OPSEARCH (2016) 53:225–247
DOI 10.1007/s12597-015-0228-3

APPLICATION ARTICLE

An objective approach of balanced cricket team selection using binary integer programming method

Dibyojyoti Bhattacharjee[1] **· Hemanta Saikia**[2]

Accepted: 23 September 2015 / Published online: 10 October 2015

Abstract Selecting a balanced playing XI in the game of cricket with the right mix of players of different specialization is a difficult decision making problem for the team management. To make the process more objective, optimization techniques can be applied to the process of selection of players from a given squad. Such an exercise has two dimensions. First, a suitable tool for performance measurement of cricketers needs to be defined. Secondly, for selecting a balanced team of XI players, an appropriate objective function and some constraints need to be formulated. Since the captain gets an obvious inclusion in the cricket team, the area specialization of the captain influences the selection of other ten positions in the playing XI. This study attempts to select the optimum balanced playing XI from a squad of players given the specialization of the captain using binary integer programming. To validate the exercise, data from the fifth season of the Indian Premier League has been used.

Keywords Cricket · Decision making · Integer programming · Optimization · Performance analysis · Sports

1 Introduction

Cricket is an outdoor game played in an oval or circular ground, where the interaction between the willow (i.e., bat) and the leather ball, takes place on a 22-yard hard surface of earth called the pitch. The game is governed by certain rules and regulations between two teams each comprising of eleven players. Unlike other sports, there are different

✉ Hemanta Saikia
h.saikia456@gmail.com

Dibyojyoti Bhattacharjee
djb.stat@gmail.com

[1] Department of Statistics, Assam University, Silchar 788011 Assam, India

[2] School of Business, Kaziranga University, Jorhat 785006 Assam, India

versions of cricket. The versions can be classified as unlimited overs cricket (Test matches) and limited overs cricket (One-day and Twenty20). Test match is the oldest version of cricket, which is generally a five-day affair. When a team bats, the other team bowls and is called an 'innings'. Here each of the two sides shall bat for at least one innings, but a maximum of two innings. The team scoring the most number of runs at the end of both their innings emerges as winner. In case of incomplete innings, even at the end of five days of play the match ends in a draw. In limited overs cricket, each team has a chance to bat and to bowl only once. The bowling team bowls the specified number of overs and the batting team tries to score maximum possible runs. In the process of scoring runs quickly, the batsmen are exposed to the risk of getting out. Since the batsmen have to bat in pairs, the innings of a batting team comes to an end if ten wickets are lost; otherwise, the batting team can continue batting until the end of the specified over limit. After completion of the batting team's innings, the opponent team (the team that bowled first) gets the opportunity to bat. The team batting second, attempts to score more runs than the team that batted first, without losing all the wickets and within the specified over limit. The team that has scored the most number of runs at the end of the match stands the winner.

Among the different skills required to become a cricketer, batting and bowling are undoubtedly the prime skills in the game of cricket. Those players who can perform reasonably well, both with the bat and the ball are called all-rounders [1]. A reasonably balanced cricket team is a blend of players drawn from diverse expertise. Precisely, a balanced cricket team demands the services of the captain, a wicket keeper, a few specialist batsmen including openers, all-rounders and specialist bowlers (both fast bowlers as well as the spinners). Consequently, a lot of subjectivity is involved in the selection of a cricket team [2]. Depending on the strengths and weaknesses of the opponent, pitch and weather conditions, the combination of players' viz. number of spinners, all-rounders, specialist batsmen and fast bowlers are decided. Though, the numbers are relatively easier to be agreed upon by the selectors/team management, the players who shall actually make it to the playing XI may call for a lengthy discussion. The paper explains an objective method of selecting the optimized balanced playing XI, from a host of potential players, considering the cricketing requirements and players' recent performances. Following Salas et al. [3], a balanced team may be defined as: "A team that comprises of a set of independent team members, each of whom possesses unique and expert level of knowledge, skills and experience related to the task which is deemed to be assigned to that member. The members of the dream team are expected to adapt, coordinate and cooperate as a team and thereby producing sustainable and repeatable team functioning at an optimal or near-optimal level of performance."

Irrespective of the format of cricket (Test, ODI (One-day International) and Twenty20) the process of selecting a playing XI is almost similar. The selectors, generally a group of former cricketers, choose a squad of fifteen cricketers (generally) from different specialization based on the recent performance of a host of cricketers, in domestic and international matches for a given series or tournament. Out of the selected cricketers one is named as the captain. When the series/tournament actually starts, the captain has to select the ultimate playing XI before each match, from the squad of players named by the selectors. Since in cricket, there is no concept of non-playing captain so the captain is always included in the playing XI. Thus, the captain is entrusted with the responsibility to select the other ten players of the playing XI.

To make the process of cricket team selection more objective, the paper proposes an optimization technique to select the best team from the squad of available cricketers. The proposed optimization technique takes due care of the area of specialization of the captain as well as other players. Furthermore, it addresses the issue of the area specialization of the captain in the playing XI and how his specialization influences the selection of other ten players of the team. The approach provides an objective user interface for the team management regarding the inclusion of players in the playing-eleven. It can also be used for recreational use e.g., selection of fantasy cricket sides or an optimum 'All-Star' team at the end of a series featuring several teams.

2 Review of literature

A cricket match generates huge amount of data and so it is not surprising to find several quantitative research works on cricket. This section focuses on some of the significant works related to optimized team selection in cricket. Kamble et al. [4] presented a selection procedure using analytical hierarchical process to choose a subset of players from a universal set of cricketers comprising of batsmen, bowlers, all-rounders and wicket keepers. Two other works addressing the same issue are Lemmer [2] and Ahmed et al. [5]. While Lemmer [2] used integer programming to reach the solution, Ahmed et al. [5] used evolutationary multi-objective optimization to choose the cricket team. Portfolio analysis was used by Barr and Kantor [6] to determine the set of batsmen who are supposed to be more suitable for a given one-day squad. Gerber and Sharp [7] proposed an integer programming technique in order to select a limited over squad of 15 players instead of a playing XI. The method included collecting data from 32 prominent South African cricketers to select the ODI squad. Extending the same idea, Lourens [8] selected an optimal Twenty20 South African cricket side based on performance statistics of a host of players who participated in the SA domestic Pro20 cricket tournament. Using integer programming, Brettenny [9] selected players for a fantasy league cricket team under certain pre-specified budgetary constraints but with a progressive approach. This optimal team, at each stage of the tournament, considered the performance of available cricketers till the previous match. Though most authors have used the binary integer programming tool for the purpose of the optimized team selection, they used different tools for measuring the performance of cricketers. Some authors have used the traditional statistics like batting average, strike rate, etc. for quantifying performance of cricketers, while others tried to combine such traditional measures to a refined statistic to evaluate players' performance. Van Staden [10] using contour plot measured the performance of batsman, bowlers and all-rounders using data from the first edition of the Indian Premier League. Using the ratio of runs scored to the resources consumed by a player, Beaudoin and Swartz [11] defined a new measure of performance for both batsman and bowlers. Fuzzy logic and stochastic models can be gainfully applied to model the uncertainties involved in the measure of performance of cricketers especially in case of smaller number of matches. Singh et al. [12] and Damodaran [13] are two relevent works in this regard. Authors like Lourens [8] and Brettenny [9] combined/ compared different refined measures in the process of optimal team selection. The need of applying such optimization models in team selection is explained in Boon and Sierksma [14]. Though the paper is related to team selection in soccer and volleyball, it is an example of how transportation problem can be modelled for this purpose. Das [15] used binary integer programming to create optimal sequences of teams in fantasy sports leagues, with special reference to cricket. The work takes into consideration the multiple aspects of such team selection. However, none of the aforementioned works assume the captain as a permanent member of the team, which the current research does. A details discussion on which follows.

