
Scheduling Softball Series in the Rocky Mountain Athletic Conference

Author(s): Marjorie Cone Saur, Kaleigh Starr, Mark Husted and Alexandra M. Newman

Source: *Interfaces*, May-June 2012, Vol. 42, No. 3, Special Issue on Analytics in Sports, Part II: Sports Scheduling Applications (May-June 2012), pp. 296-309

Published by: INFORMS

Stable URL: <https://www.jstor.org/stable/23254871>

REFERENCES

Linked references are available on JSTOR for this article:

https://www.jstor.org/stable/23254871?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

INFORMS is collaborating with JSTOR to digitize, preserve and extend access to *Interfaces*

Scheduling Softball Series in the Rocky Mountain Athletic Conference

Marjorie Cone Saur

Rio Tinto Kennecott Utah Copper, marjorie.saur@riotinto.com

Kaleigh Starr

Division of Economics and Business, Colorado School of Mines, Golden, Colorado 80401,
kstarr@mymail.mines.edu

Mark Husted

Department of Electrical Engineering and Computer Science, Colorado School of Mines,
Golden, Colorado 80401, markhusted@gmail.com

Alexandra M. Newman

Division of Economics and Business, Colorado School of Mines, Golden, Colorado 80401,
newman@mines.edu

The Rocky Mountain Athletic Conference is a Division II National Collegiate Athletic Association conference that offers, *inter alia*, women's softball. Within the conference, four-game series are played against every other conference team according to a temporally constrained schedule. Manually generated schedules result in imbalances, such as breaks of multiple home or away series and away-series season openers and closers for the same team, and fail to mitigate weather-related series disruptions. Our integer-programming-based schedules eliminate these problems while ensuring that all requisite series are played. In this paper, we present a 40-game schedule; we do not present 36- and 44-game schedules, which are nearly equivalent. For its 2011 softball season, the Rocky Mountain Athletic Conference adopted the 40-game schedule from these three schedules.

Key words: scheduling; sports applications; integer programming.

History: This paper was refereed. Published online in *Articles in Advance* November 28, 2011.

The National Collegiate Athletic Association (NCAA), originally the Intercollegiate Athletic Association of the United States (IAAUS), dates back to 1906. The NCAA is a voluntary organization through which the nation's colleges and universities (hereafter loosely referred to as schools) govern their athletics programs. The NCAA, which manages 36 sports, comprises three divisions—I, II, and III—depending on the amount of athletic scholarship funding available for student athletes. Division I athletics teams receive the most scholarship funding; Division III athletics teams receive none. Currently, about 350 schools belong to Division I, 280 belong to Division II, and 450 belong to Division III. Within each division, schools (and the associated sports they offer) are divided into conferences with approximately 10 schools each, usually based on geographi-

cal proximity. For example, the Pacific Ten Conference (Division I) includes mainly large, public West Coast and southwestern schools.

George Hancock invented softball in Chicago on Thanksgiving Day in 1887 to keep baseball players in shape during the winter months, hence the original name "indoor baseball." The game was played using a rolled-up boxing glove as a ball; it was soft enough to allow players to use their bare hands to field the ball. The bat for the first softball game was a broomstick. Hancock developed a bat and ball a short time later, and the first rules of softball were published in 1889 under the name "indoor-outdoor." The game quickly spread around the country, and in 1926 became known as softball. By the 1930s, versions of softball were being played, mostly outdoors, throughout the United States and Canada. Although

originally slow pitch, as the sport developed, both slow pitch and fast pitch became recognized with official rules. Today softball is popular in a number of nations and is played by millions worldwide.

Women’s fast-pitch softball is an NCAA sport offered in all divisions (I, II, and III). At the time of this writing, the Rocky Mountain Athletic Conference (RMAC) is a Division II conference comprising 14 schools in three states: Colorado (10 members), Nebraska (2 members), and New Mexico (2 members). The RMAC, headquartered in Colorado Springs, was established in 1909 and fields 10 men’s sports and 9 women’s sports. Formerly known as the Colorado Faculty Athletic Conference (1909–1910) and the Rocky Mountain Faculty Athletic Conference (1910–1967), the RMAC is the oldest conference in Division II and the fifth-oldest athletic conference in the United States. Twelve RMAC schools field softball teams, which are divided into East and West Divisions (not to be confused with the same name used for the overall organization of the NCAA). The East Division comprises Chadron State College, the Colorado School of Mines, Metropolitan State College of Denver (Metro State College), Regis University, the University of Colorado–Colorado Springs, and the University of Nebraska–Kearney. The West Division comprises Adams State College, Colorado State University–Pueblo, Fort Lewis College, Mesa State College, New Mexico Highlands University, and Western New Mexico University (see Figure 1).

These RMAC varsity softball teams compete against one another in the regular season to advance to the RMAC postseason conference tournament via one of the eight highest conference (i.e., regular-season) records. The regular-season softball schedule requires that each team play every other team in the conference exactly once. The schedule is temporally constrained, that is, it includes as many time slots (weekends) in which to play as series of games to be played (or, it does not include bye weekends), making it difficult to create a balanced schedule. As a result, manually generated schedules often produce undesirable characteristics; for example, the number of home and away series for a given school could be unequal or the number of consecutive away weekends for any given team in any given season could be numerous. We seek to eliminate these undesirable characteristics

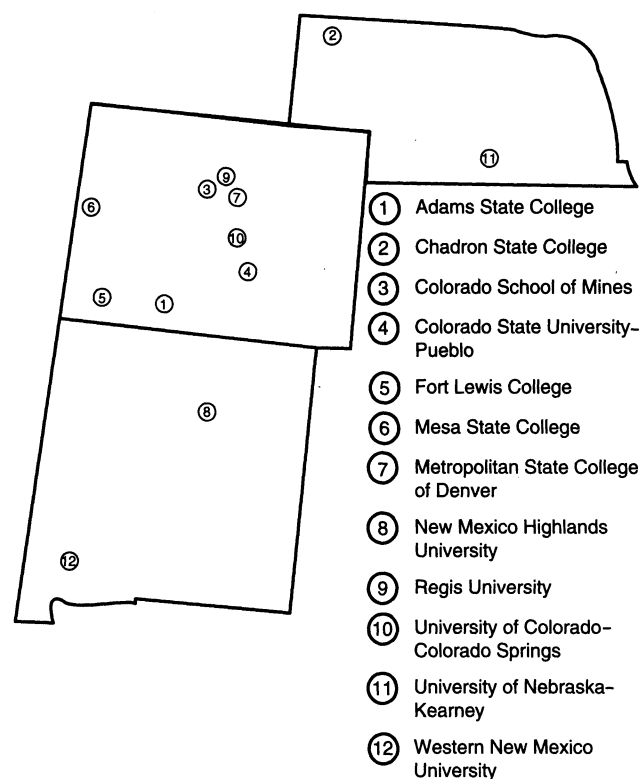


Figure 1: The RMAC consists of 12 schools that offer women’s intercollegiate softball programs; these schools are geographically dispersed over the states of Nebraska, Colorado, and New Mexico.

while ensuring that each RMAC team plays the requisite series. This paper is organized as follows. In the *Literature Review* section, we review other applications of operations research (OR) to athletics, especially sports scheduling. In the *Problem Statement* section, we further explain the regular-season scheduling problem and our method for addressing it. The *Results* section provides our schedules and explains their improvements relative to previous schedules. *Conclusions and Extensions* concludes the paper and mentions extensions to the existing work. Appendix A shows the scheduling formulation and its explanation, Appendix B provides our detailed featured solution and its proof of optimality, and Appendix C provides information regarding how we categorize schools when generating the schedules.

