ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Computers & Operations Research ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮

computers & operations research

www.elsevier.com/locate/dsw

Assignment of swimmers to dual meet events

Maciek Nowak[a,*], Marina Epelman[b], Stephen M. Pollock[b]

[a] *Department of Management, Marketing and Logistics, Georgia Southern University, P.O. Box 8154, Statesboro, GA 30460, USA*

[b] *Industrial and Operations Engineering, University of Michigan, 1205 Beal Avenue, Ann Arbor, MI 48109, USA*

Abstract

Every fall, thousands of high school swimming coaches across the country begin the arduous process of preparing their athletes for competition. With a grueling practice schedule and a dedicated group of athletes, a coach can hone the squad into a cohesive unit, poised for any competition. However, oftentimes all preparation is in vain, as coaches assign swimmers to events with a lineup that is far from optimal. This paper provides a model that may help a high school (or other level) swim team coach make these assignments. Following state and national guidelines for swim meets, we describe a binary integer model that determines an overall assignment that maximizes the total number of points scored by the squad based on the times for swimmers on the squad and for the expected opponent.

Keywords: Programming; Integer; Industries; Recreation/sports; Swimming

1. Introduction

Every fall, thousands of high school swimming coaches across the country begin the arduous process of preparing their athletes for competition. With a grueling practice schedule and a dedicated group of athletes, a coach can hone the squad into a cohesive unit, poised for any competition. However, oftentimes all preparation is in vain, as coaches assign swimmers to events with a lineup that is far from optimal. Prior to each meet, a coach must decide which athletes will compete in which events. In a sport where

* Corresponding author. Fax: +1 912 871 1523.

E-mail addresses: mnowak@georgiasouthern.edu (M. Nowak), mepelman@umich.edu (M. Epelman), pollock@umich.edu (S.M. Pollock).

0305-0548/$ - see front matter
doi:10.1016/j.cor.2004.09.028

every point counts, these decisions are extremely important. Making one poor assignment may cost the team a victory. Coach Denny Hill, winner of 20 state championships at Ann Arbor (Michigan) Pioneer High School and former national swim coach of the year, says, “a lot of times we wonder why a coach used the swimmers that they did.”

Each meet consists of multiple events in various disciplines, such as freestyle and backstroke, with points awarded in each event based on placement. Swimmers can compete in individual or relay events. However, each swimmer is restricted in the number of events he/she can race, due to both meet rules and physical limitations.

Determining which athletes to assign to which events is a difficult task, often taking years to master. Analyzing the individual performances of a squad of 60 swimmers in order to determine which two or three should compete in one event, while keeping in mind the 10 other events can be next to impossible. Similar to how a veteran basketball coach determines the best matchups for his squad, but on a much larger scale, most older swimming coaches make their lineups based on a gut instinct that can only be developed with years of experience. They know which swimmers have the best opportunity to excel in certain events, and how many events a swimmer can compete in without a reduction in performance due to exhaustion. They can do this because they have the knowledge that comes from experiencing countless scenarios in competition.

Coach Hill believes that the main cause of poor assignments is that many young coaches have neither the time nor the experience to create a competitive lineup. Moreover, in practice the complexity of this assignment task overwhelms the possible “game theoretic” nature of the underlying problem. In particular, even if a coach is fully aware of the opponent’s assignment of swimmers and their performance capabilities, the challenge of creating a good (much less an “optimal”) assignment without the help of a formal model is daunting.

This paper provides a model that may help a high school (or other level) swim team coach to assign swimmers to events at a meet. Although there may be many different objectives (maximizing the number of points won, maximizing the probability that the team wins, providing an opportunity for swimmers to qualify for later meets, etc.), this model only addresses the goal of maximizing the number of points won by a team. Because determining the optimal placement of swimmers in events is a constrained assignment problem, a binary integer program is used for the model formulation [1].

Operations research has been used extensively in sports for quite some time, including early statistical analysis of baseball by Lindsey [2,3] and Freeze [4]. More specifically, integer programming has found a use in sports, primarily as a tool in scheduling games over a season. Examples of this include Nemhauser and Trick [5] in scheduling college basketball; Bean and Birge [6] in scheduling the National Basketball Association; Ferland and Fleurent [7] in scheduling the National Hockey League; and Cain [8] for major league baseball. More recently, integer programming has been used by Adler et al. [9] to determine when a major league baseball team has been eliminated from playoff contention.

