Integer Optimisation for the Selection of a Fantasy League Cricket Team

by:
Warren Brettenny

Submitted in fulfilment of the requirements for the degree of
Magister Scientiae
in the
Faculty of Science
at the
Nelson Mandela Metropolitan University

January 2010

Supervisor: Mr G.D. Sharp
Co-Supervisors: Mr D.G. Friskin & Prof J.W. Gonsalves

Abstract

Sports fans often scrutinise the team selection strategies employed by their favourite team's coach or selection panel. Many of these fans believe that they can perform the selection process far better than those tasked with the responsibility. Fantasy leagues, provide a platform for fans to test their hand at this selection procedure. Twenty20 cricket is a new and exciting form of cricket and has become very popular in recent years.

This research focuses on bringing these concepts together by proposing a binary integer program to determine a team selection strategy for fantasy league cricket. This is done in a Twenty20 setting. The approach used in this study focuses on evaluating the effectiveness of previously developed performance measures in a fantasy league setting. Adjustments to these measures are made and new measures are proposed. These measures are then used to select a fantasy league team using a prospective approach. This is done to provide fantasy league participants with a mathematical procedure for fantasy league team selection.

Keywords: Fantasy leagues, Twenty20 cricket, optimisation, binary integer programming, team selection.

Acknowledgements

I would like to thank my supervisors, Gary Sharp and David Friskin, for all their patience and understanding throughout this year (and last year). Your constant support and encouragement has been a great help and I can not thank either of you enough. John Gonsalves, thank you for your valued input.

To my fiancé Amy. Your love, kindness and patience is the sole reason that I was able to complete this work.

My family, Mom, Dad and Kevan, I am so grateful to you. Your support and encouragement through this year and all through my studies has been a driving factor in all that I have achieved so far.

To the external examiner, Prof Tim Swartz, for his valued input and suggestions.

To my Lord and Saviour, my Heavenly Father. Thank You for Your guidance and love that I could not possibly live without.

For their financial support, I wish to expressly thank the following institutions:

- The National Research Foundation (NRF),
- The Nelson Mandela Metropolitan University (NMMU) RCD Department, and
- The Dormehl-Cunningham Trust.

Thank you.

Contents

List of Figures

List of Tables

List of Performance Measures

Introduction and Objective

Whether a participant or a supporter, sport forms an important part of modern day society. Participants who excel in an activity can often pursue the sporting discipline professionally and those who are less capable but still actively engaged can remain amateur participants at many different levels. Sporting fans can be both spectators and participants. Fans with limited sporting prowess can still be active supporters. For these fans, fantasy league participation provides an ideal platform to voice their ideas and test their knowledge.

It seems that when it comes to team selection, at both a domestic and an international level, controversy is unavoidable. One only needs to turn to the back page of ones local newspaper to discover this. As an example recent sport headlines discussed the selection of South African football striker Benni McCarthy in the national team (Amato, 2009). Hardly a day goes by in the sporting world where a question doesn't arise about a certain manager's tactics or recent player performance.

It often seems that the supporting public have an opinion on everything, but especially the team selection tactics used by a coach or manager. The public frequently believe that they can perform the selection procedure better than those tasked with the responsibility, and thus the team selection strategies of professional sports coaches and managers are subject to enormous scrutiny.

Supporters claiming they know how to better select a team than the professional coaches and managers have not, until recently, been able to test or prove their claims. The opportunity to do so has presented itself in the form of a "fantasy league." A fantasy league provides the participants with an idea of the team selection problems facing professional coaches. As such, fantasy leagues encourage thorough research and planning by their participants (Elaine Allen, Kustov and Recck, 2007).

The introduction of fantasy leagues onto the Internet increased the availability of these leagues to the public (Vichot, 2009). There are now numerous fantasy leagues available for football (soccer), cricket, rugby and baseball, to name just a few. These online fantasy leagues can attract large numbers of participants. For example, a fantasy league based on the FA English Premier League has over 2 000 000 participants (Official EPL Fantasy League, 2009). These leagues are also taken very seriously as the prizes for winning are often substantial. A closer look at fantasy leagues and how they work is provided in Section 1.2.

Summers, Swartz and Lockhart (2007) developed a mathematical strategy to draft players in hockey pools. Their research provides a detailed method for the construction and selection of fantasy league teams in hockey pools. Their publication is a pioneering study into the strategies and methodologies involved in fantasy league team selection.

The focus of this study, however, is a fantasy league based on a Twenty20 cricket competition, namely the 2008 Indian Premier League. The aim of this research is to use mathematical and statistical techniques in order to facilitate the fantasy league team selection process for Twenty20 cricket.

Chapter 1 provides a literature review which investigates previous statistical research into cricket. Included in this chapter is a brief explanation of the history and game of cricket as well as the history and concept of fantasy leagues. Chapter 2 describes the fantasy league under consideration in this study and the data collected for analysis. Chapter 3 provides the methodology used in this study. The results of the described methodology are provided and discussed in Chapter 4. Chapter 5 concludes the study and provides recommendations for further work.

Chapter 1

Literature Review

1.1 Cricket

Cricket is a traditional English bat and ball sport which is steeped in history. It is believed by some that the first reference to the game was made as early as the 13$^{\text{th}}$ century (Leach, 2007). However, the first official evidence of the game being played is dated in the 1550s (Norridge, 2008). By the 18th century the game of cricket boasted its first official set of rules (Clarke, 1998). The game of cricket is now played across the globe with the International Cricket Council (ICC) having around 120 members (Cricket Archive, 2009). The major cricket playing nations (full ICC members) are Australia, Bangladesh, England, India, New Zealand, Pakistan, South Africa, Sri Lanka, West Indies and Zimbabwe (Cricket Archive, 2009).

The general overview of a limited overs match is discussed with reference to Preston and Thomas (2002). The game is played between two teams, each consisting of eleven players. During the course of a match each team has the chance to bat and to bowl. When the one team bats, the other team will bowl (and vice versa), this is known as an "innings". The objective of the team batting first is to score as many runs as possible. In so doing the batsmen are exposed to the risk of losing their wicket, also termed being "dismissed" or "out" (Preston and Thomas, 2002). The bowling team aims to dismiss the batting team whilst restricting the number of runs scored. Once this innings is completed, the team who batted initially will bowl and vice versa. The total number of runs scored by the team who batted first is the target for the team batting second (Preston and Thomas, 2002). Typically, the team that has scored the most runs after the game has concluded is deemed the winner (Preston and Thomas, 2002). Exceptions to this include games which have been influenced by rain, where a mathematical procedure can be used to determine the winner.

Each innings is divided into intervals of at least six balls or deliveries, these intervals are called "overs" (Preston and Thomas, 2002). A ball (delivery) is completed when a bowler runs towards the pitch and delivers the ball to the batsman, who has taken guard at the opposite end of the pitch. The bowler will complete the over that was started, and no bowler can bowl two consecutive overs. Once an over is completed the bowling team swaps sides, and the next bowler's over will be bowled from the opposite end of the pitch. All the members of the bowling team who are not bowling are strategically positioned around the field in order to assist the bowler (known as "fielders").

Batsmen bat in pairs, each positioned on opposite ends of the pitch during any given delivery. The batsman to which the bowler is delivering the ball (i.e. facing the bowler) is said to be on "strike", while the other batsman is positioned on the "non-strikers" (bowlers) end (Clarke and Norman, 1999). The batsman who is on strike takes guard in front of three wooden stumps on top of which are two bails, this is know as the wicket. Only the batsman who is on strike when a ball is delivered can score runs. Runs are scored when the batsman strikes the ball using their bat. It must be noted however that a batsman need not run when the ball is struck, and any run taken when the ball has not struck the bat will be credited to the team's total, but not the batsman's. Runs are scored after a delivery when both batsmen run the length of the pitch (i.e. change ends). The decision to complete a run is made by the batsmen depending on the position of the ball after being struck. Should the ball be hit to the boundary (known as "hitting a boundary") while having touched the ground, the batsman is credited with four runs; if the ball is struck over the boundary (not touching the ground) the batsman is credited with six runs (Clarke, 1998). Provided a boundary is not hit, the batsmen can complete as many runs as they choose in the time it takes for the ball to be returned to either of the wickets (on each end of the pitch). Should the batsmen not complete an intended run by the time the ball returns to the wicket and the wickets are broken (bails dislodged), the batsman who is running toward the end with the broken wicket is dismissed (known as a *run out*).

The other common methods of dismissal are (Britannica Online Encyclopedia: Cricket - Methods of Dismissal, 2009): bowled - this occurs when the wickets are broken by the ball during the course of a delivery; LBW (leg before wicket) - which occurs if a delivery (ball) which is on its way to break the wicket is intercepted by the batsmen using their body (i.e. not their bat); and caught - which occurs when the ball is struck by the batsman and caught by a fielder without touching the ground. Other less common methods of dismissal have been omitted, the interested reader is referred to Britannica Online Encyclopedia[1] for a complete list of dismissal methods.

The game of cricket is played in two basic formats namely double innings matches and single innings (limited overs) matches. The double innings match is the oldest, and arguably the purest, form of the game. These matches consist of four innings, played over four days for South African domestic competitions and five days for international competitions (known as Test matches). Each team gets to bat (and bowl) twice, and a winning result is only achieved by a team if they score more runs than their opposition (combined over both innings) *and* dismiss all the opposition batsmen within the time allocated to the match. Should the latter condition not be met, then the match results in a draw.

Limited overs cricket began domestically in England in 1963, and in 1971 the first international game was played (Cricinfo: A Brief History of Cricket, 2009). The main difference between test cricket and limited overs cricket is that the number of overs each team gets to bowl is predetermined. The most familiar form of the limited overs game is the One Day International (ODI) which is named as such since the game is scheduled to be completed in a single day. This form of the game allows each team to bowl 50 overs (with each bowler bowling a maximum of 10 overs), and the objective is to score more runs than the opposition in the allotted overs. The

[1] http://www.britannica.com/EBchecked/topic/142911/cricket/214754/Methods-of-dismissal. (Accessed 8 December 2009)

bowling team need not dismiss the entire batting team as a result will be obtained once both teams have completed the overs allocated to them.

Initially spectator interest was high, however over time crowd attendances began to decrease and in 2003 the cricket authorities in England proposed a new format to combat this (Weaver, 2008). This adaptation, commonly known as Twenty20 cricket, is the format that this study investigates.

1.1.1 Twenty20 Cricket

Twenty20 cricket is a highly entertaining and crowd friendly short form of cricket. Since its inception in 2003, Twenty20 cricket has gained tremendous support worldwide.

Similar to the introduction of the one day (limited over) game 40 years earlier, the first Twenty20 competition was held domestically in England in 2003. The game itself is much shorter than the one day game. The format allows for a twenty overs per innings contest and is scheduled to last just three hours. Before the game was introduced a thorough market survey was conducted (Shilbury, Westerbeek, Quick and Funk, 2009). The aim of the survey was to gauge the public opinion of the game of cricket, and to determine the expected public response to the introduction of the new format (All Out for No Loss, 2007).

The results of the survey were favourable to the introduction of the new format, and on the 13$^{\text{th}}$ July 2003, the first Twenty20 game was played (Weaver, 2008). The rules of a Twenty20 game are, for the most part, the same as those of the regular one day game, however there are some differences. These differences include (Watson, 2008):

- Each bowler can bowl a maximum of one fifth of the allocated overs. For matches not interrupted by rain this maximum is 4.
- If a "front-foot no ball" (which occurs when a bowler, in his delivery stride, oversteps the popping crease) is bowled, the batting team is awarded a run. Furthermore, the delivery is re-bowled and this delivery is designated a "free-hit", where the batsman now can only be dismissed if he is run out, hits the ball twice or handles the ball.
- Fielding restrictions:
  - During the power-play (the first 6 overs), no more than 2 fielders can field outside the 30-yard circle.
  - After the power-play, no more than 5 fielders can field outside the 30-yard circle.
  - At any time, a maximum of 5 fielders are allowed on the leg side.

A further difference between the two limited over formats, is that should a Twenty20 match be completed (i.e. not abandoned for any reason) then a win/lose result will always be obtained. Should a match be affected by rain then a mathematical procedure is implemented to determine a winner (Duckworth and Lewis, 1998). Should a match be tied, that is, should both teams score the same amount of runs in their respective innings, two methods have been used to decide a winner. Originally a "bowl out" was implemented. A "bowl out" consists of five bowlers from each team each bowling a single ball at an unguarded wicket. Whichever team hits the wicket the most in their five attempts is deemed the winner. Should the scores still be tied, the match is decided using a sudden death approach (Watson, 2008). The "bowl out" has since been replaced by a "super over". This study only considers the "bowl out" as a tie decider, and thus the "super over" is not discussed. The interested reader is referred to the 2009 ICC Standard Twenty20 International Match Playing Conditions for the rules of this approach.[2].

This adaptation of the game has become increasingly popular, and given that it is still a "young" format, little research has been conducted into this form of cricket. Two Twenty20 competitions held between September 2007 and June 2008, namely the Inaugural ICC World Twenty20 and the 2008 Indian Premier League respectively, form the basis of this research.

1.2 Fantasy Leagues

Fantasy sport has millions of participants worldwide. The appeal of the game presumably finds its roots in the competitive nature of many sports supporters (Elaine Allen, Kustov and Recck, 2007). Fantasy sport provides a platform for sport followers to test their ability at one of the more controversial aspects of most sports, team selection. Whether it be in casual conversation or public debate, the team selection strategies of many sports teams finds itself under scrutiny. Fans of sports teams often believe that their personal team selection is best, and fantasy sport provides these supporters with an opportunity to prove themselves in a simulated setting.

1.2.1 Concept and History

The concept of fantasy sport has its origin as early as the 1940's (after World War II), but the true growth of the game only occurred in the 1980's (Vichot, 2009).

In 1980 Daniel Okrent introduced the "Rotisserie League" for baseball. Up to this point similar leagues had only considered statistical information from past seasons. The Rotisserie league was an innovative concept as the game now consisted of the participant "drafting" players into his fantasy league team at the start of the season. Okrent's Rotisserie league presented the participants with some of the team selection problems and concerns that were faced by real baseball managers, such as player unavailability due to injury. As such, the league required the participants to make predictions about player performance and to adjust their teams according to the circumstances presented to them (Vichot, 2009). Being a magazine writer/editor, Okrent wrote an article for Inside Sports in March 1981. This article outlined the rules of the Rotisserie league game. Owing to this publication and baseball's keen interest in statistics the fantasy league movement began to grow (Vichot, 2009). This growth has continued and with the introduction of leagues onto the Internet, includes leagues based on several sports all across the world. A fantasy league trade group, the Fantasy Sport Trade Association (FSTA), was formed in 1998 (Vichot, 2009). Today the FSTA represents more than 110 member companies and has a market size in the region of 27 million American adults (FTSA, 2009).

The game today has many forms, a common form is now discussed. A fantasy league participant is given a budget in order to purchase players for their team. The players purchased

[2] http://static.icc-cricket.yahoo.net/ugc/documents/DOC_1F113528040177329F4B40FE47C77AE2 _1254317640264_993.pdf. (Accessed: 9 December 2009)

are given a price according to their ability with superior players having higher prices and weaker players having lower prices. The selected team must have a certain formation according to the rules of the sport as well as the rules of the specific fantasy league. The general concept is that all teams should consist of players in all positions of the sport. For example, in a fantasy league dedicated to football (soccer), a participant wouldn't be allowed to select only forwards or only goalkeepers. A team consisting of goalkeepers, defenders, midfielders and forwards would need to be selected, in an acceptable formation[3].

The players are then observed, across a competition or season, and allocated points according to their performance in each game. Players are awarded points for what they did well and deducted points for what was done poorly. The players in a fantasy league participant's team are given points after each game, and the performance of the entire fantasy league team is determined as the sum of the points of all the players in the team. This is done cumulatively through an entire competition or season. A fantasy league participant can make adjustments to their team to maximise their cumulative total, provided these changes do not violate the formation and budget constraints of the league. The fantasy league participant with the highest cumulative total at the end of the competition or season is deemed the winner of the league. Owing to the large number of participants, the prizes for winning such a league are often substantial.

1.3 Statistics in Cricket

The interval nature (i.e. ball by ball, over by over) of the cricket game allows it to be analysed efficiently. Statistics plays an important role in cricket, for example the derivation of a statistical method for determining a winner in rain interrupted matches (Duckworth and Lewis, 1998). Since the introduction of the Duckworth-Lewis method supporters are more aware of the role of statistics in cricket. Basic statistics, however, have been used for many years in order to gauge a player's performance. These statistics, as well as some recently developed statistics, are defined and discussed in the next section.

1.3.1 Traditional Performance Measures

The traditional methods of measuring the ability of a player use common statistical concepts such as the mean or average. The following statistics are traditionally used to measure a cricketers ability.

Batting

The batting average is arguably the most well known and widely used performance measure in the modern game. It is the number of runs scored by a batsman per dismissal, and is defined as

$$AVE_{BAT}^{t} = \frac{\text{total number of runs scored}}{\text{total number of times dismissed}}. \tag{1.1}$$

[3]http://fantasy.premierleague.com/M/rules.mc (Accessed: 19 November 2009), discusses an example of the formation settings used in a fantasy league.

This statistic estimates the batman's ability to score runs. A higher value indicates a more desirable player. The superscript t indicates that this is the traditional definition of the statistic.

Another important statistic which is often referred to in the modern game is the batsman's strike rate. The strike rate is defined as the number of runs a batsman scores per 100 balls faced. This statistic is defined as

$$SR^t_{BAT} = \frac{\text{total number of runs scored}}{\text{total number of balls faced}} \times 100. \tag{1.2}$$

This statistic estimates how quickly the batsman scores runs which is an important requirement in the short (limited over) form of the game. In limited over matches a player with a higher strike rate is often preferred to a player with a lower strike rate. This measure is considered less important in Test match cricket.

Bowling

The two most widely used statistics for bowlers, namely the bowling average and economy rate, are similar to those calculated for batsmen. The bowling average is defined as

$$AVE^t_{BWL} = \frac{\text{total number of runs conceded}}{\text{total number of wickets taken}}. \tag{1.3}$$

This statistic estimates the average number of runs a bowler will concede per wicket taken. A low value for this statistic is desirable.

The economy rate is the bowling equivalent of the batting strike rate. The economy rate of a bowler is defined as

$$ECON^t_{BWL} = \frac{\text{total number of runs conceded}}{\text{total number of balls bowled}} \times 6. \tag{1.4}$$

The economy rate estimates the average number of runs a bowler will concede during a given over. The economy rate is of more importance in the limited over game and it is preferable for the value of this measure to be low.

The last measure in this section is known as the bowling strike rate. This measure differs noticeably from the batting strike rate and is defined as

$$SR^t_{BWL} = \frac{\text{total number of balls bowled}}{\text{total number of wickets taken}}. \tag{1.5}$$

The bowling strike rate is defined as the expected number of balls bowled per wicket taken by a given bowler. Intuitively, it is preferable for this measure to be low.

It must be noted that high values are preferable for the batting measures (Equations 1.1, 1.2) and low values of the bowling measures (Equations 1.3, 1.4 and 1.5) are preferred. This fact plays a role in some of the recently developed measures that will be investigated in the forthcoming sections.

Although these five statistics are commonly used, some analysts claim that they don't always give a fair reflection of a player's true ability. In order to determine more useful measures of a player's performance, alternative measures have been proposed. These techniques are discussed in the sections that follow.

1.3.2 The Combined Bowling Rate (CBR): Lemmer (2002)

In order to rate bowlers, Lemmer (2002) devised a measure known as the combined bowling rate (CBR). The CBR combines the three traditional bowling statistics, namely the bowling average, the bowling economy rate and the bowling strike rate.

Lemmer (2002) used the harmonic mean to combine the above three statistics. The justification for the harmonic mean is referred to the 1954 text by Kenney and Keeping. According to Lemmer(2002), they suggested that the harmonic mean be used to find the average of ratios, provided that the numerator is considered fixed and the denominator as variable. Since all of AVE^t_{BWL}, $ECON^t_{BWL}$ and SR^t_{BWL} are ratios this method was investigated. To simplify, the following variables are introduced

$$\begin{aligned} r &= \text{total number of runs conceded,} \\ w &= \text{total number of wickets taken, and} \\ b &= \text{total number of balls bowled.} \end{aligned}$$

The traditional bowling statistics can now be defined as

$$\begin{aligned} AVE^t_{BWL} &= \frac{r}{w}, \\ ECON^t_{BWL} &= \frac{r}{\frac{1}{6}b}, \text{ and} \\ SR^t_{BWL} &= \frac{b}{w}. \end{aligned}$$

It is observed that the bowling average and the bowling economy rate have the same numerator (r). Thus using the harmonic mean to find the average of these two statistics is valid. In order to include the bowling strike rate it is necessary to adjust SR^t_{BWL} to have the same numerator as AVE^t_{BWL} and $ECON^t_{BWL}$, namely r. Lemmer (2002) proposed the following adjustment to the bowling strike rate

$$SR^t_{BWL} = \frac{b}{w} = \frac{r}{r} \times \frac{b}{w} = \frac{rb}{rw} = \frac{r}{\frac{rw}{b}}. \tag{1.6}$$

Using the definition in Equation 1.6 the bowling strike rate can now be included in the calculation of the harmonic mean. The harmonic mean of a random sample $X_1, \ldots, X_n$ is defined as

$$\bar{X}_{harm} = \frac{n}{\sum\limits_{i=1}^{n} \frac{1}{X_i}}. \tag{1.7}$$

Thus the combined bowling rate (CBR) was defined by Lemmer (2002) as

$$CBR = \frac{3}{\frac{1}{AVE^t_{BWL}} + \frac{1}{ECON^t_{BWL}} + \frac{1}{SR^t_{BWL}}} = \frac{3r}{w + \frac{1}{6}b + \frac{rw}{b}}. \tag{1.8}$$

Lemmer (2002) thereafter defined a classification system (using this measure) to rate the performances of bowlers, with lower values of the CBR indicating more preferable bowlers. It was also demonstrated that this measure provided good results for ranking bowlers.

1.3.3 A Batting Performance (BP) Measure: Lemmer (2004)

Lemmer (2004) proposed a measure to assess the performance of a batsman in limited over cricket. This measure accounted for the average, consistency and strike rate of a batsmen.

The traditional measure for assessing the ability of a batsman is the batting average. This measure gives each innings an equal weighting and in so doing assumes that the batsman's recent scores have the same impact as the scores at the beginning of their career. Lemmer (2004) argued that this was not a realistic approach to the calculation of a batsman's current batting average. It was suggested that an exponentially weighted average (AVE^{exp}_{BAT}) be used in which the most recent score has weight w, the second most recent score has weight $0.96w$, the third most recent score would have weight 0.96^2w and so forth. The author chose the value of 0.96 in accordance with the weights used by PriceWaterhouseCoopers in their 2002 ratings of world cricketers. Suppose that the scores of a batsman are given as $X_1, X_2, \ldots, X_n$. In this case X_n indicates the batsman's most recent score, X_{n-1} represents the batsman's second most recent score and so forth. The exponentially weighted average (AVE^{exp}_{BAT}) is then defined as

$$AVE^{exp}_{BAT} = wX_n + 0.96wX_{n-1} + 0.96^2wX_{n-2} + \cdots + 0.96^{n-1}wX_1 \tag{1.9}$$

where w is calculated so that the the sum of all the weights equals 1.

