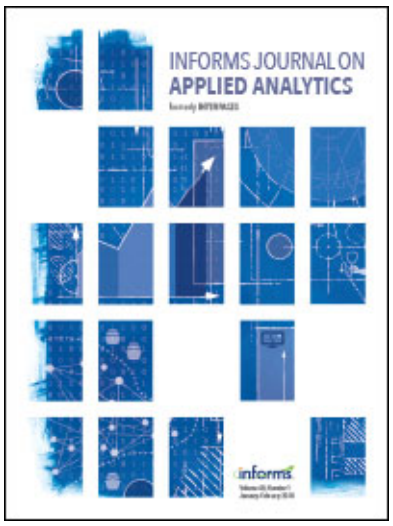


INFORMS Journal on Applied Analytics

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

Scheduling the Valley Baseball League

Elizabeth L. Bouzarth, Benjamin C. Grannan, John M. Harris, Kevin R. Hutson

To cite this article:
Elizabeth L. Bouzarth, Benjamin C. Grannan, John M. Harris, Kevin R. Hutson (2022) Scheduling the Valley Baseball League. INFORMS Journal on Applied Analytics 52(2):189-197. https://doi.org/10.1287/inte.2021.1076

http://pubsonline.informs.org/journal/inte

INFORMS JOURNAL ON APPLIED ANALYTICS
Vol. 52, No. 2, March–April 2022, pp. 189–197
ISSN 2644-0865 (print), ISSN 2644-0873 (online)

Scheduling the Valley Baseball League

Elizabeth L. Bouzarth,[a] Benjamin C. Grannan,[b] John M. Harris,[a] Kevin R. Hutson[a]

[a] Department of Mathematics, Furman University, Greenville, South Carolina 29613; [b] Department of Business and Accounting, Furman University, Greenville, South Carolina 29613

Contact: liz.bouzarth@furman.edu (ELB); ben.grannan@furman.edu, https://orcid.org/0000-0002-1154-4703 (BCG); john.harris@furman.edu (JMH); kevin.hutson@furman.edu (KRH)

Received: June 2, 2020
Revised: November 18, 2020
Accepted: December 4, 2020
Published Online in Articles in Advance: November 8, 2021

https://doi.org/10.1287/inte.2021.1076

Abstract. The Valley Baseball League is a collegiate summer baseball league with 11 teams playing 42 games between June and August. Producing a schedule has proved challenging for league management, because of the odd number of teams as well as a large number of league mandates and team-specific requests. As an amateur league with limited financial resources, there is pressure to best leverage the league schedule to maximize revenue for each team. A schedule generated by mathematical programming was enthusiastically adopted for the 2020 season and significantly outperformed previous manually produced schedules.

History: This paper was refereed.

Keywords: sports scheduling • binary linear programming • prescriptive analytics

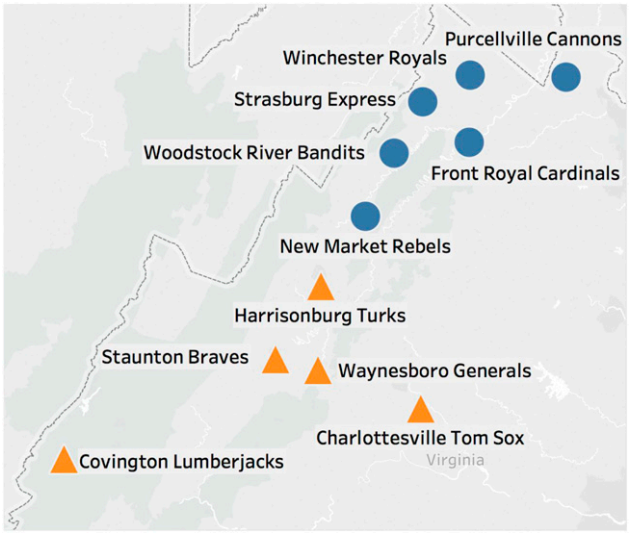
The Valley Baseball League (VBL) is a collegiate summer baseball league that was founded in 1897 and currently includes 11 teams located throughout the Shenandoah Valley in Virginia. Players in sanctioned collegiate summer baseball leagues have attended at least one year of college and must have some eligibility left in their collegiate career. The *Collegiate Baseball Newspaper* (2018) has identified over 40 distinct collegiate summer baseball leagues representing all regions of the United States and fielding players from all levels of competitive collegiate baseball programs. One purpose of collegiate summer baseball leagues is to offer development opportunities for current college baseball players who aspire to play at the next (professional) level. This is the motivation behind the VBL's motto—"Gateway to the majors"—and the league's claim to have "produced well over 1,000 professional baseball players" (Valley League Baseball 2018).

The VBL is divided into a North Division and a South Division. The North Division contains the Strasburg Express, Woodstock River Bandits, New Market Rebels, Winchester Royals, Purcellville Cannons, and Front Royal Cardinals. The South Division includes the Waynesboro Generals, Charlottesville Tom Sox, Covington Lumberjacks, Staunton Braves, and Harrisonburg Turks (see Figure 1). Teams in the South Division, having an odd number of teams, must play each other a total of six times, whereas teams in the North Division play each other between five and six times a season. Every team plays an opponent from the other division three times during the season.