Every cricket team has a captain who takes most of the decisions himself in consultation with the senior players and coach. The captain is selected for a given series/tournament and occupies a place in the playing XI in all the matches of the series. In case of injury or otherwise, if the captain fails to continue, a new captain is named who enjoys the same benefit. Generally, the captain is a senior player of the team who can have any type of specialization. The captain may be a fast bowler or a wicket keeper batsman or an all-rounder and so on. Since, the captain acquires a permanent position in the team the expertise of the captain shall be a dependent factor while selecting the other members to the playing XI. To simplify, the following example is forwarded. Suppose the captain of a cricket team is a wicket keeper who is also a middle order batsman. His presence in the team restricts the inclusion of any other wicket keeper in the playing XI. Thus, the other members of the team shall comprise of specialist bowlers, two opening batsmen, other specialist batsmen, all-rounders but probably not another wicket keeper. Another team may have their captain who is a specialist batsman and opens the innings as well. Thus, such a team shall require another opener. To sum up, with the captain getting an automatic inclusion into the playing XI, the decision about filling up the remaining ten places is reliant on the cricketing expertise of the captain. None of the aforesaid works picked up the optimization issue taking into consideration the cricketing skills of the captain. Also, the above discussion clarifies that, since the captains of the different teams shall have different expertise, a unique integer programming problem for all the teams is to be modelled. Though the optimization problem will have the same objective function, teams would have different constraints as the requirements of each team shall vary depending on the cricketing expertise of their captain. The above backdrop sets an ideal ground for undertaking the current enquiry. The work tends to develop a single binary integer programming, taking into consideration the skills of the captain, tournament restrictions (if any) and cricketing requirements and end up in selecting the best balanced team from a set of available players.

3 Methodology

The methodology used for performing the task can be broadly classified into two sections. The first section introduces the performance measure that is used to the recent feats of the cricketers in the matches they played in and the second section deals with the optimization model.

3.1 Performance measurement

Batting average and strike rate are mostly used to measure the performance of the batsman while bowling average, economy rate and bowlers' strike rate are used to measure the performance of the bowlers. But it is widely recognized that such statistics have severe limitations in assessing the true abilities of a player's performance [16].

Further, Lewis [16] mentioned that the traditional measures of performances do not allow combination of abilities of batting and bowling as they are based on incompatible scales. To overcome this limitation the following performance measure is proposed.

The performance measure of the i^{th} player is given by,

$$S_i = w_1 S_{i1} + w_2 S_{i2} + \delta_i \tag{1}$$

where

$$\delta_i = \begin{cases} (w_3 S_{i3})^{a_i} + (w_4 S_{i4})^{1-a_i} - 1, & \textit{if } i^{th} \textit{ player is either a bowler or wicket keeper} \\ 0 \quad , & \textit{if } i^{th} \textit{ player is neither a bowler nor wicket keeeper} \end{cases}$$

where a_i is an indicator variable with,

$$a_i = \begin{cases} 1, & \textit{if } i^{th} \textit{ player is a bowler} \\ 0, & \textit{if } i^{th} \textit{ player is a wicket keeper} \end{cases}$$

with S_{i1}=Performance score for batting, S_{i2}=Performance score for fielding, S_{i3}= Performance score for bowling and S_{i4}=Performance score for wicket keeping. The weights w_i (i=1, 2, 3 and 4) are determined using the Iyenger-Sudershan method. The weights in this method are determined is such a manner that the variance of each of the performance scores viz. batting, fielding, bowling and wicket-keeping, are stabilized on multiplication with the corresponding weights. This denies the dominance of a particular performance score in the index over the other scores. The details of the method can be read in Iyenger and Sudershan [17]. On computation it is found that the weights corresponding to batting, fielding, bowling and wicket-keeping are 0.21659, 0.29793, 0.22666 and 0.2588 respectively.

3.1.1 Batsman's performance measure (S_{i1})

All performance measures of the batsman take into consideration the number of runs scored by the batsman. The runs scored by a batsman, in a given match, depends on the bowling strength of the opposition, condition of the pitch, availability of resources of the batting team in terms of overs and wickets, etc. In addition to that, 50 runs scored by a batsman in a match where the team total is 150, is more valuable compared to the same number of runs against the same opposition when the team scores more than 300 runs. Considering all these factors, Lemmer [18] derived a technique that can convert the runs scored in a match to the adjusted runs based on the match condition and opposition's bowling strength. These adjusted runs are then used to define the batting performance measure S_{i1}. Thus, following Lemmer [18] the batting performance measure is defined by,

$$BP_i = (x_{1i} + (2.1 - 0.005 \times x_{2i}) \times x_{3i})/n_i \tag{2}$$

where

n_i number of innings played by the i^{th} batsman

x_{1i} sum of adjusted runs in the innings in which the i^{th} batsman was out
x_{2i} sum of adjusted runs in the innings in which the i^{th} batsman was not out
x_{3i} average of the not out score of the i^{th} batsman obtained from the adjusted runs

The adjusted runs scored by the i^{th} player in the j^{th} match is denoted by T_{ij} and is defined by,

$$T_{ij} = R_{ij}\left(SR_{ij}/MSR_j\right)^{0.5} \tag{3}$$

where R_{ij} is the runs scored by the i^{th} batsman in the j^{th} match.

$$SR_{ij} = \text{Strike rate of the } i^{th} \text{ batsman in the } j^{th} \text{ match} = \frac{R_{ij}}{B_{ij}} \times 100 \tag{4}$$

where B_{ij} is the number of balls faced by the i^{th} batsman in the j^{th} match and strike rate of the j^{th} match (MSR_j) for all the batsmen of both teams

$$MSR_j = \frac{\text{Total no. of runs scored in the match}}{\text{Total no. of balls bowled in the entire match}} \times 100 = \frac{R_j}{B_j} \times 100 \tag{5}$$

Thus, this measure compares the performance of a batsman in relation to his peers who also participated in the match. The data requirement for this measure includes individual as well as team performances from all the matches in which the i^{th} player has batted. The batting performance score (BP_i) thus obtained is then standardized by the average value of BP across all batsmen,

$$S_{i1} = \frac{BP_i}{Avg_i\,(BP_i)} \tag{6}$$

3.1.2 Fielding performance measure (S_{i2})

The different factors that are considered to quantify the fielding performance of a player are number of catches taken by the player as a fielder and the number of run-outs caused by the player in the series. The number of catches and run-outs are added to get the number of dismissals contributed by a given fielder (D_i). Accordingly the dismissal rate of fielders is defined as

$$D_i = \frac{\text{Total no. of dissimals by the } i^{th} \text{ player}}{\text{No. of matches played}} \tag{7}$$

The dismissal rate of fielders (D_i) thus obtained is then standardized by the average value of D across all fielders,

$$S_{i2} = \frac{D_i}{Avg_i\,(D_i)} \tag{8}$$

3.1.3 Bowler's performance measure (S_{i3})

Lemmer [19] proposed a bowling performance measure called the combined bowling rate (*CBR*) which is the harmonic mean of three traditional bowling statistics viz. bowling average, economy rate and bowling strike rate. If R be the total number of runs conceded by a bowler, W is the total number of wickets taken by a bowler and B is the total number of balls bowled by a bowler in a series of matches. Then the traditional bowling statistics can now be defined as,

$$\text{Bowling average } = \frac{R}{W}, \text{ Economy rate } = \frac{R}{B/6}, \text{ Bowling strike rate } = \frac{B}{W}$$

To bring parity in the numerator of the above factors, a pre-requisite of the harmonic mean, the bowling strike rate was adjusted by Lemmer [19] as follows,

$$\text{Bowling strike rate } = \frac{B}{W} = \frac{B}{W} \times \frac{R}{R} = \frac{RB}{RW} = \frac{R}{RW/B} \tag{9}$$

Thus, the combined bowling rate (*CBR*) defined by Lemmer [19] as

$$\begin{aligned} CBR &= \frac{3}{\frac{1}{bowling\,average} + \frac{1}{economy\,rate} + \frac{1}{bowling\,strike\,rate}} \\ &= \frac{3R}{W + (B/6) + W \times \frac{R}{B}} \end{aligned} \tag{10}$$