Literature Review

Applications of OR in sports are numerous and span at least the last four decades. As such, it is difficult to

give a comprehensive overview of even the most relevant work; we therefore provide some highlights and references to survey articles. OR outside the subfield of optimization has been used to assess and characterize team and player performance. However, we focus our literature review on optimization models, which have been used, *inter alia*, to schedule resources and regular and postseason play. Evans (1988) and Farmer et al. (2007) schedule umpires for the American Baseball League and professional tennis, respectively. Considerations in both models, and in umpire scheduling models in general, include, *inter alia*: (1) umpire travel costs, (2) the number of times an umpire is scheduled to referee a team or player within the same season, and (3) an umpire's skill level and other characteristics; for example, in dual-gender sports such as tennis, a referee's gender is a consideration.

Although the abovementioned applications are important and receive attention in the academic literature, game scheduling receives the most attention. One aspect of game scheduling is regular-season play. Easton et al. (2003) produce a minimum-distance schedule for a group of teams, where each team must play the other teams in the group, either at home or away, a given number of times. The authors use a combination of integer programming and constraint programming to solve the master and pricing problems, respectively, within their branch-and-price methodology. Trick (2001) presents a general two-phase framework in which home and away requirements of the teams are only considered in the second phase; this framework is applicable for situations in which restrictions on home and away patterns do not exist. Urban and Russell (2003) use constraint programming to consider scheduling games in which both competing teams are away. Their approach yields schedules for as many as 30 teams.

Specific applications are connected with both professional and college-level teams. Durán et al. (2007) and Croce and Oliveri (2006) schedule professional soccer leagues in Chile and Italy, respectively, taking into account issues of fairness to the players and benefits to the fans. The latter work borrows from the seminal work of Nemhauser and Trick (1998), who schedule regular-season games for a major U.S. basketball conference using a three-phase procedure.

Henz (2001) uses constraint programming to improve the solution time of this approach.

Authors also consider end-of-season tournaments in which a subset of teams involved in regular-season play contest for a national title; examples include Smith et al. (2006) and Smith (2009) for NCAA basketball and baseball, respectively. Specifically, these models consider how teams are paired and the locations at which games occur for postseasonal play.

Our model proposes a regular-season, temporally constrained schedule for a conference within women's Division II softball. With current hardware and software, we are able to obtain a variety of good schedules without using special optimization techniques (cf. Nemhauser and Trick 1998). Our model is unique in that it considers a variety of RMAC-specific constraints (see the *Problem Statement* section). For example, in contrast to Trick (2001), home and away patterns are of great importance; similarly, in contrast to Urban and Russell (2003), one of two competing teams is generally at home on any given weekend. Although fan preferences are of less importance than in Durán et al. (2007), Croce and Oliveri (2006), and Nemhauser and Trick (1998), we emphasize fairness to players relative to driving times and balanced home and away schedules. These considerations arise because of the academic pressures of the schools involved. Furthermore, because women's softball has a lower profile than that of professional sports and many other collegiate sports, it and the associated concerns relating to the specific schedule structures have received minimal attention. We refer the reader to a detailed, annotated bibliography concentrating principally on sports scheduling with a discussion of a few other sporting applications (Kendall et al. 2010); softball is mentioned only cursorily. In the following section, we describe our model in detail.

Problem Statement

According to the NCAA rule book, softball teams are allowed to begin playing games on February 1 of each year. However, the RMAC does not begin conference play until the last weekend in February. This leaves approximately three weeks at the beginning of the season during which RMAC teams may engage in play against teams in their region (but not

in their conference). This regional play contributes toward regional standings, which are very important for each team in the RMAC; they determine the number of RMAC teams that obtain a bid to the end-of-season regional tournament. By allowing teams the opportunity to compete against other conferences to improve regional standings, the RMAC as a whole can improve its national recognition and accreditation. However, preseason regional play leaves only 10 weekends for conference play during which 12 teams must face one another.

We refer to our proposed schedule, which the RMAC has adopted for 2011 play, as the *featured schedule*, and the schedule used in 2010 as the 2010 *actual schedule*. These schedules differ in how they address this mismatch between the number of RMAC teams that must play each other and the number of weekends in which conference play can occur. To play all the teams in the conference, the 2010 actual schedule offered two solutions that worked in tandem: (1) the schedule designated pairs of sister schools that play two games during two different weekdays of the regular season (i.e., four games); and (2) the schedule incorporated “pod play”—a weekend within the season when four teams meet at a single location and each team plays two games against two other teams in that same location, rather than the usual one-versus-one (four-game) match-up.

We change the featured schedule structure by eliminating weekday sister-school play. By eliminating weekday play, our schedule allows softball coaches to schedule regular weekly practice around which student athletes can accommodate their classes. Furthermore, weekday play had been conducted in an inconsistent manner. For example, different pairs of sister schools would play their games in different weeks; some schools would play two series of two games each in different weeks and on different fields, whereas other schools would play all four games in one “meeting” on the same field to avoid traveling twice in the same season. Except for sister-school play elimination, we retain all current scheduling rules. Specifically, we require that each RMAC team play every other RMAC team exactly once. We enforce this requirement as follows: (1) either a team plays an opponent in one four-game weekend *regular* series or (2) a team plays two games against two different

opponents in a four-game weekend *pod* series. Additionally, because conference tournament standings are based on east and west divisional records, we reserve the four-game regular series for play within a team’s own (east or west) division, because a four-game regular series provides more information for determining the better team within the series than the pod play, which has fewer (i.e., two) games against the same team. The RMAC wishes to have more information regarding the relative strength of two teams within the same division to enable it to make a better judgment about conference tournament standings. Therefore, we allow only teams that are not within the same (east or west) division to play each other in a pod series. During the pod weekend, three schools host a pod in which four teams play. Clearly, teams can only play each other in a pod weekend if they both take part in the same pod; additionally, the school hosting a pod must play in that pod.

Our model balances home and away series by allotting each team in the RMAC exactly four weekends of home play if hosting a pod and exactly five weekends of home play otherwise. In the remaining weekends, the team is a visitor. This allows each team the opportunity to play the same number of opponents on its home field.