Many factors contribute to creating a schedule for a collegiate baseball league that ensures different measures of fairness for the league. In 2017, the VBL lost a franchise, reducing its size from 12 to 11 teams. With an odd number of teams, each day of the season will contain a day off for one of the teams. The VBL faced difficulties making a fair schedule as it tried to equally distribute coveted days off while incorporating other league-wide and team-specific scheduling requests. For example, many league teams have lists of dates in which their home facility is unavailable or when they strongly prefer to play a home game. Furthermore, there are economic fairness and revenue-generating considerations that go into the schedule as well. By strategically assigning high-profile, impactful games to days of the week with a capacity for higher attendance, the league schedule is the greatest opportunity for collegiate summer baseball leagues to increase revenue. Better games on better days proliferate ticket and concession sales as well as advertising partnerships.

Major League Baseball (MLB) supplies a yearly grant to sanctioned collegiate summer baseball leagues. Despite the association with MLB and the previous adoption of the computer-driven approach to MLB league scheduling, there was no awareness by the VBL front office of the opportunity to improve scheduling via advanced analytics. Before 2017, a manual scheduling process took place over several weeks, consuming many employee hours of work to create a schedule that the league and teams considered acceptable. It was the norm for team representatives to vote down several iterations of a manually produced schedule because of any number of complaints stemming from fairness, quality, ignored team

Figure 1. (Color online) The 11 Teams of the Valley Baseball League



Notes. The VBL includes 11 teams located throughout the Shenandoah Valley in Virginia. Teams are separated into North and South Divisions (shown in circles and triangles, respectively).

requests, unacceptable travel, and so on. Starting in 2017, the authors began to automate the schedule-making process using an integer-programming approach.

The rest of this paper describes how the Valley Baseball League now creates a better schedule with more constraints and drastically reduced creation time, and it is organized as follows. The previous work in the literature on sports scheduling is presented in the Literature Review section, including examples from baseball and other sports. The Method section describes the mathematical programming approach to schedule games for the Valley Baseball League. Lastly, we summarize the results and benefits of this work in the final section.

Literature Review

Many applications of prescriptive analytics have been applied to sports scheduling. Traditionally, across many sports, nonanalytical scheduling usually was completed by hand and often relied on reducing the complexity of the scheduling process by creating round-robin or double round-robin schedules (Rasmussen and Trick 2008, Kendall et al. 2010). Creating double round-robin schedules could be simplified by using the same schedule back-to-back but switching the home and away designation for the second half of the season (Goossens and Spieksma 2009, 2012). As technology advanced, schedulers employed more computer-based analytics approaches to solve sports scheduling problems. Most of these approaches use integer programming, although other approaches, such as constraint programming (Henz 2001) and simulated annealing (Wright 2006), have been proposed as well. Kendall et al. (2010) gives an overview of many of these analytical scheduling techniques.

Most of the sports scheduling literature is differentiated by sport and by the specifics of the constraint set involved. Typical constraints can be divided into venue considerations, broadcast considerations, team considerations, and league fairness considerations. Venue considerations include geographic and stadium limitations (Nemhauser and Trick 1998, Bartsch et al. 2006, Della Croce and Oliveri 2006, Duran et al. 2007) or venue use constraints (Kostuk and Willoughby 2012). These typically involve stadiums or arenas in which different teams play home games, or multiuse venues, leaving schedulers to work around concerts and other events in order to schedule home games. Adding constraints to ensure the availability of broadcast units (Duran et al. 2007) to local venues on the date of the game has also been proposed. Although venue and broadcast constraints usually eliminate certain days from consideration in making a schedule, created schedules often impact teams differently. Teams complain that schedules are unfair in a variety of ways, so a number of scheduling constraints have been developed to not only help ensure fairness but also to help maximize revenue and fan interest. These team considerations include limiting consecutive-day games (Duran et al. 2007), limiting consecutive games against top-ranked teams (Froncek 2001), and strategically scheduling rivalry games (Nemhauser and Trick 1998). Other league fairness considerations include ensuring equal numbers of home and away games, not scheduling too many short-rest games back-to-back (Kurt et al. 2015), and limiting abundant travel distances (Kendall 2008, Grabau 2012).

Baseball and softball scheduling has a rich history in the literature (Russell and Leung 1994, Grabau 2012, Saur et al. 2012, Trick et al. 2012). Baseball scheduling, especially, has all of the venue, team, and league constraints combined with a lengthy season, which contributes to a massive number of variables in an integer program. Large-size scheduling, such as that found in Major League Baseball, often requires the scheduler to simplify the scheduling process first and then solve smaller problems. Some techniques have been proposed to handle these large-scale problems, such as scheduling breaks first (Rasmussen 2008), scheduling round-robin tournaments within smaller subsets or phases of a season (Trick 2003), or solving the problem with smaller numbers of constraints and making postschedule adjustments (Kendall 2008). On a smaller scale, local baseball and softball leagues may have fewer games to schedule but deal with issues that the major leagues do not, such as limited budgets that impact travel, no dedicated home fields, and coaches training multiple teams

(Grabau 2012). Specific leagues have geographic and weather-related constraints to consider, which include limiting early season games due to lengthy winter weather seasons (Saur et al. 2012) or front-loading schedules so as to leave open dates later in the season to account for rainouts throughout the season. The Valley Baseball League has many of these traditional baseball/softball scheduling issues, as well as some that are unique to the league, as we outline in the next section.

Method

To help the Valley Baseball League produce the best possible schedule, a binary linear programming model is developed. The binary variables of the model represent each potential game by the home team and the away team, as well as the day of the season in which the game occurs. The dimensions of the variables allow for the construction of specific constraints that reflect schedule requirements and restrictions. For example, specifying the home team in each game allows for enforcement of a balance of the frequency of home games as well as control over distribution of home games in matchups. Likewise, recording the day of a game allows for control over consecutive road games and off-days.