Later, Lemmer [20] improved the *CBR* to an adjusted measure called *CBR** which is more appropriate for quantifying bowling performance for small number of matches. The adjusted combined bowling rate (*CBR**) for the i^{th} bowler is given by,

$$CBR_i^* = 3R_i^{'} / \left[W_i^* + (B_i/6) + W_i^* \left(R_i^{'} / B_i \right) \right] \tag{11}$$

where

B_i = number of balls bowled by the i^{th} bowler
W_i^* = sum of weights of the wickets taken by the i^{th} bowler
R_i' = sum of adjusted runs (RA_{ij}) conceded by the i^{th} bowler in the j^{th} innings = $\sum_{j=1}^{n_i} RA_{ij}$

$$RA_{ij} = R_{ij} \left(RPB_{ij} / RPBM_j \right)^{0.5} \tag{12}$$

where, $RPB_{ij} = \frac{\text{Runs conceded by the } i^{th} \text{ player in the } j^{th} \text{ match}}{\text{Balls bowled by the } i^{th} \text{ player in the } j^{th} \text{ match}}$

$$RPBM_j = \frac{\text{Total runs scored in the } j^{th} \text{ match}}{\text{Total no. of balls bowled by the } j^{th} \text{ match}}$$

The measure is noteworthy because of the following two inherent issues. The factor, ($RPB_{ij}/RPBM_j$) considers the match situation in which the i^{th} bowler delivered and the factor W_i^* which refuses to give equal importance to all the wickets taken by the bowler but weights them differently based on their batting position. The detailed discussion and the different values of W_i^* is available in Lemmer [20]. The combined bowling rate has a negative dimension (i.e., lower the value the better is the bowler). So to bring parity with the batting performance, the CBR^* is inverted and is standardized by the average value of inverse CBR^* across all the bowlers, i.e.,

$$S_{i3} = \frac{1/CBR_i^*}{\underset{i}{Avg}\left(1/CBR_i^*\right)} \tag{13}$$

3.1.4 Wicket keeper's performance measure (S_{i4})

For measuring the performance of a wicket keeper two factors are considered. They are (i) Dismissal rate (ii) Bye runs conceded (rate). According to Narayanan [21], dismissal rate of a wicket keeper is defined as the number of dismissals (stumping and catches) per match. Here the term 'match' refers only to those matches where the player under consideration kept wicket for his team.

$$\text{Dismissal Rate } \left(D_i'\right) = \frac{\text{Total number of dismissals by the } i^{th} \text{ player}}{\text{No. of matches in which } i^{th} \text{player kept wickets}} \tag{14}$$

The rate in which bye runs were conceded is defined as,

$$\text{Byes Rate } \left(B_i'\right) = \frac{\text{Total bye runs conceded by the } i^{th} \text{ player}}{\text{No. of matches in which } i^{th} \text{ player kept wickets}} \tag{15}$$

It is obvious that while 'dismissal rate' has a positive dimension (i.e., positively related to the skill of the player) 'bye runs conceded' have a negative dimension. The lesser the byes rate better is the wicket keeper unlike that of the dismissal rate. Thus, instead of B_i' to bring parity between the two rates ($1/B_i'$) is considered. However, in order to combine these two measures viz. D_i' and ($1/B_i'$) into a single measure, it is necessary to standardize them. The standardized values may be defined as follows:

$$D_i = \frac{D_i'}{\underset{i}{Avg}\left(D'\right)} \quad \text{and} \quad B_i = \frac{1/B_i'}{\underset{i}{Avg}\left(1/B_i'\right)}$$

Thus, $D_i \times B_i$ can be considered as a performance measurement of wicket keeper, but in order to ensure that dismissal measure and bye rate are comparable, scale adjustment of B_i is necessary by raising a real number α to the exponent of B_i such that standard deviation of D_i and that of B_i^{α} is exactly same. The value of α can be determined by any iterative method (see, Lemmer [22] for details). Briefly speaking the method can be discussed as follows.

Let S_1=standard deviation of D_i and S_2=standard deviation of B_i with $S_1 > S_2$. Now, the first transformation of B_i values is defined as $B_i^{(1)}$ using the relation $B_i^{(1)}=(B_i)^{\alpha}$ where $\alpha = \frac{S_1}{S_2}$. This transformation of B_i values shall provide with a variance $S_2^{(1)}$ of $B_i^{(1)}$ which is closer to S_1 compared to S_2. This process is repeated again. That is we define, $B_i^{(2)}=(B_i^{(1)})^{\alpha}$ where $\alpha = \frac{S_1}{S_2^{(1)}}$. This transformation of $B_i^{(1)}$ values shall provide with a variance $S_2^{(2)}$ of $B_i^{(2)}$ which is closer to S_1 compared to $S_2^{(1)}$ where $\alpha = \frac{S_1}{S_2^{(2)}}$. This is repeated several times. The iteration stops at the j^{th} stage when the standard deviation $S_2^{(j)}$ is approximately equal to S_1. Subsequently, $B_i^{(j)}$ values shall be used in (16) in place of B_i and $\alpha = \frac{S_1}{S_2^{(j)}}$.

So, $D_i \times B_i^{\alpha}$ is reached as a performance measure of the wicket keeper, but this measure considers both the factors viz. dismissal measure and bye rate as equally important. But as a dismissal leads to loss of resources of the opponent team so it shall get relatively more importance compared to bye runs conceded. Thus, the factor $D_i \times B_i^{\alpha}$ is reformulated with a weighted product so that the relative importance of the two factors can be quantified. This leads to the definition of WK_i as,

$$\begin{aligned} WK_i &= D_i^{\beta} \times \left(B_i^{\alpha}\right)^{1-\beta}, \qquad 0<\beta<1 \\ &= \left(\frac{D_i}{Avg(D_i)}\right)^{\beta} \times \left(\left[\frac{1/B_i'}{Avg\left(1/B_i'\right)}\right]^{\alpha}\right)^{1-\beta}, \qquad 0<\beta<1 \end{aligned} \tag{16}$$

The value of β determines the relative importance of the factors and acts as a balance between the dismissal measure and bye rate. The number of bye runs conceded by a wicket keeper also depends on the quality of bowling and the activity is relatively less important than dismissals. Narayanan [21] allocated 5 points to byes conceded and 40 to dismissals, making the later 8 times more important than saving bye runs. However, such an allocation is subjective. Even the author himself did not find enough scientific reasoning for such allocation. In the absence of sufficient literature and difference in expert opinion it is difficult to converge to an objective value of β. However, Pareto ordering for multi-objective comparision is used to find out that value of β which has maximum compatibility. It shall in general be accepted that conceding of bye runs are less important than the dismissal of a batsman. Thus, the values of β are allowed to vary from 0.5 through 0.9 with an increment of 0.05. A value of β=0.5 gives equal importance to both the dismissal measure and bye rate and β=0.9 makes dismissal measure nine times more important than the bye rate. Different values of β leads to changes in the corresponding values of WK_i and so the ranks of the wicket keepers under consideration change markably. Using Pareto ordering that set of ranks which has the maximum compatibility with the other rankings are determined (*c.f.* Chakrabarty and Bhattacharjee [23] for detailed discussion of the Pareto ordering method).

Let, β takes the values $\beta_1, \beta_2, \ldots, \beta_m$

$R_i^{\beta_j}$ = Rank of the i^{th} wicket keeper for a given value of $\beta=\beta_j$ (say)

$d_i^{\beta_j,\beta_p}$ = Square of difference between ranks of the i^{th} wicket keeper for $\beta=\beta_j$ and $\beta=\beta_p=\left(R_i^{\beta_j}-R_i^{\beta_p}\right)^2$

D_{β_j} = Overall distance of ranks when $\beta=\beta_j$ with other values of β $=\sum_{\substack{p=1 \\ p\neq j}}^{m}\left(R_i^{\beta_j}-R_i^{\beta_p}\right)^2$

Thus, the compatibility score corresponding to $\beta=\beta_j$ is given by $\overline{D}_{\beta_j}$ as defined in (17). The measure is the average distance of ranks of $\beta=\beta_j$ with other β values. Lesser the compatibility score of a given index more is the compatibility of that index with a set of similar other indices.

$$\overline{D}_{\beta_j}=\frac{D_{\beta_j}}{m-1} \tag{17}$$

That value of β (=β_j say) for which $\overline{D}_{\beta_j}$ is minimum has the maximum compatability related to the other values of β. Accordingly, the most compatable value of β is identified and is replaced in (16) for subsequent analysis.