Teams either start their conference play at home and end the season away or vice versa, with the remaining home and away series generally falling on alternating weekends. If the home (away) series do not fall on alternating weekends, we penalize the repetition. Specifically, we penalize the occurrences of home-home breaks within a school’s schedule. Our objective function seeks to minimize the penalties incurred from violating this single, elasticized constraint. We disallow any school from having more than one home-home break.

The home series are allocated to prevent any team from having too many consecutive away series. Our model differentiates “close” and “far” series relative to each individual school with an RMAC softball program. For any given school, we establish “far” teams based on the greatest gap in reasonable driving time. Traveling to consecutive away series tires a team physically, results in more missed class time in a short time frame, and impairs practice quality. Teams playing their third consecutive away weekend could

Rank	From Adams State to:	Driving time (hours:minutes)	Distance determination
1st closest	Colorado State University–Pueblo	2:07	Close
2nd closest	University of Colorado–Colorado Springs	2:48	Close
3rd closest	Fort Lewis College	2:55	Close
4th closest	Colorado School of Mines	3:39	Close
5th closest	Metro State College	3:44	Close
6th closest	New Mexico Highlands University	3:46	Close
7th closest	Regis University	3:47	Close
8th closest	Mesa State College	4:44	Far
9th closest	Western New Mexico University	7:41	Far
10th closest	Chadron State College	8:22	Far
11th closest	University of Nebraska–Kearney	8:41	Far

Table 1: Schools that are “far” from Adams State, relative to other RMAC schools, are determined based on their relative driving distance.

be viewed as having a disadvantage against the home team both because of stress caused by scrambling to make up missed academic work and because of time spent in buses or vans rather than out on the field. Our schedules eliminate back-to-back far-away traveling weekends; they also eliminate three consecutive weekends of away travel, regardless of the distance of travel.

Table 1 demonstrates how we choose corresponding “far” schools for one school, Adams State. We first rank driving time from Adams State to each other RMAC school in ascending order. We use driving time as the metric because virtually all schools drive to competitions. Given the relative driving times, we subjectively and manually search for a break in the list such that at least approximately half of the schools are considered close, but that the break is significant enough to differentiate “close” and “far” schools. Although a significant difference (nearly an hour) of driving time occurs between the University of Colorado–Colorado Springs and the Colorado School of Mines, placing the dividing line between these two schools would leave only three schools in the “close” category, which might overly constrain our model. Therefore, we continue to search down the list for another significant difference in driving time. We find such a difference (one hour) for Adams State between Regis University and Mesa State College; therefore, we place the dividing line after Regis University, resulting in seven “close” schools.

Our characterization does not adhere to symmetry properties. That is, school *A* could be “far” from

school *B*, whereas school *B* is close to school *A*. This asymmetry is necessary because of the uneven geographical distribution of the schools. In particular, those schools located in Nebraska and New Mexico are relatively far away from most other schools. However, to label almost all other schools as “far” away from a Nebraska or New Mexico school would likely render the model with the associated constraint regarding back-to-back far-away travel infeasible. At best, this constraint would then have to be elasticized; however, the elasticization could lead to imposing unfair travel distances on Colorado schools. Nebraska and New Mexico have other conferences, and those schools in the RMAC have elected to be in the conference, aware of the travel distances involved. The Colorado schools already in the conference would not have allowed the more distant schools to join had this resulted in a disproportionate increase in the Colorado schools’ travel time. However, we do designate approximately the same number of “close” and “far” schools for each team. On average, each school is associated with about seven “close” schools and four “far” schools; this corresponds to an average “close” driving time of 251 minutes and a “far” driving time of 560 minutes. We provide more details concerning the possibilities for and drawbacks of more equitably distributing driving times in Appendix C.

To further strengthen our model, we add weather considerations. Specifically, early-season play is often disrupted by poor (usually cold, snowy) conditions. Schools less affected by poor weather include Colorado State University–Pueblo, Fort Lewis College

(because of close access to an alternate, warm venue), Mesa State College, Metro State College, New Mexico Highlands University, Regis University, and Western New Mexico University. We refer to these schools as “warm-weather schools.” (We grant that this is a relative designation!) We then place a constraint in the model ensuring that two of the first four series are played at warm-weather schools.

In the event of a weather disruption, games within a series can only be rescheduled for consecutive days (e.g., Friday–Saturday or Sunday–Monday) surrounding the original weekend (Saturday–Sunday) in which the games were scheduled. If the games are not played, they cannot be moved to different weekend slots but, per current RMAC rules, are canceled; this results in some teams playing fewer games than others during the regular season. Although a cancellation is undesirable because the win-loss percentage for a given team is based on fewer games, and a team may unfairly benefit or be penalized depending on the likely winner of the canceled games, it is not practical to make up the game because of the lack of time available. (Student athletes have classes during the week.) Therefore, we do not consider game rescheduling in our model.

Results

We implement our integer programming model using the mathematical programming language AMPL (AMPL 2001) and solve it with the mixed-integer programming algorithm offered by CPLEX 12.1 (IBM 2009). We set a time limit of two hours and 15 minutes, during which we gather five solutions using the algorithm’s “solution-pool” feature, which stores feasible solutions as it conducts its search. All solutions possess the same constraint violation, namely, seven penalties representing seven schools with a home-home break. We presented these schedules to the RMAC commissioner and his associates, who noticed a characteristic of some schedules, which we had not considered. Specifically, Western New Mexico University is a long drive from almost any other RMAC school. Those schools who traveled to Western New Mexico University in the 2010 season place a high premium on not having to travel there in the 2011 season (i.e., two years in a row). Therefore, from two schedules, we arbitrarily choose one that minimizes

the number of teams that travel to Western New Mexico University and also traveled there in 2010; this results in three schools’ teams traveling there in two consecutive years.

Ideally, the RMAC commissioner and his associates would have mentioned the desire to avoid repeat travel to Western New Mexico University before we obtained results from our optimization model, and we would have considered this desire explicitly in the model. For example, we might have minimized repeat travel to Western New Mexico University subject to the constraint that no more than seven schools possess a home-home break in their schedules. Ex post facto, we ran this model and found a solution in which only one school repeated travel to Western New Mexico University from 2010 to 2011. (This is the minimum number because seven teams traveled to Western New Mexico University in 2010, and Western New Mexico University hosted five teams in 2011, either as a regular four-game series or as part of a pod.) Although the ex post facto schedule might have been better in retrospect, we offer the following observations: (1) the three-team repeat travel to Western New Mexico University is a relatively minor problem in that it only affects an unnecessary two teams, and only in one year. In 2012, schedules can be constructed based on (partially) “mirroring” the schedules from the 2011 season, precluding unnecessary repeat travel. (2) RMAC administrators did not evaluate the schedule with only one repeat visit to Western New Mexico University. This schedule might have highlighted a different problem, not explicitly considered in our optimization model, yet not present in our proposed schedule. Therefore, we cannot say with certainty whether RMAC administrators would have considered the schedule with only one repeat visit to Western New Mexico University as preferable to our proposed schedule. In any case, as we demonstrate in the remainder of this section, the schedule that the RMAC adopted is a vast improvement over previous schedules. We compare the featured schedule against historical schedules, specifically, the 2010 actual schedule.