Despite this emphasis on the “major” sports, such as basketball, baseball and hockey, the potential impact of optimization on “minor” sports can be significant. There are currently over 228,000 athletes registered with USA Swimming, along with over 9200 coaches. Each of these coaches, whether a 20 year veteran or someone new to the job, could use help in creating a lineup that provides his/her team with the best chance of winning. In the year 2000 alone, over 2000 new coaches were registered by USA Swimming. As the sport of swimming continues to expand and the coaching ranks grow, there is an increasing need for a tool that can be used both as an aid to a decision maker, and as a time saver.

2. Guidelines

The size of a *squad* (the group of athletes eligible to compete) may vary widely among teams. We assume (without loss of generality) that the squad has enough swimmers to fill all the events. The *roster* (the assignment of swimmers to events) is generally fixed at the beginning of the meet (although in practice swimmers may be re-assigned just prior to an event). Therefore, we assume that the allocation of swimmers is completed prior to the first event and no changes are made during the meet. Although swim teams can compete in a variety of meets, we focus on a dual meet, in which two teams compete.

The assignment of swimmers is limited by constraints currently in effect in many high school competitions:

- each team can enter at most three swimmers in any one individual (non relay) event. Thus each individual event usually has six swimmers (three from each team), unless a team fails to (or chooses not to) enter three;
- each team can provide at most three “entries” in each relay event, where an entry is agroup of four swimmers;
- each swimmer can enter at most four events;
- each swimmer can enter at most two individual events;
- in a relay event, a team cannot be awarded points for more than two finishing places.

Coaches make decisions about which events their swimmers should enter using information such as their event times, their capabilities of performing a particular sequence of events, their performance in competition compared to training, etc. The coaches also have some limited information about the opposing team’s swimmers. We assume that all assignments are made based on an estimated time for each swimmer on “our” squad (hereafter referred to as “the squad”) for each event. These estimates can be made using information from earlier meets, training sessions or previous seasons. They may also be adjusted depending on how a coach believes a swimmer will compete under certain conditions, such as swimming two consecutive events.

In reality, of course, the actual event times are random variables and any information about past performance can only lead to estimates of parameters (such as means and variances) of their associated probability distributions. We ignore this randomness in the analysis that follows for two reasons. Firstly, the model itself allows for reasonable sensitivity analysis with respect to variability in times. Secondly, although a formulation that accounts for randomness would involve a fairly straightforward extension of our model, the effort needed to solve the resulting problem might not be balanced by any improvement in the performance of the resulting rosters.

Opponent’s times are also estimated using previously observed times from past meets, which are generally available to a coach. An opponent’s roster is then estimated using these times, and is assumed to be set prior to determining the squad assignments. This assumption, of course, neglects the game-theoretical aspects of the problem, but allows for a first cut at a roster, and (according to experienced coaches) provides an extremely useful formulation and solution. Although it would be prefer-able to be able to report on the full game-theoretical analysis, such an analysis is beyond the scope of this paper.

3. Model formulation

Our model compares the times for swimmers on the squad to those of the opponent, and determines an overall assignment that maximizes the total number of points.

3.1. Definitions

- $A = \{1, 2, \ldots, E\}$ is the set of all events. Typically $E = 11$ in a high school dual meet (diving is generally a 12th event, but is not included in this model).
- I is the set of individual events, and R is the set of relay events, where $I \cap R = \emptyset$ and $I \cup R = A$.
- t_{ij} is the estimated time for swimmer i in event j, $j \in A$.
- $a(1)_j$, $a(2)_j$ and $a(3)_j$ are the best, second best and third best times, respectively, of the opposing squad in event j. Because of increased accuracy in timing, it is reasonable to assume that there are no ties in a race, so that no two times are equal and $a(1)_j < a(2)_j < a(3)_j$ for all j.
- x_{ij}, y_{ij}, and z_{ij} are assignment variables that indicate whether swimmer i on the squad competes in event j and, if so, has the best, the second best, or the third best time on the squad, respectively, so that