Table 1, shows the computational details of the method discussed above based on data from IPL V. Looking at the ranks one shall find that KD Kartik made a gradual decrease in his ranks which started at rank 5 ($\beta=0.5$) and finished at rank 11 ($\beta=0.9$). KD Kartik is found to be most sensitive to changing values of β. DH Yagnik had a rank of six when $\beta=0.9$ but from $\beta=0.8$ towards lesser values of β he reached a rank of one. MS Bisla has the minimum variation in ranks (i.e., either 2 or 3) over the different values of β. AB de Villiers ranks changed from 9 to 7 and then to 8. Based on the above discussed Pareto ordering, the most compatable ranking is provided by β=0.7, as that value of β provides the minimum value of $\overline{D}_{\beta_j}$ (c.*f.* Table 1). So, the values of WK_i corresponding to β=0.7 is considered for further analysis.

Based on the data of IPL V, the values of D_i and B_i are computed. Using an iterative method it is found that for α=0.8857, the standard deviations of D_i and B_i^{α} are same. Thus, S_{i4} takes the final shape of,

$$WK_i=\left(\frac{D_i}{Avg(D_i)}\right)^{0.7}\times\left(\left[\frac{1/B_i'}{Avg\left(1/B_i'\right)}\right]^{0.8857}\right)^{0.3} \tag{18}$$

The wicket keeping performance score (WK_i) thus obtained is then standardized by the average value of WK across all keepers, i.e.,

$$S_{i4}=\frac{WK_i}{\underset{i}{Avg}\ (WK_i)} \tag{19}$$

Now, for a given player, (as often referred to him as the i^{th} player) the values of S_{i1}, S_{i2} and S_{i3} or S_{i4} are computed using (6), (8), (13) and (19). The values are then replaced in (1) to get the performance score of the i^{th} player (i.e., S_i).

Table 1 Ranks of wicket keepers for different values of β in (14) with α=0.8857

Player	β=0.5		β=0.55		β=0.6		β=0.65		β=0.7		β=0.75		β=0.8		β=0.85		β=0.9	
	WK_i	Rank	WK_i	Rank	WK_i	Rank	WK_i	Rank	WK_i	Rank	WK_i	Rank	WK_i	Rank	WK_i	Rank	WK_i	Rank
NV Ojha	0.66	7	0.75	7	0.83	6	0.91	6	0.97	6	1.02	6	1.07	5	1.10	5	1.12	4
BB McCullum	0.75	4	0.88	3	1.01	3	1.13	3	1.25	2	1.37	2	1.47	2	1.57	1	1.66	1
MS Dhoni	0.78	3	0.86	4	0.92	4	0.97	5	1.01	5	1.04	5	1.05	6	1.05	6	1.04	5
N Saini	0.67	6	0.78	5	0.89	5	1.00	4	1.11	4	1.21	4	1.30	4	1.38	3	1.45	2
RV Uthappa	0.35	12	0.41	12	0.46	12	0.51	12	0.55	12	0.59	11	0.63	11	0.66	11	0.69	10
MS Bisla	0.84	2	0.96	2	1.06	2	1.16	2	1.24	3	1.31	3	1.37	3	1.41	2	1.44	3
SP Goswami	0.38	11	0.45	11	0.51	10	0.58	10	0.64	10	0.70	10	0.76	9	0.81	8	0.86	8
AB de Villiers	0.56	8	0.61	8	0.66	8	0.70	8	0.73	8	0.75	7	0.76	8	0.77	9	0.76	9
KD Karthik	0.75	5	0.78	6	0.79	7	0.78	7	0.77	7	0.74	8	0.71	10	0.67	10	0.62	11
PA Patel	0.43	10	0.47	10	0.51	11	0.54	11	0.56	11	0.57	12	0.58	12	0.58	12	0.58	12
AC Gilchrist	0.43	9	0.50	9	0.57	9	0.63	9	0.69	9	0.74	9	0.79	7	0.83	7	0.87	7
DH Yagnik	5.38	1	4.56	1	3.78	1	3.09	1	2.47	1	1.95	1	1.53	1	1.18	4	0.90	6
$\overline{D}_{\beta_j}$	43.25		29.25		21		19.5		19.25		20.25		28		38.75		65.75	

3.2 The optimization model

The optimization technique used for team selection is a binary (0–1) integer-programming problem and the solution to the problem is attained using the Solver add-in available in Microsoft Excel. Suppose the final XI is to be composed in such a manner that there are at least four specialist batsman (including two openers), one wicket keeper, at least two fast bowlers, at least one spinner and at least one all-rounder. The selection needs to have exactly six bowling options available in the playing XI, including the all-rounder(s). One wicket keeper-batsman and at least four specialist batsmen shall be accommodated in the team. Out of the wicket keeper, all-rounder(s) and specialist batsmen in the team, two batsmen should have the capability of opening the innings. In the entire process of selection, the captain gets automatically selected in the playing XI, so the actual decision is to be taken for the remaining ten positions of the team, considering the expertise of the captain himself. This is to be done to avoid over representation or under representation of a particular skill in the playing XI. Let a set of binary coefficients be defined. These coefficients are used in setting the objective function and the constraints.

$\theta_i = 1$ (0) if the i^{th} player is selected for the playing XI (Otherwise)
$b_i = 1$ (0) if the i^{th} player is an opening batsman (Otherwise)
$c_i = 1$ (0) if the i^{th} player is a specialist batsman but not an opener (Otherwise)
$d_i = 1$ (0) if the i^{th} player is a spinner (Otherwise)
$e_i = 1$ (0) if the i^{th} player is a fast bowler (Otherwise)
$f_i = 1$ (0) if the i^{th} player is a wicket keeper (Otherwise)
$g_i = 1$ (0) if the i^{th} player is an all-rounder (Otherwise)

Since the captain is known and is already a member of the team, so there are 10 more places to be filled up from a collection of k players (say) of different expertise. Therefore, the first constraint is,

$$\sum_{i=2}^{k} \theta_i = 10 \quad \text{where } (i = \ 1 \text{ represents the captain}) \tag{20}$$

The constraint in (20) ensures that there are exactly 10 players in the team excluding the captain. The other constraints can be defined, only after knowing the expertise of the captain. Here attempt is made to design the remaining constraints in such a manner that the model can be generalized for any type of expertise of the captain. To do that two column vectors $\boldsymbol{l}$ and $\boldsymbol{p}$ are defined (*c.f.* Tables 2 and 3). A player in a cricket team may be either an opening batsman or middle order batsman or spinner or fast bowler or wicket keeper batsman or an all-rounder. For each of the expertise, we attach a column vector l_j such that $l_j'=(x_1,x_2,x_3,x_4,x_5)$, where the x_i's are binary variables. The values of l_j' corresponding to the different expertise are given below.

Now, based on the expertise of the captain another column vector $\boldsymbol{p}$ is defined. Such that, $\boldsymbol{p'}=(y_1, y_2, y_3, y_4, y_5)$. The values of y's are once again binary variables and is related to the expertise of the captain where the suffix 1, 2, 3, 4 and 5 represents opener, middle-order batsman, spinner, fast bowler and wicket keeper respectively. Table 3 explains how different expertise of the captain is notified for the column vector $\boldsymbol{p}$.

Table 2 Constraints for different expertise of a player

Constraint on	$j =$	l_j'
Opener	1	(1, 0, 0, 0, 0)
Middle-order batsman	2	(0, 1, 0, 0, 0)
Spinner	3	(0, 0, 1, 0, 0)
Fast bowler	4	(0, 0, 0, 1, 0)
Wicket keeper	5	(0, 0, 0, 0, 1)
All-rounder	6	(1, 1, 1, 1, 0)

It shall be noted that a captain may possess two different abilities at the same time. The vector ***p*** is defined, keeping in mind the dual ability that a captain might possess.