Objective

The success of collegiate summer baseball leagues, like most amateur organizations, hinges on thin profit margins amid limited financial resources. The Valley Baseball League has identified days during the season, called *prime days*, on which attendance is generally high, especially when a home team is playing a rival from the same division. These days are generally Friday and Saturday nights during the season and the last three days of the season with postseason implications. Therefore, the objective in this model is to maximize the number of intradivisional games played on prime days. This objective gives each team a fair shot at maximizing revenue.

League Constraints

The following constraints shape the basic structure of the yearly schedule and are mandated by the league office. These constraints establish schedule fairness for each team.

- The 2020 season runs from May 29 to July 23, and each team plays 42 games during the season, consisting of 21 home and 21 away games.
- Each team plays only one game in a given day and must have at least seven home games that occur on Fridays or Saturdays (except Charlottesville and Waynesboro, which have fewer because they play a Fourth of July game).
- There must be at least five games on each defined prime date, as well as the dates requiring a full schedule, such as opening day. The league office also mandates a list of days in which no games are to be scheduled, including dedicated rain makeup days (teams work together to reschedule postponed games) and the All-Star break. If a team is off the day before a mandated off-day, then that team must be scheduled to play a game the day immediately after the league-wide off-day to interrupt the number of consecutive idle days.

Divisional Constraints

Having an odd number of teams in a league/division adds some complication to the model. The following list of constraints captures the complete inter- and intradivisional matchup frequency in the unbalanced Valley Baseball League with two divisions. Groups of rivalry teams are used to identify how to distribute extra games.

- Teams in the South Division, the division with an odd number of teams, must play each other a total of six times, with an equal number of three home games for each team.
- Every team plays an opponent from the other division three times during the season, with either one or two home games for each team.
- Teams in the North Division will play six games against members of their own group and five games against members of the other group of northern teams. The number of home games must be balanced within the North Division; therefore, each group 1 and group 2 team has exactly three home games against teams in their own group, and the home splits between groups are either two or three games. Group 1 consists of Purcellville, Winchester, and Front Royal, whereas group 2 contains New Market, Strasburg, and Woodstock.

Team-Specific Constraints

Each team in the VBL can submit requests to the league in fulfilling the schedule. For the most recent season, the constraints include the following:

- Each team has a list of dates in which their home field is unavailable to host a collegiate baseball game.
- Charlottesville and Waynesboro play a Fourth of July game in Charlottesville.
- There are specific weeks, one in June and one in July, when Winchester does not want to play at Charlottesville or Covington.

Pattern Constraints

The remaining list of constraints enforces the spacing and frequency of games played during the season. Examples of restrictions include the time between repeat

matchups as well as spacing of off-days throughout the season.

- No team can play more than 10 consecutive days without an off-day.
- No team can have more than one consecutive off-day during the season. The exception to this rule is that at least one team must be off the day before and after a mandated league-wide off-day due to the odd number of teams.
- No team can have more than two consecutive home games or away games during the season.
- Each team must play between two and four home games and road games during each week (seven days) of the season.
- Each team must play at least six games during weeks 1 and 2 of the season; at least four games during weeks 3 and 4; at least three games in weeks 5 and 6 (there are more league-mandated off-days during these weeks); and at least four games in weeks 7 and 8.
- Each pair of teams cannot play each other more than once in any three-day period during the season.
- Teams will not be scheduled to play opponents on back-to-back road trips that are ranked within their top-three furthest distances.
- Each team must have between three and four home games on Sundays during the season.

Results, Benefits, and Conclusion

In the fall of 2019, the model described in the previous section was solved using the Gurobi software on a laptop computer equipped with an Intel Core i7 processor, 1.8 GHz, and 8 GB of RAM. With a run time under three hours, a schedule was generated that satisfied all of the aforementioned constraints with a 1% optimality gap. The model can be run to a 0% optimality gap in a few minutes without the constraints enforcing an equitable number of weekend home games for each team. Data parsing was done in Python to translate the schedule into a Microsoft Excel spreadsheet, as requested by the league office. Figures 2 and 3 show the first month of the 2020 season, with prime days, intradivision games, and off-days indicated for North and South teams, respectively. A generated 2020 Valley Baseball League schedule was unanimously approved by league stakeholders in November of 2019; however, the league was forced to cancel the 2020 season due to the COVID-19 pandemic.

Although some of the scheduling constraints introduced in the previous sections were addressed in seasons prior to 2018, the iterative effort used to create the schedules did not allow for all constraints to be satisfied. This meant that the schedules in these seasons presented challenges to some teams but not