Note that the solutions we generate subscribe to the new scheduling paradigm. In past schedules, most schools were associated with a (relatively close) sister school, which they generally played twice during

the week (although this varied between sister-school teams). We replace this weekday play with an extra four-game weekend series at the beginning of the season. Hence, our featured schedule contains only weekend play, and one more such weekend (i.e., 10 weekends) than the 2010 actual schedule, which contained only nine weekends of play. The change in schedule structure makes a perfect comparison between the actual and featured schedules impossible when counting home and away breaks and total number of home and away series, and when assessing weather-related play. Therefore, we compare both actual and featured schedules only with respect to weekend (nonsister-school) play (i.e., play common to both schedules). Despite, or because of, the change in the schedule structure, we are able to generate good, consistent solutions that balance home and away series, consider weather, and eliminate the ad hoc nature in which sister-school play was conducted.

Although solutions can be compared across many dimensions, we chose the following characteristics that we deem most important:

- The total number of home (away) series each team plays
- The total number of home (away) opponents each team plays
- The greatest number of home (away) series played in successive weekends, termed a “home (away) break”
- The greatest number of away series played against “far” schools in successive weekends, termed a “far-away break”
- Whether or not each team either starts the season with conference play at home or ends it at home
- Whether or not each school plays two of its first four series at warm-weather schools.

We note that the actual 2010 schedule includes one case in which the difference between the number of home and away series for a given team is three; for the other schools, the number of home and away series is relatively even (given that there are nine series other than sister-school play). However, the 2010 actual schedule appears much less balanced if we consider the number of home and away opponents, which differs from the number of home and away series because of pod play. We “count” hosting

a pod as hosting two home series, because the team hosting the pod is playing two different opponents on its home field. With the actual schedule, only one school has a perfectly even balance, whereas most schools have a four-six split. In the extreme case, one school plays only 3 of 10 opponents at home (outside of sister-school play). Our featured schedule has all teams playing exactly 5 opponents at home and 6 opponents away. The three teams that host the pod series play four weekends at home and six weekends away. The total number of series (and opponents) for each school is one more in Table 3 than in Table 2 (i.e., Table 3 depicts a schedule with one more weekend game than Table 2) because we do not count weekday sister-school play (which accounts for one series and one opponent) in the 2010 actual schedule; in the 2011 featured schedule, all series occur on weekends.

Another problem with the actual schedules is the order in which home and away series are played (e.g., multiple away series or far-away series on consecutive weekends, which are countered at other points in the schedule with multiple home series). The 2010 actual schedule has all but one team playing at least two consecutive weekends at home, with two teams playing three consecutive weekends at home. To counterbalance the home breaks, each team in the 2010 actual schedule plays at least two consecutive away series, and five teams play three consecutive away series. Of the away breaks, four teams play two far-away teams in a row in the 2010 actual schedule (including one team that starts its season this way). By contrast, our featured schedule has seven teams playing two consecutive weekends at home and the other five teams playing no more than one home series on consecutive weekends. In Appendix B, we show that our featured schedule is optimal given our other (hard) constraints, specifically, those that enforce pod play and those that require that conference play either start or end at home for each school. Finally, no team in our featured schedule plays more than one far-away series on two consecutive weekends.

Although softball players are accustomed to the low-profile nature of their sport, it is relevant and fair to consider aspects of the schedule that might increase the sport’s profile and player morale. Starting or ending conference play at home provides an

Characteristic	Quantitative assessment for each team											
	1	2	3	4	5	6	7	8	9	10	11	12
Number of home series	4	4	4	5	4	5	5	4	4	4	3	5
Number of away series	5	5	5	4	5	4	4	5	5	5	6	4
Number of home opponents	4	4	4	6	4	6	5	4	4	4	3	6
Number of away opponents	6	6	6	4	6	4	5	6	6	6	7	4
Longest home break	2	2	3	3	2	2	2	2	2	2	1	2
Longest away break	3	2	2	3	2	2	2	2	3	2	3	3
Longest “far-away” break	1	1	2	1	1	1	1	2	2	1	2	1
One of first and last series at home?	Y	N	Y	Y	N	Y	Y	Y	Y	Y	Y	Y
Warm-weather series in first four weeks	1	2	2	3	4 [†]	3	3	2	4	1	3 [†]	2

Table 2: The table shows characteristics of the actual 40-game schedule. We summarize important characteristics of the 2010 RMAC schedule (minus sister-school play) for each team; each column corresponds to a team with the number given in Figure 1.

[†]These schools each played one “warm-weather series” at a neutral location.

opportunity for athletics departments to give their teams schoolwide attention. In the actual schedule, two teams play their first and last conference series away, whereas in the featured schedule, no team does. We summarize these results in Tables 2 and 3.

Characteristic	Quantitative assessment for each team											
	1	2	3	4	5	6	7	8	9	10	11	12
Number of home series	5	5	5	5	4	5	5	5	5	4	4	5
Number of away series	5	5	5	5	6	5	5	5	5	6	6	5
Number of home opponents	5	5	5	5	5	5	5	5	5	5	5	5
Number of away opponents	6	6	6	6	6	6	6	6	6	6	6	6
Longest home break	2	2	1	2	1	1	2	2	2	1	1	2
Longest away break	2	2	1	2	2	1	2	2	2	2	2	2
Longest “far-away” break	1	1	1	1	1	1	1	1	1	1	1	1
One of first and last series at home?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Warm-weather series in first four weeks	2	2	2	4	2	3	4	3	4	2	2	2

Table 3: The table shows characteristics of the featured 40-game schedule. Of the five solutions with seven penalties, we choose one based on attributes that are not considered in the integer program; that is, we minimize the number of schools traveling to Western New Mexico University (a long drive for any team) in consecutive years (i.e., 2010 and 2011). We summarize important characteristics of this solution for each team in the RMAC. Each column corresponds to a team with the number given in Figure 1.

Conclusions and Extensions

We have applied standard integer programming modeling techniques to determine a schedule for NCAA Division II women’s softball; this schedule is far preferable to the 2010 actual schedule. Although, with thoughtful mathematical formulation our program does not necessitate any special decomposition algorithms to enable it to find solutions in a reasonable amount of time, the schedule(s) we obtain contain characteristics that are nearly impossible to include in manually generated solutions.