Accordingly, the following possible constraints are formulated in addition to (20) to select an optimal cricket team.

$$\sum_{i=2}^{k} \theta_i b_i = 2 - l_1' p \tag{21}$$

The constraint (21) ensures that the team has exactly two opening batsmen. In case the captain is an opening batsman then the team needs one more opener otherwise two openers are to be selected. This issue is negotiated by $\boldsymbol{l_1'p}$. When the captain is an opener it may be noted that (*c.f.* Tables 2 and 3),

$$\begin{aligned} l_1'p &= (1,\ 0,\ 0,\ 0,\ 0)\,(1,\ 0,\ 0,\ 0,\ 0)' = 1, \\ &= 0 \text{ otherwise} \end{aligned}$$

The same would happen even if the captain is simultaneously a wicket keeper as well as an opening batsman or an all-rounder with ability to open the innings.

$$\sum_{i=2}^{k} \theta_i c_i \geq 2 - l_2' p \tag{22}$$

Table 3 Constraints based on expertise of the captain

Expertise of the captain	p'	Expertise of the captain if all-rounder	p'
Opener	(1, 0, 0, 0, 0)	Fast bowler and opener	(1, 0, 0, 1, 0)
Middle-order batsman	(0, 1, 0, 0, 0)	Spinner and opener	(1, 0, 1, 0, 0)
Spinner	(0, 0, 1, 0, 0)	Fast bowler and middle-order batsman	(0, 1, 0, 1, 0)
Fast bowler	(0, 0, 0, 1, 0)	Spinner and middle-order batsman	(0, 1, 1, 0, 0)
Wicket keeper and opener	(1, 0, 0, 0, 1)		
Wicket keeper and Middle-order batsman	(0, 1, 0, 0, 1)		

The constraint (22) ensures that the team has at least two middle-order batsmen. In case the captain is a middle-order batsman then the team needs at least one more middle-order batsman otherwise at least two middle-order batsmen are to be selected. This issue is negotiated by $\boldsymbol{l_2'p}$. When the captain is a middle-order batsman then it may be noted that (*c.f.* Tables 2 and 3),

$$\begin{aligned} l_2'p &= (0,\ 1,\ 0,\ 0,\ 0)\ (0,\ 1,\ 0,\ 0,\ 0)' = \ 1, \\ &= 0,\ \text{otherwise} \end{aligned}$$

The same would happen even if the captain is simultaneously a wicket keeper as well as a middle-order batsman or an all-rounder who is a sound middle order batsman.

$$\sum_{i=2}^{k} \theta_i d_i \geq 2 - l_3' p \tag{23}$$

The constraint (23) ensures that the team has at least two spin bowlers. In case the captain is a spinner then the team needs at least one more spinner otherwise at least two spinners are to be selected. This issue is negotiated by $\boldsymbol{l_3'p}$. When the captain is a spinner then it may be noted that (*c.f.* Tables 2 and 3),

$$\begin{aligned} l_3'p &= (0,\ 0,\ 1,\ 0,\ 0)\ (0,\ 0,\ 1,\ 0,\ 0)' = 1, \\ &= 0,\ \text{otherwise} \end{aligned}$$

The same would happen even if the captain is an all-rounder who can spin the ball.

$$\sum_{i=2}^{k} \theta_i e_i \geq 2 - l_4' p \tag{24}$$

The constraint (24) ensures that the team has at least two fast bowlers. In case the captain is a fast bowler then the team needs at least one more fast bowler otherwise at least two fast bowlers are to be selected. This issue is negotiated by $\boldsymbol{l_4'p}$. When the captain is a fast bowler then it may be noted that (*c.f.* Tables 2 and 3),

$$\begin{aligned} l_4'p &= (0,\ 0,\ 0,\ 1,\ 0)\ (0,\ 0,\ 0,\ 1,\ 0)' = \ 1, \\ &= 0,\ \text{otherwise} \end{aligned}$$

The same would happen even if the captain is an all-rounder with fast bowling ability.

$$\sum_{i=2}^{k} \theta_i f_i \geq 1 - l_5' p \tag{25}$$

The constraint (25) ensures that the team has at least one wicket keeper. In case the captain is a wicket keeper then the team generally does not need any other wicket keeper otherwise at least one wicket keeper is to be selected. This issue is negotiated by $\boldsymbol{l_5'p}$. The vector $\boldsymbol{p}$ has two possible values viz. (1, 0, 0, 0, 1)′ or (0, 1, 0, 0, 1)′. The former one is used when the wicket keeping captain is an opener and the later when the wicket keeping captain is a middle-order batsman. Thus, it may be noted that (*c.f.* Tables 2 and 3),

$$l_6'p = (0,\ 0,\ 0,\ 0,\ 1)\ (1,\ 0,\ 0,\ 0,\ 1)' \text{or}\ (0,\ 0,\ 0,\ 0,\ 1)\ (0,\ 1,\ 0,\ 0,\ 1)' = 1,$$
$$= 0,\ \text{otherwise}$$

$$\sum_{i=2}^{k} \theta_i g_i \geq 2 - l_6' p \tag{26}$$

The constraint (26) ensures that the team has at least one all-rounder. In case the captain is an all-rounder then the team may or may not employ any other all-rounder otherwise at least one all-rounder needs to be selected to bring balance in the team. This issue is negotiated by $\boldsymbol{l_6'p}$. The vector p related to all-rounder can assume any one of the following possible values viz. (1, 0, 0, 1, 0)′ or (1, 0, 1, 0, 0)′ or (0, 1, 0, 1, 0)′ or (0, 1, 1, 0, 0)′. The vector $\boldsymbol{p}$ depends on whether the captain is an all-rounder by the virtue of being a fast bowler and opening batsman or spinner and opening batsman or fast bowler and middle-order batsman or spinner and middle-order batsman respectively (*c.f.* Table 3). Thus, it may be noted that (*c.f.* Tables 2 and 3),

$$l_6'p = (1,\ 1,\ 1,\ 1,\ 0)\ \left[(1,\ 0,\ 0,\ 1,\ 0)' \text{or}\ (1,\ 0,\ 1,\ 0,\ 0)\ \text{or}\ (0,\ 1,\ 0,\ 1,\ 0)\ \text{or}\ (0,\ 1,\ 1,\ 0,\ 0)\right] = 2,$$
$$= 1,\ \text{otherwise}$$

Thus, if the captain is an all-rounder the constraint may or may not select any other all-rounder but if the captain is not an all-rounder then the model shall pick up at least one all-rounder in the optimum team.

Generally, most of the captains these days prefer to take the field with a sixth bowling option including the all-rounders especially in case of Twenty20 match. In limited overs cricket, the maximum number of overs that can be bowled by a bowler is fixed. For example, in a 50-overs-a-side match a bowler can bowl a maximum of 10 overs and in case of Twenty20 matches it is only four. Thus, it is mandatory that the fielding team needs to employ at least five bowlers (including all-rounders) in a complete innings. The constraints discussed earlier shall take care of this restriction. However, the authors feel that an optimum team shall have six bowling options (including the all-rounders). Accordingly, a constraint is proposed.

$$\sum_{i=2}^{k} \theta_i (d_i + e_i + g_i) \geq 6 - \left(l_3' + l_4'\right) p \tag{27}$$