Figure 2. (Color online) North Division Teams

2020 VBL schedule

	FrontRoyalCardinals	NewMarketRebels	PurcellvilleCannons	StrasburgExpress	WinchesterRoyals	WoodstockRvrBndts
29-May	host StrasburgExpress	off day	at HarrisonburgTurks	at FrontRoyalCardinals	host WoodstockRvrBndts	at WinchesterRoyals
30-May	at WoodstockRvrBndts	at StrasburgExpress	at WinchesterRoyals	host NewMarketRebels	host PurcellvilleCannons	host FrontRoyalCardinals
31-May	off day	host WaynesboroGenerals	host CharlottesvilleTomSox	at WoodstockRvrBndts	at StauntonBraves	host StrasburgExpress
1-Jun	host NewMarketRebels	at FrontRoyalCardinals	off day	off day	at CovingtonLumberjacks	at HarrisonburgTurks
2-Jun	at WinchesterRoyals	host CharlottesvilleTomSox	host WoodstockRvrBndts	at StauntonBraves	host FrontRoyalCardinals	at PurcellvilleCannons
3-Jun	host WaynesboroGenerals	host StauntonBraves	at WinchesterRoyals	host CharlottesvilleTomSox	host PurcellvilleCannons	host CovingtonLumberjacks
4-Jun	at CharlottesvilleTomSox	at StrasburgExpress	host CovingtonLumberjacks	host NewMarketRebels	off day	off day
5-Jun	host StrasburgExpress	host WinchesterRoyals	at WoodstockRvrBndts	at FrontRoyalCardinals	at NewMarketRebels	host PurcellvilleCannons
6-Jun	at HarrisonburgTurks	at PurcellvilleCannons	host NewMarketRebels	off day	host WoodstockRvrBndts	at WinchesterRoyals
7-Jun	host CharlottesvilleTomSox	off day	at WaynesboroGenerals	at HarrisonburgTurks	at CovingtonLumberjacks	host StauntonBraves
8-Jun	at StrasburgExpress	at StauntonBraves	host WoodstockRvrBndts	host FrontRoyalCardinals	host WaynesboroGenerals	at PurcellvilleCannons
9-Jun	at CovingtonLumberjacks	at WinchesterRoyals	off day	at WoodstockRvrBndts	host NewMarketRebels	host StrasburgExpress
10-Jun	off day	host PurcellvilleCannons	at NewMarketRebels	host StauntonBraves	at CharlottesvilleTomSox	at WaynesboroGenerals
11-Jun	host WaynesboroGenerals	host CovingtonLumberjacks	host WoodstockRvrBndts	at CharlottesvilleTomSox	host HarrisonburgTurks	at PurcellvilleCannons
12-Jun	at WinchesterRoyals	at WoodstockRvrBndts	host StrasburgExpress	at PurcellvilleCannons	host FrontRoyalCardinals	host NewMarketRebels
13-Jun	host PurcellvilleCannons	at StrasburgExpress	at FrontRoyalCardinals	host NewMarketRebels	at WoodstockRvrBndts	host WinchesterRoyals
14-Jun	host HarrisonburgTurks	host CovingtonLumberjacks	at CharlottesvilleTomSox	host StauntonBraves	off day	at WaynesboroGenerals
15-Jun	off day	off day	off day	off day	off day	off day
16-Jun	at StrasburgExpress	host WoodstockRvrBndts	at WaynesboroGenerals	host FrontRoyalCardinals	at HarrisonburgTurks	at NewMarketRebels
17-Jun	at StauntonBraves	off day	host HarrisonburgTurks	at CovingtonLumberjacks	at WoodstockRvrBndts	host WinchesterRoyals
18-Jun	off day	at HarrisonburgTurks	off day	off day	host CovingtonLumberjacks	host CharlottesvilleTomSox
19-Jun	at PurcellvilleCannons	at CharlottesvilleTomSox	host FrontRoyalCardinals	host WinchesterRoyals	at StrasburgExpress	off day
20-Jun	host WoodstockRvrBndts	host StrasburgExpress	at WinchesterRoyals	at NewMarketRebels	host PurcellvilleCannons	at FrontRoyalCardinals
21-Jun	at NewMarketRebels	host FrontRoyalCardinals	host WaynesboroGenerals	host CharlottesvilleTomSox	host StauntonBraves	host CovingtonLumberjacks
22-Jun	off day	off day	off day	off day	off day	off day
23-Jun	host PurcellvilleCannons	at StauntonBraves	at FrontRoyalCardinals	at WoodstockRvrBndts	off day	host StrasburgExpress
24-Jun	host StauntonBraves	at PurcellvilleCannons	host NewMarketRebels	host WaynesboroGenerals	at HarrisonburgTurks	host CharlottesvilleTomSox
25-Jun	off day	off day	off day	off day	off day	at CovingtonLumberjacks

Notes. The first month of the accepted 2020 Valley Baseball League schedule. Prime days, intradivision games, and off-days are highlighted. This season was cancelled due to the COVID-19 pandemic.

