South African Journal for Research in Sport, Physical Education and Recreation, 2006, 28(2): 81-90.
Suid-Afrikaanse Tydskrif vir Navorsing in Sport, Liggaamlike Opvoedkunde en Ontspanning, 2006, 28(2): 81-90.
ISSN: 0379-9069

SELECTING A LIMITED OVERS CRICKET SQUAD USING AN INTEGER PROGRAMMING MODEL

Hannah GERBER* & Gary D. SHARP**
** School of Statistics and Actuarial Science, University of the Witwatersrand, Johannesburg, Republic of South Africa*
***Department of Mathematical Statistics, Nelson Mandela Metropolitan University, Port Elizabeth, Republic of South Africa*

ABSTRACT

An integer programming model was developed to select an one day international (ODI) cricket squad of 15 players. To develop the method, batting, bowling and fielding ability, which are measured differently, had to be placed onto the same scale, so an ability-indexing technique was used. This paper describes the ability-indexing and integer programme used and discusses the results of an empirical study conducted using the statistics of 32 South African cricket players.

Key words: Cricket; Squad selection; Integer programme; Optimization.

INTRODUCTION

The game of cricket is no longer just a recreational pastime, the advent of the professional game has resulted in it becoming a career. Young, promising players commit many hours of their time to develop and hone the necessary skills to compete at the highest and thus most lucrative level. Cricketers spend considerable time and effort to develop their fitness, their mental ability (Slogrove *et al*., 2002) and their tactical strategic choices (Preston & Thomas, 2000), all skills necessary of a professional player.

Whether or not professional cricket players make a success of their chosen profession is all to often dependent on factors beyond their control. Considerations such as leadership roles within administration bodies can impact on the professionalism of the sport (Ristow *et al*., 1999), and hence on the success of a cricket player. Similarly, team selection is beyond a player’s control, selection is a task assigned to a group of convenors who on occasion have to select teams based on conditions other than merit (Press Trust of India, 2005). This is an arduous and extremely difficult task, fraught with tangible subjectivity and open to criticism from media (Emslie, 2005), spectators and players alike. Preston and Thomas (2002) provide a simple description of limited overs cricket and those less familiar with the game are referred to their paper.

This study proposes an integer programming model, which can be used for team selection. The model removes subjectivity from the decision making process and reduces the likelihood of a player’s failure due to influences beyond their control. The model was developed to select a squad of 15 players rather than an 11-person team, as this has become the norm for many international touring cricket sides.

METHODOLOGY

Background

A cricket team consists of 11 players with players specializing in batting, bowling, fielding, wicket keeping or a combination of these disciplines (all-rounders). A batsman is considered to be a valuable resource if the average number of runs scored per game, hereafter referred to as batting average, is relatively high. Similarly a bowler is considered to be a valuable resource if the average number of runs conceded per over, hereafter referred to as economy rate, is relatively low. In this study, fielding is considered as a specialist position and a fielder is considered as a valuable resource if the average number of dismissals per game, defined as dismissal rate, is relatively high. These statistics are the primary source of player ability.

There are alternative methods for determining the ability of a cricket player. Barr and Van den Honert (1998) propose that batsmen's ability be measured by an adjusted batting average, which includes a measure of consistency. Lemmer (2002, 2004) provides a ranking technique for batting and bowling ability. However this study consigned itself to the afore-mentioned batting average, economy rate and dismissal rate. Furthermore the frequently used statistic, bowler's strike rate, defined as the number of balls bowled per wicket taken is ignored, as it was felt that this statistic is more applicable to the longer version of the game.

Requirements of the Integer Programming Model

To model team selection using an integer programme (IP) requires that all abilities, batting, bowling and fielding be measured on the same scale. To clarify this issue, consider the batting average and economy rate of players. The batting averages of cricketers (usually in the 30's or 40's), are measured as runs per innings (excluding not outs) whilst bowlers economy rates (preferably below 4.5) are measured as runs per over. These measurements are scaled differently and when optimizing an IP model, the objective function requires that the measurements are valued on the same scale. Considering these two abilities, the IP model would preferentially select the larger numerical values (assuming a maximization problem), over the smaller ones, thus batsmen would be chosen in preference to bowlers.

To address this, the IP model was developed in two stages. Initially ability-indexing techniques were used to calculate coefficients for cricketing abilities (batting, fielding or bowling) so that they would be comparable and thereafter the objective function and the constraints were developed.