The constraint (27) ensures that the team has not less than six bowling options including the all-rounders. In case the captain is an all-rounder or a bowler then the team needs to have at least five more bowling options. But if the captain is a batsman (opener or middle-order) or wicket keeper then not less than six bowling options including the all-rounders are necessary. This issue is negotiated by $(l_3'+l_4')\boldsymbol{p}$. The vector $\boldsymbol{p}$ is related to the captain and can take any value laid down in Table 3, depending on the expertise of the captain. Values of the vectors $l_3'$and l_4' are provided in Table 2. It can be seen that the term $(l_3'+l_4')\boldsymbol{p}$ results to 1 if the captain is an all-rounder (any expertise) or a bowler (any type either fast or spin) and 0 otherwise. However, this constraint is optional. If a captain is confident on the performance of his bowlers he may not pick up a sixth bowling option. But the authors feel that in these days of power cricket with lots of Twenty20 cricket around a team needs to have an additional bowling option. This shall provide a protection to the captain in case one of the regular bowlers goes for lots of runs. All these constraints from (20) to (27) shall be used while the optimization function is given in (28). The issue is to maximize Z, where,

$$Z=\sum_{i=2}^{k}\theta_i S_i \tag{28}$$

4 Data consideration

In order to validate the model, the data is collected from the fifth season of Indian Premier League (IPL). The IPL is the first franchisee based cricket tournament initiated by the Board of Control for Cricket in India (BCCI), where reputed international players team up with the upcoming Indian talent. In the fifth season of IPL (IPL-V), nine teams participated. The teams were named after Indian cities or states (provinces) but were owned by franchisees. The teams are named as Chennai Super Kings (CSK), Deccan Chargers (DC), Delhi Daredevils (DD), Kings XI Punjab (KXIP), Kolkata Knight Riders (KKR), Mumbai Indians (MI), Pune Warriors India (PWI), Rajasthan Royals (RR) and Royal Challengers Bangalore (RCB). The teams played each other twice in a home and away basis. At the end of the league the top four teams qualifies for the play-offs. The play-offs comprise of three matches including the final. A total of 76 matches are scheduled but 74 matches were actually played. Two matches were abandoned because of bad weather. The different teams have captains with different specialization and each team played a considerable number of matches in a span of 54 days. Hence, data from this tournament is used to select the optimal teams from each of the nine squads.

For measuring the performance of cricketers, players who have performed in a larger number of games should be considered. The actual quality of a player may not be properly judged from one or two games. The effects of outstanding or poor, single performances are smoothed over the larger number of games [16]. The nature of professional sport ensures that the majority of individuals will experience sufficient match-play to enable this type of methodology to be deployed [24]. Different authors depending on the length of the series and the format of the game have considered different benchmarks for a player to be included in the study. The criteria shall be determined in such a way that along with sufficient number of matches played, more number of cricketers can be considered in the training sample. Also there shall be comparable representation of batsmen and bowlers from each team. After several trials the following criteria are decided.

Those batsmen who had played (i) at least three innings in IPL V (ii) faced at least 50 deliveries in IPL V and (iii) had a batting average≥12 are considered. Bowlers who had played (i) at least three matches in IPL V (ii) bowled at least 10 overs (60 deliveries) in

IPL V and (iii) dismissed at least 3 batsmen are included in the training sample. Given the above-mentioned selection criteria for batsmen, 5 each from CSK, RCB and RR, 6 each from DC, KXIP and MI, 7 each from DD and PWI and 4 from KKR are selected. The total number of batsmen accordingly included in the training sample is 51. With the bowling criteria mentioned above, 55 bowlers are included with a breakup of 7 each from RCB, PWI and DC; 6 each from DD and KKR while MI contributed 4 bowlers and RR contributed 8 bowlers rest of the teams contributed 5 bowlers each. For selection of wicket keepers, only those who have kept wickets for at least five matches are considered. A total of 13 wicket keepers are selected. Those players who are selected both as bowler and as batsman under the given criteria are termed as all-rounders. This has led to the selection of 19 all-rounders. Table 4 shows the squad size of the different teams that participated in IPL along with the number of selected players and the ability of the captains of the corresponding teams.

In Table 4, figures in parenthesis in column 2 and 3 represent (number of foreign players, number of Indian players). Though KC Sangakkara is a wicket keeper as well as batsman but in IPL V, he did not keep wickets for his team. Similarly, R Dravid generally bats at number three but in IPL V, he opened all the innings for his team.

To compute the performance statistics, match-wise information of IPL V for the selected batsmen (runs scored, balls faced, number of innings, out/not out etc.), bowlers (runs conceded, balls bowled, wickets taken, position of wickets in the batting order, matches played, etc.) and wicket keepers (catches taken, stampings, matches played, bye runs conceded, etc.) are collected from http://www.espncricinfo.com/indian-premier-league-2012/. In the IPL, there is a restriction that the playing XI of a given team shall not have more than four foreign players. To implement this regulation in the optimization another binary coefficient h_i is defined.

h_i=1 (0), if the i^{th} player is a foreigner (Otherwise)

Table 4 Squad size and information about the captain of the teams in IPL V

Team	Number of players		Captain	Captain's ability
	In the squad	Selected for optimization		
Chennai Super Kings	25 (10,15)	14 (6, 8)	MS Dhoni	Wicket keeper and middle order batsman
Deccan Chargers	34 (9, 25)	15 (6, 9)	KC Sangakkara	Opening batsman
Delhi Daredevils	32 (10,22)	15 (5, 10)	V Sehwag	Opening batsman
Kolkata Knight Riders	26 (10, 16)	15 (7, 8)	G Gambhir	Opening batsman
Kings XI Punjab	32 (12,20)	15 (6,9)	AC Gilcrist	Wicket keeper and opening batsman
Pune Warriors India	33 (12, 21)	17 (8,11)	SC Ganguly	Opening batsman
Mumbai Indians	34 (11, 23)	14 (6, 8)	Harbhajan Singh	All-rounder
Rajasthan Royals	30 (8, 22)	17 (6,11)	R Dravid	Opening batsman
Royal Challengers Bangalore	31 (12,19)	14 (6, 8)	DL Vettori	Spinner

Source: The data in the table is compiled from www.espncricinfo.com

Accordingly, another constraint is defined. The constraint will not allow the selection of more than 4 foreign players in the playing XI.

$$\sum_{i=1}^{k} \theta_i h_i \leq 4 \quad \text{(if the captain of the team is an Indian)} \tag{29a}$$

or

$$\sum_{i=1}^{k} \theta_i h_i \leq 3 \quad \text{(if the captain of the team is a foreign player)} \tag{29b}$$

Thus, the formulation of selecting the optimized balanced playing XI under the cricketing and IPL restrictions considering the specialty of the captain is achieved. In the following section the optimization for the different teams are attained.

5 Results and discussions

Using Solver add-in of Microsoft Excel the objective function (28) is maximized for each team separately subject to the given constraints. The constraints shall take a note that a balanced squad of XI players is selected, given the IPL restriction that a playing XI cannot have more than four foreign players and the expertise of the captain. Table 5 provides the optimum playing XI of the nine teams. The first name is that of the captain and gets an entry into the team because of his position and the other positions are decided by the optimization technique. The strength of the optimum team (maximum value of Z) for each of the IPL squad is provided in the seventh column of Table 6. Amongst the different IPL squads the optimum team of KKR has the maximum score