The consistency of a batsman has been investigated in several research papers. Barr and Van den Honert (1998) investigated batting consistency by the use of the geometric coefficient (G). This measure was based on match scores in test match cricket. They calculated the geometric coefficient as the inverse of the coefficient of variation (CV) in the sample of match scores. Lemmer and Nel (2001) then suggested that the inverse of the geometric coefficient (G) i.e. the coefficient of variation be used as a measure of the consistency of batsmen. It was subsequently noticed that the CV was not an accurate measure of a batsman's consistency as low not out scores and scores high above the average adversely influenced the standard deviation and thus the CV (Lemmer, 2004). This could lead to inaccurate indications of a batsman's consistency. Lemmer (2004) mentioned that this error was corrected by using a different measure, the adjusted coefficient of variation (CV_{adj}). To calculate CV_{adj}, Lemmer (2004) explained that the traditional batting average (AVE^t_{BAT}) as well as an adjusted standard deviation are used. The adjusted standard deviation was calculated similarly to the traditional standard deviation, however, not out scores and scores above the average were disregarded. The adjusted coefficient of variation (CV_{adj}) was then defined as the adjusted standard deviation divided by the batting average (provided that both measures existed). Lemmer (2004) defined the consistency coefficient (CC) as $CC = \frac{1}{CV_{adj}}$. This definition changed the direction of the measure to allow for the interpretation that the higher values of CC now corresponded to the more consistent batsmen. Barr and Van den Honert (1998) used the geometric coefficient (G) to adjust the batting average by multiplying the two values together. This was done without standardising as the consistency of a batsman was indicated by the deviation of G from the benchmark value of 1 (Lemmer, 2004). By multiplying the two measures together Barr and Van den Honert (1998) were able to scale a batsman's average up or down according to their level of consistency. In order to incorporate a batsman's consistency into their batting performance measure the value of CC needed to be standardised. Lemmer (2004) standardised the consistency coefficient by dividing it by its average. This ratio (C_{st}) was defined as

$$C_{st} = \frac{CC}{\overline{CC}} \tag{1.10}$$

where $\overline{CC} = \frac{1}{n}\sum_{i=1}^{n} CC_i$.

Lemmer (2004) reasoned that in limited overs cricket the batting strike rate (SR^t_{BAT}) was an important measure of a batsman's ability and should be included. In order to incorporate the strike rate into the batting performance measure it too had to be standardised. This was accomplished using a method similar to that used in Equation (1.10). The standardised strike rate (SR_{st}) was thus defined as

$$SR_{st} = \frac{SR^t_{BAT}}{\overline{SR^t_{BAT}}} \tag{1.11}$$

where $\overline{SR^t_{BAT}} = \frac{1}{n}\sum_{i=1}^{n}\left(SR^t_{BAT}\right)_i$.

Lemmer (2004) then proposed that the batting performance (BP) be calculated as a product of the above three measures. The problem with this, however, was that the standard deviation of SR_{st} differed to that of C_{st}. In order to correct for this Lemmer (2004) used the following technique to equate the standard deviations.

Let the standard deviation of SR_{st} be given by $\sigma\left(SR_{st}\right)$ and the standard deviation of C_{st} be given by $\sigma\left(C_{st}\right)$. Suppose further that SR_{st_i} is the SR_{st} value for player i. Then for all values of i the following iterative transformation is made

$$SR^{[k]}_{st_i} = \left(SR^{[k-1]}_{st_i}\right)^{\frac{\sigma(C_{st})}{\sigma\left(SR^{[k-1]}_{st}\right)}} \tag{1.12}$$

where k indicates the number of iterations performed and $SR^{[0]}_{st_i} = SR_{st_i}$. The process is repeated until the standard deviation of $\sigma\left(SR^{[k]}_{st_i}\right)$ is equal to that of $\sigma\left(C_{st}\right)$ to the required level of accuracy (usually 3 to 4 decimal places). The process requires 4 to 5 iterations to achieve this accuracy.

The values of $SR^{[k]}_{st_i}$, where k represents the final iteration, are denoted as SR^{tr}_{st}.

The batting performance measure (BP) proposed by Lemmer (2004) was thus calculated as:

$$BP = AVE^{exp}_{BAT} \times C_{st} \times SR^{tr}_{st}. \tag{1.13}$$

Barr and Van den Honert (1998) used a measure similar to this one, however Lemmer (2004) scaled the batman's average using a consistency and a strike rate measure, whereas Barr and Van den Honert's (1998) measure only used a consistency measure.

1.3.4 A Second Batting Performance Measure: Barr and Kantor (2004)

Barr and Kantor (2004) proposed an alternative measure for comparing and selecting batsmen in limited overs cricket. They argued that, owing to the limited number of balls available in a limited overs cricket match, a batman's strike rate along with his batting average should be considered when establishing a performance measure. A graphical method, similar to the representation of risk versus return in portfolio analysis, was introduced by plotting the batting strike rate on the y-axis and the batsman's probability of going out ($P\,(\text{Out})$) on the x-axis. In financial markets an improvement in expected returns is associated with larger risk. In cricket it was argued that any improvement in the batting strike rate would be associated with an increased probability of being dismissed (Barr and Kantor, 2004). This is intuitive since if a batsman attempts to score runs more quickly, there is a greater risk of being caught, run out and bowled.

Barr and Kantor (2004) defined the probability of a batsman being dismissed on any given ball as

$$P\,(\text{Out}) = \frac{\text{total number of times dismissed}}{\text{total number of balls faced}}. \tag{1.14}$$

It was observed that

$$\frac{SR^t_{BAT}}{P\,(\text{Out})} = \frac{\frac{\text{total number of runs scored}}{\text{total number of balls faced}}}{\frac{\text{total number of times dismissed}}{\text{total number of balls faced}}} = \frac{\text{total number of runs scored}}{\text{total number of times dismissed}} = AVE^t_{BAT}.$$

Plotting the strike rate on the vertical axis and $P\,(\text{Out})$ on the horizontal axis thus gives $AVE^t_{BAT} = \frac{SR^t_{BAT}}{P(\text{Out})} = \frac{y}{x}$. As a result, any straight line passing through the origin represents a certain batting average (Barr and Kantor, 2004). If we consider Figure 1.1 the straight lines through the origin (L, M and N) represent groups of batsmen with equal batting averages. Horizontal lines indicate batsmen of equal strike rates and vertical lines indicate batsmen of equal $P\,(\text{Out})$(Barr and Kantor, 2004).

The measure proposed by Barr and Kantor (2004) was a weighted product of the batting strike rate and the batting average. This measure denoted as BK was given as

$$\begin{aligned} BK &= \left(SR^t_{BAT}\right)^{\alpha}\left(AVE^t_{BAT}\right)^{1-\alpha}, \\ &= y^{\alpha}\left(\frac{y}{x}\right)^{1-\alpha}, \\ &= \frac{y}{x^{1-\alpha}} \end{aligned} \tag{1.15}$$

where $0 \leq \alpha \leq 1$ is a parameter of the balance between the batting strike rate and batting average (Barr and Kantor, 2004). Higher values of BK indicate greater batsman ability and are preferred to lower values. If $\alpha = 1$ the measure places no emphasis on the batting average and setting $\alpha = 0$ places no emphasis on batting strike rate (Barr and Kantor, 2004). It was conjectured that $\alpha = \frac{1}{2}$, thus weighting both measures equally. The authors noted that plotting curves of the form

$$y = cx^{\frac{1}{2}}$$

yield isoquants which represent batsmen of equal ability, indicated in Figure 1.1 as R, S and T. All batsmen falling on the same isoquant have the same BK value. The BK value is thus maximised by "*selecting batsmen according to the highest isoquant on which they lie*" (Barr and Kantor, 2004).

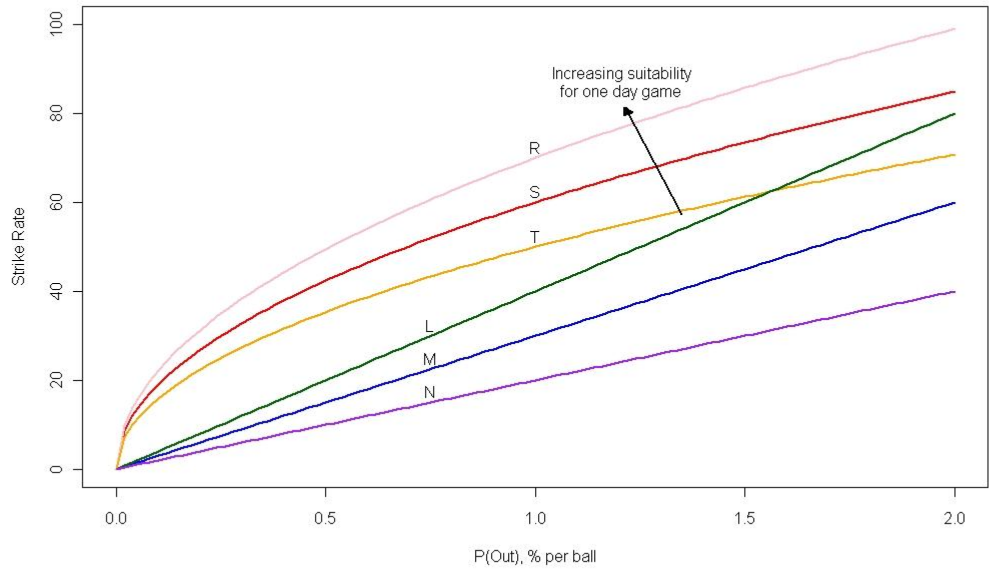

Figure 1.1: Barr and Kantor (2004): A graphical display of the curves of equal suitability of batsman to the one-day game

Barr and Kantor (2004) used the statistics from the 2003 ICC Cricket World Cup to successfully indicate the usefulness of their proposed measure.

1.3.5 Integer Programming Method: Gerber and Sharp (2006)

Gerber and Sharp (2006) proposed an integer programming technique in order to select a limited overs cricket squad. The method included collecting the data for 32 prominent South African cricketers and using this data to facilitate team selection for the South African One Day International (ODI) cricket squad. It is important to note that the method provided was used for the selection of a *squad* of 15 cricketers, not a *team* of 11.

Gerber and Sharp (2006) proposed that measures be developed to give an indication of each player's batting, bowling, fielding, all rounder and wicket keeping ability. The relevant measure (or index) would then be applied to each player in their area of expertise (i.e. the batting index would be calculated for a specialist batsman etc.).

The method proposed to measure the batting ability (i.e. the batting index) was defined as

$$BT_i^{GS} = \left(\frac{\text{batting average of batsman } i}{\text{sum of all the batting averages of all the batsmen}} \right) \times \text{number of specialist batsmen} \tag{1.16}$$

which compared the batting average of the given batsman to that of all the other batsmen.

Since the integer program used was a maximisation problem, it required that cricketers with higher values for the given performance measures be more desirable than players with lower values. In the case of bowlers, however, a lower economy rate implies a more desirable player.

Gerber and Sharp (2006) overcame this obstacle by using the following methodology.

Let

$$v_i = \left[k - \left(\frac{\text{economy rate of specialist bowler } i}{\text{sum of all the economy rates of all the specialist bowlers}}\right)\right] \tag{1.17}$$

where the constant k was chosen to be the smallest positive integer such that

$$k > \left(\frac{\text{economy rate of specialist bowler } i}{\text{sum of all the economy rates of all the specialist bowlers}}\right).$$

Introducing this constant ensured that more desirable bowlers (i.e. bowlers with low economy rates) had higher values of v_i. The same strategy used to determine the batsman's index was then used to calculate the bowling index. This measure was given as

$$BL_i^{GS} = \left(\frac{v_i}{\text{sum of all } v_i}\right) \times \text{number of specialist bowlers.} \tag{1.18}$$

The fielding ability of player i was measured using the fielding index, defined as

$$FLD_i^{GS} = \left(\frac{\text{dismissal rate of fielder } i}{\text{sum of the dismissal rates of all the fielders}}\right) \times \text{number of specialist fielders} \tag{1.19}$$

where the dismissal rate of a player was defined as the average number of fielding dismissals (i.e. run outs, catches) made by the player per match. Gerber and Sharp (2006) argued that the use of this measure alone as a selection criterion may be unnecessary, since most selectors would prefer to include a specialist bowler or batsman rather than a fielder.

All rounders can be defined as players who perform well in differing roles in the team. For example, a player who can both bat and bowl is considered an all rounder. Gerber and Sharp (2006) suggested that all rounders be divided into four categories: (1) cricketers who can both bat and bowl; (2) cricketers who can both bat and field; (3) cricketers who can both bowl and field; and (4) cricketers who can bat, bowl and field. The category that will be of interest in this study is (1) Players who can both bat and bowl. The all-rounder ability index for each player was defined as:

$$ALR_i^{GS} = \left(\frac{\text{sum of relevant player index values}}{\text{sum of that category of index values}}\right) \times \text{number of players in all rounder category.} \tag{1.20}$$

The wicket keeping ability was defined similarly to that of the fielding category (Equation 1.19) and is given as

$$WKR_i^{GS} = \left(\frac{\text{dismissal rate of keeper } i}{\text{sum of the dismissal rates of all the keepers}}\right) \times \text{number of specialist keepers.} \tag{1.21}$$

In order to select the squad a binary integer programming technique was implemented, and the model was treated as one of maximisation (Gerber and Sharp, 2006). The model was flexible and could include as many ability measures as determined by the team selectors. For the purpose of illustration, only the above mentioned ability measures are shown in this review. The relationship between a player and their respective ability measure can be illustrated in a two way matrix as shown in Table 1.1.

		Abilities					
		Bat	Bowl	Field	All Round 1	...	Wicket Keeper
		1	2	3	4	...	8
Player	1	a_{11}	a_{12}	$\cdots$	$\cdots$	$\cdots$	a_{18}
	2	a_{21}	a_{22}	$\cdots$	$\cdots$	$\cdots$	a_{28}
	$\vdots$	$\vdots$	$\vdots$				
	n	a_{n1}	a_{n2}				a_{n8}

Table 1.1: Two-way representation matrix of players and abilities

In Table 1.1 the value a_{ij} is defined as the index measure of player i, for ability j. The number of abilities in Table 1.1 can be expanded to include any number of measures.

Gerber and Sharp (2006) defined the decision variables for the integer program as

$$x_{ij} = \begin{cases} 1 & \text{if player } i \text{ is selected for ability } j \\ 0 & \text{if player } i \text{ is not selected for ability } j. \end{cases} \tag{1.22}$$

The objective function of the integer program was maximise

$$Y = \sum_{j=1}^{k} y_j \tag{1.23}$$

where k represents the total number of ability measures considered and

$$y_j = \sum_i a_{ij} x_{ij}. \tag{1.24}$$

Each y_j thus represents the ability index for ability j.

The constraints of the model could be changed according to the requirements of the selection. Some of the necessary constraints are shown below (Gerber and Sharp, 2006),

$$\begin{aligned} \sum_i \sum_j x_{ij} = m \quad &: \quad \text{to ensure exactly } m \text{ players are selected} \\ \sum_j x_{ij} \leq 1 \quad &: \quad \text{to ensure a player is selected only once} \\ x_{ij} \in \{0, 1\} \quad &: \quad \text{to satisfy definition of the decision variable} \end{aligned}$$

Additional constraints, subject to the requirements of the selectors, could easily be added to this model. As an example, consider including a constraint to ensure that at least z_b batsmen are selected. The constraint may be expressed as

$$\sum_i x_{i1} \geq z_b. \tag{1.25}$$

In this constraint the value $j = 1$ is included as this number corresponds to the batting ability listed in Table 1.1. Additional constraints concerning other abilities may also be included. This integer programming technique is used extensively in this research.

1.3.6 Comparison of Cricketers using Graphical Displays: Van Staden (2009)

Van Staden (2009) mentioned that a study by Kimber in 1993 proposed a method to graphically represent all three traditional bowling measures. It was noted that although this method gained little widespread use, it is a simple and powerful tool for the comparison of bowlers. Van Staden (2009) expanded the use of these graphical displays to compare the batting and all-round performance of cricketers. These graphical displays and comparisons were used on the cricketers who took part in the 2008 Indian Premier League.

For the purpose of this paper, van Staden (2009) redefined the economy rate of a bowler to be the number of runs conceded per k balls. The new measure, termed $ECON_k$, was defined as

$$ECON_k = k \times \frac{\text{total number of runs conceded}}{\text{total number of balls bowled}}. \tag{1.26}$$

It can easily be seen that $ECON_6$ is the traditional economy rate defined in Section 1.3.1. From Equation 1.26 and the traditional measures of a bowler's average and strike rate the following hyperbolic relation was found to exist (van Staden, 2009)

$$ECON_k \times SR^t_{BWL} = k \times AVE^t_{BWL}. \tag{1.27}$$

Van Staden (2009) plotted SR^t_{BWL} on the y-axis and $ECON_k$ (with $k = 100$) on the x-axis, by doing this all three traditional bowling measures could be represented on a single graph. This is done by adding hyperbolic contours to the graph, which represented AVE^t_{BWL}. Van Staden (2009) mentioned that, although $k = 100$ was used, any logical value of k (such as $k = 6$) could be used. An example of the graph obtained is given in Figure 1.2.

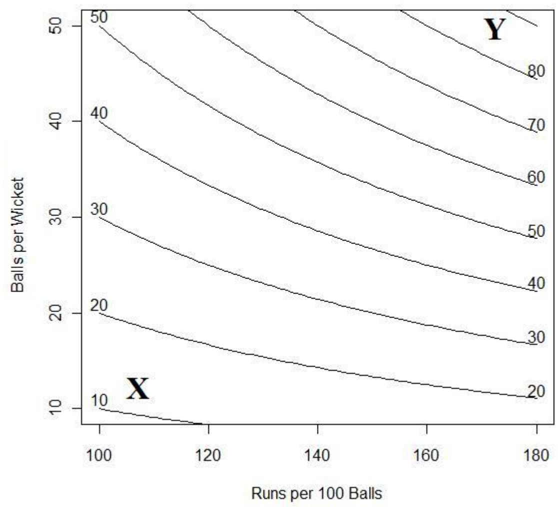

Figure 1.2: Van Staden (2009): Graph used for the comparison of cricketers

Van Staden (2009) emphasised that a bowler would ideally prefer to "*maximise the number of wickets taken and minimise the number of runs conceded*". This would result in better bowlers having lower values of AVE^t_{BWL}, SR^t_{BWL} and $ECON_k$. Low SR^t_{BWL} values correspond to low values on the y-axis of Figure 1.2 similarly low $ECON_k$ values correspond to low values on the x-axis of Figure 1.2. Thus, cricketers with better bowling statistics would appear in the lower left area of Figure 1.2 as indicated by "X" and bowlers with poor bowling statistics would appear in the area indicated by "Y". The partitions of the graph created by the contours aid in the comparison of bowlers by indicating the bowling average. That is, all bowlers lying on the same contour have equal bowling averages. The use of a graph such as Figure 1.2 provides a convenient and efficient graphical method to compare all the traditional bowling measures simultaneously.

In order to compare batsmen, van Staden (2009) suggested an approach similar to that of Barr and Kantor (2004). Barr and Kantor (2004) evaluated batsmen using isoquants created when plotting $P\,(\text{Out})$ against SR^t_{BAT}. The approach used by van Staden (2009) plotted the inverse of $P\,(\text{Out})$, i.e. the balls faced by a batsmen per dismissal, on the y-axis and SR^t_{BAT} on the x-axis, thus essentially swapping the axes of the graph used by Barr and Kantor (2004). This also allows the graph in Figure 1.2 to be used for the comparison of batsman. The contours of the plot, when considering batsmen, indicate the traditional batting average. The main difference when plotting batsmen on Figure 1.2 is that batsmen appearing closer to the top right area of the plot (indicated as "Y") are preferred to those lying towards the bottom left area (indicated as "X"). Since batsmen and bowlers use the same set of axes for comparison, van Staden (2009) suggested that, by showing a players batting and bowling measures on the same graph plot, one could assess the all round ability of a cricketer. The approach detailed by van Staden (2009) is a convenient and highly effective method for the comparison of cricketers.

1.3.7 Other Statistical Research into Cricket

Arguably the most notable statistical research into the game of cricket was the study conducted by Duckworth and Lewis (1998). They determined a method to reset targets in rain interrupted limited overs cricket matches in the paper entitled "*A fair method of resetting the target in interrupted one-day cricket matches*". De Silva, Pond & Swartz (2001) extended the use of the Duckworth-Lewis method to determine the magnitude of victory in one day cricket matches. This interesting and convenient method is particularly useful for breaking ties in tournament standings.

Beaudoin and Swartz (2003) developed a measure to assess the performance of batsmen and bowlers. Their measure involved taking into account the resources used (according to the Duckworth-Lewis method) by both batsmen and bowlers in one-day cricket. This measure developed, called the runs per match (RM), is sensible and provided good results. Although this measure provides a good indication of a cricketer's ability, it is omitted from this study as the data required for its calculation were not always available.

The use of dynamic programming in cricket has been used in several research papers specifically Clarke (1988) and Clarke and Norman (1999). Dynamic programming was also used as a method to reset the target in rain interrupted games by Preston and Thomas (2002), although the Duckworth-Lewis method is still the method of choice of the ICC.

Research into accurate simulation in cricket has recently been concluded, with good results. Swartz, Gill, Beaudoin and de Silva (2006) dealt with finding the optimal batting order in one day cricket by the use of simulation techniques. Swartz, Gill and Muthukumarana (2009) investigated the simulations of entire one day cricket matches. This was done by using historic data to determine probabilities of various outcomes on each ball of a match. It was noted that the simulations provided reasonable results and it was envisioned that this technique could be used to answer various questions regarding the game of cricket (Swartz, Gill and Muthukumarana, 2009). These papers show that the use of computer simulations in cricket research can be beneficial.

It is clear that cricket is a highly researched and publicised topic, specifically in the field of statistics. The research provided in this dissertation endeavours to add to this body of knowledge.

1.4 Optimisation

Mathematical optimisation refers to the techniques involved in finding the "best" solution to a given problem, provided that it can be expressed mathematically (Snyman, 2005). This optimal or "best" solution is found using various computational techniques, and is usually subject to certain constraints. Most decision making problems require the identification of three main components. Taha (2003) suggests that these three components are identified by answering the following questions:

1. What are the decision options or choices?
2. What restrictions are present when making the decision?
3. What value or outcome must be optimised when evaluating the alternatives?