$$x_{ij} = \begin{cases} 1 & \text{if } i \text{ is assigned to event } j \text{ and has the best time on the squad in event } j, \\ 0 & \text{otherwise,} \end{cases}$$

$$y_{ij} = \begin{cases} 1 & \text{if } i \text{ is assigned to event } j \text{ and has the second best time on the squad in event } j, \\ 0 & \text{otherwise,} \end{cases}$$

$$z_{ij} = \begin{cases} 1 & \text{if } i \text{ is assigned to event } j \text{ and has the third best time on the squad in event } j, \\ 0 & \text{otherwise.} \end{cases}$$

  If j is a relay event, $x_{ij} = 1$ if swimmer i is assigned to the team of four swimmers that has the best time among the squad's relay teams in event j, etc.
- $r(1)_j$, $r(2)_j$, and $r(3)_j$ are the best, second best and third best (i.e., last) times, respectively, realized by the squad in event j. Since four swimmers are assigned to each relay event, the sum of their individual times is the realized time for the event.

3.2. Natural constraints and relations among the variables

Since each event must have three entries from a squad,

$$\sum_i x_{ij} = 1 \quad \text{for all } j \in I, \qquad \sum_i x_{ij} = 4 \quad \text{for all } j \in R,$$

$$\sum_i y_{ij} = 1 \quad \text{for all } j \in I, \qquad \sum_i y_{ij} = 4 \quad \text{for all } j \in R,$$
$$\sum_i z_{ij} = 1 \quad \text{for all } j \in I, \qquad \sum_i z_{ij} = 4 \quad \text{for all } j \in R. \tag{1}$$

Each swimmer can enter at most four events, with at most two of those being individual events. Therefore, the following constraints apply:

$$\sum_{j \in I} (x_{ij} + y_{ij} + z_{ij}) \leqslant 2 \quad \text{for all } i, \tag{2}$$

$$\sum_{j \in A} (x_{ij} + y_{ij} + z_{ij}) \leqslant 4 \quad \text{for all } i.$$

Finally, a swimmer can only place once in an event, leading to

$$x_{ij} + y_{ij} + z_{ij} \leqslant 1 \quad \text{for all } i, j. \tag{3}$$

The realized times in an event are a function of which swimmers swim in the event, and their estimated times. This is expressed by the relations

$$r(1)_j = \sum_i t_{ij} x_{ij} \quad \text{for all } j \in A,$$
$$r(2)_j = \sum_i t_{ij} y_{ij} \quad \text{for all } j \in A,$$
$$r(3)_j = \sum_i t_{ij} z_{ij} \quad \text{for all } j \in A. \tag{4}$$

In order to force consistency in the placement order, we use constraints

$$r(1)_j + \varepsilon \leqslant r(2)_j \quad \text{for all } j \in A,$$
$$r(2)_j + \varepsilon \leqslant r(3)_j \quad \text{for all } j \in A, \tag{5}$$

where ε is a very small positive number.

The realized times in each event are compared to the opponent's times to determine the outcome of the various events. For example, if

$$r(1)_j < r(2)_j < a(1)_j < a(2)_j < a(3)_j < r(3)_j, \tag{6}$$

then the squad would be awarded points for first, second and sixth place in event j.

The goal of the squad is to maximize the total points won during the meet. To address this objective, we define the indicator variables w_{jlmn} for $1 \leqslant l < m < n \leqslant 6$, that specify the outcome of event j, where

$$w_{jlmn} = \begin{cases} 1 & \text{if, in event } j, \text{ the squad receives places } l, m \text{ and } n, \\ 0 & \text{otherwise.} \end{cases}$$

For an individual event j, $w_{jlmn} = 1$ if and only if the squad's best swimmer receives place l, the second best swimmer receives place m, and the third best swimmer receives place n; interpretation for relay events is analogous.