Table 5 Optimum balanced playing XI of the teams under IPL restriction

CSK	DC	DD	KKR	KXIP	PWI	MI	RCB	RR
MS Dhoni ✋	KC ✈ ∏ Sangakkara	V Sehwag ∏	G Gambhir ∏	AC ✈ Gilchrist ✋	SC Ganguly ∏	Harbhajan Singh ∏⦸	DL Vettori ⦸ ✈	R Dravid ∏
M Vijay ∏	PA Reddy ∏	M Jayawardene ∏ ✈	DB Das ∏	A Mahmood ∏ ⦸ ✈	M Manhas ∏	AT Rayudu ∏	CH Gayle ∏ ✈	AM Rahane ∏
MEK Hussey ∏ ✈	JP Duminy ∏ ✈	KP ∏ ✈ Pietersen	LR Shukla ∏	Gurkeerat Singh ∏	MK Pandey ∏	DR Smith ∏ ✈	MA Agarwal ∏	BJ Hodge ∏ ✈
S Badrinath ∏	S Dhawan ∏	Y Nagar ∏	MK Tiwary ∏	Mandeep Singh ∏	SPD Smith ∏ ✈	RG Sharma ∏	SS Tiwary ∏	OA Shah ∏ ✈
SK Raina ∏	Anand Rajan ⦸	AB Agarkar ⦸	L Balaji ⦸	SD Chitnis ∏	AB Dinda ⦸	SR Tendulkar ∏	V Kohli ∏	A Chandila ⦸
DE Bollinger ⦸ ✈	A Ashish Reddy ⦸	M Morkel ⦸ ✈	R Bhatia ⦸	AD Mascarenhas ⦸ ✈	M Kartik ⦸	MM Patel ⦸	AB McDonald ⦸ ✈	Pankaj Singh ⦸
BW ✈ Hilfenhaus ⦸	A Mishra ⦸	S Nadeem ⦸	SP Narine ⦸ ✈	Harmeet Singh ⦸	MN Samuels ∏⦸ ✈	PP Ojha ⦸	HV Patel ⦸	SK Trivedi ⦸
R Ashwin ⦸	DW Steyn ⦸ ✈	UT Yadav ⦸	JH Kallis ∏⦸ ✈	P Awana ⦸	R Sharma ⦸	SL Malinga ⦸ ✈	P ⦸ Parameswaran	J Botha ∏⦸ ✈
SB Jakati ⦸	V Pratap Singh ⦸	VR Aaron ⦸	I Abdulla ⦸	P Kumar ⦸	WD Parnell ⦸ ✈	JEC Franklin ∏⦸ ✈	Z Khan ⦸	SR Watson ∏⦸ ✈
DJ Bravo ∏⦸ ✈	DT Christian ∏⦸ ✈	IK Pathan ∏⦸	YK Pathan ∏⦸	RJ Harris ⦸ ✈	AD Mathews ∏⦸ ✈	KA Pollard ∏⦸ ✈	Vinay Kumar ∏⦸	STR Binny ∏⦸
RA Jadeja ∏⦸	PA Patel ✋	NV Ojha ✋	BB McCullum ✈✋	PP Chawla	RV Uthappa ✋	KD Karthik ✋	AB de Villiers ✋ ✈	DH Yagnik ✋

✋ -Wicket Keeper ⦸- Bowler ∏⦸- All-rounder ∏ -Batsman ✈- Foreign Player

Table 6 Teams' performance, average and optimum team strength in IPL V

Team (1)	Matches played (2)	Won (3)	Percentage of winning (4)	Average no. of players common with optimum team (5)	Team strength		Difference between average and optimum strength (8)
					Average (6)	Optimum (7)	
CSK	18	10	55.56	9	4.0997	4.6276	0.5279
MI	17	10	58.82	9	3.9939	4.0257	0.0318
RCB	15	8	53.33	9	4.2695	4.5336	0.2641
KKR	17	12	70.59	9	4.5637	4.975	0.4113
DD	18	11	61.11	8	3.9727	4.0699	0.0972
DC	15	4	26.67	8	3.2509	3.4776	0.2267
PWI	16	4	25	8	3.5101	3.7968	0.2867
KXIP	16	8	50	7	3.5752	3.7423	0.1671
RR	16	7	43.75	7	2.8957	3.7786	0.8829

(i.e., 4.975) followed by CSK (4.6276) and RCB (4.5336). The least optimum value of Z is attained by DC with a score of 3.4776.

Sometimes, based on the strength of the opposition or playing conditions a given team may need to include more number of batsmen or more spinners or more all-rounders etc. Such issues can be handled by adjusting the constants in the right side of the constraints (Eqs. (21) to (27)). However, it is obvious that in no case the constraint (20) shall be violated.

It is noticed that in none of the matches the teams played exactly with the optimum balanced playing XI as displayed in Table 5. However, each team had many players in the playing XI common with the optimum teams in the different matches. Table 6 shows the position of the other teams in relation to their respective optimum teams. On an average CSK, MI, RCB and KKR had nine players in common with their respective optimum teams. The average team strength of each of the franchisees is computed. The team strength of a given team in a given match is attained by adding the values of S_i for all players of the team who made into the playing XI in that match. This is obtained for all the matches played by the team in the tournament and averaged over all the matches to attain the average team strength. It may be noted that the average team strength of some of the franchisees are much different from the strength of the corresponding optimum teams like RR, DD, MI and DC (*c.f.* Table 6: Column 8). The minimum difference between average team strength and optimum team is recorded for CSK followed by KKR. It is interesting to note that both CSK and KKR reached the finals of IPL V; this may be attributed to the fact that they had a balanced team and in most cases they fielded in a team relatively close to the optimum team. However, such claim needs verification for subsequent seasons of IPL. But one has to remember that choosing the optimum team is not the only key to success as the squad needs to be balanced and shall contain superlative performers under each expertise.

Using the same model one can pick up an 'All-Star' IPL team for that season. An 'All-Star' team is actually a fantasy team that is formed by the experts/fans/sports journalists after a tournament is over by taking players from all the teams. This practice is also common in several other team sports like football, hockey, basketball, baseball, etc. In cricket, the ‘All-Star’ team is formed with performers of the tournament from different expertise viz. batting, bowling or wicket keeping. Such ‘All-Star’ teams are only fantasy teams, and such a team never actually plays any match, yet selection of a player in an ‘All-Star’ team is recognition of player’s performance in the tournament. The ‘All-Star’ team is subjectively selected by experts/fans/sports journalists after the tournament. However, the current model shall objectively select the ‘All-Star XI’. Since the current model needs the captain to be named earlier, so the captain of the champion team KKR (i.e., G Gambhir) is considered as the captain of the ‘All-Star XI’ of IPL V. With Gambhir being an opening batsman, one place for the opener in the team is filled up by him. Gambhir’s selection as captain shall generate $p' = (1, 0, 0, 0, 0)$. For selecting the ‘All-Star XI’ the optimization function remains the same as (28). But with G Gambhir being named as the captain of the ‘All Star’ team the other constraints changes to,

$$\sum_{i=2}^{k} \theta_i b_i = 1 \qquad \text{\# to ensure that team has exactly two opening batsmen including Gambhir}$$

$$\sum_{i=2}^{k} \theta_i c_i \geq 2 \qquad \text{\# to ensure that team has at least two middle order batsmen}$$

$$\sum_{i=2}^{k} \theta_i d_i \geq 2 \qquad \text{\# to ensure that team has at least two spin bowlers}$$

$$\sum_{i=2}^{k} \theta_i e_i \geq 2 \qquad \text{\# to ensure that team has at least two fast bowlers}$$

$$\sum_{i=2}^{k} \theta_i f_i \geq 1 \qquad \text{\# to ensure that team has at least one wicket keeper}$$

$$\sum_{i=2}^{k} \theta_i g_i \geq 1 \qquad \text{\# to ensure that team has at least one all} - \text{rounder}$$

$$\sum_{i=2}^{k} \theta_i (d_i + e_i + g_i) \geq 6 \qquad \text{\# to ensure that team has at least six bowling options}$$

Since, Gambhir is an Indian player so constraint (29a) shall be used to handle the restriction on the number of foreign players to be included in any IPL team.

$$\sum_{i=1}^{k} \theta_i h_i \leq 4 \qquad \text{\# to ensure that team has a maximum of 4 foreign players}$$

The ultimate team is provided in Table 7. The total team strength of the ‘All-Star’ team is 6.6616, which is higher than the team strengths of the individual optimum teams (*c.f.* Table 6, Column 7).

The ‘All-Star’ team has two openers (one of whom is the captain himself and the other one can also keep wickets), two fast bowlers, two spinners, two all-rounders, three middle order batsman which includes a wicket keeper as well. Amongst the IPL teams, KKR contributes five players and CSK three players to the ‘All-Star’ team while RR, KXIP and PWI contributes one player each. There is no representation to the ‘All-Star XI’ from DC, DD, RCB and MI. In terms of the countries, there are seven Indian

Table 7 The All Star team of IPL season V

Player's name	Team	Score	Country	Type
G Gambhir	KKR	0.5953	India	Captain and Opener
R Ashwin	CSK	0.4268	India	Spinner
DJ Bravo	CSK	0.8844	West Indies	All rounder
MS Dhoni	CSK	0.6392	India	Batsman and keeper
SPD Smith	PWI	0.8455	Australia	Batsman
L Balaji	KKR	0.4821	India	Fast Bowler
R Bhatia	KKR	0.3394	India	Fast Bowler
BB McCullum	KKR	0.9736	New Zealand	Opening Batsman and keeper
SE Marsh	KXIP	0.5626	Australia	Batsman
A Chandila	RR	0.4723	India	Spinner
YK Pathan	KKR	0.4404	India	All rounder
Total Team Strength ($\sum S_i$)	6.6616			

players in the 'All-Star XI', two from Australia and one each from West Indies and New Zealand.