The answers to the above questions provide the decision maker with what is needed to solve an optimisation problem. These are (Lourens, 2009)

1. *The decision variables*: These are the variables which, at the end of the optimisation procedure, provide the decision maker with the answers which they sought.
2. *The constraints*: The constraints are equations which define the restrictions imposed on the decision variables. These are expressed in terms of the decision variables.
3. *The objective function*: This defines the outcome or value which must be optimised in terms of the decision variables.

Once the above three components are identified, the problem is ready to be solved.

The optimisation models discussed are the Linear Programming (LP) model and the Binary Integer Programming (BIP) model.

1.4.1 The Linear Programming (LP) Model

Taha (2003) explains that the LP model is used for optimisation models with "*strict linear objective and constraint functions*". Furthermore, all the decision variables in a LP model are real-valued, and as such can take on any value on the real line (subject to the constraints). This is the primary difference between the Linear and the Integer Programming techniques.

An example of an n variable LP model is formulated in the following way.

The *objective function*:

$$\max\ Z = a_1x_1 + a_2x_2 + \ldots + a_nx_n$$

The *constraints*:

$$\begin{array}{rcl}
c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n & \leq & b_1 \\
c_{21}x_1 + c_{22}x_2 + \ldots + c_{2n}x_n & \leq & b_2 \\
\vdots & \vdots & \vdots \\
c_{k1}x_1 + c_{k2}x_2 + \ldots + c_{kn}x_n & \leq & b_k \\
x_1, x_2, \ldots, x_n & \geq & 0
\end{array}$$

In order to determine the optimal values of the *decision variables*, $x_1, x_2, \ldots, x_n$, a computational technique known as the simplex method is often used. This method is discussed in detail in Taha (2003). The solution for this problem is thus given as the values of $x_1, x_2, \ldots, x_n$ which maximise the objective function, subject to the given constraints. The example above is a simple representation of the LP model and can easily be adjusted to suit the requirements of a given problem.

1.4.2 The Binary Integer Programming (BIP) Model

Integer Programming (IP) models have the same formulation as traditional LP models, however, the decision variables are integer-valued rather than real-valued (Taha, 2003). This added constraint is necessary for many practical problems where the values of the decision variables may not contain fractions such as "number of people" or "yes-no" type decisions (Taha, 2003). The "yes-no" decision variable x_i is often formulated as

$$x_i = \begin{cases} 1 & \text{if object } i \text{ is selected ("yes")} \\ 0 & \text{if object } i \text{ is not selected ("no").} \end{cases}$$

If all the decision variables are of this type the model is referred to as a Binary Integer Programming (BIP) model (Lourens, 2009).

IP models are more difficult to solve than LP models. Taha (2003) states that there does not exist a computer code which will consistently solve integer programming problems. Of the two methods described in Taha (2003), it is noted that the Branch-and-Bound method is the more reliable method used to find the optimal solution to an integer programming problem. Taha (2003) also mentions that most commercial software use the Branch-and-Bound method for solving integer programming problems.

The Branch-and-Bound (B&B) Algorithm

The Branch-and-Bound algorithm was developed by A. Land and G. Doig in 1960 (Taha, 2003). The method initially solves the IP problem as a traditional LP problem (Lourens, 2009), ignoring the added integer constraint. Once solved, the values of the decision variables are considered and if any of the values are non-integers, the feasible region is split by adding extra constraints and the problem is re-assessed. This procedure continues until all the integer decision variables are integer-valued.

To explain the concept, consider the following example: Suppose that the shaded region indicated in Figure 1.3 indicates the feasible region for a given maximisation linear programming problem.

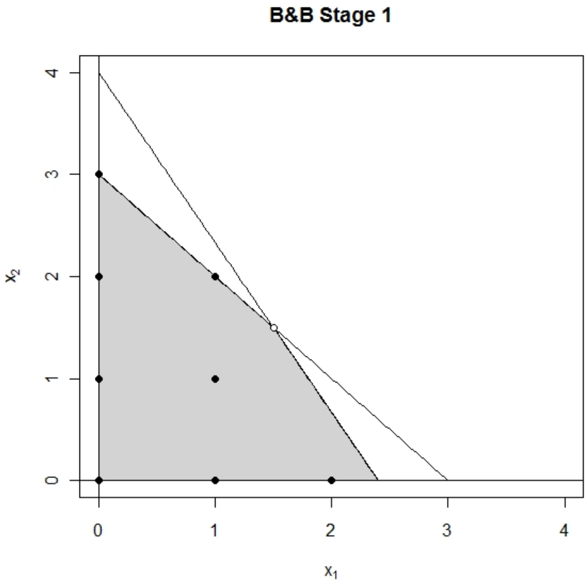

Figure 1.3: Initial Branch-and-Bound Stage: Solving a single LP problem (Taha, 2003)

Initially, the Branch-and-Bound method will solve the problem as a simple LP problem. This results in the optimal decision of $x_1 = x_2 = 1.5$, indicated on the graph as a white point. The black points on the graph indicate all the integer values within the feasible region. Given that this is actually an IP problem, it is desired that x_1 and x_2 are integers, thus the black points indicate the feasible points for the integer variables.

The Branch-and-Bound method, now adds the following constraints to the model:

$$x_1 \leq 1 \quad \text{and} \quad x_1 \geq 2$$

thus removing the optimal point (non-integer) from consideration. Since the above restrictions are mutually exclusive, they are not included in the same linear program. Thus two separate linear programs are run and the solutions for each are obtained (Taha, 2003). The feasible regions for these programs are given in Figure 1.4.

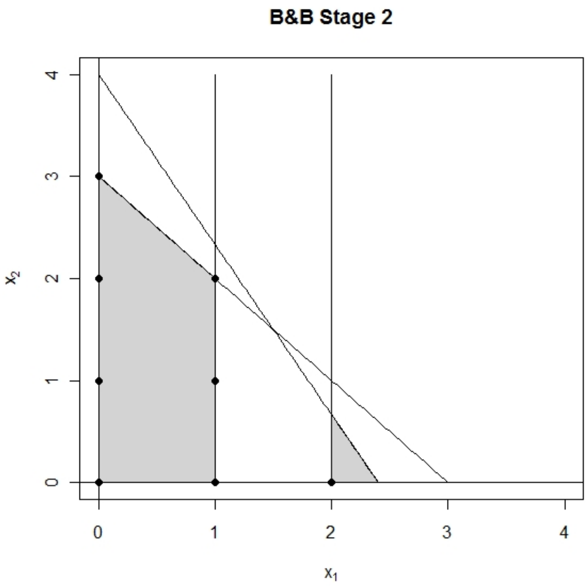

Figure 1.4: Second Branch-and-Bound Stage: Solving two separate LP problems (Taha, 2003)

The method then goes on to check the values of x_1 and x_2 for both linear programs. If the values are integers for both feasible regions, the most desirable of the two is chosen (according to the value of the objective function). If not, the method is repeated until the optimal integer valued solution is found.

In summary, this chapter reviews the game of cricket by providing a brief discussion of the rules and history of the game. The history and concept of fantasy sport is then discussed. This is followed by an investigation of relevant statistical research into the game of cricket. The chapter is concluded with a discussion of mathematical optimisation techniques (specifically LP and IP models).

Chapter 2

Fantasy League and Data

2.1 The Fantasy League Game

The fantasy league on which this dissertation is based uses the fantasy league on the CricInfo website for the 2008 Indian Premier League (CricInfo IPL Fantasy League, 2008). This fantasy league was chosen as the CricInfo website is widely used by researchers (Lemmer, 2003 & 2004; Gerber and Sharp, 2006). Furthermore, the rules and scoring methods defined for this league were intuitive.

2.1.1 Fantasy League Rules

The CricInfo website categorised each of the players who took part in the 2008 Indian Premier League competition into one of four distinct categories. These categories were: batsmen, bowlers, all rounders and wicket keepers. Each player was then given a price ranging from 50 000 to 150 000 currency units (hereafter referred to as "units"), with higher prices corresponding to historically more desirable players and lower prices to less desirable players.

There are a number of rules mentioned on the CricInfo website to define the restrictions on the fantasy league. The following are the rules and restrictions used for the fantasy league setting of this study.

- The total cost of each team is restricted to a maximum of 1 000 000 units.
- Each team must consist of 11 players.
- Each team must be made up in one of the following ways
  - 4 Batsmen, 4 Bowlers, 2 All Rounders and 1 Wicket Keeper, or
  - 4 Batsmen, 3 Bowlers, 3 All Rounders and 1 Wicket Keeper.
- A maximum of 20 changes are allowed to be made to the team during the course of the competition.
- A cricketer can only be selected for the category in which they have been assigned (by the website). For example, a cricketer classified by the website as a wicket keeper can not be selected in a team as a batsman or a bowler.

An example of two possible fantasy league teams, satisfying the restrictions above, is provided in Table 2.1.

	Team 1			**Team 2**	
	Batsmen	**Price**		**Batsmen**	**Price**
1	M Hayden	90 000	1	M Hayden	90 000
2	G Smith	95 000	2	R Ponting	90 000
3	S Fleming	85 000	3	S Fleming	85 000
4	M ul-Haq	80 000	4	M ul-Haq	80 000
	Bowlers	**Price**		**Bowlers**	**Price**
1	D Steyn	90 000	1	A Kumble	85 000
2	S Warne	95 000	2	C Vaas	80 000
3	A Kumble	85 000	3	N Bracken	85 000
4	C Vaas	80 000			
	All Rounders	**Price**		**All Rounders**	**Price**
1	J Oram	105 000	1	S Pollock	95 000
2	D Bravo	75 000	2	M Hafeez	75 000
			3	D Bravo	75 000
	Wicket Keeper	**Price**		**Wicket Keeper**	**Price**
1	A Gilchrist	110 000	1	MS Dhoni	150 000
	TOTAL COST	990 000		**TOTAL COST**	990 000

Table 2.1: Example of possible fantasy league cricket teams

Team 1 includes 4 batsmen, 4 bowlers, 2 all rounders and 1 wicket keeper whilst Team 2 includes 4 batsmen, 3 bowlers, 3 all rounders and 1 wicket keeper. An interesting observation to note is that both teams have the same total cost but differ considerably in composition. The difference in these teams illustrates the difficulty faced by fantasy league participants in deciding on their team.

Once selected a player can be removed or replaced at the discretion of the fantasy league participant. A fantasy league participant would ideally make changes to their team in order to include the best players through the course of the competition. These changes are allowed so long as none of the restrictions are violated. Should a fantasy league participant feel that no changes are necessary, this is acceptable.

Other fantasy league rules may differ from these rules but in principle team composition and a limiting value of team cost are the generic restrictions.

2.1.2 Fantasy League Scoring

Each cricketer is allocated points based on their performance in every match in which they participated. The aim of a fantasy league participant is to select the cricketers who score the highest points throughout the competition. The fantasy league participant with the highest number of points at the end of the competition is deemed the winner.

Tables 2.2 to 2.4 summarise the points allocation that was used by the CricInfo Fantasy League for the 2008 Indian Premier League. The tables demonstrate that most aspects of a cricketer's performance are accounted for in a given match. This shows that the fantasy league score obtained by the cricketer provides a reasonable approximation of their influence on the match.

Event (Batting)	**Points**
Runs Scored	1
Six Hit (Bonus)	2
Dismissed for a Duck	-10
Provided Runs Scored ≥ 20 *the following bonus points apply*	
$0 \leq SR^t_{BAT} < 50$	-30
$50 \leq SR^t_{BAT} < 75$	-20
$75 \leq SR^t_{BAT} < 100$	-10
$100 \leq SR^t_{BAT} < 150$	0
$150 \leq SR^t_{BAT} < 180$	10
$180 \leq SR^t_{BAT} < 200$	20
$200 \leq SR^t_{BAT}$	30
Milestone Bonus	
On reaching 25 runs	10
On reaching 50 runs	20
On reaching 75 runs	40
On reaching 100 runs	80

Table 2.2: CricInfo IPL Fantasy League batting points allocation table

To indicate how Table 2.2 is used, consider the following example. A cricketer hits 3 sixes while scoring 60 runs off 40 balls in a given IPL match. The number of fantasy league points scored for this is calculated as follows.

- 1 point for each run scored (60) $= 60 \times 1 = 60$ points,
- 2 points for each six hit (3) $= 3 \times 2 = 6$ points,
- $SR^t_{BAT} = \frac{\text{number of runs scored}}{\text{number of balls faced}} \times 100 = \frac{60}{40} \times 100 = 150$, thus scoring 10 points, and
- 20 points for reaching 50 runs.

The strike rate is considered in this case since the cricketer scored 20 or more runs. The fantasy league score of the cricketer for their batting is thus $60 + 6 + 10 + 20 = 96$ points.

Table 2.3 provides the points allocation for bowling.

Event (Bowling)	**Points**
Dismissing a Batsman, an All Rounder or a Wicket Keeper	25
Dismissing a Bowler	10
Maiden over	40
Provided Overs Bowled ≥ 2 *the following bonus points apply*	
$0 \leq ECON^{t}_{BWL} < 3$	30
$3 \leq ECON^{t}_{BWL} < 5$	20
$5 \leq ECON^{t}_{BWL} < 7$	10
$7 \leq ECON^{t}_{BWL} < 9$	0
$9 \leq ECON^{t}_{BWL} < 11$	-10
$11 \leq ECON^{t}_{BWL} < 14$	-20
$14 \leq ECON^{t}_{BWL}$	-30
Milestone Bonus	
On taking 2 wickets	10
On taking 3 wickets	20
On taking 4 wickets	40
On taking 5 wickets	80

Table 2.3: CricInfo IPL Fantasy League bowling points allocation table

Suppose the cricketer in the previous example bowled 4 overs, conceding 40 runs and taking 1 wicket. The wicket taken was that of an opposition all rounder. The cricketer is awarded the following points.

- 25 points for each batsman, all rounder or wicket keeper dismissed. One all-rounder dismissed = 25 points.
- $ECON^{t}_{BWL} = \frac{\text{number of runs conceded}}{\text{number of balls bowled}} \times 6 = \frac{40}{24} \times 6 = 10$, thus -10 points are scored.

The economy rate is included in this case since the bowler bowled 2 or more overs. The fantasy league score for the cricketer's bowling is thus $25 - 10 = 15$ points. Table 2.4 is used to calculate the fantasy league score of a cricketer for fielding.

Event (Fielding)	**Points**
Taking a catch	15
Stumping	15
Run Out (direct)	30
Run Out (indirect, per person)	10

Table 2.4: CricInfo IPL Fantasy League fielding points allocation table

Table 2.4 is used similarly to Table 2.2 and Table 2.3. A direct run out is one where only one fielder is involved in the run out. The fielder would then be awarded 30 points. If other fielders are in involved, for example a wicket keeper, it would then be considered an indirect run out, and each participating fielder would be awarded 10 points. In addition 50 bonus points are scored by a cricketer who wins the man-of-the-match award.

Suppose that the cricketer mentioned in the previous two examples also took 1 catch and won the man-of-the-match award. The cricketer would then score an additional $15 + 50 = 65$ points.

For each game, every cricketer who played is awarded fantasy league points in each of the batting, bowling, fielding and man-of-the-match categories. The sum of the points (across these categories) is the total fantasy league score of a cricketer for the game.

In our example, the cricketer's total fantasy league score would be

$$\begin{aligned}&(\text{Batting Points}) + (\text{Bowling Points}) + (\text{Fielding Points}) + (\text{Man-of-the-Match Points})\\&= (96) + (15) + (15) + (50)\\&= 176 \text{ points.}\end{aligned}$$

2.2 Data

2.2.1 The Collection of Game Data

Data were collected for two Twenty20 competitions held between September 2007 and June 2008. The two competitions were the 2007 ICC World Twenty20 held in South Africa in September and the inaugural Indian Premier League played in India from April to June 2008.

Match scorecards for each match of the two competitions were obtained from the ESPN CricInfo (www.cricinfo.com) website and captured into *Microsoft Excel* (Microsoft Corporation, 1985 - 2006). There were a total of 27 matches played in the 2007 ICC World Twenty20 of which one was abandoned. The 2008 Indian Premier League consisted of 59 matches of which one was abandoned.

The data captured for batting were the number of runs scored; the number of balls faced, the number of 6's hit and the number of 4's hit. For bowling the data captured were the number of overs bowled; the number of maidens bowled; the number of runs conceded; the number of wickets taken; the number of no balls bowled and the number of wides bowled. For fielding the data captured were the number of catches; the number of run-outs and the number of stumpings. These data were collected for each cricketer who took part in either of the two competitions. The data were collected from the website between 17 April 2008 to 11 August 2008.

2.2.2 The Collection of Player Data

A complete limited overs career history for each player who took part in either of the competitions was obtained from the Player Oracle on the CricketArchive website (www.cricketarchive.com/cgi-bin/ask_the_player_oracle.cgi). The data were collected from the website between 6 September 2008 and 28 January 2009. The data included in the analysis were for the period up to and including 1 June 2008 so as to coincide with the conclusion of the 2008 Indian Premier League.

The ListA and Twenty20 data were collected for each player. ListA classified matches include all domestic and international matches where the number of overs per innings is between 40 (domestic) and 50 (international) i.e. traditional limited over games. Twenty20 type matches include all domestic and international 20 over matches (ICC Rules and Regulations, 2008).

The data collected for each player include the following: the number of runs scored (and method of dismissal); the number of overs bowled; the number of maidens bowled; the number of runs conceded; the number of wickets taken; the number of catches taken and the number of stumpings made. Career batting strike rates were calculated from the game data extracted from the CricInfo website for Twenty20 matches. ListA batting strike rates were estimated using data collected from the CricInfo website on 7 September 2008. When a ListA strike rate was not available for a player it was reasoned that the players ODI strike rate would be adequate. Each match was then classified as either a domestic or an international match. The data were captured into *Microsoft Excel* (Microsoft Corporation, 1985 - 2006) and then exported into *R 2.8.1* (The R Foundation for Statistical Computing, 2008) for analysis.

A limitation to this method of data collection is that the strength of the opposition is not taken into account. As such, good scores against weaker teams are considered to be as relevant as good scores against strong teams. This is an area that is open to further research.

In summary, this chapter provides an overview of the rules and nature of a fantasy league cricket game. The origin and nature of the data collected for analysis purposes are also discussed.

Chapter 3

Methodology

The aim of this dissertation is to provide a prospective approach to facilitate the selection and management of a fantasy league team. The fantasy league rules on which this research is based is the 2008 IPL fantasy league as found on the CricInfo website. A prospective approach implies that only historic or past data are considered. Thus for the selection of the initial fantasy league team only data prior to the beginning of the 2008 IPL is considered. As the IPL progresses more recent player data will become available. These data are then used to make adjustments to the initial team. These adjustments are made in accordance with the restrictions imposed on the fantasy league (Section 2.1.1).

Initially mathematical techniques are developed to determine a strategy to select a starting fantasy league team. In order to maximise the total fantasy league score, a strategy is developed to make team changes through the course of the competition. To establish these selection strategies the following methodology is used.

3.1 Preliminary Analysis

3.1.1 Fantasy League Point Distribution

To acquire prior information about the points scored in a fantasy league it is necessary to analyse a similar competition. The fantasy league based on the 2008 IPL is the first of its kind. As such, no information regarding the distribution of fantasy league points between player categories (i.e. batsmen, bowlers, all rounder and wicket keepers) and across cricketing disciplines (i.e. batting, bowling and fielding) is available. Given that this information could be used in the development and modification of performance measures for fantasy league tournaments, this section describes how the 2007 ICC World Twenty20 tournament was used as a basis for data analysis and interpretation.

The scoring methodology described in Section 2.1.2 is applied to the 2007 ICC World Twenty20 and fantasy league scores for each cricketer are calculated. These fantasy league scores are evaluated in order to provide insight into the structure of a Twenty20 fantasy league competition. The data obtained from this competition is believed to be relevant since it was held only seven months prior to the 2008 IPL. Adding to this there are a significant number of cricketers who participated in both tournaments. It thus seems reasonable that this tournament can provide relevant data.

The format of the ICC World Twenty20 is different to that of the IPL. The ICC World Twenty20 consists of two round robin stages followed by a semi-final and a final. The IPL has a league formation with each team playing every other team twice (home and away). The top four teams then play a knockout in a semi-final and a final to determine the winner. As a result, fantasy league scores within the ICC World Twenty20 competition are not comparable to those within the IPL competition. To avoid this, only fantasy league score information relating to the tournament in its entirety is used. That is, for the purposes of this study the entire tournament is considered as a single event and individual games within the competition are not considered. The distribution of fantasy league points across the various player categories and cricketing disciplines is calculated using the data from this tournament. The results of this analysis are provided in Section 4.1.

3.1.2 Linear Regression Modelling

The fantasy league scoring methodology described in Section 2.1.2 provides a single value which quantifies a cricketer's overall contribution to a match. This fantasy league score is of great importance in this study. It is desired to include cricketers who tend to score high fantasy league scores and omit those who don't. To gauge the expected fantasy league score of a cricketer, it is necessary to determine fantasy league scores for all the games in which they played. Ideally this would be done using the scoring methodology of Section 2.1.2 to determine the fantasy league scores for each of the cricketer's games throughout their career. However, the career data collected for each cricketer does not include the number of balls faced by a batsmen, the number of run outs (direct or indirect) or the number of batsmen, bowlers, all rounders or wicket keepers dismissed by a bowler. As such the fantasy league scores in these cases need to be estimated.

It is decided to use a linear regression model to estimate these fantasy league scores. The linear regression model is fitted using the data from the 2007 ICC World Twenty20 and the 2008 IPL. Since the IPL is a domestic competition and the ICC World Twenty20 is an international competition it is argued that this model would be representative of both the domestic and international game. The model is then used to estimate fantasy league scores for each cricketer in every limited overs game in which they played (i.e. ListA and Twenty20). Both formats of the game are included as this arguably provides the most comprehensive information regarding a cricketer's fantasy league scoring ability.

The fantasy league scores estimated by this model are then used to develop fantasy league specific performance measures. The results of the fitted linear regression models are provided in Section 4.2.

3.2 Development and Adaptation of Performance Measures

The performance measures discussed in Chapter 1 provide insight and methods for the ranking and comparison of cricketers. These measures were developed to rank and compare cricketers under traditional circumstances i.e. not within a fantasy league.

The aim of this research is to provide a selection strategy for a fantasy league team. Therefore several of these performance measures are modified in order for them to better suit the scenario under consideration. It is believed that these adapted measures will provide more accurate results than the original measures. Adaptations are made using the results and observations from the points distribution analysis discussed in Section 3.1.2. Furthermore, several performance measures using the estimated fantasy league scores (using the linear model) are also developed.

These measures, along with the measures detailed in Section 1.3.2 and Section 1.3.4 are considered in the selection of a fantasy league team.