Table 1
Points for individual and relay events

Place	1	2	3	4	5	6
Points (individual event)	6	4	3	2	1	0
Points (relay event)	8	4	2	0	0	0

This outcome indicator can be used to construct constraints on the realized times $r(\cdot)_j$. For the example given in (6) above, in which the squad places first, second and sixth in event j, $w_{j126}=1$, and the condition $r(2)_j < a(1)_j$ (the squad's second best swimmer is faster than the opponents best swimmer) must hold. This can be enforced by using the constraint

$$r(2)_j + \varepsilon \leqslant a(1)_j w_{j126} + (1 - w_{j126})M, \tag{7}$$

where M is a very large number. Thus, if $w_{j126} = 1$, then $r(2)_j < a(1)_j$ as desired. However, if the outcome does not occur, so that $w_{j126} = 0$, then $r(2)_j$ is essentially unconstrained (it is bounded above by a very large number). Similarly, for the example in (6) to hold, $a(3)_j < r(3)_j$ must be true for the squad's third best swimmer to place sixth. The constraint

$$r(3)_j \geqslant a(3)_j w_{j126} + \varepsilon \tag{8}$$

ensures this, since $r(3)_j$ is, essentially, either bounded from below by ε or by $a(3)_j$, depending on the outcome.

For each event j, every feasible combination of l, m and n has constraints similar to those in (7) and/or (8) for each of the three swimmers in the event.

Finally, the constraint

$$\sum_{1 \leqslant l < m < n \leqslant 6} w_{jlmn} = 1 \quad \text{for all } j \in A \tag{9}$$

guarantees that each event will have exactly one outcome.

3.3. Objective function

The reward for each outcome is determined by adding the points for each place for each event. The point structure used for this model is the one used in Michigan High School competition, as shown in Table 1.

Let g_{jlmn} be the reward for receiving the lth, mth and nth places in event j. For example, $g_{j126} = 10$ if j is an individual event and $g_{j126} = 12$ for a relay event. This parameter incorporates the fact that a team cannot be awarded points for more than two finishing places in a relay event. For example, $g_{j123} = 12$ for a relay event, since the points for the third place finish are not included. The squad's total score in the meet is

$$T = \sum_{j,l,m,n} g_{jlmn} w_{jlmn}. \tag{10}$$

The problem then is to maximize T with respect to the assignment variables x_{ij}, y_{ij}, z_{ij} and the variables $r(1)_j$, $r(2)_j$, $r(3)_j$, and w_{jlmn}.

3.4. Integer programming formulation

For convenience in notation, we define (for all j) $a(s)_j = 0$ if $s \leqslant 0$ and $a(s)_j = M$ (an arbitrarily large number) if $s \geqslant 4$. The integer program formulation of the rostering problem can be now stated as

$$\max_{x_{ij}, y_{ij}, z_{ij}, r(1)_j, r(2)_j, r(3)_j, w_{jlmn}} \sum_{j,l,m,n} g_{jlmn} w_{jlmn} \tag{11}$$

subject to

$$\sum_i t_{ij} x_{ij} = r(1)_j, \; \sum_i t_{ij} y_{ij} = r(2)_j, \; \sum_i t_{ij} z_{ij} = r(3)_j \quad \text{for all } j \in A,$$

$$r(1)_j + \varepsilon \leqslant r(2)_j, r(2)_j + \varepsilon \leqslant r(3)_j \quad \text{for all } j \in A,$$

$$\sum_{1 \leqslant l < m < n \leqslant 6} w_{jlmn} = 1 \quad \text{for all } j \in A,$$

$$\sum_i x_{ij} = 1, \; \sum_i y_{ij} = 1, \; \sum_i z_{ij} = 1 \quad \text{for all } j \in I,$$

$$\sum_i x_{ij} = 4, \; \sum_i y_{ij} = 4, \; \sum_i z_{ij} = 4 \quad \text{for all } j \in R,$$

$$\sum_{j \in I} x_{ij} + y_{ij} + z_{ij} \leqslant 2 \quad \text{for all } i,$$

$$\sum_{j \in A} x_{ij} + y_{ij} + z_{ij} \leqslant 4 \quad \text{for all } i,$$

$$x_{ij} + y_{ij} + z_{ij} \leqslant 1 \quad \text{for all } i, j,$$

$$\varepsilon + r_j(1) \leqslant a(l)_j w_{jlmn} + M(1 - w_{jlmn}) \quad \text{for all } j \in A \text{ and } 1 \leqslant l < m < n \leqslant 6,$$