Figure 3. (Color online) South Division Teams

2020 VBL schedule

	CharlottesvilleTomSox	CovingtonLumberjacks	HarrisonburgTurks	StauntonBraves	WaynesboroGenerals
29-May	at CovingtonLumberjacks	host CharlottesvilleTomSox	host PurcellvilleCannons	host WaynesboroGenerals	at StauntonBraves
30-May	off day	host StauntonBraves	at WaynesboroGenerals	at CovingtonLumberjacks	host HarrisonburgTurks
31-May	at PurcellvilleCannons	at HarrisonburgTurks	host CovingtonLumberjacks	host WinchesterRoyals	at NewMarketRebels
1-Jun	host WaynesboroGenerals	host WinchesterRoyals	host WoodstockRvrBndts	off day	at CharlottesvilleTomSox
2-Jun	at NewMarketRebels	off day	at WaynesboroGenerals	host StrasburgExpress	host HarrisonburgTurks
3-Jun	at StrasburgExpress	at WoodstockRvrBndts	off day	at NewMarketRebels	at FrontRoyalCardinals
4-Jun	host FrontRoyalCardinals	at PurcellvilleCannons	at StauntonBraves	host HarrisonburgTurks	off day
5-Jun	at StauntonBraves	host WaynesboroGenerals	off day	host CharlottesvilleTomSox	at CovingtonLumberjacks
6-Jun	host CovingtonLumberjacks	at CharlottesvilleTomSox	host FrontRoyalCardinals	at WaynesboroGenerals	host StauntonBraves
7-Jun	at FrontRoyalCardinals	host WinchesterRoyals	host StrasburgExpress	at WoodstockRvrBndts	host PurcellvilleCannons
8-Jun	host HarrisonburgTurks	off day	at CharlottesvilleTomSox	host NewMarketRebels	at WinchesterRoyals
9-Jun	off day	host FrontRoyalCardinals	host StauntonBraves	at HarrisonburgTurks	off day
10-Jun	host WinchesterRoyals	at HarrisonburgTurks	host CovingtonLumberjacks	at StrasburgExpress	host WoodstockRvrBndts
11-Jun	host StrasburgExpress	at NewMarketRebels	at WinchesterRoyals	off day	at FrontRoyalCardinals
12-Jun	at CovingtonLumberjacks	host CharlottesvilleTomSox	off day	host WaynesboroGenerals	at StauntonBraves
13-Jun	host StauntonBraves	host HarrisonburgTurks	at CovingtonLumberjacks	at CharlottesvilleTomSox	off day
14-Jun	host PurcellvilleCannons	at NewMarketRebels	at FrontRoyalCardinals	at StrasburgExpress	host WoodstockRvrBndts
15-Jun	off day	off day	off day	off day	off day
16-Jun	host CovingtonLumberjacks	at CharlottesvilleTomSox	host WinchesterRoyals	off day	host PurcellvilleCannons
17-Jun	at WaynesboroGenerals	host StrasburgExpress	at PurcellvilleCannons	host FrontRoyalCardinals	host CharlottesvilleTomSox
18-Jun	at WoodstockRvrBndts	at WinchesterRoyals	host NewMarketRebels	host WaynesboroGenerals	at StauntonBraves
19-Jun	host NewMarketRebels	at WaynesboroGenerals	host StauntonBraves	at HarrisonburgTurks	host CovingtonLumberjacks
20-Jun	at StauntonBraves	host HarrisonburgTurks	at CovingtonLumberjacks	host CharlottesvilleTomSox	off day
21-Jun	at StrasburgExpress	at WoodstockRvrBndts	off day	at WinchesterRoyals	at PurcellvilleCannons
22-Jun	off day	off day	off day	off day	off day
23-Jun	host HarrisonburgTurks	at WaynesboroGenerals	at CharlottesvilleTomSox	host NewMarketRebels	host CovingtonLumberjacks
24-Jun	at WoodstockRvrBndts	off day	host WinchesterRoyals	at FrontRoyalCardinals	at StrasburgExpress
25-Jun	host WaynesboroGenerals	host WoodstockRvrBndts	off day	off day	at CharlottesvilleTomSox

Notes. The first month of the accepted 2020 Valley Baseball League schedule. Prime days, intradivision games, and off-days are highlighted. This season was cancelled due to the COVID-19 pandemic.

others. One such challenge noted during those seasons was the existence of 10-day stretches with no rest days. Between 2013 and 2017, there were an average of 20 such instances over all teams per season, with the 2015 season having 42 instances of long stretches of games without breaks.

Another challenge identified in seasons prior to 2018 was the frequency of back-to-back road games where the two away cities were at long distances from the traveling teams' cities. Because teams in this league do not stay overnight on road trips, such back-to-back occurrences meant two consecutive days and nights of longer-than-average travel distances. Between 2013 and 2017, the league averaged 10 instances per season of this undesired situation. One major benefit of the model described in this paper is that these issues can be addressed. Since 2018, the first season for which this model was used in schedule creation, all 10-game stretches and consecutive long-distance road games have been eliminated.

As is indicated by the objective function of the model, the primary goal was to increase the number of intradivision matchups that take place on the league's prime days. Table 1 gives the percentage of such games that took place on those days for each season from 2013 to 2020. Note that the 2014 season is not included, because, for that season, the league did not separate teams into divisions.

As the table indicates, the percentages have risen above 95% in the seasons for which the schedule has been produced via math programming. Increases such as this have significant economic impact for individual teams and the league. The Executive Committee of the Valley Baseball League estimates that fans and revenues have increased by 40%–45% on the prime dates (Leonard 2020).

A sensitivity analysis highlights the effect of constraint parameter values on solution quality. Two important schedule features were analyzed: the number

Table 1. Percentage of Prime Day Games Scheduled to Hold Intradivision Games by Year

Year	% Prime day games containing intradivisional matchups
2013	72.97
2015	69.89
2016	67.73
2017	87.50
2018	95.51
2019	95.35
2020	96.70

of active days in a row before an off-day and when two teams play each other too frequently. During model development, league decision makers identified the acceptable thresholds for these features. Interestingly, changes to the right-hand-side values in these nonbinding constraints had no effect on the objective function until eventually causing model infeasibility. For example, the maximum number of days played in a row without a break can be reduced from 10 to 6 for each team without consequence, whereas an extreme value of only allowing five games followed by a mandated break proved impossible given the other constraints. Furthermore, teams can avoid playing each other again for up to eight days under the same schedule quality; however, pushing the replay restriction to 9 and 10 days results in a 12.6% loss in objective function value and model infeasibility, respectively.

An alternative objective function explores the maximum number of home games possible on Friday or Saturday for each team. After removing the constraint to enforce an equitable minimum of seven weekend home games for each team (with the exception of five for Charlottesville due to a special holiday game request), each team can be isolated in the objective function and receive a maximum of 13 weekend home games. The distribution of weekend home games is heavily lopsided without league-wide minimums, and some teams have as little as two weekend home games during the entire season. However, returning a minimum tally of seven weekend home games for every team results in a maximum value of seven games for any one team in the objective function.