6 Conclusion

In cricket, the team management goes for a subjective selection of players but obviously with an aim of offering the best possible balanced XI from an available squad of players. The method discussed here provides an objective way of selecting players taking into account the cricketing requirements, tournament rules and specialization of the captain. This model of selecting the optimum XI needs extension for accommodating issues like strength of the opponent, change of captaincy of a team due to injury or otherwise, pitch and weather condition etc. The optimum selection tool discussed in this paper is ahead of other such existing tools as it can take into consideration the capability of the captain and consider the need of fast bowlers, spinners, openers and the middle order batsmen separately in a team. As Twenty20 format of cricket is highly scoring, so a team shall have some additional bowling options. Such options may come handy when any (or more) of the regular bowler(s) concedes lot of runs. The current tool even takes care of that by including at least six bowling options.

An extension to this research endeavor can be used for selecting a team for the next match when a tournament is in progress; considering the values of performance statistics of cricketers up to the previous match. This shall make the optimum team dynamic. At the end of each match, the performance related statistics of the players can be fed into the system and a fresh optimum team can be generated. Thus, the optimum team can change after every game of a tournament, providing a basis of selecting the playing XI for the next match by the team management. Though the numerical example is related to Twenty20 format of cricket, yet the method of team selection can well be applied to the longer formats of the game like 50 overs cricket and test matches with necessary modification to the performance statistics of cricketers and the set of constraints. Besides cricket the model with necessary modification in performance measurement and the set of constraints can be applied in other sports like soccer, American football, ice and field hockey and several other team sports and even to the selection of project teams. Also, instead of considering performance of cricketers as crisp quantities- stochastic or fuzzy techniques can be applied in modeling the uncertainties involved in performance measurement, which can be viewed as a future extension of this work.

However, such models are yet to replace the subjective aspects like strength of the opponent, pitch/ground conditions, experience of players, format of the game etc. which are kept in mind by the team management while selecting the playing XI for a match. But such optimization models are a means of support which the team management needs to include, for identifying the right person for the right place. Though this exercise shall provide the best team considering the previous performance of the players, from the set of available players, yet it is not necessary that the best team on paper is always the best in their performance. Cricket is a game of uncertainties and Statistics is a science of averages. So such studies are only expected to provide better result than subjective selection. The optimization tools are widely used in decision making in several other fields like business, policy formation, traffic control etc. but their application in sports is still at its infancy. Though researchers have applied several such optimization models for decision making in sports yet they are hardly applied by the end users viz. team management, coaches, captains, etc. To bridge the gap, in one hand the team management in different sports should be exposed to the capability of such optimization tools and on the other hand the researchers have to overcome a lot of resistance (in terms of quantification) owing to the subjectivity involved in such team selection.

Acknowledgments The authors would like to thank Mr. Dhriti S. Gupta and Mr. Pulak Bora, Assistant Professors of Business Communication, School of Business, The Assam Kaziranga University, Jorhat-785006 (Assam) for their help in English language editing of the manuscript.

References

1. Saikia, H., Bhattacharjee, D.: On classification of all-rounders of the Indian premier league (IPL): a Bayesian approach. Vikalpa: J. Decis. Mak. **36**, 51–66 (2011)
2. Lemmer, H.H.: Team selection after a short cricket series. Eur. J. Sport Sci. (2011). doi:10.1080/17461391.2011.587895
3. Salas, E., Rosen, M.A., Burke, C.S., Goodwin, G.F., Fiore, S.M.: The making of a dream team: when expert teams do best. In: Ericsson, K.A., Charness, N., Feltovich, P.J., Hoffman, R.R. (eds.) The Cambridge Handbook of Expertise and Expert Performance, pp. 439–453. Cambridge University Press, New York (2006)
4. Kamble, A., Rao, R., Kale, A., Samant, S.: Selection of cricket players using analytical hierarchy process. Int. J. Sports Sci. Eng. **5**, 207–212 (2011)
5. Ahmed, F., Jindal, A., Deb, K.: Multi-objective optimization and decision making approaches to cricket team selection. Appl. Soft Comput. (2012). doi:10.1016/j.asoc.2012.07.031
6. Barr, G.D.I., Kantor, B.S.: A criterion for comparing and selecting batsmen in limited overs cricket. J. Oper. Res. Soc. **55**, 1266–1274 (2004)

7. Gerber, H., Sharp, G.: Selecting a limited overs cricket squad using an integer programming model. S. Afr. J. Res. Sport Phys. Educ. Recreat. **28**, 81–90 (2006)
8. Lourens, M.: Integer optimization for the selection of a twenty20 cricket team. Nelson Mandela Metropolitan University. http://statssa.gov.za/isi2009/ScientificProgramme/IPMS/0738.pdf (2010). Accessed 21 June 2010
9. Brettenny, W.: Integer optimisation of the selection of a fantasy league cricket team. Nelson Mandela Metropolitan University. http://dspace.nmmu.ac.za:8080/jspui/bitstream/10948/…/W20%B.pdf (2010). Accessed 21 December 2010
10. van Staden, P.J.: Comparison of cricketers' bowling and batting performances using graphical displays. Curr. Sci. **96**, 764–766 (2009)
11. Beaudoin, D., Swartz, T.: The best batsmen and bowlers in one-day cricket. S. Afr. Stat. J. **37**, 203–222 (2003)
12. Singh, G., Bhatia, N., Singh, S.: Fuzzy cognitive maps based cricket player performance evaluator. Int. J. Enterp. Comput. Bus. Syst. **1**, 1–15 (2011)
13. Damodaran, U.: Stochastic dominance and analysis of ODI batting performance: the Indian cricket team, 1989–2005. J. Sports Sci. Med. **5**, 503–508 (2006)
14. Boon, B.H., Sierksma, G.: Team formation: matching quality supply and quality demand. Eur. J. Oper. Res. **148**, 277–292 (2003)
15. Das, D.: Moneyballer: an integer optimization framework for fantasy cricket league selection and substitution. http://debarghyadas.com/files/IPLpaper.pdf (2014). Accessed 11 July 2015
16. Lewis, A.: Towards fairer measures of player performance in one-day cricket. J. Oper. Res. Soc. **56**, 804–815 (2005)
17. Iyenger, N.S., Sudarshan, P.: A method of classifying regions from multivariate data. Econ. Pol. Wkly **5**, 2048–2052 (1982)
18. Lemmer, H.H.: An analysis of players' performances in the first cricket twenty20 world cup series. S. Afr. J. Res. Sport Phys. Educ. Recreat. **30**, 73–79 (2008)
19. Lemmer, H.H.: The combined bowling rate as a measure of bowling performance in cricket. S. Afr. J. Res. Sport Phys. Educ. Recreat. **24**, 37–44 (2002)
20. Lemmer, H.H.: A method for the comparison of the bowling performances of bowlers in a match or a series of matches. S. Afr. J. Res. Sport Phys. Educ. Recreat. **27**, 91–103 (2005)
21. Narayanan, A.: Test wicket keepers - an analysis. The ESPNcricinfo Blog. http://blogs.cricinfo.com/itfigures/archives/wicketkeepers/ (2008). Accessed 4 March 2010
22. Lemmer, H.H.: A measure for the batting performance of cricket players. S. Afr. J. Res. Sport Phys. Educ. Recreat. **26**, 55–64 (2004)
23. Chakrabarty, N., Bhattacharjee, D.: Aggregation and weightage issues concerning composite index development: experience with digital divide in Asian countries. Asia Pac. J. Libr. Inf. Sci. **2**, 24–37 (2012)
24. Bracewell, P.J., Ruggiero, K.: A parametric control chart for monitoring individual batting performances in cricket. J. Quant. Anal. Sports **5**, 1–19 (2009)