$$\varepsilon + a(l-1)_j w_{jlmn} \leqslant r(1)_j \quad \text{for all } j \in A \text{ and } 1 \leqslant l < m < n \leqslant 6,$$

$$\varepsilon + r_j(2) \leqslant a(m-1)_j w_{jlmn} + M(1 - w_{jlmn}) \quad \text{for all } j \in A \text{ and } 1 \leqslant l < m < n \leqslant 6,$$

$$\varepsilon + a(m-2)_j w_{jlmn} \leqslant r(2)_j \quad \text{for all } j \in A \text{ and } 1 \leqslant l < m < n \leqslant 6,$$

$$\varepsilon + r_j(3) \leqslant a(n-2)_j w_{jlmn} + M(1 - w_{jlmn}) \quad \text{for all } j \in A \text{ and } 1 \leqslant l < m < n \leqslant 6,$$

$$\varepsilon + a(n-3)_j w_{jlmn} \leqslant r(3) \quad \text{for all } j \in A \text{ and } 1 \leqslant l < m < n \leqslant 6,$$

$$x_{ij}, y_{ij}, z_{ij} \in \{0, 1\} \quad \text{for all } i, j,$$

$$w_{jlmn} \in \{0, 1\} \quad \text{for all } j \in A \text{ and } 1 \leqslant l < m < n \leqslant 6.$$

4. Results and analysis

The model was tested using data obtained from Westminster Academy (High School) in Atlanta, GA. The men's squad has 17 swimmers who are eligible to compete in 11 events. An extension was made to the formulation in order to allow for a medley relay event, such that the times for the swimmers are stroke specific. The resulting IP for a problem of this size has 781 variables and 1530 constraints, requiring 3.5 s to solve on a Sun SPARC station using CPLEX version 8.1.0 [10].

The data we used for our analysis comes from a dual meet in which the Westminster men's team lost 97–73. The predicted times used for the squad are found in Table 2, and the predicted opponent's times in Table 3. The medley relay times are split into each stroke (free, back, breast, and butterfly) for the squad. The 200 and 400 freestyle relay times (200 FR and 400 FR) are the split times for each swimmer, while the time for the opponent is the cumulative total. These predictions were made based on performance in previous events. There is some variation in times between swimmers for most events. This variation is common, particularly at the high school level, and it should give a coach some room for error in making predictions. The sensitivity of the code to these errors is further discussed below.

Using our model, we created an assignment of swimmers that would have scored 89 out of 170 possible points, more than the 86 points required for a victory. The assignments, places and points for the 11 events determined using the model are shown in Table 4. The actual results from the meet are in Table 5. Note that due to the limited roster size, there were only two entries in each of the relay events in the actual meet. This does not have an effect on the results, as only the top two relay teams on each squad score points. The first three events listed are relays and the assignment numbers indicate the swimmer number.

The two tables show the impact that a few minor changes can have on the outcome of a meet. For example, if swimmer 2 were entered in event 7 instead of event 5, he would most likely have won the event and contributed an additional four points to the team total. However, most of the improvements result in minor changes in points. For example, shifting swimmer 1 from event 4 to event 5 would have allowed the team to gain two points, as they could have finished the event in third and fourth, instead of third and sixth place. This indicates that the lineup is, in general, effective, which is not surprising considering the veteran coach at Westminster. Some of the differences may be attributed to the coach hoping for a swimmer to qualify for the state meet in an event that he does not generally swim, or giving a swimmer a rest in an event. Yet, the overall improvements that can be made to a good lineup are an indicator of the potential for improvement to a bad lineup.

In order to analyze the model's sensitivity to values of the times, we calculated how the squad would perform against an opponent that is faster than expected. Keeping the original optimized squad assignments as shown in Table 4, all opponent's times were lowered by a fixed percentage and the points that the squad would have received against the "faster" opponent were calculated. If the opponent's times were 1% lower than anticipated, the roster would have only scored 80 points, losing the meet, yet still scoring seven more points than the actual result. If the opponent's times were 2% lower, the roster would have scored 75 points, still better than the actual meet result!