As in all arenas, it is difficult to gain acceptance of a computer-generated schedule, which many people tend to view with suspicion. Emphasizing both the benefits of the featured schedule—which were absent in previous-year schedules—and also the schedule’s objectivity (hence, fairness) relative to each school helps establish credibility for the schedule’s implementation. In this paper, we present a 40-game schedule that allows for nonconference (i.e., regional) play in the first three weeks of the season. This gives teams the chance to gain status outside of the conference for postseason tournament play. With 10 remaining weekends for conference play and with weekday play precluded, our featured schedule includes one pod to ensure that all RMAC teams see each other exactly once (weather permitting!) during the regular season. Some (more competitive) schools wish to see a more aggressive schedule with two weekends of pod play, which allows an entire month of pre-RMAC games. Other schools are content simply to participate in the conference, and prefer conference play in a more relaxed manner over early-season regional play. As such, these schools would like to see a 44-game RMAC schedule comprising 11 weeks of a 4-game series against a single team on each weekend, leaving only two weeks for early-season regional play. A third camp of schools is interested in the 40-game compromise that we present here. Clearly, fewer pods make the schedule easier for us to obtain, but this relaxation is counterbalanced by the desire for more attention to be given to weather-related scheduling, especially because longer seasons begin earlier in the year. We presented the RMAC coaches and administrators with one featured schedule for each of these three alternatives, thereby retaining the decision makers as part of the process. In June 2010, RMAC school

representatives voted to accept our 40-game schedule for 2011 play.

The RMAC commissioner has pointed out key characteristics, regardless of the number of games incorporated, such that our featured schedules were accepted: (1) we construct our schedule in a systematic and reproducible manner; and (2) the way in which we construct our schedule “removes the personalities from the process.” These aspects assist in forming a consensus regarding which schedule to adopt because decision makers no longer view schedules as whimsically created or biased toward particular schools. The Colorado School of Mines athletics director points out that for years coaches and administrators have been under the impression that the scheduling problems simply could not be fixed. He argues that using sound mathematical techniques precludes this thought process from persisting in the RMAC.

Although the featured schedule has its benefits, it also has some potential limitations: certain schools may have preferred home weekends for recruiting or related activities; however, no school indicated these preferences to us; hence, we do not currently consider such limitations. In subsequent years, we could easily incorporate such preferences. The RMAC commissioner has indicated a desire to create schedules in four-year cycles, thereby allowing each school to host a pod once every four years. Finally, teams may join or leave the RMAC, necessitating schedule updates. However, for the present, the following facts preclude these details from needing to be considered: (1) softball fields generally are dedicated for softball use; unlike basketball gymnasiums, they are not used for multiple types of events; and (2) we have no concrete information regarding the schools that will constitute the RMAC in the future. In addition, each school has a different travel budget; however, the featured schedule balances costs for each school without consideration of any given school’s budget. We could easily modify this, but deem it inappropriate in light of our attempt to capture a fair schedule for each school; in this context, it would not be reasonable to require teams from more affluent schools to travel more often and (or) farther.

With the current considerations and corresponding detail contained in the model, instances run fast

enough to preclude us from considering alternate formulations or solution techniques. However, we are aware that constraint programming can be a useful and efficient modeling paradigm for scheduling (Henz 2001). Were our solution times deemed excessive either with the current model or with a future model (e.g., if more schools with softball programs were to join the conference), we might wish to consider constraint programming or methods of decomposing the problem (Nemhauser and Trick 1998). However, this is beyond the scope and necessity of our current project.

Appendix A. Formulation

Sets

- T : set of teams,
- T^P : set of teams that host pods,
- T^E : set of teams in the East Division,
- T^W : set of teams in the West Division,
- T^C : set of cold-weather teams,
- F_i : set of teams “far” from team i ,
- S : set of all slots (weekends),
- S^n : set of nonpod slots (i.e., weekends 1, 2, 3, 4, 5, 6, 8, 9, 10).

Variables

- $x_{ijk} = \begin{cases} 1 & \text{if team } i \text{ plays at team } j \text{ in nonpod slot } k, k \in S^n \\ 0 & \text{otherwise,} \end{cases}$
- $y_{ijp} = \begin{cases} 1 & \text{if team } i \text{ plays team } j \text{ in a pod hosted by team } p \\ 0 & \text{otherwise,} \end{cases}$
- $z_{ip} = \begin{cases} 1 & \text{if team } i \text{ plays in a pod hosted by team } p \\ 0 & \text{otherwise,} \end{cases}$
- $w_p = \begin{cases} 1 & \text{if team } p \text{ hosts a pod} \\ 0 & \text{otherwise,} \end{cases}$

subject to

$$\sum_{j \in T: i \neq j} x_{ijk} + \sum_{j \in T: i \neq j} x_{jik} = 1 \quad \forall i \in T, k \in S^n, \quad (1)$$

$$\sum_{k \in S^n} (x_{ijk} + x_{jik}) + \sum_{p \in T^P} y_{ijp} = 1 \quad \forall i \in T, j \in T \ni i \neq j, \quad (2)$$

$$\sum_{j \in T: i \neq j} (x_{jik} + x_{ji, k+1}) \leq 1, \quad \forall i \in T, k \in S: k < 10, k \neq 6, 7, \quad (3)$$

$$\sum_{j \in T: i \neq j} x_{jik} + w_i \leq 1 \quad \forall i \in T, k \in S: k = 6, \quad (4)$$

$$w_i + \sum_{j \in T: i \neq j} x_{ji, k+1} \leq 1 \quad \forall i \in T, k \in S: k = 7, \quad (5)$$