Ability coefficients

The proposed indexing method requires that the coefficients used in the IP objective function only be calculated for those players that are considered as specialists in their ability, i.e. batting, bowling, fielding, keeping wicket or as all-rounders. The following indexing method was used to calculate the specialist ability coefficients:

Batting ability

$$Bat\ index\ for\ player\ i = \left(\frac{Batting\ average\ of\ batsman}{Sum\ of\ all\ the\ batting\ averages\ of\ all\ the\ batsmen} \right) \times number\ of\ specialist\ batsmen$$

This equation relates the batting average of one batsman to the mean batting average of all batsmen. The resulting index coefficient will be comparable with the other cricketing abilities.

Bowling ability

Integer optimization requires that the objective function either be maximized or minimized (Taha, 2003). To setup the objective function in a maximization form requires that all ability indices be comparable. In cricket an increasing batting average is fine for a maximization problem, whilst a decreasing economy rate is fine for a minimization problem. To model these abilities in a single maximization function requires that we multiply the bowling ability by –1. Thus we consider the negative of the bowling ability as shown below. Furthermore, since these coefficients are negative (economy rate has to be minimized), a constant, which is sufficiently large, is added to ensure all coefficients are positive.

$$\text{Let} \quad v_i = \left[k - \left(\frac{\text{Economy rate of specialist bowler } i}{\text{Sum of all the economy rates of all the specialist bowlers}} \right) \right]$$

The constant k, was chosen as the smallest positive integer such that $\left[k - \left(\frac{\text{Economy rate of specialist bowler } i}{\text{Sum of all the economy rates of all the specialist bowlers}} \right) \right] > 0$

Then the same normalizing strategy that was used for batsmen is used to determine the bowling index for individual bowlers.

$$\text{Bowl index for player } i = \left(\frac{v_i}{\text{sum of all } v_i} \right) \times \text{number of specialist bowlers}$$

Fielding ability

$$\text{Field index for player } i = \left(\frac{\text{Dismissal rate of fielder}}{\text{Sum of all the dismissal rates of all fielders}} \right) \times \text{number of specialist fielders}$$

This equation relates the dismissal rate of one fielder to the mean dismissal rate of all fielders. The resulting index coefficient will be comparable with the other cricketing abilities.

All-rounder ability

In developing the general model we have chosen to classify four classes of all-rounders. The first category is for players who can both bat and bowl. Category two includes players who can both bat and field, category three includes players who can both bowl and field and category four includes players who can bat, bowl and field. This is a generalization approach and some selection panels may classify all-rounders differently.

$$\text{All} - \text{round index for player } i = \left(\frac{\text{sum of individual players index values}}{\text{sum of that category index values}} \right) \times \text{number of players in all} - \text{round category}$$

For each of the four categories under consideration, the index values of individual abilities where summed and then normalized as per the method used for batting indices.

Keeping wicket ability

$$Wicket\ keep\ index\ for\ player\ i = \left(\frac{Dismissal\ rate\ of\ keeper}{Sum\ of\ all\ the\ dismissal\ rates\ of\ all\ keepers} \right) \times number\ of\ specialist\ keepers$$

This equation relates the dismissal rate of one wicketkeeper to the mean dismissal rate of all wicketkeepers. The resulting index coefficient will be comparable with the other cricketing abilities.

Model definition

To standardize, the IP model was treated as one of maximization. Convenors may have different interpretations of the necessary requirements for a squad. To allow for this flexibility the model was set up to consider *p* possible abilities, as illustrated in Table 1.

TABLE 1: **TWO-WAY REPRESENTATION MATRIX OF PLAYERS AND ABILITIES**

		Abilities							
		Bat	Bowl	Field	All Round 1	All Round 2			p^{th} ability
		1	2	3	4	5			*p*
Player	1	$a_{1,1}$	$a_{1,2}$	…	…	…	…		$a_{1,p}$
	2	$a_{2,1}$							
	:								
	n	$a_{n,1}$							$a_{n,p}$

The decision variables are defined as

$$x_{ij} = \begin{cases} 1 & if\ \ player\ i\ is\ selected\ \ for\ ability\ \ j \\ 0 & if\ \ player\ i\ is\ not\ selected\ \ for\ ability\ \ j \end{cases}$$

The objective function is stated as maximize, $Z = z_1 + z_2 + z_3 + z_4 + \ldots + z_p$
Where the terms in the expression are

$z_1 = \sum_i a_{i1}x_{i1}$ Maximizing ability index for the batsmen (ability 1).