3.2.1 Lemmer's CBR Adaptation (*FCBR*)

Lemmer (2002) suggested the combined bowling rate (CBR) as a measure of the bowling performance of a cricketer. This measure is reviewed in Section 1.3.2 and defined in Equation 1.8. The definition used is

$$CBR = \frac{3r}{w + \frac{1}{6}b + \frac{rw}{b}}.$$

In order to modify this measure for fantasy league purposes, it is necessary to consider the fantasy league score allocation to various bowling outcomes (Table 2.3). It can be seen that the economy rate of a bowler as well as the wicket taking ability of a bowler are important as they account for a large portion of the point allocation. Measures of these abilities are already present in the current formulation of the CBR. The one attribute of a bowler's performance which scores very highly is their ability to bowl a maiden over[1]. There are 40 points allocated to bowling a maiden, this is the second highest points allocation, with only a 5 wicket haul of more value. The ability of a bowler to bowl a maiden must thus also be included.

It is proposed that a new measure, the maiden rate (MDN), be introduced. The maiden rate is defined as

$$MDN = \frac{o}{m} = \frac{\frac{1}{6}b}{m} \tag{3.1}$$

where $o =$ the number of overs bowled and $m =$ the number of maidens bowled. The maiden rate is interpreted as the number of overs bowled per maiden. The values of MDN range between $[1, \infty)$, and the measure can only be calculated once a maiden is bowled. The lower the value of the measure, the better the ability of the bowler. For example, a player with a maiden rate of 1 is a player who bowls a maiden every over (synonymous with an economy rate of zero). A player with a maiden rate of 2 bowls a maiden for every 2 completed overs, and so forth.

The modified CBR which includes the maiden rate (MDN), is termed the "fantasy combined bowling rate" ($FCBR$) and defined as

$$\begin{aligned} FCBR &= \frac{4}{\frac{1}{AVE^t_{BWL}} + \frac{1}{ECON^t_{BWL}} + \frac{1}{SR^t_{BWL}} + \frac{1}{MDN}} \\ &= \frac{4r}{w + \frac{1}{6}b + \frac{rw}{b} + \frac{\frac{1}{6}rb}{m}}. \end{aligned} \tag{3.2}$$

It is proposed that using this measure will assist in the selection of bowlers for a fantasy league team.

[1] A maiden over is a bowled when a bowler completes an over without conceding any runs.

3.2.2 Barr and Kantor Bowling Measure

Barr and Kantor (2004) developed a batting measure denoted as BK (in Equation 1.15) and defined as

$$BK = \left(SR^t_{BAT}\right)^{\alpha} \left(AVE^t_{BAT}\right)^{1-\alpha}.$$

This measure gives an indication of the batting ability of a cricketer. Sharp, Brettenny, Lourens, Gonsalves and Stretch (2009) use an adaptation of this measure to evaluate bowling performance. Similar to that of Sharp et. al. (2009) the following bowling performance measure is defined

$$BK_{BWL} = \left(ECON^t_{BWL}\right)^{\beta} \left(AVE^t_{BWL}\right)^{1-\beta} \tag{3.3}$$

where $0 \leq \beta \leq 1$. This measure gives an indication of the bowlers ability to take wickets (AVE^t_{BWL}) and to limit the number of runs conceded ($ECON^t_{BWL}$). The value of β indicates the perceived importance of the bowling economy rate with respect to the bowling average. Since low values of AVE^t_{BWL} and $ECON^t_{BWL}$ indicate a good bowler, it is clear that low values of BK_{BWL} are preferred.

3.2.3 Adjusted Barr and Kantor Batting Measure

To develop a performance measure to be used in a fantasy league setting, an adjustment to the methodology of Barr and Kantor (2004) is proposed. Barr and Kantor (2004) discuss a method to combine the batting average and strike rate of a batsmen. The fantasy league used in this study only considers the number of runs scored by a cricketer in a match and does not account for "out" or "not out" batting scores or even whether the cricketer batted or not. Thus the batting average used by Barr and Kantor (2004) may provide misleading results in a fantasy setting as many "not out" scores can inflate this statistic. To determine the number of runs scored by a cricketer per match in which they played an adjusted batting average is proposed. This measure is termed the mean batting score ($\overline{BT}$) of a cricketer and is defined as

$$\overline{BT} = \frac{\text{total number of runs scored}}{\text{total number of matches played}}. \tag{3.4}$$

This measure is used since the fantasy league score of a player is based on their participation in a given match (i.e. if a batsman doesn't bat in a given match his fantasy league score for batting is zero). The above measure allocates a score of 0 to a player who participated in a match but did not bat. In so doing the $\overline{BT}$ measure will increase for batsmen who tend to bat more often (for example: opening batsmen). This is ideal for a fantasy league team as a participant would want to select batsmen who tend to bat more frequently (thus increasing their opportunity to score points). The adjusted Barr and Kantor batting measure is now defined as

$$BK_{BAT} = \left(SR^t_{BAT}\right)^{\alpha} \left(\overline{BT}\right)^{1-\alpha} \tag{3.5}$$

where $0 \leq \alpha \leq 1$.

3.2.4 Gerber and Sharp Adaptation

The approach used by Gerber and Sharp (2006) to gauge a cricketer's ability is outlined in Section 1.3.5. The batting ability of cricketer i is measured using the ratio of his batting average to the mean batting average of all batsmen (given in Equation 1.16 as BT_i^{GS}). The bowling ability of a cricketer is similarly calculated using the bowling economy rate, however this measure (given in Equation 1.18 as BL_i^{GS}) is adjusted so that high values of the measure indicate more desirable bowlers. The approach used by Gerber and Sharp (2006) is unique in that it provides a researcher with a method to gauge the fielding ability of a cricketer (given in Equation 1.19 as FLD_i^{GS}). This is useful as a combination of these measures can be used to assess the overall ability of a cricketer. In this study the calculation of BT_i^{GS} is adapted to fit a fantasy league scenario. This is done by calculating BT_i^{GS} using a mean batting score ($\overline{BT}$) rather than the batting average as in Equation 1.16.

It is decided that a linear combination should be used in order to combine all the above measures into a single overall performance measure. In other words, for cricketers listed as specialist batsmen, the measure is defined as

$$BAT_i = a_1 BT_i^{GS} + a_2 BL_i^{GS} + a_3 FLD_i^{GS} \tag{3.6}$$

where i indicates the i^{th} batsman. Since batsmen aren't selected for their bowling ability, BL_i^{GS} is not calculated for batsmen and a mean value of $BL_i^{GS} = 1$ is used.

For bowlers, the following measure is defined

$$BWL_i = b_1 BT_i^{GS} + b_2 BL_i^{GS} + b_3 FLD_i^{GS} \tag{3.7}$$

where i indicates the i^{th} bowler. Similar to the justification used for batsmen, bowlers are not selected for their batting ability, BT_i^{GS} is not calculated and a mean value of $BT_i^{GS} = 1$ is used.

The measure for wicket keepers is defined as

$$WKR_i = c_1 BT_i^{GS} + c_2 BL_i^{GS} + c_3 FLD_i^{GS} \tag{3.8}$$

where i indicates the i^{th} wicket keeper. As with batsmen, a value of $BL_i^{GS} = 1$ is used since wicket keepers do not traditionally bowl.

For all rounders, the measure is defined as

$$ALR_i = d_1 BT_i^{GS} + d_2 BL_i^{GS} + d_3 FLD_i^{GS} \tag{3.9}$$

where i indicates the i^{th} all rounder.

The values of a_i, b_i, $c_{i,}$ and d_i where $i \in \{1, 2, 3\}$ are determined using the results obtained from the analysis 2007 ICC World Twenty20. These values are calculated as the ratio of the fantasy league points scored in each discipline (i.e. batting, bowling and fielding) to the total number of points scored in that discipline. This is done for each player category (i.e. batsman, bowler, all rounder and wicket keeper). These values are calculated and provided in Section 4.1.

It must be noted that the standard deviations of the indices (BT_i^{GS}, BWL_i^{GS} and FLD_i^{GS}) are equated (to the fourth decimal place) using the method described in Lemmer (2004) and detailed in Section 1.3.3. The base value used is the standard deviation of FLD_i^{GS} for each category as this is the only index common to all measures. This was done in order to standardise the results.

3.2.5 Measures using the Estimated Fantasy League Scores

To evaluate the fantasy league performance of cricketers, the estimated fantasy league scores using the linear model (Section 3.1.2) are used. Basic descriptive statistics, such as the mean and median, are used to assess the fantasy performance of players.

To gauge the current fantasy performance of cricketers, measures which focus on (or allocate a higher weight to) the more recent scores are used. The measures used in this study are the m-point moving average denoted (MA_m), the exponentially weighted average denoted (AVE^{exp}), the mean and the median.

Suppose that $Y_1, Y_2, \ldots, Y_n$ denote the estimated fantasy league scores for a cricketer who has played n limited overs (ListA and Twenty20) matches. Suppose further that the Y_n denotes the most recent score, Y_{n-1} denotes the second most recent score and so on. This would imply that Y_1 denotes the estimated fantasy league score of the cricketer in their first official limited overs match.

The m-point Moving Average

The m-point moving average is defined as

$$MA_m = \frac{Y_n + Y_{n-1} + \cdots + Y_{n-m+1}}{m}$$

for $n \geq m$. This method thus only considers the last m matches in which the cricketer took part. The measure is thus defined as the arithmetic mean of the estimated fantasy league scores obtained for the cricketer in their last m matches. If $n < m$, the arithmetic mean is used for the available data. The measure is defined as

$$F_{MA_m} = \begin{cases} \frac{Y_n + Y_{n-1} + \cdots + Y_1}{n} & \text{if } n < m \\ \frac{Y_n + Y_{n-1} + \cdots + Y_{n-m+1}}{m} & \text{if } n \geq m. \end{cases} \tag{3.10}$$

This measure gives an indication of the current performance (or "form") of the cricketer. For the purpose of this study it is decided to use $m \in \{5, 10\}$. These values are arbitrarily chosen to represent the most recent form ($m = 5$) and an extended period of form ($m = 10$) of the cricketer.

The Exponentially Weighted Average

Using the definition of the exponentially weighted average is defined in Section 1.3.3, the exponentially weighted average of the estimated fantasy league score is defined as

$$F_{EWA} = AVE^{exp} = wY_n + \alpha w Y_{n-1} + \alpha^2 w Y_{n-2} + \cdots + \alpha^{n-1} w Y_1 \tag{3.11}$$

where $0 \leq \alpha \leq 1$ and w is defined as in Section 1.3.3. Setting $\alpha = 1$ would result in the arithmetic mean thus giving every observation an equal weight. Should $\alpha = 0$ only the most recent observation is considered. An α value between 0 and 1 results in higher weightings for more recent data points. In accordance with Lemmer (2004) and the PriceWaterhouseCoopers ratings system of 2002 it is decided to use $\alpha = 0.96$.

The Mean and Median

The mean and median of the estimated fantasy league scores are also considered as performance measures for each of the cricketers. The fantasy mean is defined as

$$F_{mean} = \frac{1}{n}\sum_{i=1}^{n} Y_i. \tag{3.12}$$

The fantasy median is defined as

$$F_{med} = \begin{cases} Y^{\left[\frac{n+1}{2}\right]} & \text{if } n \text{ is odd} \\ \frac{Y^{\left[\frac{n}{2}\right]} + Y^{\left[\frac{n+2}{2}\right]}}{2} & \text{if } n \text{ is even} \end{cases} \tag{3.13}$$

where $Y^{[i]}$ is the i^{th} order statistic of $Y_1, Y_2, \ldots, Y_n$.

These statistics are created in order to provide measures which are specifically related to fantasy league cricket, rather than cricket in general. By using these measures it is hoped that the fantasy league ability of a cricketer may be better estimated.

3.3 Evaluation of Performance Measures

To evaluate the usefulness of the performance measures in Section 3.2 a fantasy league setting is defined for the 2007 ICC World Twenty20 tournament. These measures are evaluated to determine which is the best measure for establishing a starting line up for the 2008 IPL fantasy league. The measures are also used to assist with team changes as the tournament progresses. It must be noted that evaluating the performance measures using the 2007 ICC World Twenty20 is not an ideal approach. Since the 2008 IPL was held in a different country and thus under different playing conditions, the results have limitations. However, since the 2008 IPL was the first of its kind no other information is available and as such the 2007 ICC World Twenty20 data is used.

The usefulness of a measure is determined using the coefficient of correlation. This statistic provides a method for ranking the performance measures. This ranking is based on the correlation between the values of the performance measures for a cricketer and the total number of fantasy league points scored by the cricketer. The performance measures are then ranked according to the value of the coefficient of correlation. High values of the coefficient of correlation imply that "good" values of the performance measure associate with high fantasy league scores. The performance measures with the highest estimated correlation are thus selected to be used in the objective function of the BIP model used to select the fantasy league team (Section 3.4). It is envisioned that optimising this function of the performance measures will result in maximising the total fantasy league score.

A number of performance measures are evaluated for each player category (i.e batsmen, bowlers, all rounders and wicket keepers). The results of this evaluation are provided and discussed in Section 4.3.2.

3.3.1 The Coefficient of Correlation (ρ)

To assess the usefulness of a performance measure the coefficient of correlation (ρ) is used. The sample estimate of the coefficient of correlation is defined in Mendenhall and Sincich (2003) as

$$\begin{aligned}\hat{\rho} &= \frac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}} \\ &= \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2\sum(y_i-\bar{y})^2}}\end{aligned} \tag{3.14}$$

were $\hat{\rho}$ is the measure of the strength of the linear relationship between the variables x and y. The range of $\hat{\rho}$ is given as $-1 \leq \hat{\rho} \leq 1$, with values near 1 indicating a strong positive linear relationship and values near -1 indicating a strong negative linear relationship between the variables. Values of $\hat{\rho}$ near zero indicate no linear relationship is present.

In order to assess the usefulness of a performance measure, the coefficient of correlation between the values of the performance measure (for each player in each category) and the total fantasy league score of the player is calculated. The total fantasy league score is used instead of the average fantasy league score for each player as there are a number of cricketers who only played in one or two games, and these values skewed the results.

For all but three measures (CBR, $FCBR$ and BK_{BWL}) values of the coefficient of correlation near 1 are desired. This indicates that cricketers who have high values of the performance measure would score highly in a fantasy league. For CBR, $FCBR$ and BK_{BWL} where low values of these measures indicate good cricketers, it is desired that the values of the coefficient of correlation be near -1.

It must be noted that all performance measures for the 2007 ICC World Twenty20 are calculated using data from prior to the commencement of the competition. This is done in order to investigate the usefulness of a performance measure to select a fantasy league team using data only available at the start of the competition.

The best performance measure is chosen for each category according to their $\hat{\rho}$ value, and these performance measures are then used to select the team for the 2008 IPL.

Optimisation of the Barr and Kantor Measures using $\hat{\rho}$

In order to select the best values of α for the BK_{BAT} measures, as well as the best values of β for the BK_{BWL} measures the value of the estimated correlation coefficient is used.

The value of α is chosen (for each player category) by using *Microsoft Excel's Solver*. *Solver* is used to maximise the estimated correlation (measured by $\hat{\rho}$) between the BK_{BAT} values and the total fantasy league scores by changing the value of α. This is done in order to select the most appropriate value of α for the fantasy league selection process.

It is clear that if the value of α is 0 or 1 the BK_{BAT} measure is equivalent to the mean batting score or batting strike rate respectively. Thus, should the optimal value of α be one of these end points, the measure is no longer a Barr and Kantor measure. To prevent this an added restriction of $0.1 \leq \alpha \leq 0.9$ is imposed. This adjustment ensures that each measure ($\overline{BT}$ and SR^t_{BAT}) contributes at least 10% towards the calculation of the BK_{BAT} measure. (The value of 10% is chosen arbitrarily). This optimisation procedure is run for each category for both ListA and Twenty20 games (separately).

A similar procedure is followed to select the β values for the BK_{BWL} measures.

3.3.2 List of the Performance Measures used

The performance measures assessed are listed below and denoted by the terms in brackets.

- The mean of the estimated (and standardised) fantasy league scores (denoted F_{mean}).
- The median of the estimated (and standardised) fantasy league scores (denoted F_{med}).
- The 5 point moving average of the estimated (and standardised) fantasy league scores (denoted $F_{MA_{10}}$).
- The 10 point moving average of the estimated (and standardised) fantasy league scores (denoted $F_{MA_{10}}$).
- The exponentially weighted average of the estimated (and standardised) fantasy league scores (denoted F_{EWA}).
- Lemmer's batting performance measure for ListA games (BP_{LtA}) and for Twenty20 games (BP_{Tw20}).
- Lemmer's combined bowling rate for ListA games (CBR_{LtA}) and for Twenty20 games (CBR_{Tw20}).
- The Fantasy combined bowling rate for ListA games ($FCBR_{LtA}$) and for Twenty20 games ($FCBR_{Tw20}$).
- Adapted Gerber and Sharp Measure for batsmen for ListA games (BAT_i^{LtA}) and for Twenty20 games (BAT_i^{Tw20}).
- Adapted Gerber and Sharp Measure for bowlers for ListA games (BWL_i^{LtA}) and for Twenty20 games (BWL_i^{Tw20}).
- Adapted Gerber and Sharp Measure for all rounders for ListA games (ALR_i^{LtA}) and for Twenty20 games (ALR_i^{Tw20}).
- Adapted Gerber and Sharp Measure for wicket keepers in ListA games (WKR_i^{LtA}) and for Twenty20 games (WKR_i^{Tw20}).
- The adjusted Barr and Kantor batting selection criterion for ListA games (BK_{BAT}^{LtA}) and for Twenty20 games (BK_{BAT}^{Tw20}).

- The adapted Barr and Kantor measure for bowlers for ListA games (BK^{LtA}_{BWL}) and for Twenty20 games (BK^{Tw20}_{BWL}).

Since every cricketer is classified into one of four categories (batsman, bowler, all rounder, wicket keeper), to assess the performance of each individual player different measures are used.

Table 3.1 indicates the performance measures evaluated for each player category (as determined on CricInfo IPL Fantasy League (2008)).

Measure	**Batsmen**	**Bowlers**	**All Rounders**	**Wicket Keepers**
F_{mean}	★	★	★	★
F_{med}	★	★	★	★
$F_{MA_{10}}$	★	★	★	★
F_{MA_5}	★	★	★	★
F_{EWA}	★	★	★	★
BP_{LtA}	★		★	★
BP_{Tw20}	★		★	★
CBR_{LtA}		★	★	
CBR_{Tw20}		★	★	
$FCBR_{LtA}$		★	★	
$FCBR_{Tw20}$		★	★	
BAT_i^{LtA}	★			
BAT_i^{Tw20}	★			
BWL_i^{LtA}		★		
BWL_i^{Tw20}		★		
ALR_i^{LtA}			★	
ALR_i^{Tw20}			★	
WKR_i^{LtA}				★
WKR_i^{Tw20}				★
BK^{LtA}_{BAT}	★		★	★
BK^{LtA}_{BAT}	★		★	★
BK^{LtA}_{BWL}		★	★	
BK^{Tw20}_{BWL}		★	★	

Table 3.1: Performance measures which were assessed for each player category

3.4 Formulation of a Binary Integer Programming Model

The problem investigated in this dissertation can be stated as follows:

A team of *eleven players* must be chosen. Each player falls within a designated category (batsman, bowler, all rounder, wicket keeper) and is given a price. The team must be chosen *according to a predetermined formation* provided in Section 2.1.1. The total value of the team cannot exceed *1 000 000 units*. The tournament is partitioned into several *fantasy stages*. A maximum number of *s changes* can be made to this team during the course of the tournament. The objective is to maximise the fantasy league score by making the appropriate team selection changes at the start of each stage. In so doing the total number of points across the entire tournament will be maximised.

The problem is a BIP problem, as the decision to be made at each stage of the tournament is whether or not a player should be selected. The BIP technique is useful as it allocates a value of 1 to the decision variable if the player should be selected and a value of 0 if not selected.

In order to formulate the BIP model, the following three components must be defined

- The decision variables;
- The constraints; and
- The objective function.

3.4.1 The Decision Variables

Suppose that there are n players available for selection during the tournament which consists of q fantasy stages. The decision vector $\underline{X}$ $(nq \times 1)$ is defined as

$$\underline{X}' = \begin{bmatrix} x_1 & x_2 & \cdots & x_{nq} \end{bmatrix}$$

where $x_{i+(k-1)n}$ is the decision variable for player i for stage k and $k \in \{1, 2, \ldots, q\}$. The decision variable is thus defined as

$$x_{i+(k-1)n} = \begin{cases} 1 & \text{if player } i \text{ is selected for stage } k \\ 0 & \text{otherwise.} \end{cases}$$

3.4.2 The Constraints

The three constraints imposed on a fantasy league team address the following limitations imposed by the rules of the league.

- Budget;
- Player and team composition; and
- Changes.

Before the constraints are formulated the following two vectors are defined:

1. The "zero" vector, denoted $\underline{\mathbf{0}}$ $(n \times 1)$ is defined as:

$$\underline{\mathbf{0}}' = \begin{bmatrix} 0 & 0 & \cdots & 0 \end{bmatrix}$$

2. The "one" vector, denoted $\underline{\mathbf{1}}$ $(n \times 1)$ is defined as:

$$\underline{\mathbf{1}}' = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$$

Budget Constraints

The "budget" constraints address the budget limitation imposed on a fantasy league team. The constraint ensures that the total value of the players selected does not exceed 1 000 000 units.

Define the price of player i as p_i. The price vector $\underline{P}$ $(n \times 1)$ is thus defined as

$$\underline{P}' = \left[\begin{array}{cccc} p_1 & p_2 & \cdots & p_n \end{array} \right].$$

Given that the prices of the players are fixed throughout this fantasy league, this vector remains constant.

The budget must not be exceeded at any stage of the tournament, it is thus necessary to include a budget constraint for each of the q stages. The budget vector for stage k denoted $\underline{F}_k$ $(nq \times 1)$ is now defined as

$$\underline{F}_k' = \left[\begin{array}{ccccccc} \overbrace{\underline{\mathbf{0}}'}^{1^{\text{st}} \text{ stage}} & \cdots & \overbrace{\underline{\mathbf{0}}'}^{(k-1)^{\text{th}} \text{ stage}} & \underbrace{\underline{P}'}_{k^{\text{th}} \text{ stage}} & \overbrace{\underline{\mathbf{0}}'}^{(k+1)^{\text{th}} \text{ stage}} & \cdots & \overbrace{\underline{\mathbf{0}}'}^{q^{\text{th}} \text{ stage}} \end{array} \right]. \tag{3.15}$$

The budget constraints can now be expressed as

$$\text{Constraint 1}: \quad \underline{F}_k'\underline{X} \leq 1\ 000\ 000 \quad \forall\ k \in \{1, 2, \ldots, q\}.$$

These constraints ensure that the budget of 1 000 000 units is not exceeded at any stage of the competition.