Similarly, we analyzed how the squad would perform if its swimmers were less competitive than predicted. The opponent's times were held constant while the squad's times were increased. An increase of 1% resulted in a score of 80 points by the lineup in Table 4, the same score as when the opponent's times were lowered by 1%. The team would have lost the meet, but the score would have been much closer than the actual outcome. Increasing the squad's times by 2% also produced the same result as when the opponent's times were decreased by 2%, with a score of 75 points. It should be noted, however,

Table 2
Predicted times for squad swimmers

Swimmer	200 FR	400 FR	200 Free	200 IM	50 Free	100 Fly	100 Free	500 Free	100 BA	100 BR	Medley Relay—Stroke Splits			
1	27.5	56.65	120.24	134.12	27.5	62.14	56.65	326.28	0	0	27.5	31.53	38.54	29.82
2	23.62	56.49	0	119.85	23.62	56.41	0	0	56.36	0	23.62	26.79	33.62	27
3	29.74	72.34	154.28	168.5	29.74	87.54	72.34	411.82	0	91.23	29.74	0	0	0
4	25.86	0	120.22	0	25.86	60.61	0	0	68.27	0	25.86	0	0	27.81
5	29.77	0	156.69	0	29.77	90.88	0	0	85.3	89.69	29.77	0	0	0
6	26.39	0	115.27	123.98	26.39	0	53.64	298.02	0	0	26.39	31.3	37.65	29.82
7	29.25	72.35	169.35	0	29.25	0	72.35	0	103	87.3	29.25	0	0	0
8	23.34	51.12	105.44	0	23.34	0	51.12	281.59	0	0	23.34	29.15	33.98	29.97
9	27.38	61.27	0	145.35	27.38	72.16	61.27	0	0	75.05	27.38	34.59	36.59	30.36
10	24.51	56.24	0	153.48	24.51	74.11	55.45	0	0	0	24.51	34.81	37.78	32.18
11	26.65	57.48	130.33	0	26.65	0	57.48	0	69.28	0	26.65	30.6	43.61	35.23
12	31.71	72.58	149.6	177.18	31.71	0	73.02	0	82.3	0	31.71	37.11	0	0
13	24.76	54.84	117.21	0	24.76	0	54.84	348.18	0	0	24.76	35.13	37.2	30.81
14	25.51	0	0	0	25.51	65.95	0	0	0	68.43	25.51	0	0	0
15	25.12	0	0	0	25.12	0	0	0	0	67.62	25.12	0	0	0
16	28.47	59.96	123.15	0	28.47	0	59.96	339.81	0	0	28.47	39.39	44.51	37.35
17	24.07	52.3	0	136.8	24.07	57.11	51.95	0	57.7	0	24.07	27.01	32.96	26.36

Zero entries indicate that the swimmer does not participate in this event.

Table 3
Predicted times for opponent

Swimmer	200 FR	400 FR	200 Free	200 IM	50 Free	100 Fly	100 Free	500 Free	100 BA	100 BR	200 MR
1	95.6	210.3	108.26	131.51	22.3	59.3	50.44	307.88	53.51	70.21	104.86
2	101.22	243.36	125.55	145.57	23.24	61.31	54.19	372.08	58.82	70.84	109.61
3	103.6	266.15	127.36	146.46	25.04	68.78	54.85	600	66.04	71.93	115.74

Table 4
Optimized results for Westminster men's squad

Event	Assignment			Places	Points
	1	2	3		
1	{2,8,10,17}	{4,13,14,15}	{1,5,6,9}	1,4,6	8
2	{1,12,13,17}	{3,8,10,11}	{7,14,15,16}	2,3,6	6
3	{2,4,9,10}	{3,11,12,16}	{5,6,7,13}	3,5,6	2
4	8	4	16	1,3,4	11
5	6	1	9	1,3,4	11
6	10	13	14	3,4,6	5
7	2	17	4	1,2,4	12
8	6	13	11	2,4,6	6
9	8	1	16	1,3,4	11
10	2	17	11	2,3,6	7
11	15	14	9	1,2,6	10