$z_2 = \sum_i a_{i2}x_{i2}$ Maximizing ability index for the bowlers (ability 2).

$z_3 = \sum_i a_{i3}x_{i3}$ Maximizing ability index for the fielders (ability 3).

$z_4 = \sum_i a_{i4}x_{i4}$ Maximizing ability index for the all-rounders in category 1 (ability 4).

:

$z_p = \sum_i a_{ip}x_{ip}$ Maximizing ability index for the ability *p*.

Constraints

The constraints are dependent on the objectives of the team selectors. If a convenor requires that five all-rounders be selected, then this constraint would be included in the model. The constraints were generalized for the model so that convenors have the flexibility to set their own requirements.

The following three constraints must be included in the model

$\sum_i \sum_j x_{ij} = k$; to ensure exactly *k* players are selected for the squad.

$\sum_j x_{ij} = 1 \ for\ all\ i$; to ensure that a player is only selected once.

$x_{ij} = 0 \ or\ 1$; to satisfy the definition of the decision variable.

The following constraints or variations thereof may be included in the model

$\sum_i x_{i1} \geq y_1$; to ensure sufficient batsmen (y_1) are selected.

$\sum_i x_{i2} \geq y_2$; to ensure sufficient bowlers (y_2) are selected.

$\sum_i x_{i3} \geq y_3$; to ensure sufficient fielders (y_3) are selected.

$\sum_i x_{i4} \geq y_4$; to ensure sufficient category 1 all-rounders (y_4) are selected.

:

$\sum_i x_{ip} \geq y_p$; to ensure sufficient players (y_p) requiring ability *p* are selected.

Additional constraints may be added at the convenors discretion. One such case could be the inclusion of a captain who quite possibly may be unable to make the squad solely on ability but is required for their captaincy and strategic skills. There are several variations to these types of constraints and one such case is shown in the illustrated model.

ILLUSTRATION OF THE MODELLING PROCESS (SOUTH AFRICAN ODI SQUAD)

Data

The ability measures were calculated from player statistics available from the cricinfo website. The statistics economy rate, batting average and number of dismissals per game for 32 South African cricket players were obtained for the period up to and including September 2003. For players with less than 15 international games first class statistics were used whilst statistics obtained as a Protea were used for players with 15 or more international games. The ability indices calculated as discussed above are shown in Table 2. The rows of the table identify the player *i* (*i*=1, 2, …, 32) and the columns of the table identify the ability *j* (*j*=1, 2, …, 9).

TABLE 2: **TWO-WAY REPRESENTATION MATRIX OF SOUTH AFRICAN ODI PLAYERS AND ABILITIES**

		Ability (j)								
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	Players (i)	**Bat**	**Bowl**	**Field**	**Bat/ Bowl**	**Bat/ Field**	**Bowl/ Field**	**B/B/F**	**Wick**	**Bat/W**
1	Abrahams, S.		1.02							
2	Adams, P.R.		0.99							
3	Bacher, A.M.	1.00								
4	Benkenstein, D.M.	0.52								
5	Boje, N.	0.81	0.98		0.94					
6	Boucher, M.V.	0.74							1.35	1.05
7	Dawson, A.C.		1.05							
8	Dippenaar, H.H.	1.23								
9	Donald, A.A.		1.02							
10	Elworthy, S.		0.99							
11	Gibbs, H.H.	1.05		0.95		0.98				
12	Hall, A.J.	0.69	1.00	0.96	0.88	0.81	0.97	0.92		
13	Kallis, J.H.	1.30	0.95	0.98	1.18	1.12	0.96	1.13		
14	Kemp, J.M.	0.78	0.97	0.98	0.92	0.86	0.97	0.95		
15	Kirsten, G.	1.20								
16	Klusener, L.	1.27	0.95		1.16					
17	Langeveldt, C.K.		1.02							
18	McKenzie, N.D.	1.13								
19	Nel, A.		1.05							
20	Ntini, M.		1.01							
21	Ontong, J.L.		1.02	1.17			1.09			
22	Peterson, R.J.		0.98							
23	Pollock, S.M.	0.70	1.06		0.92					
24	Prince, A.G.	0.78								
25	Rudolph, J.A.	1.48		0.79		1.12				
26	Smith, G.C.	1.10		1.17		1.11				
27	Stewart, E.L.R.								0.85	
28	Telemachus, R.		0.95							
29	Van Jaarsveld, M.	1.13								
30	Van Wyk, M.N.	1.09							0.80	0.95
31	Willoughby, C.M.		1.02							
32	Zondeki, M.		0.95							