$$\begin{aligned}
 \sum_{i \in T: i \neq j} (x_{ij1} + x_{i,j,10}) &= 1 \quad \forall j \in T, & (6) \\
 y_{ijp} &= y_{jip} \quad \forall i \in T, j \in T: i \neq j, p \in T^p, & (7) \\
 \sum_{j \in T: i \neq j} y_{ijp} &= 2z_{ip} \quad \forall i \in T, p \in T^p, & (8) \\
 \sum_{j \in T, k \in S^n: i \neq j} x_{jik} + 2w_i &= 5 \quad \forall i \in T, & (9) \\
 \sum_{p \in T^p} w_p &= 3, & (10) \\
 \sum_{i \in T} z_{ip} &= 4w_p \quad \forall p \in T^p, & (11) \\
 \sum_{p \in T^p} z_{ip} &= 1 \quad \forall i \in T, & (12) \\
 w_p &\leq z_{ip} \quad \forall i \in T, p \in T^p: i = p, & (13) \\
 \sum_{j \in T^E: i \neq j} y_{ijp} &= 0 \quad \forall i \in T^E, p \in T^p, & (14) \\
 \sum_{j \in T^W: i \neq j} y_{ijp} &= 0 \quad \forall i \in T^W, p \in T^p, & (15) \\
 \sum_{j \notin T^C, k \in S^n: k < 5} x_{ijk} &\geq 2 \quad \forall i \in T^C, & (16) \\
 \sum_{j \in T, k \in S: i \neq j \text{ and } k < 5} x_{jik} + \sum_{j \notin T^C, k \in S^n: i \neq j \text{ and } k < 5} x_{ijk} &\geq 2 \quad \forall i \notin T^C, & (17) \\
 \sum_{j \in F_i: i \neq j} (x_{ijk} + x_{ij,k+1}) &\leq 1 \quad \forall i \in T, k \in S: k \neq 6, 7; k < 10, & (18) \\
 \sum_{p \in F_i} z_{ip} + \sum_{j \in F_i: i \neq j} x_{ijk} &\leq 1 \quad \forall i \in T, k \in S: k = 6, & (19) \\
 \sum_{p \in F_i} z_{ip} + \sum_{j \in F_i: i \neq j} x_{ij,k+1} &\leq 1 \quad \forall i \in T, k \in S: k = 7, & (20) \\
 \sum_{j \in T: i \neq j} (x_{jik} + x_{ji,k+1} + x_{ji,k+2}) &\geq 1 \quad \forall i \in T, k \in S: k = 1, 2, 3, 4, 8, & (21) \\
 w_i + \sum_{j \in T: i \neq j} (x_{jik} + x_{ji,k+1}) &\geq 1 \quad \forall i \in T, k \in S: k = 5, & (22) \\
 w_i + \sum_{j \in T: i \neq j} (x_{jik} + x_{ji,k+2}) &\geq 1 \quad \forall i \in T, k \in S: k = 6, & (23) \\
 w_i + \sum_{j \in T: i \neq j} (x_{ji,k+1} + x_{ji,k+2}) &\geq 1 \quad \forall i \in T, k \in S: k = 7. & (24)
 \end{aligned}$$

Constraints (1) require that in each nonpod weekend, a team plays exactly one other team, either at home or away. Constraints (2) require that each team plays every other team once during the season, either at home, away, or in a pod (at home or away). Constraints (3)–(5) penalize two home series in a row, either two regular series, series in which a hosted pod is preceded by a home series, or in which a hosted pod is followed by a home series, respectively. Constraints (6) ensure that a team either starts its

season at home or ends it at home. Constraints (7) ensure that if team i plays team j in a pod, then team j plays team i in the same pod. Constraints (8) guarantee that a team plays two other teams in each pod slot. Constraints (9) force there to be exactly five home series for each team, where hosting a pod counts as two home series. (This forces each team to play exactly five opponents at home.) Constraints (10) require that in the pod weekend, three schools host a pod. Constraints (11) ensure that a pod comprises exactly four teams. Constraints (12) require that a team must play in a pod in the pod weekend. Constraints (13) insist that if a team hosts a pod, then it must play in that pod. Constraints (14) and (15) preclude teams from the same division, east or west, respectively, from playing each other in a pod slot. Constraints (16) force schools in cold-weather areas to play at least two of their first four series away at schools in warm-weather areas. Constraints (17) force schools in warm-weather areas to play at least two of their first four series either at home or away at schools in warm-weather areas. Constraints (18)–(20) ensure that no team plays two consecutive “far-away” series. Constraints (21)–(24) require that each team play at home at least once over any consecutive three-weekend series. The four different variations of this constraint address the different slots in which the pod series lies relative to the regular four-game series.

The objective is to minimize the unweighted sum of the penalties, where the elasticized constraints, or those with penalties for their violation, are constraints (3)–(5), denoted by a dot over their relational operators; Brown et al. (1997) provide a discussion of and notation for such constraints. These constraints allow two-game home breaks only if a penalty is incurred. Explicitly stated, a binary variable in our code, e_{ik} , assumes a value of 1 if team i plays at home both in slot k and in slot $k + 1$, and 0 otherwise. In the objective, we minimize the sum over all teams and slots of this variable. We use an additional constraint ($\sum_k e_{ik} \leq 1 \quad \forall i$) to disallow more than one penalty to be acquired by the same team.

We omit these details, including the elastic variable definition and constraints limiting the use of the elastic variable, in our formulation because an objective gives the reader the sense that there is a predetermined, important goal. We approached our scheduling problem with no such goal in mind. Rather, we sought to find a feasible schedule, which equitably distributes the games among all RMAC schools. The objective we adopt arose out of necessity; without the relaxation, the model is infeasible, and relaxing other constraints is either impossible from the RMAC scheduling perspective, or less desirable.

After the CPLEX presolve, the relevant instance of our 40-game model contains just over 700 constraints, nearly 1,900 variables (of which most are binary), and 13,270 nonzeros in the “A matrix” (i.e., matrix of left-hand side coefficients). To obtain feasible solutions, we tune CPLEX as follows. We use the primal simplex algorithm at the root node and at the

subproblem nodes of the branch-and-bound tree. Regarding the solution pool specifically, we ask for up to five feasible solutions to be stored at any given time, retaining those with the best objective function values if more than five feasible solutions are found. We employ a parameter setting to find more feasible solutions—specifically, up to five—after running the optimization algorithm. Finally, we restrict the solutions collected to possess objective function values within five (penalty) units of that of the best solution found. With these settings and on a 1.5 GHz Ultra SPARC IIIi Unix workstation with 2 GB RAM, CPLEX requires less than a minute to find the first feasible solution, which has a total penalty of 9, and fewer than 500 seconds to find an optimal solution (i.e., a solution with a penalty of 7). Overall, we allowed the model to run for 8,100 seconds and collected five solutions, all with a penalty of 7.

Appendix B. Featured 40-Game Schedule

Our schedule contains seven penalties (i.e., seven home-home breaks), which we show here to be the optimal number of penalties. In each weekend except that in which pod-series play occurs (i.e., week 7), six home games and six away games are played. Pod-series play occurs in week 7, and 3 of the 12 teams host a pod. Because of this and the fact

that constraints (21)–(24) ensure that no team plays three consecutive away series, 3 of the teams that play away in week 7 also play away in week 6; additionally, 3 of the teams away in week 7 also play away in week 8, where the away teams in week 6 not hosting a pod in week 7 do not include any of the away teams in week 8. Hence, in the weekends before, during, and after the pod, there are 6, 9, and 6 teams, respectively, that play away series. This suggests the following (partial) patterns for teams not hosting pods (see Table B.1): (1) a “perfect” schedule in which home and away games occur in alternating weekends (see column 2); (2) a two-game away break occurring before the pod weekend with a home-game season opener (see column 3); (3) a two-game away break occurring before the pod weekend with a home-game season closer (see column 4); and (4) a two-game away break occurring after the pod weekend, which necessarily implies a home-game season opener (see column 5) or a home-home penalty (see column 6). At most, 3 teams can have the perfect schedule given in column 2, because not all 9 teams away in the pod weekend can play at home before and after the pod; both before and after the pod, three teams would be lacking opponents! Therefore, at least 6 teams must have schedules such as those occurring in columns 3, 4, 5, and 6. However, in these schedules, no