One additional benefit of scheduling with optimization models is flexibility. When schedules are created by hand or with other methods that are less efficient, it is difficult to adjust constraints as they arise. Changes, even small ones, can lead schedule makers back to square one. With this model, a new or adjusted constraint can be addressed by adding or editing a line or two of code. The flexibility also allows for other potential adjustments, such as the addition of a team, the restructuring of divisions, changes in season length, and so on. Additionally, one could imagine the league using the model to create hypothetical schedules as it investigates the potential economic impact of various decisions or modifications.

In conclusion, math programming has been enthusiastically adopted as a scheduling tool by the league office of the Valley Baseball League. Since turning to advanced analytics to generate schedules, revenues are up and complaints about the schedule are down. Team owners are confident going into each new season knowing the path to the playoffs is equitable across the league. Collegiate baseball players benefit from consistent rest days and the reduction of long consecutive road trips by team bus that are unnecessary. Despite being smaller in revenue and impact than Major League Baseball, the Valley Baseball League has caught up to the professional ranks and also benefits from a computer-driven approach to scheduling.

Acknowledgments

The authors thank James McCumber and John Leonard for their willingness to review numerous iterations of schedules and to serve as liaisons to the league office. Additional acknowledgment is due to undergraduate research students Taylor Edens, John Michael Boswell, Sophia Callanan, Joseph Lastovic, and Caleb Roberson for their important contributions to the 2018–2020 official Valley Baseball League schedules.

Appendix. Binary Linear Programming Formulation

The following decision variables, sets, and parameters are required to formulate and solve the binary linear programming model for the 2020 Valley Baseball League season.

Decision Variables

The decision variables are defined to be $x_{ijk} = 1$ if team i is home to team j on the kth day of the season and 0 otherwise. Controlling which team is home and away on specific days allows for both team requests and league mandates.

Sets

T: all teams in the league
S: teams in the South Division
N: teams in the North Division
$G1$: North Division teams that play more frequently
$G2$: North Division teams that play more frequently, where $G1 \cap G2 = \emptyset$
D: days during the season
P: prime days during the season
W: weekend (Friday and Saturday) days during the season
Su: Sundays during the season
F: days during the season that require a full slate of games to be played
M: strategic idle days across the entire league
A: days of the All-Star break
C: days when Winchester does not want to play on the road against Covington or Charlottesville
K_d: days of the dth week of the season, for $d = 1, \ldots, \lceil |D|/7 \rceil$

Parameters

g_d: minimum number of games each team must play during the d th week of the season, for $d = 1, \ldots, \lceil |D|/7 \rceil$
δ_{ij}: equals 1 if team i and team j are in the same division and 0 otherwise
γ_{ik}: equals 1 if team i cannot play a home game on day k and 0 otherwise
d_{ij}: rank order of travel distance in miles for each team i to play at opponent j, where rank 1 indicates the farthest opponent of team i
k_4: day of July 4th

$$\max \sum_{i \in T} \sum_{j \in T} \sum_{k \in P} x_{ijk} \delta_{ij} \tag{A.1}$$

$$\text{subject to} \quad \sum_{j\in T}\sum_{k\in D} x_{ijk} + x_{jik} = 42 \quad \forall i \in T \tag{A.2}$$

$$\sum_{j\in T}\sum_{k\in D} x_{ijk} = 21 \quad \forall i \in T \tag{A.3}$$

$$\sum_{k\in D} x_{iik} = 0 \quad \forall i \in T \tag{A.4}$$

$$\sum_{j\in T} x_{ijk} + x_{jik} \le 1 \quad \forall i \in T,\ k \in D \tag{A.5}$$

$$\sum_{j\in T}\sum_{k\in W} x_{ijk} \ge 7 \quad \forall i \in T \tag{A.6}$$

$$\sum_{i\in T}\sum_{j\in T} x_{ijk} \ge 5 \quad \forall k \in F \tag{A.7}$$

$$\sum_{i\in T}\sum_{j\in T}\sum_{k\in M\cup A} x_{ijk} = 0 \tag{A.8}$$

$$\sum_{j\in T} x_{ijk} = 0 \quad \forall i \in T,\ k \in D \text{ such that } \gamma_{ik} = 1 \tag{A.9}$$

$$\sum_{k\in D} x_{ijk} + x_{jik} = 6 \quad \forall i,\ j \in S \text{ such that } j \ne i \tag{A.10}$$

$$\sum_{k\in D} x_{ijk} = 3 \quad \forall i,\ j \in S \text{ such that } j \ne i \tag{A.11}$$

$$\sum_{k\in D} x_{ijk} + x_{jik} = 6 \quad \forall i,\ j \in G1 \text{ such that } j \ne i \tag{A.12}$$

$$\sum_{k\in D} x_{ijk} + x_{jik} = 6 \quad \forall i,\ j \in G2 \text{ such that } j \ne i \tag{A.13}$$

$$\sum_{k\in D} x_{ijk} + x_{jik} = 5 \quad \forall i \in G2,\ j \in G1 \tag{A.14}$$

$$\sum_{k\in D} x_{ijk} = 3 \quad \forall i,\ j \in G1 \text{ such that } j \ne i \tag{A.15}$$

$$\sum_{k\in D} x_{ijk} = 3 \quad \forall i,\ j \in G2 \text{ such that } j \ne i \tag{A.16}$$