Table 5
Actual results for Westminster men's squad

Event	Assignment			Places	Points
	1	2	3		
1	{6,8,10,17}	{9,14,15,16}	No entry	1,4	8
2	{1,6,8,17}	{3,7,11,16}	No entry	2,4	4
3	{1,10,11,14}	{2,9,12,15}	No entry	3,6	2
4	8	1	12	1,3,6	9
5	6	9	2	1,3,6	9
6	10	15	14	3,5,6	4
7	17	10	3	1,5,6	7
8	8	11	7	2,5,6	5
9	6	1	16	1,3,4	11
10	17	11	13	3,5,6	4
11	15	14	9	1,2,6	10

that even a 1% or 2% difference from estimated times might not be too common. Competitive swimmers are consistent enough so that even the most inexperienced coaches can predict how well their team will perform within an overall margin of error around 1%.

If a coach believes that the opponent may actually swim faster than predicted, the opponent's times can be lowered and a new roster can be determined using the model. Lowering the opponent's times by 1% and re-rostering resulted in a roster that would have scored 85 points (a tie) against the "faster" opponent, and the same 85 points if the opponent's times were not faster, but actually those in Table 3. Similarly, if the opponent's times were lowered by 2%, re-rostering would result in a score of 83 points against the "faster" opponent, as well as against the times in Table 3. The scores 85 and 83 are better than those found pitting the original roster against the faster opponent. However, they result in a loss or tie even if the opponent does not swim faster than predicted. This may raise an interesting dilemma for a coach: does one use a roster that will win if the opponent swims as predicted, or one that will perform well if the opponent swims faster than predicted, but will not necessarily win, even if the prediction is correct? This question cannot be answered for the coach; however, the model allows the coach to choose from a variety of strong options representing numerous potential scenarios.

5. Conclusions

Our model finds an optimal assignment of swimmers to events. It can be useful as a decision aid for novice coaches overwhelmed by the prospect of creating a full lineup, or as a timesaver for more experienced coaches who would like to spend more time coaching rather than mulling over various possible swimmer lineups. Determining the top swimmer in an event is generally trivial; it is the second and third positions that are most difficult to fill, particularly in relay events. While placing first in an event may be a psychological boost for the team, a meet is often won by the lower tier swimmers. As Pete Higgins, national swim coach of the year at Westminster Academy in Atlanta, GA attests, "we can get first place in an event, but we're not going to be winning anything if our other guys get fifth and sixth." The model can take a lot of the guesswork out of deciding which swimmers to place in the crucial lower positions.

Straightforward extensions to the model might include:

- allowing for performance degradation if a swimmer competes in two consecutive events. This may be accounted for by increasing the time in the second event by some percentage.
- incorporating a prohibition (as some coaches prefer) against swimming in two consecutive events,
- allowing (as is the case in some states) four entries per team in an event.

The model provides a tool for the future analysis of strategies that may be used as a rule of thumb when creating lineups, such as what to do with a swimmer who is the best on the team in the majority of events. Finally, it may be used to assign athletes to events in track and field competitions, which have similar constraints and objectives.

References

[1] Nemhauser GL, Wolsey LA. Integer and combinatorial optimization. New York: Wiley; 1988.
[2] Lindsey GR. Statistical data useful for the operation of a baseball team. Operations Research 1959;7:197–207.
[3] Lindsey GR. An investigation of strategies in baseball. Operations Research 1963;11:477–501.
[4] Freeze RA. An analysis of baseball batting order by monte carlo simulation. Operations Research 1974;22:728–35.
[5] Nemhauser GL, Trick MA. Scheduling a major college basketball conference. Operations Research 1998;46:1–8.

[6] Bean JC, Birge JR. Reducing traveling costs and player fatigue in the national basketball association. Interfaces 1980;10: 98–102.
[7] Ferland JA, Fleurent C. Allocating games for the national hockey league using integer programming. Operations Research 1993;41:649–54.
[8] Cain Jr WO. Optimal strategies in sports. Chapter A computer assisted heuristic approach used to schedule the major league baseball clubs. Amsterdam: North-Holland; 1977, pp. 32–41.
[9] Adler I, Erera A, Hochbaum DS, Olinick EV. Baseball, optimization and the world wide web. Interfaces 2002;32:12–22.
[10] ILOG, Inc. CPLEX 8.0 Reference manual. Paris, France, 2002.