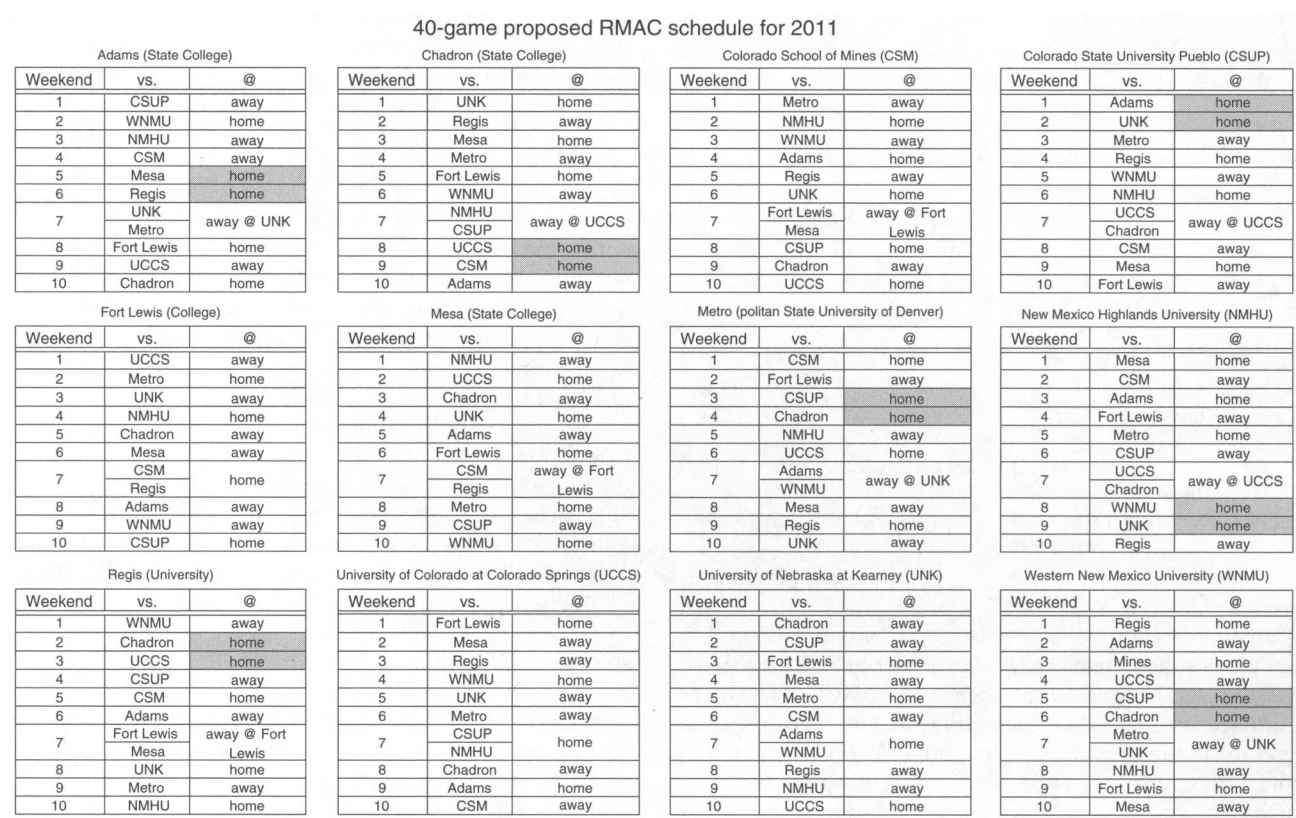


Figure B.1: Our featured 40-game schedule shows the seven penalty occurrences marked in gray.

Week	Teams with away pods					Teams hosting pods	
	"Perfect" schedule	Away series before pod; home opener	Away series before pod; home closer	Away series after pod; home opener	Away series after pod; home closer	Home opener	Home closer
1	A	H	A	H	A	H	A
2	H	*	*	*	*	*	*
3	A	*	*	*	*	*	*
4	H	*	*	*	*	*	*
5	A	H	H	*	*	*	*
6	H	A	A	H	H	*	*
7	A	A	A	A	A	H	H
8	H	H	H	A	A	*	*
9	A	*	A	H	H	*	*
10	H	A	H	A	H	A	H

Table B.1: There are several feasible home-away patterns (with degrees of freedom marked with an asterisk) given our scheduling rules, which include pod play in week 7, no away-away-away breaks, no bye weekends, and either a conference opener or closer at home, but not both. The bold row denotes the pod weekend.

matter how we populate the asterisks (while adhering to the rules that there can be no more than two consecutive away games, and that each team must play another in a weekend either at its home field or the opponent’s home field), we incur two consecutive home games. Therefore, we conclude that there must be at least six penalties in our objective function, all of which are derived from six schools that do not host a pod. However, one can see many feasible schedules for those schools hosting pod play in which consecutive home games are avoided, in part because these schools play only four, rather than five, home “weekends” (including pod weekends). See columns 7 and 8 of Table B.1.

The question then remains as to whether the schools that do not host pods can incur more than the six penalties described above. An equivalent question is whether the three schools that do not host a pod and do not play an away series either before or after the pod weekend can all have perfect schedules. We answer this by considering the following. Each pod has four teams that play games against two of the three other teams present. Each pod also includes two teams from each division. Teams in the same division do not play each other in the pod. Two of these three schools in question (regarding their ability to all have a perfect schedule) reside in the same division and one resides in a different division. The two from different divisions can possess the same home and away patterns and still play each other during the season by facing each other in the pod. The two from the same division cannot. To play each other, one team must have a home game (in a week other than week 7) when the other has an away game. However, the team that must trade an away series for a home series (or vice versa) from the perfectly balanced schedule to play a team in its own division creates a schedule with the seventh home-home break (i.e., seventh penalty). Hence, a 40-game schedule with seven penalties is optimal.

Appendix C. “Close” and “Far-Away” Designations

We give the average and standard deviation of each school’s driving time, relative both to the schools designated as “close” to a given school, and those designated as “far away” from a given school (see Table C.1). All times are given in the format hours:minutes. The two Nebraska schools, Chadron State College and the University of Nebraska-Kearney, and one New Mexico school, Western New Mexico University, possess the longest average driving times, whereas schools closer to the metro Denver area possess relatively short average driving times. Shorter average driving times to the close schools tend to equate to shorter average driving times to the “far-away” schools. Note, however, that standard deviations for the three schools with very long “close” average driving times tend to be much lower than those for the Denver metro schools with short average “close” driving times. The standard deviations for the “far-away” schools tend to be in the same range for all schools.