$$2 \le \sum_{k\in D} x_{ijk} \le 3 \quad \forall i \in G2,\ j \in G1 \tag{A.17}$$

$$2 \le \sum_{k\in D} x_{ijk} \le 3 \quad \forall i \in G1,\ j \in G2 \tag{A.18}$$

$$\sum_{k\in D} x_{ijk} + x_{jik} = 3 \quad \forall i \in S,\ j \in N \tag{A.19}$$

$$1 \le \sum_{k\in D} x_{ijk} \le 2 \quad \forall i \in S,\ j \in N \tag{A.20}$$

$$\sum_{j\in T}\sum_{a\in\{0,1,\ldots,10\}} x_{ij(z+a)} + x_{ji(z+a)} \le 10 \quad \forall i \in T,\ z \in 1,\ldots,|D|-10 \tag{A.21}$$

$$\sum_{j\in T}\sum_{a\in\{0,1\}} x_{ij(z+a)} + x_{ji(z+a)} \ge 1 \quad \forall i \in T,\ z \in Z \tag{A.22}$$

$$\sum_{j\in T}\sum_{a\in\{0,1,2\}} x_{ij(z+a)} \le 2 \quad \forall i \in T,\ z \in 1,\ldots,|D|-2 \tag{A.23}$$

$$\sum_{i\in T}\sum_{a\in\{0,1,2\}} x_{ij(z+a)} \le 2 \quad \forall j \in T,\ z \in 1,\ldots,|D|-2 \tag{A.24}$$

$$\sum_{j\in T} x_{ij(z-1)} + x_{ji(z-1)} + x_{ij(z+1)} + x_{ji(z+1)} \ge 1 \quad \forall i \in T,\ z \in M \tag{A.25}$$

$$\sum_{j\in T} x_{ij(z-1)} + x_{ji(z-1)} + x_{ij(z+2)} + x_{ji(z+2)} \ge 1 \quad \forall i \in T,\ z = \min(A) \tag{A.26}$$

$$2 \le \sum_{j\in T}\sum_{k\in K_d} x_{ijk} \le 4 \quad \forall i \in T,\ d = 1,\ldots,\lceil |D|/7 \rceil \tag{A.27}$$

$$2 \le \sum_{j\in T}\sum_{k\in K_d} x_{jik} \le 4 \quad \forall i \in T,\ d = 1,\ldots,\lceil |D|/7 \rceil \tag{A.28}$$

$$\sum_{j\in T}\sum_{k\in K_d} x_{jik} \ge g_d \quad \forall i \in T,\ d = 1,\ldots,\lceil |D|/7 \rceil \tag{A.29}$$

$$3 \le \sum_{j\in T}\sum_{k\in Su} x_{ijk} \le 4 \quad \forall i \in T \tag{A.30}$$

$$\sum_{a\in\{0,1,2\}} x_{ij(z+a)} + x_{ji(z+a)} \le 1 \quad \forall i,\ j \in T,\ z \in 1,\ldots,|D|-2 \tag{A.31}$$

$$\sum_{a\in\{0,1\}}\sum_{j\in T:d_{ij}\le 3} x_{ji(z+a)} \le 1 \quad \forall i \in T,\ z \in 1,\ldots,|D|-1 \tag{A.32}$$

$$\sum_{k\in C}\sum_{j\in\{\text{Covington},\text{Charlottesville}\}} x_{j(\text{Winchester})k} = 0 \tag{A.33}$$

$$x_{(\text{Charlottesville})(\text{Waynesboro})k_4} = 1 \tag{A.34}$$

$$x_{ijk} \in \{0,1\} \quad \forall i,\ j \in T,\ k \in D \tag{A.35}$$

The objective function (A.1) maximizes the number of intradivisional games that are scheduled on prime days throughout the season to boost fan attendance and revenues. Constraint (A.2) requires 42 total games for each team, and each team has exactly 21 home and away games during the season because of constraint (A.3). Constraints (A.4) and (A.5) eliminate self-games and playing more than one game in a calendar day, respectively, for each team. Constraint (A.6) gives each team in the league at least seven home games on the Friday and Saturday dates to boost attendance. Constraint (A.7) ensures a full slate of five games (the maximum in a league with 11 teams) on specific days of the season, such as opening day or before a league-wide holiday. Constraint (A.8) schedules no games on strategic days during the season to build in makeup contests and no games during the All-Star break. Team-specific requests to not play a home game due to unavailable facilities are incorporated in constraint (A.9).

Next, the inter- and intradivisional matchups are created for the Valley Baseball League, accounting for the two divisions and an odd number of teams. For the five teams in the South Division, constraints (A.10) and (A.11) establish six total games against each divisional opponent that are split evenly across home locations. The six teams in the North Division will play six games against members of their own group (constraints (A.12) and (A.13)) and five games against members of the other group of North teams (constraint (A.14)). The number of home games must be balanced within the North Division; therefore, constraints (A.15) and (A.16) give each group 1 and group 2 team exactly three home games against teams in their own group, whereas constraints (A.17) and (A.18) balance the home splits between groups to be either two or three games. Lastly, interdivisional matchups are structured to be three total games (constraint (A.19)) with a fair tally of one or two home games for each South team (constraint (A.20)). Constraints (A.19) and (A.20) together imply one or two home games for each North team.