Modelling a South African ODI squad

IP Model

To illustrate the flexibility of the model ten constraints were used. Three constraints were included as a matter of necessity, six constraints were included as a matter of choice and one constraint was included to illustrate model flexibility. The constraints were chosen for illustration purposes only and do not reflect policy of any organizing cricket associations. Several variations of these constraints are possible.

Objective function

Maximize, $Z = z_1 + z_2 + z_3 + z_4 + \ldots + z_9$

The following constraints were included as a matter of necessity

1. $\sum_i \sum_j x_{ij} = 15$; to ensure exactly 15 players were selected for the squad.
2. $\sum_j x_{ij} = 1 \textit{ for all } i$; to ensure that a player is only selected once.
3. $x_{ij} = 0 \text{ or } 1$; to satisfy the definition of the decision variables.

The following constraints were included as a matter of choice

4. $\sum_i x_{i1} \geq 4$; to ensure that there were at least four specialist batsmen.
5. $\sum_i x_{i2} \geq 4$; to ensure that there were at least four specialist bowlers.
6. $\sum_i x_{i3} \geq 1$; to ensure that there was at least one specialist fielder.
7. $\sum_i x_{i4} \geq 2$; to ensure that there were at least two specialist batting/bowling all-rounders.
8. $\sum_i \sum_{j=4,5,6,7} x_{ij} \geq 3$; to ensure that there were at least three specialist all-rounders (all categories).
9. $\sum_i \sum_{j=8,9} x_{ij} \geq 1$; to ensure that there was at least one specialist wicketkeeper.

The following constraint was included to select a captain

10. $\sum_j x_{26j} = 1$; to ensure that Graeme Smith was selected for the squad.

The choice constraints provide the squad with at least six batsmen (constraints 4 & 7), at least six bowlers (constraints 5 & 7), at least one specialist fielder and at least one wicketkeeper. The inclusion of a captaincy constraint ensures that the person chosen as captain elect was selected. The constraints were motivated by considering that for an 11-person ODI team at least five bowlers, at least four recognized batsmen and a specialist wicketkeeper are needed. Selection convenors may have alternative interpretations of what an ODI team requires and thus may opt for variations of the chosen constraints. For example, a specialist spinner, a specialist opening batsmen and/or a specialist strike bowler constraints may be included in the model.

RESULTS

The IP model had 68 variables and 38 constraints. To determine the solution to the proposed model, *Solver*, a linear optimization "*Excel Add In*" package was used. To confirm the results, both authors independently of each other ran the same programme and obtained the same solution. The illustration model yielded a feasible optimal solution. This is not always the case, often an IP model is over constrained and no feasible solution is possible. Goal programming is an option that could be considered if this occurs.

The squad members selected by the IP model were:

Batsmen

H.H. Dippenaar, G. Kirsten, N.D. McKenzie, J.A. Rudolph, M. van Jaarsveld and M. van Wyk

Bowlers

A. Dawson, C.K. Langeveldt, A. Nel and S.M. Pollock

All-rounders

J.H. Kallis (batting and bowling), L. Klusener (batting and bowling) and G.C. Smith (batting and fielding)

Wicketkeeper

M.V. Boucher

Fielder

J.L. Ontong

DISCUSSION

Interestingly, in this example the captaincy constraint would not have influenced the squad selection, as Graeme Smith would have been selected solely on ability. The inclusion of a fielding constraint was perhaps unnecessary, as convenors may prefer to include an extra specialist bowler or all-rounder, however this constraint was used to illustrate how easy it is to include additional constraints in the optimization model.

Alternative criteria for selecting an ODI squad could be used. Team selection criteria like merit, development and transformation constraints could be included at the convening panels discretion and motivations for the use thereof provided. Squad selections in these cases may differ depending on the constraints included but the optimal squad for the constraints used will still be obtained. As an example, the IP was recalculated after removing constraint 7 (at least two specialist batting/bowling all-rounders). The IP squad selection showed one player change and one ability change. This constraint change meant that the batsman, M. van Wyk was replaced by J.H. Kallis, and the all-rounder J.H. Kallis (bat/bowl) was replaced by H.H. Gibbs (bat/field).