The statistics in Table C.1 show that schools in Nebraska and Western New Mexico University are a relatively long drive from all other schools. There might be an inclination to try to make average driving times for “close” and “far-away” schools equitable among all the RMAC participants. However, these schools have the least number of schools considered close to them, which lowers what might be an even higher average driving time to the “close” schools.

Our categorization of “close” and “far-away” schools serves to implement the constraint in which we do not permit travel to two “far-away” schools in two consecutive weeks. Note that we have two “perfectly balanced” schedules (see Appendix B) in which home and away games alternate. Two of the three schools (as long as they are not in the same division) with the highest average driving time to the “close” and “far-away” schools could be given these schedules; the other school could be given a schedule that has a single away-away break. In the latter case, it would be

School	Close schools			"Far-away" schools		
	Average driving time (hh:mm)	Standard deviation (hh:mm)	No.	Average driving time (hh:mm)	Standard deviation (hh:mm)	No.
Adams State College	3:34	0:44	7	7:20	2:16	4
Chadron State College	6:16	0:46	6	12:12	2:54	5
Colorado School of Mines	2:14	2:00	6	7:03	2:37	5
Colorado State University–Pueblo	2:09	0:54	6	7:08	1:45	5
Fort Lewis College	5:24	1:16	8	10:55	2:16	3
Mesa State College	5:08	0:48	7	10:23	1:17	4
Metropolitan State College of Denver	2:07	2:01	6	6:57	2:36	5
New Mexico Highlands University	4:57	0:56	8	10:06	1:17	3
Regis University	2:12	2:04	6	6:58	2:37	5
University of Colorado– Colorado Springs	2:01	1:21	6	7:10	1:51	5
University of Nebraska– Kearney	6:03	0:44	6	11:56	2:55	5
Western New Mexico University	8:10	1:59	5	13:25	2:47	6

Table C.1: For each school, we give the average and standard deviation of driving time for the schools that are considered “close” and “far away” from the given school. We also note the number of schools in each of these two categories.

possible to constrain one of the away games in this break to be at the school as close to it as a schedule would feasibly permit. In this manner, “close” and “far-away” designations for the three (or even more) schools with the greatest average driving times would be irrelevant.

However, all schools join the RMAC with the expectation that they will be treated like the other schools, and that no school that is geographically separated from the center of mass will receive special compensation for long average driving times. Implementing a policy such as the one suggested in the previous paragraph would create unfairness among schools, overly constrain schedules, and preclude multiyear schedules in which pod hosting should be rotated among schools. For these reasons, we treat the schools as equitably as possible with respect to our categorizations.

Acknowledgments

The authors would like to acknowledge the initial contributions of Stacie LaRocque-Tschacher, a Colorado School of Mines softball player who brought this problem to our attention, and her project partner, Colorado School of Mines student Erick Kennedy. Ashley Menger also assisted in coding the project. The authors thank Ryan Bracken for retrieving background material and references. They also thank Ric Wendlandt (Faculty Athletics Representative), Tom Spicer (Athletics Director), Annie Van Wetzlinga (Head Softball Coach), Dixie Cirillo (Associate Athletics Director), all of the

Colorado School of Mines, and J. R. Smith, RMAC Commissioner, for their comments on and support of this project.

References

AMPL. 2001. AMPL optimization LLC, www.ampl.com.
Brown, G., R. Dell, K. Wood. 1997. **Optimization and persistence.** *Interfaces* 27(5) 15–37.
Croce, F., D. Oliveri. 2006. Scheduling the Italian football league. *Comput. Oper. Res.* 33(7) 1963–1974.
Durán, G., M. Guajardo, J. Miranda, D. Sauré, S. Souyris, A. Weintraub. 2007. Scheduling the Chilean soccer league by integer programming. *Interfaces* 37(6) 539–552.
Easton, K., G. Nemhauser, M. Trick. 2003. Solving the travelling tournament problem: A combined integer programming and constraint programming approach. E. Burke, P. De Causmaecker, eds. *Proc. 4th Internat. Conf. Practice Theory Automated Timetabling (PATAT 2002), Lecture Notes in Computer Science*, Vol. 2740. Springer, Heidelberg, Germany, 100–109.
Evans, J. 1988. A microcomputer-based decision support system for scheduling umpires in the American Baseball League. *Interfaces* 18(6) 42–51.
Farmer, A., J. Smith, L. Miller. 2007. Scheduling umpire crews for professional tennis tournaments. *Interfaces* 37(2) 187–196.
Henz, M. 2001. Scheduling a major college basketball conference—Revisited. *Oper. Res.* 49(1) 163–168.
IBM. 2009. ILOG CPLEX. Accessed April 1, 2010, <http://www-01.ibm.com/software/integration/optimization/cplex>.
Kendall, G., S. Knust, C. Ribeiro, S. Urrutia. 2010. Scheduling in sports: An annotated bibliography. *Comput. Oper. Res.* 37(1) 1–19.

- Nemhauser, G., M. Trick. 1998. Scheduling a major college basketball conference. *Oper. Res.* 46(1) 1–8.
- Smith, J. 2009. Organization of the NCAA baseball tournament. *IMA J. Management Math.* 20(2) 213–232.
- Smith, J., B. Fraticelli, C. Rainwater. 2006. A bracket assignment problem for the NCAA men's basketball tournament. *Internat. Trans. Oper. Res.* 13(3) 253–271.
- Trick, M. 2001. A schedule-then-break approach to sports timetabling. E. Burke, W. Erben, eds. *Proc. 3rd Internat. Conf. Practice Theory Automated Timetabling (PATAT 2000), Lecture Notes in Computer Science*, Vol. 2079. Springer, Heidelberg, Germany, 242–253.
- Urban, T., R. Russell. 2003. Scheduling sports competitions on multiple venues. *Eur. J. Oper. Res.* 148(2) 302–311.

Joel Smith, Commissioner, Rocky Mountain Athletic Conference (RMAC), writes: "The Rocky Mountain Athletic Conference (RMAC) provides schedules for many of its team sports. This annual exercise is complex and often provides what some would consider suboptimal schedules which have not been screened for travel distances or home/away balance.

"The Colorado School of Mines (CSM) student group consisting of Marjorie Cone Saur, Kaleigh Starr, and Mark Husted, along with their faculty adviser, A. Newman, has recently provided some analysis of our conference softball schedule. In doing so, the CSM research group has considered all schools with equal objectivity, systematically scheduling games according to pre-determined sets of rules that eliminate both bias and imbalance. The output from the CSM research group has been well received by our membership.

"They proposed three alternatives: a 36-game schedule, a 40-game schedule and a 44-game schedule, which satisfy the basic schedule structures that schools are contemplating. The RMAC accepted the 40-game schedule at its June meeting.

"The CSM research group has provided the RMAC with optimized schedules that should enhance the student athlete experience and provide a fair and even game distribution for our members."