The spacing and pattern of scheduled games is presented next. In constraint (A.21), teams will play no more than 10 total games in a row. Teams cannot have more than one consecutive off-day in constraint (A.22) with the exception of mandatory league-wide off-days. No team will play more than two consecutive home or road games in constraints

(A.23) and (A.24), respectively. Additionally, no team is idle both before and after a league-wide off-day; constraint (A.25) addresses both sides of the built-in weather make-up days, whereas constraint (A.26) covers before and after the All-Star break. Each team has between two and four home games and road games each week of the season due to constraints (A.27) and (A.28), as well as at least g_d total games each week of the season because of constraint (A.29) to prevent pooling. The number of home games on Sundays is balanced between three and four games for each team in constraint (A.30). In constraint (A.31), no two teams can play each other more than once every three days during the season. Teams are protected from consecutive road trips against the third (or worse) ranked furthest opponents in constraint (A.32). Lastly, in constraint (A.33), Winchester cannot play at Covington or Charlottesville on June 23–June 26 or July 7–July 10. Waynesboro travels to play in Charlottesville for the Fourth of July holiday in constraint (A.34). Binary decision variables are established in constraint (A.35).

References

Bartsch T, Drexl A, Kroger S (2006) Scheduling the professional soccer leagues of Austria and Germany. *Comput. Oper. Res.* 33(7):1907–1937.

Collegiate Baseball Newspaper (2018). Summer collegiate leagues. Retrieved January 15, 2018, http://baseballnews.com/summer-collegiate-leagues/.

Della Croce F, Oliveri D (2006) Scheduling the Italian football league: An ILP-based approach. *Comput. Oper. Res.* 33(7): 1963–1974.

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

Froncek D (2001) Scheduling the Czech national basketball league. *Congressus Numerantium* 153:5–24.

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

Goossens D, Spieksma F (2012) Soccer schedules in Europe. *J. Scheduling* 15(5):641–651.

Grabau M (2012) Softball scheduling as easy as 1-2-3 (strikes you're out). *Interfaces* 42(3):310–319.

Henz M (2001) Scheduling a major college basketball conference—revisited. *Oper. Res.* 49(1):163–168.

Kendall G (2008) Scheduling English football fixtures over holiday periods. *J. Oper. Res. Soc.* 59(6):743–755.

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

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

Kurt M, Karwan M, Pandey NK, Cunningham K (2015) Alleviating competitive imbalance in NFL schedules: An integer programming approach. *MIT Sloan Sports Analytics Conf.* (MIT Sloan Sports Analytics, Boston), 1–10.

Leonard J (2020) Information regarding increase in fans and revenue provided via personal communication with the authors, April 13.

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

Rasmussen RV (2008) Scheduling a triple round robin tournament for the best Danish soccer league. *Eur. J. Oper. Res.* 185(2): 795–810.

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

Russell RA, Leung JM (1994) Devising a cost effective schedule for a baseball league. *Oper. Res.* 42(4):614–625.

Saur MC, Starr K, Husted M, Newman AM (2012) Scheduling softball series in the Rocky Mountain Athletic Conference. *Interfaces* 42(3):296–309.

Trick MA (2003) *Integer and Constraint Programming Approaches for Round Robin Tournament Scheduling* (Springer, Berlin).

Trick MA, Yildiz H, Yunes T (2012) Scheduling Major League Baseball umpires and the traveling umpire problem. *Interfaces* 42(3):232–244.

Valley League Baseball (2018). Valley League Baseball homepage. Retrieved January 15, 2018, http://www.valleyleaguebaseball.com/view/.

Wright MB (2006) Scheduling fixtures for Basketball New Zealand. *Comput. Oper. Res.* 33(7):1875–1893.

Verification Letter

John Leonard, Media Relations Director, Valley Baseball League, P.O. Box 1127, 182 East Lee Highway, New Market, VA 22844, writes:

"The Valley Baseball League is a nonprofit summer collegiate league in the Shenandoah Valley in Virginia. With 11 teams currently in the league, developing a game schedule with many parameters is a huge nightmare. These parameters include teams not traveling or hosting games two nights in a row, having an equal number of home games on prime weekend dates, and spacing out each team's off-days, even though one team needs to be off each and every evening. Other parameters include avoiding playing the same team two days in succession, maximizing divisional matchups on prime dates, and honoring one team's wish to host a game on the Fourth of July holiday.

"The research team from Furman University has saved the league from this logistical nightmare with their optimized schedule program. In taking so many different parameters into account and optimizing the league schedule, the schedule makers have saved the league an estimated 40 or more hours of frustration and helped each team maximize attendance and profits during the season.

"I was personally amazed as well at how few complaints the owners of the teams have reviewing the computer-generated schedule. The produced schedule satisfied every single requirement that the league sent to the research team.

"Overall, the Valley League is clearly better off, thanks to the work of the Furman faculty and students."

Elizabeth L. Bouzarth is an associate professor of mathematics at Furman University. She earned her PhD in mathematics from the University of North Carolina at Chapel Hill in 2008. As an applied mathematician, she has many research interests ranging from sports analytics to computational fluid dynamics with applications to biology, with projects often involving undergraduate collaborators.

Benjamin C. Grannan is an assistant professor of business analytics at Furman University. Ben earned his PhD from Virginia Commonwealth University in 2014 and has published research with military, healthcare, and sports analytics applications.

John M. Harris is a professor of mathematics at Furman University. He earned his PhD in mathematics at Emory University in 1995. His current research interests include sports analytics, recreational mathematics, and graph theory.

Kevin R. Hutson is a professor of mathematics at Furman University. Kevin received his PhD in mathematical sciences in 2002 from Clemson University. His scholarship is at the interface of operations research and computer science and in recent years has centered around network optimization and sports analytics. These interests have led him to interesting consulting opportunities with the NCAA and ESPN.