Player and Formation Constraints

The "player" constraints ensure that the total number of players in the fantasy league team equals 11 at all stages of the tournament. As with the budget constraint it is necessary to include the player constraint at each stage of the tournament. The player vector for stage k, denoted $\underline{E}_k$ $(nq \times 1)$, is defined as

$$\underline{E}_k' = \left[\begin{array}{ccccccc} \overbrace{\underline{\mathbf{0}}'}^{1^{\text{st}} \text{ stage}} & \cdots & \overbrace{\underline{\mathbf{0}}'}^{(k-1)^{\text{th}} \text{ stage}} & \underbrace{\underline{1}'}_{k^{\text{th}} \text{ stage}} & \overbrace{\underline{\mathbf{0}}'}^{(k+1)^{\text{th}} \text{ stage}} & \cdots & \overbrace{\underline{\mathbf{0}}'}^{q^{\text{th}} \text{ stage}} \end{array} \right].$$

The player constraints are now formulated as

$$\text{Constraint 2}: \quad \underline{E}_k'\underline{X} = 11 \quad \forall\ k \in \{1, 2, \ldots, q\}$$

which ensures that 11 players are selected at each stage of the tournament.

In order to include the constraints on the formation of a team as required by the rules of the fantasy league, the following variables are defined.

$$d_i = \begin{cases} 1 & \text{if player } i \text{ is a batsman} \\ 0 & \text{otherwise} \end{cases} \tag{3.16}$$

$$c_i = \begin{cases} 1 & \text{if player } i \text{ is a bowler} \\ 0 & \text{otherwise} \end{cases} \tag{3.17}$$

$$b_i = \begin{cases} 1 & \text{if player } i \text{ is a wicket keeper} \\ 0 & \text{otherwise} \end{cases} \tag{3.18}$$

$$a_i = \begin{cases} 1 & \text{if player } i \text{ is an all rounder} \\ 0 & \text{otherwise} \end{cases} \tag{3.19}$$

These variables are defined for all $i \in \{1, 2, \ldots, n\}$.

The number of batsmen selected for a fantasy league team at each stage of the competition must be 4. In order to include this as one of the constraints the following batting vector, denoted $\underline{BT}$ $(n \times 1)$, is defined as

$$\underline{BT}' = \left[\begin{array}{cccc} d_1 & d_2 & \cdots & d_n \end{array} \right].$$

The *batsmen* vector for stage k, denoted $\underline{D}_k$ $(nq \times 1)$, is defined similarly to Equation 3.15 as

$$\underline{D}_k' = \left[\begin{array}{ccccccc} \overbrace{\underline{\mathbf{0}}'}^{1^{\text{st}}\text{ stage}} & \cdots & \overbrace{\underline{\mathbf{0}}'}^{(k-1)^{\text{th}}\text{ stage}} & \underbrace{\underline{BT}'}_{k^{\text{th}}\text{ stage}} & \overbrace{\underline{\mathbf{0}}'}^{(k+1)^{\text{th}}\text{ stage}} & \cdots & \overbrace{\underline{\mathbf{0}}'}^{q^{\text{th}}\text{ stage}} \end{array} \right].$$

Using a similar approach, the bowling, wicket keeping and all rounder vectors (denoted $\underline{BL}$ $(n\times1)$, $\underline{WK}$ $(n \times 1)$ and $\underline{AR}$ $(n \times 1)$ respectively) are defined as

$$\begin{aligned} \underline{BL}' &= \left[\begin{array}{cccc} c_1 & c_2 & \cdots & c_n \end{array} \right]; \\ \underline{WK}' &= \left[\begin{array}{cccc} b_1 & b_2 & \cdots & b_n \end{array} \right]; \\ \underline{AR}' &= \left[\begin{array}{cccc} a_1 & a_2 & \cdots & a_n \end{array} \right] \end{aligned}$$

respectively. Similar to formulation the batsmen vector, the *bowler* vector of stage k, denoted $\underline{C}_k$ $(nq \times 1)$, is defined as

$$\underline{C}_k' = \left[\begin{array}{ccccccc} \overbrace{\underline{\mathbf{0}}'}^{1^{\text{st}}\text{ stage}} & \cdots & \overbrace{\underline{\mathbf{0}}'}^{(k-1)^{\text{th}}\text{ stage}} & \underbrace{\underline{BL}'}_{k^{\text{th}}\text{ stage}} & \overbrace{\underline{\mathbf{0}}'}^{(k+1)^{\text{th}}\text{ stage}} & \cdots & \overbrace{\underline{\mathbf{0}}'}^{q^{\text{th}}\text{ stage}} \end{array} \right].$$

The *wicket keeper* vector for stage k, $\underline{B}_k$ $(nq \times 1)$, denoted is defined as

$$\underline{B}_k' = \left[\overbrace{\underline{\mathbf{0}}'}^{1^{\text{st}}\text{ stage}} \quad \cdots \quad \overbrace{\underline{\mathbf{0}}'}^{(k-1)^{\text{th}}\text{ stage}} \quad \underbrace{\underline{WK}'}_{k^{\text{th}}\text{ stage}} \quad \overbrace{\underline{\mathbf{0}}'}^{(k+1)^{\text{th}}\text{ stage}} \quad \cdots \quad \overbrace{\underline{\mathbf{0}}'}^{q^{\text{th}}\text{ stage}} \right].$$

Lastly, the *all rounder* vector for stage k, denoted $\underline{A}_k$ $(nq \times 1)$, is similarly defined as

$$\underline{A}_k' = \left[\overbrace{\underline{\mathbf{0}}'}^{1^{\text{st}}\text{ stage}} \quad \cdots \quad \overbrace{\underline{\mathbf{0}}'}^{(k-1)^{\text{th}}\text{ stage}} \quad \underbrace{\underline{AR}'}_{k^{\text{th}}\text{ stage}} \quad \overbrace{\underline{\mathbf{0}}'}^{(k+1)^{\text{th}}\text{ stage}} \quad \cdots \quad \overbrace{\underline{\mathbf{0}}'}^{q^{\text{th}}\text{ stage}} \right].$$

The number of batsmen in a fantasy league team must equal 4 at each stage of the competition. These constraints are included as

$$\text{Constraint 3}: \quad \underline{D}_k'\underline{X} = 4 \quad \forall\, k \in \{1, 2, \ldots, q\}.$$

The number of bowlers in a team can be either 3 or 4, thus the following constraints are added

$$\text{Constraint 4}: \quad 3 \le \underline{C}_k'\underline{X} \le 4 \quad \forall\, k \in \{1, 2, \ldots, q\}.$$

The wicket keeper constraints (only one wicket keeper per team) follow as

$$\text{Constraint 5}: \quad \underline{B}_k'\underline{X} = 1 \quad \forall\, k \in \{1, 2, \ldots, q\}.$$

Finally, a fantasy league team can consist of either 2 or 3 all rounders, these constraints are formulated as

$$\text{Constraint 6}: \quad 2 \le \underline{A}_k'\underline{X} \le 3 \quad \forall\, k \in \{1, 2, \ldots, q\}.$$

Constraint on the Number of Changes

For the purpose of generality, although the fantasy rules state that a maximum of 20 changes can be made, it is assumed that s changes are allowed where $s \in \{0, 1, 2, \ldots\}$.

In a tournament consisting of q stages, there are $q-1$ opportunities to make changes to the team, after stage 1, after stage 2, all the way to after stage $q-1$ (i.e. before the last stage (stage q)). In order to count the number of changes made, the following summation is required

$$\sum_{i=1}^{n(q-1)} (x_i \times x_{i+n}). \tag{3.20}$$

Since $x_i = 0$ or 1 $\forall i$, we have that $x_i \times x_{i+n} = 0$ or 1. In fact, when considering x_i and x_{i+n}, x_i indicates player i's selection during the current stage, and x_{i+n} indicates his selection in the following stage. Thus $x_i \times x_{i+n} = 1$ indicates that player i was selected for both stages, whereas if $x_i \times x_{i+n} = 0$, this may indicate that a change was made. This change could be that player i was included in the current stage and removed for the following stage, or vice versa. Clearly

$x_i \times x_{i+n} = 0$ could also indicate that player i was selected for neither stage, however this does not effect the results. The summation given in Equation 3.20 thus indicates the number of players who remained from one stage to the next, throughout the duration of the tournament.

If we consider a 2 stage tournament, Equation 3.20 becomes

$$\sum_{i=1}^{n(2-1)} (x_i \times x_{i+n}) = \sum_{i=1}^{n} (x_i \times x_{i+n})$$

this equation represents the number of players who remained in the initial team after the first stage. Clearly if no changes were made, then $\sum_{i=1}^{n} (x_i \times x_{i+n}) = 11$. If one was made then $\sum_{i=1}^{n} (x_i \times x_{i+n}) = 10$. Thus, in order to determine the number of changes made in this example, the following equation is used

$$11 \times (2-1) - \sum_{i=1}^{n(2-1)} (x_i \times x_{i+n}).$$

So, for competitions with q stages the equation to count the total number of changes made simply becomes

$$11 \times (q-1) - \sum_{i=1}^{n(q-1)} (x_i \times x_{i+n}).$$

The constraint on the number of changes that can be made is thus given as

$$\text{Constraint 7}: \quad 11 \times (q-1) - \sum_{i=1}^{n(q-1)} (x_i \times x_{i+n}) \leq s.$$

3.4.3 The Objective Function

The objective function is formulated using a retrospective approach and a prospective approach. A retrospective approach is formulated to determine the fantasy league team selection for the tournament when all the fantasy league results are available (i.e. at the end of the tournament). A prospective approach is formulated to determine a selection procedure which only uses the values of the performance measures, and no information regarding future fantasy league scores is known. This approach can be used from the start of the tournament.

Retrospective Approach

The retrospective approach is formulated by defining a score vector, denoted $\underline{S}\,(nq \times 1)$ and defined as

$$\underline{S}' = \begin{bmatrix} s_1 & s_2 & \cdots & s_{nq} \end{bmatrix}$$

where $s_{i+(k-1)n}$ denotes the fantasy league score of player i for stage k where $k \in \{1, 2, \ldots, q\}$. Clearly this information is only available at the conclusion of the tournament. The objective function for this approach is then given as

$$Z = \underline{S}'\underline{X} = s_1 x_1 + s_2 x_2 + \cdots + s_{nq} x_{nq}. \tag{3.21}$$

Since the information in $\underline{S}$ is only available at the end of the tournament, this is a retrospective approach. Maximising the objective function Z according to the constraints results in the optimal solution to the fantasy league selection problem after the results are known. It represents the best possible selection and change strategy. However this strategy is only available at the completion of the tournament, and is of little value to a fantasy league participant who requires a strategy at the commencement of the competition.

The solution of the problem using this objective function (Z) is thus viewed as the ideal scenario, and is used for comparative purposes only.

Prospective Approach

To develop a practical approach to the fantasy league team selection process, a prospective approach is necessary. This method is where at each stage of the fantasy league, only data up to that stage is available for analysis. It is thus necessary to use the performance measures and the tournament stage data to select or modify the fantasy league team. The values of the available data are then optimised using the BIP model at each stage to determine the fantasy league team at each stage.

To develop a formulation of the objective function for this approach, the following performance vector, denoted $\underline{M}$ $(nq \times 1)$, is defined

$$\underline{M}' = \left[\begin{array}{cccc} m_1 & m_2 & \cdots & m_{nq} \end{array} \right]$$

where $m_{i+(k-1)n}$ denotes the performance measure of player i calculated *prior* to the commencement of stage k where $k \in \{1, 2, \ldots, q\}$. It must be noted that the performance measures mentioned are calculated using the most recent information available. The performance measures used to construct the $\underline{M}$ vector are those which performed the best according to the value of the coefficient of correlation. The team selection process would now become a matter of ensuring that the cricketers with the best values of the chosen performance measures are selected. By optimising the sum of the performance measures (at each stage) it envisioned that the best performing players will be selected for the following stage.

To define the information available to a team selector, the possibility that a particular player is injured or unavailable for a particular stage must be accounted for. Since knowledge of these events would most likely be known to a fantasy league participant, some adjustments need to be made to the performance vector $\underline{M}$ $(nq \times 1)$.

Suppose that player i did not participate in a particular stage of the tournament, say stage t. To adjust for this the value given to $m_{i+(t-1)n}$ is zero (i.e. $m_{i+(t-1)n} = 0$). The vector containing this adjusted information is denoted $\underline{R}$ $(nq \times 1)$ and defined as

$$\underline{R}' = \left[\begin{array}{cccc} r_1 & r_2 & \cdots & r_{nq} \end{array} \right]$$

where $r_{i+(k-1)n}$ denotes the adjusted performance measure (i.e. with zero's) for player i for stage k.

The data required to calculated the performances vectors $\underline{M}$ $(nq \times 1)$ and $\underline{R}$ $(nq \times 1)$ is collected sequentially. As a result these vectors are only fully defined at the last stage of the competition. The vectors $\underline{M}$ $(nq \times 1)$ and $\underline{R}$ $(nq \times 1)$ are thus split into q separate vectors of size $n \times 1$, denoted $\underline{M_k}$ $(n \times 1)$ and $\underline{R_k}$ $(n \times 1)$ where $k \in \{1, 2, \ldots, q\}$ respectively. These vectors represent the performance measure and adjusted performance measures calculated for each stage sequentially.

The prospective approach then runs separate optimisation procedures for each stage of the competition. One procedure for each of the $\underline{M_k}$ $(n \times 1)$ and $\underline{R_k}$ $(n \times 1)$ vectors. In so doing, the method remains entirely prospective in nature and thus is of use to a fantasy league team selector.

In order to facilitate this prospective approach a few changes need to be made to Constraint 7. The constraint is given as

$$11 \times (q-1) - \sum_{i=1}^{n(q-1)} (x_i \times x_{i+n}) \leq s$$

where q is the number of stages and s is the maximum number of changes allowed. This constraint considers the problem in its entirety and limits the total number of changes made across a whole competition. This constraint can only be used if the entire performance vector, $\underline{M}$ $(nq \times 1)$, is defined. Thus rather than running a single optimisation over the competition as a whole, q separate BIP problems are solved (one for each stage of the competition).

A problem arises when the maximum number of changes over the entire competition is considered. Suppose that the maximum number of changes allowed is s and that changes can only be made after each stage. Thus if there are q stages there are only $q-1$ opportunities to make changes to the team. A simple method for calculating the number of changes allowed at each stage is by dividing the maximum number of changes allowed (s) by the number of opportunities available to make changes ($q-1$). Thus the maximum number of changes allowed to be made between consecutive stages is

$$s_i = \frac{s}{q-1}.$$

Although it is not ideal to limit the number of changes between stages (as this could result in a non-optimal overall solution), it is a solution that a prospective approach can employ. Should $s_i \notin \mathbb{Z}$, then s_i is rounded to the nearest integer and the value of s_i for the final stage is adjusted such that $\sum s_i = s$. Furthermore, should less than s_i changes be made in any given stage, then the maximum number of changes allowed for the next stage is increased by the number of unused changes in the previous stage. This procedure is continued throughout the competition.

An alternative method to determine the best way to make changes is by considering the extent to which the value of the statistics improve when a cricketer is changed. In this case small increments would result in the team staying constant across the stages, whereas large increments would justify a change in the team. This is mentioned as an area for future investigation.

The q individual integer programming problems are defined for the following periods: (1) selection of the initial team (i.e. for Stage 1); (2) selection of team for Stage 2; (3) selection of team for Stage 3; $\cdots$; and (q) selection of team for stage q. These selections are all performed prior to the start of each stage.

For the selection of the initial team (i.e. for the first stage), the number of changes made is of no concern as the changes are only monitored from the commencement of the competition. Thus, in order to select this initial team, an integer programming problem with Constraints 1 to 6 is considered. The notation for the decision vector is changed from X to Y for the prospective approach. This is done to distinguish it from the decision vector used in the retrospective approach. Let the decision vector for stage 1, denoted $\underline{Y}_1$ $(n \times 1)$, where

$$\underline{Y}'_1 = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \end{bmatrix}$$

and y_{1i} is a binary decision variable indicating whether or not player i is selected for Stage 1. The objective function can thus be written as

$$Z_1 = \underline{Y}'_1 \underline{M}_1 = m_1 y_{11} + m_2 y_{12} + \cdots + m_n y_{1n} \tag{3.22}$$

where m_i are the performance measures defined for the performance vector $\underline{M}$.

In order to determine the selected teams for the following stages (i.e. $\underline{Y}_2$, $\underline{Y}_3$, $\cdots$, $\underline{Y}_q$), $q-1$ additional integer programs are run. For these optimisation routines, the number of changes allowed is limited to s_i. In order to satisfy this requirement the following constraint is included in each of the integer programming problems

$$\text{Constraint 8}: \quad 11 - \underline{Y}'_{k-1}\underline{Y}_k \leq s_i$$

where $\underline{Y}_{k-1}$ is the solution to the previous integer programming problem.

These integer programs only consist of one stage and include the Constraints 1 to 6 as well as Constraint 8. The integer program is run to determine the decision variable $\underline{Y}_k$ (i.e. to select the team for stage k) where

$$\underline{Y}'_k = \begin{bmatrix} y_{k1} & y_{k2} & \cdots & y_{kn} \end{bmatrix}.$$

and y_{ki} is a binary decision variable indicating whether or not player i is selected for stage k where $k \in \{2, 3, \ldots, q\}$. In order to determine $\underline{Y}_k$ the objective function

$$Z_k = \underline{Y}'_k \underline{M}_k = m_{1+(k-1)n} y_{k1} + m_{2+(k-1)n} y_{k2} + \cdots + m_{n+(k-1)n} y_{kn}$$

is used, where m_i are the performance measures defined for the performance vector $\underline{M}$. The decision vector $\underline{Y}_k$ therefore indicates the prospective selection strategy for stage k.

This optimisation problem is run $q-1$ times, and ideally the prospective decision vector $\underline{Y}$ $(nq \times 1)$ defined as

$$\underline{Y}' = \begin{bmatrix} \underline{Y}'_1 & \underline{Y}'_2 & \cdots & \underline{Y}'_q \end{bmatrix}$$

will be similar to decision vector $\underline{X}$ obtained by solving the single retrospective integer program with the objective function Z and Constraints 1 to 7.

The prospective methodology described is then run for a second time. This time the performance vector $\underline{M}$ $(nq \times 1)$ information is replaced with the adjusted performance vector $\underline{R}$ $(nq \times 1)$ information. By solving this integer program an entirely prospective solution is found for the scenario when injury and availability data are included.

An illustration of the results of these approaches is presented in Chapter 4.

This chapter provides a detailed description of the methodology used in this study. The use of preliminary descriptive statistics and linear regression modelling is discussed. The methodology behind the development and adaptations of new, fantasy league specific, performance measures is provided. An approach to evaluate the given performance measures is also discussed. Finally, a detailed description of the BIP used to select the fantasy league team is provided. The development of the BIP is the major contribution of this study to the body of academic knowledge whilst the application to fantasy league team selection for cricket is a first.

Chapter 4

Results and Discussion

4.1 Analysis of the Fantasy League Point Distribution

The results discussed in this section include the data collected across all the games of the 2007 ICC World Twenty20 tournament. The results are used for the allocation of weights to the adapted Gerber and Sharp measures described in Section 3.2.4. The results also provide insight into the nature of fantasy league scoring in a Twenty20 cricket competition. To analyse the data, it is necessary to categorise each of the cricketers participating in the competition into one of the following categories: batsman; bowler; all rounder or wicket keeper. The distribution of fantasy league points across each of the categories and each of the disciplines (viz. batting, bowling and fielding) is then determined.

4.1.1 Results

A total of 24 621 fantasy league points were scored in the competition. The distribution of these points according to the respective cricketing disciplines is given in Table 4.1.

Discipline	**Points**	%
Batting	9 906	40.23
Bowling	9 255	37.59
Fielding	4 160	16.90

Table 4.1: Summary statistics of fantasy league points according to cricketing discipline in the 2007 ICC World Twenty20

The remaining 5.28% (1 300 points) were allocated to man-of-the-match awards. This table implies that in this fantasy league setting the batting ability of a player is the most important ability followed by the bowling ability and then the fielding ability of a cricketer. These results lend support to the argument that the number of batsmen selected in a fantasy league team should be high. It must however be noted that all cricketers bat (if required) while not all bowl, and this could be the reason for the slightly inflated number of points allocated to the batting discipline. To further explore the distribution of points and gain a better understanding of the results of Table 4.1, it is necessary to consider the spread of points across the different player categories. The distribution of points across the player categories is given in Table 4.2.

Category	Points	%	Ave. Score	MOTM
Batsmen	7 319	29.73	38.72	9
Bowlers	7 612	30.92	45.58	9
All Rounders	7 361	29.90	42.30	7
Wicket Keepers	2 329	9.46	34.76	1

Table 4.2: Summary statistics of fantasy league points according to player category in the 2007 ICC World Twenty20

The information given in Table 4.2 shows that bowlers in fact score the most points, followed by all rounders, batsmen and then wicket keepers. These results suggest that the number of bowlers selected in a fantasy league team should be as high as possible. The total number of points accumulated by all rounders just exceeds that of batsmen, however, the average score of an all rounder is noticeably more than that of a batsman. This illustrates that all rounders should be preferred to batsmen. Given the restrictions imposed by the fantasy league under consideration in this research, the results of Table 4.2 suggest that the following team formation is optimal: 4 Batsmen, 4 Bowlers, 2 All Rounders and 1 Wicket Keeper.

Table 4.1 shows that although the batting discipline scores the most points, batsmen themselves are not the highest scorers. Rather, the high score allocated to the batting discipline can be attributed to the fact that all cricketers bat, while not all bowl. Interestingly, batsmen and bowlers both won an equal number of man-of-the-match awards (9), and yet the average batsmen score is still less than that of all rounders who won only 7 such awards.

In conclusion for this section the joint distribution of points for cricketing discipline and player category is considered. This is achieved by combining the results of Table 4.1 and Table 4.2. These results are tabulated in Table 4.3.

	Batting	Bowling	Fielding
Batsmen	0.8231	0.0138	0.1631
Bowlers	0.0289	0.8587	0.1124
All Rounders	0.4038	0.4322	0.1640
Wicket Keepers	0.5327	-0.0088	0.4761

Table 4.3: Ratio of fantasy league points scored in cricketing discipline for each player category in the 2007 ICC World Twenty20

This table gives the ratio of points scored in each cricketing discipline for the respective player categories. It must be noted that although the man-of-the-match awards do effect the total number of points scored by a player, "man-of-the-match" is not a cricketing discipline. As such, the ratios indicated in Table 4.3 exclude points awarded in this category.

The information provided in Table 4.3, in the case of batsmen and bowlers, provides no real surprises and is what can be expected. Batsmen score the majority of their points in the batting discipline and bowlers in the bowling discipline. Interesting to note is that players listed as batsmen can bowl, although it is unlikely. This explains the 1.38% allocated to the bowling discipline of the batsmen.

The all rounders category provides an interesting observation. All rounders score more points for bowling than for batting. This implies that bowling all rounders should possibly be preferred to batting all rounders (this observation is merely noted, as information regarding the all rounder types is not readily available for this study).

Lastly, the observation that 47.61% of wicket keepers scores comes from the fielding discipline is expected as wicket keepers are historically the most prominent fielders in the game. A point of concern in Table 4.3 is that the wicket keepers have a negative ratio for the bowling category. This is explained by the fact that T Taibu of Zimbabwe, who is a wicket keeper, did not keep wicket during every game and on one occasion was called upon to bowl. The resulting fantasy league score for his bowling was a negative score (-20), indicating a poor bowling performance. Since there is a small chance that players listed as wicket keepers can bowl, this result was included.