The model does not make selection easy for a player who has performed poorly in the early years of their career. This model relies on long-term form and players who start their career off slowly, as measured by the ability index, may struggle for selection. There are two solutions to this problem; additional constraints can be included in the model or short-term statistics rather than career statistics used to measure ability. Furthermore, if a player hasn't performed well after having played 15 games for the Protea's, it is difficult to be reselected. This is a drawback as newcomers to teams are often used in unconventional roles and their statistics can be adversely affected. Yet again there are two possible solutions to this problem; in exceptional cases additional constraints can be included in the model or the period (15 games) over which the statistics are determined be extended. These are considerations that bear further investigation.

The analysis of the SA ODI model was used for illustration purposes. To select the optimum National squad, selection convenors should include all players competing in the provincial limited overs format of the game. The methodology would remain the same, but a more sophisticated optimization solver, such as *LINDO*, would be required.

CONCLUSIONS AND FURTHER RESEARCH

The model removes the subjectivity of convenors so that ability is used to determine the optimum squad. Ability is measured by the player indices and will need to be updated at regular intervals. The model is relatively simple and alternative measures for determining the ability of a player could have been used. In particular, a bowler's strike rate, measured as the number of balls bowled relative to the number of wickets taken, is considered by some to be a crucial ability. This ability is useful, if, as some believe, that losing wickets reduces a team's ability to set or reach a demanding target. This is a consideration that bears further investigation, particularly in light of the new 20-over format of the game that has generated substantial spectator interest in the last two seasons.

The model was developed for squad selection of players competing in the limited overs format of the game; the same methodology could be used for selecting a Test squad. Different ability indices may be more preferable in this longer format of the game and this is also an area, which bears further investigation. Also, in this format of the game, the choice constraints as decided by the authors may be irrelevant. All-rounder constraints may be of less importance and this constraint can be amended accordingly. Similarly, opening specialist batsmen may be necessary in this scenario. In all these cases, the relevant coach and team selectors can decide the choice constraints.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Catherine G. Logie for the language editing of the manuscript, Dr. Richard A. Stretch for his useful cricketing comments and the referees who provided valuable feedback and suggestions on the original manuscript.

REFERENCES

BARR, G.D.I. & VAN DEN HONERT, R. (1998). Evaluating batsmen's scores in test cricket. *South African Statistical Journal*, 32: 169-183.

EMSLIE, N. (2005). Call for cricket fans to support SA at St George's. *The Herald*, Friday, 18 February.

LEMMER, H.H. (2002). The combined bowling rate as a measure of bowling performance in cricket. *South African Journal for Research in Sport, Physical Education and Recreation*, 24(2): 37-44.

LEMMER, H.H. (2004). A measure for the batting performance of cricket players. *South African Journal for Research in Sport, Physical Education and Recreation*, 26(1): 55-64.

PRESS TRUST OF INDIA (2005). "SA selection policy flawed, says Rhodes." Hyperlink [http://www.indianexpress.com/full_story.php?content_id=64759]. Retrieved 19 February 2005.

PRESTON, I. & THOMAS, J. (2000). Batting strategy in limited overs cricket. *The Statistician*, 49(1): 95-106.

PRESTON, I. & THOMAS, J. (2002). Rain rules for limited overs cricket and probabilities of victory. *The Statistician*, 51(2): 189-202.

RISTOW, A.M.; AMOS, T.L. & STAUDE, G.E. (1999). Transformational leadership and organizational effectiveness in the administration of cricket in South Africa. *South African Journal of Business Management*, 30(1): 1-5.

SLOGROVE, L.; POTGIETER, J.R. & FOXCROFT, C.D. (2002). Batting related experiences of South African universities cricketers. *South African Journal for Research in Sport, Physical Education and Recreation*, 24(1): 101-111.

TAHA, H.A. (2003). *Operations research, an introduction* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Ms. Hannah Gerber: School of Statistics and Actuarial Science, University of the Witwatersrand, P.O. Box 139, Wits 2050, Johannesburg, Republic of South Africa. Tel.: +27 (0)11 717 6276 (w), +27 (0)84 869 4121 (mob), Fax. +27 (0)11 717 6285, Email: hgerber@stats.wits.ac.za

(Subject editor: Prof. E. Terblanche)