Using the result in Table 4.3 the adapted Gerber and Sharp measures detailed in Section 3.2.4 may now be fully defined. The measure for batsmen thus becomes

$$BAT_i = 0.8231 \cdot BT_i^{GS} + 0.0138 \cdot BL_i^{GS} + 0.1638 \cdot FLD_i^{GS}.$$

For bowlers

$$BWL_i = 0.0289 \cdot BT_i^{GS} + 0.8587 \cdot BL_i^{GS} + 0.1124 \cdot FLD_i^{GS}.$$

For wicket keepers

$$WKR_i = 0.5327 \cdot BT_i^{GS} - 0.0088 \cdot BL_i^{GS} + 0.4761 \cdot FLD_i^{GS}.$$

For all rounders

$$ALR_i = 0.4038 \cdot BT_i^{GS} + 0.4332 \cdot BL_i^{GS} + 0.1640 \cdot FLD_i^{GS}.$$

The results discussed in this section thus provide information necessary for the development and modification of cricket performance measures for fantasy leagues.

4.2 Linear Regression Model

A multiple linear regression model is fitted to data from the 2007 ICC World Twenty20 and the 2008 IPL. This linear model is used to estimate fantasy league scores for cricketers for games played outside these competitions. The following data is collected for each cricketer: the number of runs scored (and method of dismissal); the number of overs bowled; the number of maidens bowled; the number of runs conceded; the number of wickets taken and the number of catches taken. From this data an extra variable, named "duck", is included. A batsman scores a "duck" if he is dismissed without scoring any runs. The "duck" variable is thus a binary variable which equals 1 if the batsman scored a duck and 0 otherwise.

In order to fit a linear regression model the quantitative variables given in Table 4.4 are defined

Symbol	Variable	Type
y	Fantasy League Score	Integer
x_1	Runs Scored	Integer
x_2	Duck	Binary
x_3	Catches Taken	Integer
x_4	Overs Bowled	Real
x_5	Maidens Bowled	Integer
x_6	Runs Conceded	Integer
x_7	Wickets Taken	Integer

Table 4.4: Variables used in the linear regression model

The variables indicated in Table 4.4 are used in the linear model, as this is the only information collected for the games outside of the ICC World Twenty20 and IPL competitions (See Section 2.2.2).

Before the regression models are fitted, the data set is divided into two categories.

- Category 1: data points including the cricketers who both batted and bowled in the same match.
- Category 2: data points including the cricketers who only batted in a given match.

As a result of this distinction, it is required that two separate linear regression models be fitted. The fantasy league scores are then divided into y_1 = fantasy scores from category 1; and y_2 = fantasy scores from category 2.

The following model is used for the data points from Category 1

$$E(y_1) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 \tag{4.1}$$

the intercept parameter (β_0) is set equal to zero. This is done to ensure that all players who had no impact on a game i.e. $x_1 = x_2 = \ldots = x_7 = 0$, would score zero (as they would using the scoring method defined in Section 2.1.2).

Similarly the following model is used for data points from Category 2

$$E(y_1) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \tag{4.2}$$

with the equivalent rationalisation for setting $\beta_0 = 0$.

4.2.1 Results of the Regression Analysis

Microsoft Excel is used to perform the regression analysis and the models are estimated using the method of least squares.

Category 1

The model for Category 1 is given by

$$E(y_1) = 1.9195x_1 - 10.2564x_2 + 14.4717x_3 + 9.8940x_4 + 36.9303x_5 - 1.4858x_6 + 35.2458x_7 \tag{4.3}$$

This model is fitted using 988 data points and the results of the regression analysis indicate a very good fit. These results are tabulated in Table 4.5.

Measure	Value
Multiple R	0.9574
Adjusted R^2	0.9150
Standard Error	15.4704
p-value (F-Test)	0

Table 4.5: Results of regression analysis for Category 1

The high adjusted R^2 value of 0.915016 and the relatively low standard error indicate that this model fits the data well. The F-Test which is used to determine the utility of the model i.e. whether the model is adequate for predicting y (Mendenhall and Sincich, 2003), has a p-value of close to zero indicating that the fitted model is useful for predicting fantasy league scores. It must be noted that all the individual β parameters are significantly different from zero (using t-test). The full output of this model is given in Appendix A, Table A.1.

Category 2

The regression model fitted to the data from Category 2 is given as

$$E\left(y_2\right) = 1.7985x_1 - 8.3076x_2 + 13.609x_3 \tag{4.4}$$

This model is fitted using 873 data points and the goodness-of-fit results are given in Table 3.6.

Measure	Value
Multiple R	0.9341
Adjusted R^2	0.8712
Standard Error	16.4394
p-value (F-Test)	0

Table 4.6: Results of regression analysis for Category 2

Once again, the high value for the adjusted R^2 and the low value of the standard error indicate that the model provides a good fit. Similarly to Category 1, the F-Test for the utility of the model has a p-value of close to zero and all the individual β parameters are significantly different from zero (using t-test). The full output of this model is given in Appendix A, Table A.2.

The tabulated results show that the fitted linear models provide a satisfactory method to estimate fantasy league scores.

4.2.2 Estimation of Fantasy League Scores

In order to gauge the overall ability of a cricketer, both their ListA and Twenty20 records are assessed. Given that Twenty20 is a new format of the game, the inclusion of ListA information is justified by the limited number of Twenty20 data. Combining the two formats (ListA and

Twenty20) also gives a more comprehensive account of every cricketers overall limited overs career, both domestic and international.

As a result of the inclusion of ListA matches, using the fitted linear regression model requires extrapolation. Mendenhall and Sincich (2003) define extrapolation as "*predicting outside the experimental region*," and this is required as the nature of a ListA game is different to that of a Twenty20 game. Since the linear model used only Twenty20 data, extrapolation is inevitable. The concerns regarding extrapolation (outlined in Mendenhall and Sincich (2003)) are not an issue in this research, as the use of the linear model in the case of ListA games is purely to give a measure of a cricketer's performance in a given game for comparative purposes. Since no data was collected for fantasy leagues based on ListA matches, this is the only indication available of a player's fantasy league score. Since no ListA fantasy league points allocation is available, it is argued that the use of the linear model is an acceptable method to estimate the fantasy league scores.

In order to estimate the scores accurately, each data point is categorised into one of the following game types: ListA International (ODI); ListA Domestic; Twenty20 International and Twenty20 Domestic. An estimated fantasy league score for each data point is then calculated using the appropriate linear model. This is done for each cricketer, for every game in which they played. The mean fantasy league scores obtained for each game type is summarised in Table 4.7.

Type	**Mean Score**
ListA International (ODI)	85.6853
ListA Domestic	92.4094
Twenty20 International	43.3033
Twenty20 Domestic	52.2199

Table 4.7: Mean predicted fantasy league scores for each game type

An interesting observation regarding the results in Table 4.7 is that the standard of the game is reflected. For example, for both Twenty20 and ListA formats the average obtained for international games is lower than that obtained for domestic games. This implies that the international game is of a higher standard than the domestic game and this concurs with expectation. Furthermore, the fact that ListA (both domestic and international) games have a higher average is logical given that the game is longer, thus allowing for more maidens to be bowled and more runs to be scored. An observation worth noting is that the number of runs scored in ListA matches are higher than that of Twenty20 matches, and bowling economy rates are traditionally lower in the longer form of the game. These factors result in higher fantasy league scores in ListA matches than in Twenty20 matches.

To investigate a cricketer's entire limited over career both ListA and Twenty20 data are combined into a single data set. The fitted linear model is then used to estimate fantasy league scores for every game in which the cricketer played. For the purpose of analysis, only the estimated fantasy league scores are considered for each cricketer. Table 4.7 indicates that there is a notable discrepancy between predicted fantasy league scores for each game type. Owing to this, analysis of the estimated fantasy league scores may provide misleading results, particularly for cricketers who have played in all formats of the game. To compare the estimated fantasy league scores on a similar scale, it is necessary to standardise the results. To do this each estimated fantasy league score is divided by its respective mean score (given in Table 4.7). In other words, suppose that Y_{ij} is defined as the predicted fantasy league score of player Y in their i^{th} game (where j indicates the game type). That is $j \in \{1, 2, 3, 4\}$ where 1 indicates a ListA international match; 2 a ListA domestic match; 3 a Twenty20 international match and 4 a Twenty20 domestic match. Suppose further that M_j indicates the mean fantasy league score from each game type (listed in Table 4.7). The standardised fantasy league score for player Y in game i, denoted Y_i^{st}, is thus given as

$$Y_i^{st} = \frac{Y_{ij}}{M_j}. \tag{4.5}$$

This adjusted fantasy league score of a player is thus an indication of their fantasy league score as a proportion of the mean score for the given game type. Using this method each cricketer's entire limited overs career (both Twenty20 and ListA) is combined into a single data set, without the need to distinguish between game types. These adjusted fantasy league scores are then used in the calculation of the fantasy performance measures defined in Section 3.2.5.

4.3 Evaluation of the Performance Measures

4.3.1 Results

Each of the performance measures listed in Section 3.4.1 are evaluated using the coefficient of correlation (ρ). The results of the assessment for each player category, as defined by the CricInfo IPL Fantasy League (2008), are provided and discussed next.

Performance measures are only calculated for cricketers who played in at least 3 matches prior to the commencement of the 2007 ICC World Twenty20 tournament. Given that the Twenty20 cricket is a new format of cricket a limited number of these games have been played by the cricketers. Ideally the minimum number of games played would be a higher value, such as 5 or 10. However using these values would significantly decrease the data set available for analysis. For this reason a minimum of 3 games is chosen. This minimum number of matches needed to be met in the game type for which the performance measure was calculated. As an example, for a player to have a Twenty20 performance measure (for example BP_{Tw20}), the cricketer would need to have played in at least 3 *Twenty20* matches. This requirement is also used in the case of ListA matches.

Fantasy type measures such as F_{mean} or F_{med}, where both Twenty20 and ListA matches are combined, required a cricketer to have participated in at least 3 matches when both disciplines are combined. A cricketer is thus required to have participated in at least 3 limited overs matches (regardless of the game format).

Batsmen

The estimated coefficient of correlation ($\hat{\rho}$) between the the performance measures calculated for batsmen and their total fantasy league score in the 2007 ICC World Twenty20 is given in Table 4.8.

Performance Measure	$\hat{\rho}$	n
F_{mean}	0.4145	45
F_{med}	0.3901	45
F_{MA_5}	0.3075	45
$F_{MA_{10}}$	0.3905	45
F_{EWA}	0.4431	45
BP_{LtA}	0.3480	44
BP_{Tw20}	0.1509	29
BAT_i^{LtA}	0.4840	45
BAT_i^{Tw20}	-0.0633	30
BK_{BAT}^{LtA}	0.5230	45
BK_{BAT}^{Tw20}	0.0258	30

Table 4.8: Batsmen: Estimated correlation between performance measures and fantasy league score

Given the results in Table 4.8 the exponentially weighted average of the estimated fantasy league score (F_{EWA}) is chosen as the most useful measure for the selection of batsmen for a fantasy league team. The reasons for this decision are that the F_{EWA} is calculated using a data set which combines both Twenty20 and ListA matches (in chronological order). This is desirable since it provides an indication of the current form using both formats of the game. Furthermore, F_{EWA} assigns a higher weight to the more recent matches played by the cricketer thus this method is good for determining a cricketer's current performance. For these reasons it is argued that it is appropriate to use F_{EWA} when investigating the changes that need to be made during the course of the competition. Despite a lower $\hat{\rho}$ estimate for F_{EWA} than for BK_{BAT}^{LtA} and BAT_i^{LtA} it is reasoned that F_{EWA} has less limitations than BK_{BAT}^{LtA} and BAT_i^{LtA}.

Although the Barr and Kantor batting measure for ListA games (BK_{BAT}^{LtA}) has the best estimated correlation ($\hat{\rho} = 0.5230$), this measure is obtained using only ListA data. During the Twenty20 tournament investigated in this study, the new data that becomes available cannot be included. Therefore the measure cannot be updated after each stage of the tournament. A distinct drawback to a fantasy league scenario which requires updating after each stage. Using the BK_{BAT}^{LtA} measure in this case will provide good results for once-off team selection, however when changes are necessary this measure is not ideal. To adapt the BK_{BAT}^{LtA} measure to include Twenty20 data a combined measure is used. This combined measure is defined as

$$BK_{BAT}^{com} = \left(BK_{BAT}^{Tw20}\right)^{\gamma_1} \left(BK_{BAT}^{LtA}\right)^{1-\gamma_1} \tag{4.6}$$

where $0 \leq \gamma_1 \leq 1$. *Solver* is then used to maximise the correlation between BK_{BAT}^{com} and the total fantasy league score, by changing the value of γ_1. The maximum correlation achieved using this method is $\hat{\rho} = 0.3217$, when $\gamma_1 = 0.0071$. The correlation using the BK_{BAT}^{com} measure is now no longer the highest and the exclusion of this measure is further justified.

The adapted Gerber and Sharp measure BAT_i^{LtA} has a correlation of $\hat{\rho} = 0.4840$. Although this value is high, the measure has the same drawback as the BK_{BAT}^{LtA} measure i.e. it only considers ListA data. In order to include Twenty20 data, the corresponding Twenty20 performance measure (BAT_i^{Tw20}) is used. For this measure $\hat{\rho} = -0.0633$, indicating that high values of BAT_i^{Tw20} correspond to low fantasy league scores which is undesirable. In order to combine the two measures, a method similar to that in Equation 4.6 is required. The combined measure BAT_i^{com} is defined as

$$BAT_i^{com} = \left(BAT_i^{Tw20}\right)^{\gamma_2} \left(BAT_i^{LtA}\right)^{1-\gamma_2} \tag{4.7}$$

where $0 \leq \gamma_2 \leq 1$. Using *Solver* to maximise the correlation between BAT_i^{com} and the fantasy league score results in $\hat{\rho} = 0.1607$ when $\gamma_2 = 0$[1]. It is clear too that this combined measure does not provide a high correlation. For these reasons this measure was also excluded from consideration.

The results of the optimisation routine used for the Barr and Kantor measures are discussed in Appendix D, Section D.1.

The full set of data for this section are given in Appendix B, Tables B.1 and B.2.

Bowlers

Table 4.9 provides the estimates of the correlation coefficients ($\hat{\rho}$) between the bowling performance measures of a cricketer and their total fantasy league score for the 2007 ICC World Twenty20 competition.

Performance Measure	$\hat{\rho}$	n
F_{mean}	0.1265	39
F_{med}	0.1180	39
F_{MA_5}	0.2586	39
$F_{MA_{10}}$	0.1515	39
F_{EWA}	0.1757	39
CBR_{LtA}	-0.1706	39
CBR_{Tw20}	0.0229	30
$FCBR_{LtA}$	-0.1291	39
$FCBR_{Tw20}$	-0.0095	30
BWL_i^{LtA}	0.1848	39
BWL_i^{Tw20}	0.0110	30
BK_{BWL}^{LtA}	-0.1856	39
BK_{BWL}^{Tw20}	-0.0103	29

Table 4.9: Bowlers: Estimated correlation between performance measures and fantasy league score

The values of the estimated correlation coefficient given in Table 4.9 are poor. The highest value of $\hat{\rho}$ is obtained for F_{MA_5} where $\hat{\rho} = 0.2586$. Since Twenty20 data is included in the calculation of this measure no further adjustments are necessary. This measure only considers the last 5 matches in which the cricketer played and thus gives an indication of the cricketer's current performance. This is a useful quality for a performance measure, especially for fantasy leagues where team changes need to be made regularly. For these reasons F_{MA_5} is chosen as the preferred measure when selecting bowlers for a fantasy league team.

[1]Since $\gamma_2 = 0$ one would expect that the estimated correlation for BAT_i^{com} would be equivalent to the estimated correlation for BAT_i^{LtA}. This is not the case since BAT_i^{com} is only calculated when both the corresponding ListA and Twenty20 measures are available. Table 4.8 indicates that 45 ListA data points are used. Only 30 Twenty20 corresponding data points are available. This indicates that the estimated correlation for BAT_i^{com} uses at most 30 data points. This explains the difference between the estimated correlations for BAT_i^{com} and BAT_i^{LtA}.

An interesting observation regarding the results shown in Table 4.9 is that measures using the estimated fantasy league scores tended to perform particularly well. Furthermore, the CBR outperformed the $FCBR$ in the ListA format of the game, but the $FCBR$ outperformed the CBR for the Twenty20 format. An argument for this observation is that maiden overs, which are included in the calculation of $FCBR$, have a greater effect on the Twenty20 format of the game than the ListA format.

The results of the optimisation routine used for the Barr and Kantor measures are discussed in Appendix D, Section D.2.

The full set of data for this section are given in Appendix B, Tables B.3 and B.4.

All Rounders

Table 4.10 lists the estimates of the correlation coefficient ($\hat{\rho}$) between the performance measures and the total fantasy league score for all rounders in the 2007 ICC World Twenty20 competition.

Performance Measure	$\hat{\rho}$	n
F_{mean}	0.4576	52
F_{med}	0.4837	52
F_{MA_5}	0.5566	52
$F_{MA_{10}}$	0.4799	52
F_{EWA}	0.4563	52
BP_{LtA}	0.3358	49
BP_{Tw20}	0.1091	32
CBR_{LtA}	-0.2158	51
CBR_{Tw20}	-0.0923	35
$FCBR_{LtA}$	-0.1076	51
$FCBR_{Tw20}$	-0.0937	35
ALR_i^{LtA}	0.3589	51
ALR_i^{Tw20}	0.2833	35
BK_{BAT}^{LtA}	0.4276	49
BK_{BAT}^{Tw20}	0.3383	36
BK_{BWL}^{LtA}	-0.2254	51
BK_{BWL}^{Tw20}	-0.2158	35

Table 4.10: All Rounders: Estimated correlation between performance measures and fantasy league score

The results show that the 5-point moving average of the estimated fantasy league score, denoted F_{MA_5}, with $\hat{\rho} = 0.5566$ has the largest estimated correlation. A benefit of this measure is that F_{MA_5} was also chosen as the preferred performance measure for the selection of bowlers. This gives and indication of the consistency of this performance measure. Therefore the 5-point moving average of the estimated fantasy league score is selected as the preferred measure when selecting all rounders for a fantasy league team.

There are some notable similarities between Tables 4.9 and 4.10. In both tables all the measures using the estimated fantasy league score (i.e. F_{mean}, F_{med}, F_{MA_5}, $F_{MA_{10}}$ and F_{EWA}) have the highest estimated correlations with the total fantasy league score. Furthermore in both Table 4.9 and 4.10, the $FCBR$ outperformed the CBR in the Twenty20 format of the game, while the CBR outperformed the $FCBR$ for the ListA format. This provides additional support that maiden overs have more of an effect on the Twenty20 format of the game than the ListA format.

The results of the optimisation of the Barr and Kantor measures are provided in Appendix D, Section D.3.

The full set of data for this section are given in Appendix B, Tables B.5 and B.6.

Wicket Keepers

The final category investigated is that of wicket keepers. Table 4.11 provides the estimated correlation coefficients of the performance measure to the total fantasy league score.

Performance Measure	$\hat{\rho}$	n
F_{mean}	0.6200	17
F_{med}	0.6365	17
F_{MA_5}	0.5212	17
$F_{MA_{10}}$	0.4061	17
F_{EWA}	0.5676	17
BP_{LtA}	0.4707	17
BP_{Tw20}	0.2726	9
WKR_i^{LtA}	0.4694	17
WKR_i^{Tw20}	-0.4536	12
BK_{BAT}^{LtA}	0.5782	17
BK_{BAT}^{Tw20}	0.2861	12

Table 4.11: Wicket Keepers: Estimated correlation between performance measures and fantasy league score

The measure chosen for the wicket keeper category is F_{EWA}. The correlation estimate for this measure is $\hat{\rho} = 0.567595$. This estimate is higher than that of the selected measures for batsmen, bowlers and all rounders. The selection of this measure is justified by the drawbacks of measures with higher $\hat{\rho}$ values.

The highest estimated correlation obtained is that corresponding to the median of the estimated fantasy league score, F_{med}. The value of $\hat{\rho} = 0.6365$ indicates that this is an effective measure. The drawback of using the median is that when new data are included into the calculation of the median, it is unlikely that its value will change considerably. This will definitely be the case if the initial data set is large, and the new data set is small (as is often the case when selecting a fantasy league team). The mean of the standardised fantasy league score, F_{mean}, has the same drawback, as a small set of new data will have little effect on the value of the mean. The F_{med} and F_{mean} are thus considered as good measures for a once-off indication of a players ability, but are not ideal when considering a progressive tournament (i.e. when new data are added).

The Barr and Kantor BK_{BAT}^{LtA} measure was excluded from consideration in the wicket keeper category. The motivation for this is similar to that mentioned in the batsmen category. This is discussed in Appendix D, Section D.4.

The full set of data for this section are given in Appendix B, Table B.7.

Additional Discussion of Results

It appears that the measures based solely on Twenty20 data tended to perform poorly in comparison to the corresponding ListA measures. A possible reason for this is that the Twenty20 format of the game is new and yet to establish itself (i.e. cricketers need to adapt to a new format of the game). In addition there is a limited amount of Twenty20 data to work with, however, with the growing popularity of the game, this problem should soon be overcome. The results indicated that cricketers who perform well in the ListA format of the game tend to perform well in the Twenty20 format of the game. This indicates that good cricketers will perform well in both formats of the game. The high correlations associated with the ListA measures arguably indicate that it is important to consider ListA data when selecting a Twenty20 team. This further justifies the inclusion of the ListA data into the calculation of the F_{mean}, F_{med}, F_{MA_5}, $F_{MA_{10}}$ and F_{EWA} measures.

The α and β values for the Barr and Kantor measures (BK_{BAT} and BK_{BWL} respectively) supported the belief that for batsmen the strike rate is more important than the batting average and for bowlers the economy rate is more important than the bowling average. These results are provided and discussed in Appendix D.

In conclusion the most consistent performance measures were those based on the estimated fantasy league scores. These measures (F_{mean}, F_{med}, F_{MA_5}, $F_{MA_{10}}$ and F_{EWA}) tended to outperform other measures. This implies that the methods in Section 3.2.5 are useful measures for determining the ability of a cricketer.

In summary the performance measures listed in Table 4.12 are chosen in order to select cricketers according to player category.

Player Category	**Performance Measure**
Batsman	F_{EWA}
Bowler	F_{MA_5}
All Rounder	F_{MA_5}
Wicket Keeper	F_{EWA}

Table 4.12: Selected performance measures for each player category

An interesting observation from Table 4.12 is that when bowling is involved in a player category (i.e. bowlers and all rounders) the 5-point moving average is chosen, whereas when bowling is not involved the exponentially weighted average is preferred.

4.3.2 Software Limitations

The BIP problem to be solved in this study includes a large number of integer decision variables. A problem of this magnitude requires a suitable software package. The software available for this study was *Lingo Super v10.0* (LINDO Systems Inc, 2006). This software is effective for solving integer programming problems, however the *Super* version can only solve for a maximum of 200 integer variables. Given this limitation the illustrative problem of the algorithm is simplified. The optimisation procedure is restricted to 50 cricketers for the IPL tournament which is subdivided into 4 stages.

The 50 cricketers were made up of 15 batsmen, 15 bowlers, 15 all rounders and 5 wicket keepers. The stage breakdown is shown in Table 4.13

	Begin Date	**End Date**	**Games**
Stage 1	18 April 2008	29 April 2008	16
Stage 2	30 April 2008	9 May 2008	14
Stage 3	10 May 2008	20 May 2008	14
Stage 4	21 May 2008	1 June 2008	15

Table 4.13: 2008 IPL fantasy league stage breakdown for integer program

The 50 cricketers included in this illustrative example are the top 15 batsmen, bowlers and all rounders and 5 wicket keepers ranked according to the performance measures in Table 4.12. Given the reduced size of this fantasy league, the number of changes allowed for a team was decreased from 20 to 9.

4.3.3 The Cricketers Available for Selection

In order to select the starting eleven cricketers, the 50 top ranked cricketers according to the performance measures shown in Table 4.12 are used. The results provided in the Tables 4.14 to 4.17 are calculated using only data available prior to the commencement of the 2008 IPL.

The 15 batsmen used for the fantasy league game are given in Table 4.14.

Rank	**Name**	**Fantasy Price**	F_{EWA}
1	D Lehmann	65 000	1.4339
2	D Hussey	90 000	1.3945
3	S Tiwary	50 000	1.3524
4	K Goel	50 000	1.1475
5	S Tendulkar	125 000	1.0974
6	G Smith	95 000	1.0469
7	M Tiwary	95 000	1.0454
8	Y Singh	120 000	1.0168
9	M Hayden	90 000	1.0097
10	R Taylor	75 000	0.9953
11	B Hodge	75 000	0.9407
12	S Marsh	55 000	0.9304
13	D Ravi Teja	50 000	0.9222
14	S Ganguly	120 000	0.9171
15	V Sehwag	105 000	0.9169

Table 4.14: Top 15 batsmen according to F_{EWA}

According to the ICC rankings, this list of batsmen includes some of the best batsmen in the world. These players include G Smith, V Sehwag and S Tendulkar. The inclusion of S Marsh, a relatively unheard of player, indicates the effectiveness of using the F_{EWA} measure, as he goes on to be the top run scorer of the 2008 IPL competition.

The 15 bowlers used for the fantasy league game are provided in Table 4.15.

Rank	**Name**	**Fantasy Price**	F_{MA_5}
1	RP Singh	115 000	1.8423
2	D Fernando	75 000	1.8370
3	D Zoysa	75 000	1.7704
4	D Vettori	100 000	1.6946
5	N Bracken	85 000	1.6829
6	S Tanvir	75 000	1.6127
7	M Gony	50 000	1.4669
8	L Malinga	90 000	1.4658
9	P Ojha	50 000	1.4271
10	A Noffke	75 000	1.4001
11	G McGrath	85 000	1.3803
12	B Lee	115 000	1.3701
13	M Kartik	85 000	1.3613
14	S Akhtar	85 000	1.3585
15	P Amarnath	50 000	1.3466

Table 4.15: Top 15 bowlers according to F_{MA_5}

Several highly rated bowlers are included in this list, for example B Lee, G McGrath and L Malinga. The inclusion of S Tanvir, a relatively unknown player, indicates the usefulness of using the 5-point moving average as, similar to S Marsh the unknown batsman, S Tanvir goes on to be the highest scoring bowler in the 2008 IPL fantasy league.

The 15 all rounders used for the fantasy league game are given in Table 4.16.

Rank	**Name**	**Fantasy Price**	F_{MA_5}
1	S Pollock	95 000	2.2641
2	P Kumar	75 000	1.9512
3	S Watson	75 000	1.8738
4	D Bravo	75 000	1.7681
5	S Afridi	110 000	1.6192
6	R Jadeja	50 000	1.5286
7	J Hopes	80 000	1.5195
8	M Hafeez	75 000	1.5063
9	D Mascarenhas	80 000	1.4371
10	S Malik	95 000	1.4348
11	J Sharma	75 000	1.4235
12	M Khote	50 000	1.3942
13	J Kallis	120 000	1.2978
14	D Thornely	55 000	1.2548
15	A Nayar	50 000	1.1904

Table 4.16: Top 15 all rounders according to F_{MA_5}

This list includes some highly rated all rounders in recent years, these are S Afridi, S Pollock and J Kallis. The usefulness of this measure is indicated by the inclusion of S Watson, who goes on to be the highest scoring all rounder for the 2008 IPL fantasy league. Although this measure is effective, one notable absence from this list is S Jayasuriya, who performed poorly in matches leading up to the IPL and is thus unavailable for the selection of the fantasy league team.

The 5 wicket keepers used for the fantasy league game are given in Table 4.17.

Rank	Name	Fantasy Price	F_{EWA}
1	B McCullum	105 000	1.0379
2	A Gilchrist	110 000	0.9832
3	Y Takawale	50 000	0.8858
4	AB de Villiers	85 000	0.8851
5	MS Dhoni	150 000	0.8361

Table 4.17: Top 5 wicket keepers according to F_{EWA}

This list includes some highly rated wicket keepers such as A Gilchrist and MS Dhoni. A notable observation about the information in Table 4.18 is that B McCullum is the highest ranked wicket keeper at the beginning of the 2008 IPL. He then goes on to score a world record of 158 runs in his first game. This lends support to the use of F_{EWA} as a fantasy league team selection measure, since it emphasises the current performance of a cricketer.

The cricketers listed in the Tables 4.14 to 4.17 are used as a population from which the fantasy league team will be selected and modified (using a BIP technique) during the course of the 2008 IPL competition.

4.4 The Integer Programming Problem

The optimisation problem is solved using *Lingo v 10.0 Super*, the results of which are given in the following sections. The full set of data used for the optimisation problem is provided in Appendix C, Tables C.1 and C.2.

4.4.1 Results

Retrospective Approach (Z)

The retrospective approach involves solving a BIP problem including Constraints 1 to 7 with the objective function $Z = \underline{S}'\underline{X} = s_1x_1 + s_2x_2 + \cdots + s_{nq}x_{nq}$.

The solution to this problem (as determined by the software) is shown in Table 4.18. The results and team selections represent the best solution to the 2008 IPL fantasy league for the population of players. This strategy is determined when the tournament is completed and thus provides a target for the prospective approach. The selections and changes shown in the table represent the ideal approach for this fantasy league. The reason is that the solution assumes that every player's fantasy league points are known beforehand. In a practical fantasy league scenario this is not the case. It is extremely difficult to predict how well a player will perform in an upcoming game. Durbach and Thiart (2007) argue that there is little empirical evidence to support the perception of a batsman having periods of good and bad form. They claim it is near impossible to predict a players future performance based on recent form.

The results in Table 4.18 are thus a target for the prospective methods. The total number of points scored in this ideal scenario is 11 494. Given the unpredictability of cricket performances (Durbach and Thiart, 2007), this value should never be reached. However the aim of this dissertation is to provide a mathematical method which will optimise the selection process and ideally achieve the best result without knowing the outcome.

Stage 1			Stage 2			Stage 3			Stage 4		
Name	*Price*	*Fan Pts*	*Name*	*Price*	*Fan Pts*	*Name*	*Price*	*Fan Pts*	*Name*	*Price*	*Fan Pts*
Batsmen			**Batsmen**			**Batsmen**			**Batsmen**		
M Hayden	90 000	331	S Marsh	55 000	324	S Marsh	55 000	261	S Marsh	55 000	678
Y Singh	120 000	284	S Ganguly	120 000	210	S Ganguly	120 000	391	S Ganguly	120 000	148
V Sehwag	105 000	254	V Sehwag	105 000	364	V Sehwag	105 000	222	Y Singh	120 000	288
R Taylor	75 000	226	D Hussey	90 000	267	G Smith	95 000	217	G Smith	95 000	184
Bowlers			**Bowlers**			**Bowlers**			**Bowlers**		
G McGrath	85 000	165	G McGrath	85 000	235	G McGrath	85 000	190	D Fernando	75000	322
M Gony	50 000	230	M Gony	50 000	115	M Gony	50 000	40	M Gony	50 000	326
B Lee	115 000	214	S Tanvir	75 000	389	S Tanvir	75 000	215	S Tanvir	75 000	374
All Rounders			**All Rounders**			**All Rounders**			**All Rounders**		
S Watson	75 000	466	S Watson	75 000	275	S Watson	75 000	341	S Watson	75 000	373
S Pollock	95 000	223	S Pollock	95 000	259	S Pollock	95 000	270	J Hopes	80 000	163
D Bravo	75 000	241	D Bravo	75 000	73	D Bravo	75 000	379	D Bravo	75 000	0
Wicket Keeper			**Wicket Keeper**			**Wicket Keeper**			**Wicket Keeper**		
A Gilchrist	110 000	384	A Gilchrist	110 000	237	M Dhoni	150 000	173	M Dhoni	150 000	173
TOTAL	995 000	3 018	**TOTAL**	935 000	2 748	**TOTAL**	980 000	2 699	**TOTAL**	970 000	3 029
Changes		-	*Changes*		4	*Changes*		2	*Changes*		3

Table 4.18: Solution of integer programming problem: The retrospective approach (Z)

Prospective Approach using $\underline{M}$

This approach assumes that the fantasy league participant will have no knowledge of whether a cricketer is available to play or not. In other words the fantasy league participant will not follow the competition closely, but rather only use the values of the performance measures at each stage of the competition (to facilitate team selection). This approach is not an ideal representation of a fantasy league, as most participants will follow the news concerning the league closely, thus knowing whether a cricketer is injured or unavailable. By knowing this the participant can make the appropriate changes to their team. The results using this approach thus represent the usefulness of using only the performance measure as a selection criterion, without knowledge of a cricketers availability.

The sequential optimisation procedure described in Section 3.4.3 and using data from the performance vector $\underline{M}$ $(nq \times 1)$ is conducted in order to provide an entirely prospective approach to the fantasy league team selection problem. The results of this sequential optimisation procedure are provided in Table 4.19. (The solution with performance measures included is given in Table C.2.) This prospective approach resulted in a total fantasy league score of 5 479, which is poor result when compared to the optimal solution of 11 494.

Procedure	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Total**
Retrospective	3 018	2 748	2 699	3 029	*11 494*
Prospective ($\underline{M}$)	1 704	691	1 279	1 805	*5 479*

Table 4.20: Comparison of fantasy league points per stage: Optimal Solution and Prospective approach using $\underline{M}$

From Table 4.20 it can clearly be seen that this prospective approach performs poorly throughout the competition. The points scored in Stage 2 are particularly poor. However the points appear to increase steadily in the following two stages. A possible explanation for this is that optimisation procedure is reacting to the data collected from the competition. This results in cricketers with high values of the performance measures (i.e. cricketers who are performing well) being included in the final stages. This is evidenced by the inclusion of S Tanvir and S Marsh in the third and fourth stages respectively. S Tanvir and S Marsh won the awards for the best bowler and best batsman of the competition respectively, and their inclusion in a fantasy league team is thus ideal.

The most concerning aspect of this team selection process is the amount of zero scores found in Table 4.19. Consider in particular B McCullum, D Lehmann and N Bracken. All these players were selected for Stages 2, 3 and 4 and in every stage scored 0 fantasy league points. Since a fantasy league score of zero is unlikely, the most rational explanation for these scores is that the players were unavailable and/or injured. Since this information is often available (on websites / in the news) it would be incomplete to consider only this solution to the problem. Adjustments according to who is available need to be made in order to provide the optimisation procedure with all the necessary information. This is done in the following section.

An important aspect of this team selection method to notice is its potential. Firstly, the selection of B McCullum for the first stage is highly convenient as he scored an unbeaten 158 in his first game, traditionally selectors would probably have chosen A Gilchrist or M Dhoni. This

Stage 1			Stage 2			Stage 3			Stage 4		
Name	*Price*	*Fan. Pts*	*Name*	*Price*	*Fan. Pts*	*Name*	*Price*	*Fan. Pts*	*Name*	*Price*	*Fan. Pts*
Batsmen			**Batsmen**			**Batsmen**			**Batsmen**		
D Hussey	90 000	148	D Hussey	90 000	267	D Hussey	90 000	181	D Hussey	90 000	2
D Lehmann	65 000	18	D Lehmann	65 000	0	D Lehmann	65 000	0	D Lehmann	65 000	0
S Tiwary	50 000	19	S Tiwary	50 000	24	S Tiwary	50 000	0	M Hayden	90 000	0
K Goel	50 000	67	K Goel	50 000	21	V Sehwag	105 000	222	S Marsh	55 000	678
Bowlers			**Bowlers**			**Bowlers**			**Bowlers**		
D Zoysa	75 000	25	D Zoysa	75 000	41	S Tanvir	75 000	215	S Tanvir	75 000	374
RP Singh	115 000	187	N Bracken	85 000	0	N Bracken	85 000	0	N Bracken	85 000	0
D Fernando	75 000	0	D Fernando	75 000	0	D Fernando	75 000	50	D Fernando	75 000	322
			D Vettori	100 000	45	D Vettori	100 000	0			
All Rounders			**All Rounders**			**All Rounders**			**All Rounders**		
S Watson	75 000	466	S Watson	75 000	275	S Watson	75 000	341	S Watson	75 000	373
S Pollock	95 000	223	J Hopes	80 000	18	S Pollock	95 000	270	S Pollock	95 000	56
P Kumar	75 000	128							D Mascarenhas	80 000	0
Wicket Keeper			**Wicket Keeper**			**Wicket Keeper**			**Wicket Keeper**		
B McCullum	105 000	423	B McCullum	105 000	0	B McCullum	105 000	0	B McCullum	105 000	0
TOTAL	870 000	1 704	**TOTAL**	850 000	691	**TOTAL**	920 000	1 279	**TOTAL**	890 000	1 805
Changes		-	*Changes*		3	*Changes*		3	*Changes*		3

Table 4.19: Solution of integer programming problem: The prospective approach using $\underline{M}$

shows that the method has potential to select the right players at the right time. Furthermore, the selection of S Watson from the beginning of the tournament is another indication of the proposed method's potential. S Watson, at the time of the 2008 IPL, had been dropped from the Australian cricket team and many all rounders such as J Kallis or S Afridi would traditionally be chosen ahead of him. This prospective method selects him for its initial team and keeps him throughout the competition. At the end of the tournament S Watson scores the most fantasy league points in the entire competition.

In summary, this prospective approach selects a fantasy league team which does not perform particularly well. However, upon inspection, there are aspects of selection technique which highlight the potential of the approach.

Prospective Approach using $\underline{R}$

The adjusted performance measure vector $\underline{R}$ $(nq \times 1)$ contains information as to whether or not a cricketer played in a given stage of the competition. It is assumed that all this data is available to the fantasy league participant prior to the commencement of each stage.

This however is not an entirely accurate representation of the knowledge available to fantasy league participants, as there are certain players who aren't injured and who are available to play, but are not selected by their coach. Although most of the information regarding a player's availability and health will be available to the fantasy league participants, the unknown information is whether or not a cricketer will play in a certain game. As an example of the type of information that is available to fantasy league participants, consider the fact that the Australian and New Zealand cricket teams each had a test match series scheduled at the same time as the 2008 IPL. It was common knowledge that these cricketers would not be available for most of the competition. Furthermore, active fantasy league participants would also know when a cricketer is injured as this information is available on the fantasy league website or in the news. So, although not all the information on whether a cricketer will play or not is available beforehand, it assumed that most of it is. Thus the results obtained will represent a close approximation to the actual results (when using this method in future).

Table 4.21 gives the team selection results of the prospective approach when using the adjusted performance measure data found in the $\underline{R}$ $(nq \times 1)$ vector. (The solution with performance measures included is given in Table C.2.) The total fantasy league score obtained by this optimisation strategy is 7 510. The points scored using this approach is compared to the optimal approach (retrospective) in Table 4.22.

Procedure	**Stage 1**	**Stage 2**	**Stage 3**	**Stage 4**	**Total**
Retrospective	3 018	2 748	2 699	3 029	*11 494*
Prospective ($\underline{R}$)	1 704	1 778	1 750	2 278	*7 510*

Table 4.22: Comparison of fantasy league points per stage: Optimal Solution and Prospective approach using $\underline{R}$

Table 4.22 shows that using this prospective approach provides very consistent results. Furthermore, the high number of points scored in the last stage emphasises the ability of this routine

Stage 1			**Stage 2**			**Stage 3**			**Stage 4**		
Name	*Price*	*Fan. Pts*	*Name*	*Price*	*Fan. Pts*	*Name*	*Price*	*Fan. Pts*	*Name*	*Price*	*Fan. Pts*
Batsmen			**Batsmen**			**Batsmen**			**Batsmen**		
D Hussey	90 000	148	D Hussey	90 000	267	D Hussey	90 000	181	D Hussey	90 000	2
D Lehmann	65 000	18	G Smith	95 000	97	G Smith	95 000	217	G Smith	95 000	184
S Tiwary	50 000	19	S Tiwary	50 000	24	V Sehwag	105 000	222	V Sehwag	105 000	68
K Goel	50 000	67	K Goel	50 000	21	K Goel	50 000	16	S Marsh	55 000	678
Bowlers			**Bowlers**			**Bowlers**			**Bowlers**		
D Zoysa	75 000	25	D Zoysa	75 000	41	S Tanvir	75 000	215	S Tanvir	75 000	374
RP Singh	115 000	187	RP Singh	115 000	209	RP Singh	115 000	125	RP Singh	115 000	28
D Fernando	75 000	0	D Vettori	100 000	45	D Fernando	75 000	50	D Fernando	75 000	322
All Rounders			**All Rounders**			**All Rounders**			**All Rounders**		
S Watson	75 000	466	S Watson	75 000	275	S Watson	75 000	341	S Watson	75 000	373
S Pollock	95 000	223	S Pollock	95 000	259	S Pollock	95 000	270	S Pollock	95 000	56
P Kumar	75 000	128	P Kumar	75 000	303	P Kumar	75 000	2	S Afridi	110 000	26
Wicket Keeper			**Wicket Keeper**			**Wicket Keeper**			**Wicket Keeper**		
B McCullum	105 000	423	A Gilchrist	110 000	237	A Gilchrist	110 000	111	A Gilchrist	110 000	167
TOTAL	870 000	1 704	**TOTAL**	930 000	1 778	**TOTAL**	960 000	1 750	**TOTAL**	1 000 000	2 278
Changes		-	*Changes*		3	*Changes*		3	*Changes*		2

Table 4.21: Solution of integer programming problem using the second prospective approach

to select cricketers who are performing well in the competition. When compared to the retrospective approach this approach provides satisfactory results.

The main drawbacks of the initial prospective approach are obvious from Table 4.21. The inclusion of cricketers who are not available is no longer a concern. Clearly the results using this prospective approach are more favourable than using the initial prospective approach. This indicates that including information regarding a cricketers availability can notably improve the performance of the optimisation routine.

Table 4.21 provides some interesting observations. The prospective approach includes B McCullum in the first stage of the competition and then used A Gilchrist for the remainder of the competition. This shows that this approach is able to select the cricketers who have a good chance of performing well and to make adjustments when they are not available. For example, despite B McCullum scoring 423 points he is replaced by A Gilchrist as he is unavailable from Stage 2. The inclusion of S Watson provides strong evidence that the method is useful, and his inclusion is common to all selection strategies. The inclusion of S Tanvir and S Marsh, as discussed previously, also indicates the effectiveness of this approach. The inclusion of D Fernando in Stage 1 of the prospective approach is not an error despite his fantasy league score of zero. In fact, this illustrates the unpredictability of cricketers performance, as he had a high value for his performance measure (F_{MA_5}) going into the tournament and only played a single game in which he scored no points. This provides strong evidence that the optimal strategy given in Table 4.18 is near impossible to achieve using only historic data.

This prospective strategy involves inputting data into a simple integer optimisation problem, as the data are available, and then running it. No intuition is required, and this is a purely mathematical approach. For this prospective approach to score 7 510 points using only historic data, shows that it is a very useful tool when selecting a fantasy league cricket team.

When all three selection strategies are considered the following observations are made.

- The strategies which performed particularly well selected teams consisting of 3 bowlers and 3 all rounders. This could provide information to future fantasy league team selectors that the number of all rounders included should be high. This can be considered as a constraint of the integer programming problem.

- The pricing of players in this fantasy league did not provide much of a restriction to the team selection. This is noticed since the total amount spent in each stage rarely approaches the 1 000 000 unit cut-off. In order to make the league more competitive the hosting website might consider lowering this cut off rate. This change again can be addressed by amending the constraints of the integer program. The integer program is designed to select the best team (according to the performance measure used) under the given constraints. Although in this example the price constraint is not of much concern, for many other fantasy leagues it provides a much greater restriction. Thus the use of an integer program is a necessity.

Chapter 5

Conclusion and Further Work

The aim of this study was to provide a fantasy league participant with a strategy to facilitate the team selection process for fantasy league cricket. To do this a number of published and new performance measures were evaluated. The results of this evaluation indicated that the best measures used the estimated fantasy league scores of each cricketer, calculated prior to the beginning of each stage. In addition it was found that the measures which allocated more weight to recent player performance provided better results.

Using these findings a strategy was developed to facilitate the team selection process for a fantasy league based on the 2008 IPL tournament. This strategy involved using the performance measures to select a fantasy league team at each stage of the competition. Given the software limitations the fantasy league was restricted to 50 cricketers over 4 stages. The results of the selection process using a prospective approach provided satisfactory results when compared to the ideal scenario. As such the method described in this dissertation can be used by any fantasy league participant for the purposes of adequate fantasy league team selection.

Although this research provided some useful results, it must be noted that it is a pioneering study into fantasy league team selection for Twenty20 cricket, and as such requires further research. Given the growing popularity of both fantasy sport and Twenty20 cricket, this area of research is both new and exciting. Furthermore, with the introduction of more Twenty20 and fantasy league data, the estimation of fantasy league scores using a linear model becomes less necessary.

After an extensive study into the subject matter some recommendations for further work are noticed. These recommendations are

- A software package with the ability to solve the entire integer problem is required. This may provide more detailed insight into the "ins and outs" of fantasy league cricket.

- It was found during the course of this research that the estimated and standardised fantasy league scores of each player could be modelled well using a gamma distribution. The use of this knowledge could provide a better method to rank cricketers, as well as to provide a starting point for accurate simulations of fantasy leagues.

- Investigating the fixture list in order to select cricketers from stronger teams (when they are playing weaker teams) is a useful tactic in fantasy leagues. The inclusion of this information into the integer program could provide better results.

In conclusion, this study proposes a mathematical routine for the selection of a Twenty20 cricket team. The team is selected for a fantasy game using a prospective approach. This approach is such that it could be used, with satisfactory results, by any fantasy league participant to facilitate team selection. However, in order to test the overall accuracy of the proposed methodology, it should be tested within an externally controlled fantasy league scenario. This study provides a useful platform from which further research into this subject matter can be launched.

References

ALL OUT FOR NO LOSS (2007). History of Twenty20 cricket. 26 September 2007. [http://www.cricinus.com/2007/09/twenty20cricket-history.html]. Accessed: 27 February 2009

AMATO, C. (2009). Back to the drawing board: Bafana shake off glumness thanks to return of Parreira and McCarthy but extend goal drought. *Sunday Times*. November 15, 2009.

BARR, G.D.I. and KANTOR, B.S. (2004). A criterion for comparing and selecting batsmen in limited overs cricket. *Journal of the Operational Research Society*, **55** (12), 1266 – 1274.

BARR, G.D.I. and VAN DEN HONERT, R (1998). Evaluating batsmen's scores in test cricket. *South African Statistical Journal*, **32** (2), 169 – 183.

BEAUDOIN, D. and SWARTZ, T.B. (2003). The best batsmen and bowlers in one day cricket. *South African Statistical Journal*, **3**7 (2), 203 – 222.

BRITANNICA ONLINE ENCYCLOPEDIA: CRICKET - METHODS OF DISMISSAL (2009). [http://www.britannica.com/EBchecked/topic/142911/cricket/214754/Methods-of-dismissal]. Accessed: 8 December 2009.

CLARKE, S.R. (1988). Dynamic programming in one-day cricket – optimal scoring rates. *The Journal of the Operational Research Society*, **39** (4), 331 – 337.

CLARKE, S.R. (1998). Test Statistics. Statistics in Sport (Bennet, J., Ed.), 83 – 103, Arnold, London.

CLARKE, S.R. and NORMAN, J.M. (1999). To run or not?: Some dynamic programming models in cricket. *The Journal of the Operational Research Society*, **50** (5), 536 – 545.

CRICINFO: A BRIEF HISTORY OF CRICKET (2009). [http://www.cricinfo.com/ci/content/story/239757.html]. Accessed: 19 November 2009.

CRICINFO IPL FANTASY LEAGUE (2008). [http://fantasy.cricinfo.com/fantasy/fantasyleague/user/game_home.html?game_id=22]. Accessed: 21 April 2008.

CRICKET ARCHIVE (2009). List of ICC Member Countries. [http://www.cricketarchive.com/Archive/Countries/index.html]. Accessed: 11 December 2009

DE SILVA, B.M.; POND, G.R. and SWARTZ, T.B. (2001). Estimation of the magnitude of victory in one-day cricket. *The Australian and New Zealand Journal Of Statistics*, **43** (3), 259 – 268.

DUCKWORTH, F.C. and LEWIS, A.J. (1998). A fair method for resetting the target in interrupted one-day international cricket matches. *The Journal of the Operational Research Society*, **49** (7), 220 – 227.

DURBACH, I.N. and THIART, J. (2007). On a common perception of a random sequence in cricket. *South African Statistical Journal*, **41** (2), 161 – 187.

ELAINE ALLEN, I.; KUSTOV, K. and RECCK, G. (2007). Building a better fantasy baseball team. *Quality Progress*, (April), 24 – 29.

FTSA (2009). [http://www.fsta.org]. Accessed: 19 November 2009.

GERBER, H. and SHARP, G.D. (2006). Selecting a limited overs cricket squad using an integer programming model. *South African Journal for Research in Sport, Physical Education and Recreation*, **28** (2), 81 – 90.

ICC RULES AND REGULATIONS (2008). ICC Classification of Official Cricket, 2008. [http://l.yimg.com/t/icccricket/pdfs/classification-of-official-cricket-july-2008.pdf]. Accessed: 8 December 2009.

LEACH, J. (2007). From lads to lords: The history of cricket: 1300 – 1787. [http://www.jl.sl.btinternet.co.uk/stampsite/cricket/ladstolords/1300.html]. Accessed: 29 September 2009.

LEMMER, H.H. and NEL, J. (2001). A classification scheme for the consistency of batsmen. *South African Statistical Journal*, **35** (1), 147 – 173.

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.

LOURENS, M. (2009). Integer optimisation for the selection of a Twenty20 cricket team. Master's Dissertation, Nelson Mandela Metropolitan University. Port Elizabeth, South Africa.

MENDENHALL, W. and SINCICH, T. (2003). A Second Course In Statistics: Regression Analysis (6th Edition). Prentice Hall (Pearson Education): Upper Saddle River, New Jersey.

NORRIDGE, J (2008). Can we have our balls back, please? Penguin Books Ltd. London, England.

OFFICIAL EPL FANTASY LEAGUE (2009). [http://fantasy.premierleague.com]. Accessed: 19 November 2009.

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

SHARP, G.D.; BRETTENNY, W.J.; LOURENS, M; GONSALVES, J.W. and STRETCH, R.A. (2009). Integer optimisation for the selection of a Twenty20 cricket team. Submitted to the Journal of the Operational Research Society for review.

SHILBURY, D; WESTERBEEK, H; QUICK, S and FUNK, D. (2009). Strategic Sport Marketing (3rd Edition). Allen and Unwin: Crows Nest, Australia.

SNYMAN, J.A. (2005). Practical Mathematical Optimisation: an introduction to basic optimization theory and classical and new gradient-based algorithms. Springer Science+Business Media, USA.

SUMMERS, A.E.; SWARTZ, T.B. and LOCKHART, R.A. (2007). Optimal drafting in hockey pools. Statistical Thinking in Sports. Editors: J. Albert and R.H. Koning. Chapman & Hall/CRC (Taylor and Francis Group): Boca Raton, Florida. 2008.

SWARTZ, T.B.; GILL, P.S. and MUTHUKUMARANA, S. (2009). Modelling and simulation in one-day cricket. *The Canadian Journal Of Statistics*, **37** (2), 143 – 160.

SWARTZ, T.B.; GILL, P.S.; BEAUDOIN, D. and DE SILVA, B.M. (2006). Optimal batting orders in one-day cricket. *Computers & Operations Research*, **33** (7), 1939 – 1950

TAHA, H.A. (2003). Operations Research: An Introduction (7th Edition). Prentice Hall (Pearson Education): Upper Saddle River, New Jersey.

VICHOT, R. (2009). A history of fantasy sports and its adoption by sports journalists. [http://jag.lcc.gatech.edu/blog/2009/01/history-of-fantasy-sports-and-its-adoption-by-sports-journalists.html].
Accessed: 19 November 2009

VAN STADEN, P.J. (2009). Comparison of cricketers' bowling and batting performances using graphical displays. Scientific Correspondence. *Current Science*, **96** (6), 764 – 766.

WATSON, N.M. (2008). A descriptive analysis of the 2007 inaugural ICC Twenty20 cricket world cup in South Africa. Honours Treatise, Nelson Mandela Metropolitan University. Port Elizabeth, South Africa.

WEAVER, P. (2008). The Twenty20 revolution that turned innocent sideshow into a hunt for millions. The Guardian, 26 July 2008. [http://www.guardian.co.uk/sport/2008/jul/26/twenty20domestic.cricket1]. Accessed: 27 February 2009.

Data Collection

ESPN CRICINFO. [www.cricinfo.com]. Accessed between 17 April 2008 and 7 September 2008.

CRICKET ARCHIVE: PLAYER GURU.
[www.cricketarchive.com/cgi-bin/ask_the_player_oracle.cgi]. Accessed between 6 September 2008 and 28 January 2009.

Software Used

Microsoft Excel 2003 & 2007, Microsoft Corporation (1985 - 2006).
Lingo v 10.0 Super, LINDO Systems Inc. (2006)
R version 2.8.1 (22-12-2008), The R Foundation for Statistical Computing (2008).

Appendix A

Regression Analysis Results

The results in Table A.1 and Table A.2 provide the output of the regression analysis discussed in Section 3.2.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.957364033
R Square	0.916545892
Adjusted R Square	0.915016101
Standard Error	15.47036266
Observations	988

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	7	2578555.546	368365.1	1539.138	0
Residual	981	234784.8104	239.3321		
Total	988	2813340.356			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
Duck	-10.25640279	2.13620211	-4.80123	1.82E-06	-14.44845352	-6.064352056	-14.44845352	-6.064352056
Catches	14.47172392	0.967028824	14.96514	9.27E-46	12.57404117	16.36940667	12.57404117	16.36940667
Runs Scored	1.919533697	0.034141606	56.22271	0	1.852534726	1.986532668	1.852534726	1.986532668
Overs	9.893983468	0.554961146	17.82825	8.39E-62	8.804936102	10.98303083	8.804936102	10.98303083
Maidens	36.93031877	2.828456176	13.05671	4.85E-36	31.37979912	42.48083842	31.37979912	42.48083842
Runs Conc	-1.485840842	0.057723502	-25.7407	4.7E-112	-1.59911657	-1.372565114	-1.59911657	-1.372565114
Wickets	35.24581118	0.551886301	63.86426	0	34.16279785	36.32882452	34.16279785	36.32882452

Table A.1: Excel output for regression analysis of Category 1 data

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.934134
R Square	0.872607
Adjusted R Square	0.871165
Standard Error	16.43944
Observations	873

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	1610518.628	536839.5	1986.417	0
Residual	870	235122.0322	270.2552		
Total	873	1845640.66			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
Duck	-8.30759	2.221702798	-3.73929	0.000197	-12.66811695	-3.947070938	-12.66811695	-3.947070938
Catches	13.60901	0.74284777	18.32005	1.18E-63	12.15102749	15.06699338	12.15102749	15.06699338
Bt Runs	1.798472	0.020128737	89.34847	0	1.75896531	1.837978418	1.75896531	1.837978418

Table A.2: Excel output for regression analysis of Category 2 data

Appendix B

Performance Measure Data

The results in Tables B.1 to B.7 provide the data used to assess the performance measures mentioned Section 3.3.2. The correlations provided are discussed in Section 4.3.1.

The data in Tables B.8 to B.14 provide the data used to optimise the weighting of the Barr and Kantor Measures discussed in Section 3.3.1. The results are discussed in Appendix D.

Should a cell values be blank or contain “NaN”, “NA”, “Inf” or “-”; this indicates that the data were unavailable for consideration or that the value was not computable.

		T20 Games	ListA Games	WC Fantasy Score	Fantasy MA 10	Fantasy MA 5	Fantasy EWA	Fantasy Mean	Fantasy Median	Lemmer Batting ListA	Lemmer Batting T20	Gerber Sharp Batting T20	Gerber Sharp Batting ListA	Barr and Kantor Batting ListA	Barr and Kantor Batting T20	Gerber Sharp Combin	Barr and Kantor Combin
	TOTAL Correl	-	-	-	*0.390512*	*0.307474*	*0.443096*	*0.414504*	*0.390117*	*0.347998*	*0.1509*	*-0.06334*	*0.484034*	*0.522975*	*0.025829*	*0.160665*	*0.321659*
1	Ashraful	2	134	200	0.825808	1.317836	0.628025	0.509938	0.280095	18.8928	-	-	0.618307	44.07839	-	-	-
2	Ahmed	3	83	243	0.584524	0.615861	0.656274	0.703292	0.57186	27.77967	15.3502	0.665898	0.873875	53.0171	121.14	0.873875	53.32949
3	Siddique	0	15	147	0.577864	1.019377	0.507783	0.455221	0.270535	40.78441	-	-	0.767549	48.30973	-	-	-
4	Chowdhury	1	23	17	0.511217	0.3333	0.790656	0.902872	0.633053	11.43593	-	-	0.520934	40.80591	-	-	-
5	Nazimuddin	2	36	-8	0.918807	1.165408	0.547886	0.366952	0.173339	34.07611	-	-	0.778402	49.79949	-	-	-
6	Iqbal	2	25	81	0.142937	0.149393	0.151618	0.142389	-0.00782	24.82219	-	-	0.76905	46.0322	-	-	-
7	Pietersen	16	167	355	0.533813	0.728246	0.784055	0.902163	0.735856	49.16608	19.85453	0.898158	1.313238	62.69402	115.7857	1.313238	62.96806
8	Shah	26	245	135	0.691121	0.763258	0.860401	0.659975	0.443678	38.61114	31.38383	1.246299	1.04762	53.45768	117.6335	1.04762	53.75828
9	Solanki	20	318	109	0.642937	0.688242	0.941355	0.645655	0.422035	40.07919	51.69715	1.714226	0.999071	50.79703	116.8949	0.999071	51.09892
10	Mishra	2	32	8	0.368612	0.599484	0.480116	0.485086	0.389414	22.16787	-	-	0.944146	49.59696	-	-	-
11	Obanda	1	3	15	0.29003	0.29003	0.289048	0.29003	0.103918	-	-	-	0.596371	39.83563	-	-	-
12	Fulton	5	87	117	0.79747	0.917677	0.578605	0.525606	0.383154	28.42607	4.815597	0.662281	1.077891	53.08573	88.16767	1.077891	53.27756
13	McMillan	14	323	364	0.688566	0.655888	0.803743	0.799809	0.584356	25.62245	35.8841	1.097163	0.953684	52.20698	117.0208	0.953684	52.50744
14	Vincent	14	181	206	0.891336	0.550853	0.887637	0.656315	0.465321	32.9331	20.90817	1.046003	1.119239	49.86515	101.0033	1.119239	50.11601
15	Nazir	18	133	292	0.812424	1.370428	0.796021	0.683575	0.478745	29.02655	31.7409	1.093804	1.055278	54.72043	128.6733	1.055278	55.05409
16	McCallum	0	32	-10	0.267327	0.440311	0.2647	0.25249	0.052823	12.13607	-	-	0.753168	43.20176	-	-	-
17	Poonia	3	24	4	0.360622	0.282429	0.339463	0.375921	0.270535	8.38553	21.63108	0.206566	0.635442	42.68169	108.9628	0.635442	42.96705
18	Sheikh	0	5	-	0.097393	0.097393	0.099031	0.097393	0.054107	3.4449	-	-	0.200019	19.75099	-	-	-
19	Watts	0	86	86	0.252842	0.172385	0.411174	0.403022	0.248893	12.5175	-	-	0.653298	39.71225	-	-	-
20	Bodi	27	77	-	1.066187	1.343458	0.844601	0.769233	0.584356	22.35227	-	0.589934	0.885589	45.23009	-	0.885589	-
21	Duminy	25	53	83	0.422009	0.30577	0.686993	0.703849	0.472785	12.03165	11.14931	0.898334	0.848038	49.13838	91.7489	0.848038	49.35701
22	Mubarak	8	109	207	1.178445	0.843749	0.914619	0.824547	0.595201	14.82137	41.06905	1.835626	0.962463	46.59207	109.5087	0.962463	46.87601
23	Tharanga	3	89	88	0.537215	0.588566	0.516033	0.532271	0.343678	22.88134	6.004469	0.504713	1.064507	51.43465	81.11612	1.064507	51.60151
24	Deonarine	4	51	-	0.926313	1.216327	0.973757	0.805527	0.562714	11.80267	-	0.925911	0.69696	35.72657	-	0.69696	-
25	Morton	4	94	-	0.482937	0.387464	0.605473	0.680567	0.476894	27.17653	-	0.276095	1.187184	50.1805	-	1.187184	-

Table B.1: Data used for the assessment of the batsmen performance measures (1 of 2)

		T20 Games	ListA Games	WC Fantasy Score	Fantasy MA 10	Fantasy MA 5	Fantasy EWA	Fantasy Mean	Fantasy Median	Lemmer Batting ListA	Lemmer Batting T20	Gerber Sharp Batting T20	Gerber Sharp Batting ListA	Barr and Kantor Batting ListA	Barr and Kantor Batting T20	Gerber Sharp Combin	Barr and Kantor Combin
26	Samuels	5	149	59	0.894342	0.455154	1.017254	1.134735	0.978485	21.35131	37.69787	1.390862	0.921151	51.37504	150.4532	0.921151	51.76902
27	DS_Smith	1	63	106	0.708009	0.445817	0.714017	0.581285	0.460237	61.40971	-	-	0.973922	62.22955	-	-	-
28	Marumisa	0	11	-	0.35991	0.168371	0.312569	0.332857	0.094785	18.33504	-	-	0.757588	40.84809	-	-	-
29	Masakadza	2	66	119	1.451593	1.73408	0.989749	0.783888	0.583531	20.41316	-	-	0.935351	45.74819	-	-	-
30	Matsikenyeri	2	116	35	0.617121	0.483163	0.627672	0.559416	0.338448	19.30903	-	-	0.701256	43.66019	-	-	-
31	Sibanda	1	96	127	0.711907	0.707239	0.641871	0.499321	0.326777	21.13344	-	-	0.817273	42.69542	-	-	-
32	Hayden	2	291	395	1.368963	1.094703	1.094456	0.867656	0.711907	47.04718	-	-	1.550484	61.83062	-	-	-
33	Hodge	33	199	138	1.279357	1.183074	1.243689	1.03579	0.76291	61.66003	43.69353	2.954816	1.358207	63.36425	122.3974	1.358207	63.66154
34	M_Hussey	12	262	124	0.253252	0.34545	0.555818	0.784376	0.561451	21.73628	12.25057	2.141224	1.303034	56.27093	106.4688	1.303034	56.52662
35	Ponting	6	349	103	0.906007	0.798637	0.924475	0.741526	0.576895	47.76661	37.2194	1.922952	1.435184	60.82044	125.5897	1.435184	61.1348
36	Y_Singh	13	246	452	0.94275	1.174063	0.969482	0.898711	0.688566	48.8571	19.00205	1.024921	1.083827	58.98916	101.668	1.083827	59.21791
37	Gambhir	5	123	337	0.59827	0.276161	0.581467	0.485934	0.265321	37.28262	6.773556	0.741518	1.214248	59.04153	85.18965	1.214248	59.19563
38	Sehwag	11	240	203	1.347246	0.919644	1.073158	1.026334	0.863625	41.83981	18.6853	0.777882	1.117167	64.40571	108.0315	1.117167	64.64298
39	R_Sharma	8	27	256	1.229153	0.489708	1.012081	0.991885	0.605999	13.85141	35.94249	1.716811	1.116343	52.44147	120.6928	1.116343	52.75318
40	Uthappa	9	48	200	0.451653	0.613071	0.516366	0.555787	0.286964	45.9689	35.5825	0.829085	1.364461	65.25949	105.2682	1.364461	65.48171
41	R_Taylor	4	61	232	0.575361	0.462156	0.754074	0.806903	0.519428	25.842	38.3142	1.392639	1.398959	61.33826	139.5552	1.398959	61.69781
42	ul_Haq	15	98	371	0.761721	0.834942	0.782991	0.764732	0.595201	30.97011	63.38691	1.230768	1.345188	63.02684	102.8534	1.345188	63.24668
43	Butt	11	85	64	0.753202	0.762334	0.705134	0.783543	0.606072	29.52245	8.304414	0.581803	1.430569	60.96579	86.99302	1.430569	61.12008
44	Y_Khan	13	190	307	1.072397	1.012742	0.791499	0.702297	0.478495	25.26584	11.57474	0.799705	1.090341	52.98262	101.7849	1.090341	53.22913
45	G_Smith	19	169	176	1.031682	1.22308	1.148585	1.093897	0.939484	40.2937	62.53091	2.273299	1.547691	62.71217	118.4673	1.547691	62.99642
46	Gibbs	7	321	179	0.918477	0.865959	0.72627	0.684219	0.432335	39.61502	7.970761	0.401233	1.2029	58.39223	94.25083	1.2029	58.59133
47	Jayawardene	5	311	297	0.992002	0.886966	0.846756	0.693745	0.57186	32.40366	-	0.358902	1.12417	53.7118	73.71632	1.12417	53.83283
48	Silva	5	123	119	0.871795	1.013009	0.714088	0.585802	0.492374	22.59701	5.436407	1.079763	0.956788	48.54442	76.0762	0.956788	48.69972
49	Sarwan	4	184	104	0.808042	1.188803	0.653655	0.612496	0.412707	37.28613	14.71485	0.66698	1.208983	56.60767	82.53196	1.208983	56.75962
50	Chanderpaul	2	321	77	0.824625	0.712049	0.834493	0.798216	0.606872	37.54253	-	-	1.171164	53.1758	-	-	-
51	Bosman	23	117	-	0.336799	0.007002	0.608186	0.640789	0.482854	26.01767	-	1.480125	0.947055	56.96963	-	0.947055	-

Table B.2: Data used for the assessment of the batsmen performance measures (2 of 2)

		T20 Games	ListA Games	WC Fantasy Score	Fantasy MA 10	Fantasy MA 5	Fantasy EWA	Fantasy Mean	Fantasy Median	Lemmer CBR ListA	Lemmer FCBR ListA	Lemmer CBR T20	Lemmer FCBR T20	Gerber Sharp Bowling T20	Gerber Sharp Bowling ListA	Barr and Kantor Bowling ListA	Barr and Kantor Bowling T20
TOTAL Correl		-	-	-	*0.151538*	*0.258616*	*0.175677*	*0.126471*	*0.118033*	*-0.17056*	*-0.12957*	*0.022922*	*-0.00948*	*0.010993*	*0.184848*	*-0.1856*	*-0.01028*
1	Clark	2	119	525	1.927029	2.478104	1.371506	1.201044	1.125427	10.23338	10.6696	-	-	-	1.190751	5.230843	-
2	Hilfenhaus	5	22	-	0.907938	0.983265	0.789211	0.749268	0.779534	10.10191	10.58431	9.680141	12.90685	1.397795	1.393375	5.00347	6.850318
3	Johnson	3	41	299	1.024357	0.905639	1.009702	1.010455	0.854892	11.69609	12.85787	12.46154	16.61538	1.655808	0.664911	6.067723	6.973848
4	Tait	5	53	-	1.07603	1.22308	1.166567	1.150105	0.965883	10.64418	12.75794	9.567853	10.77464	1.261329	0.699759	5.814187	7.145619
5	Mortaza	3	76	64	1.352301	0.935337	1.497541	1.544422	1.365461	10.55108	10.43183	16.96815	22.6242	0.064088	1.144911	5.407922	10.62546
6	Rasel	2	59	162	1.130464	0.987362	0.995588	0.959598	0.907108	9.705383	9.465724	-	-	-	1.358026	4.914051	-
7	Rahman	0	14	-	1.130838	0.543235	1.130374	1.174897	1.190356	10.6928	10.96868	-	-	-	0.996073	5.528821	-
8	Anderson	19	121	115	1.544983	1.68757	1.184006	1.150819	1.091202	10.64775	10.51773	16.34493	21.79325	0.183679	0.9404	5.568576	10.23216
9	Broad	12	31	125	0.875326	1.093471	0.825865	0.800317	0.773696	12.12223	13.05569	9.925413	13.23388	2.021674	0.553569	6.316137	6.446952
10	Kirtley	30	218	2	1.033237	0.627642	1.00727	1.082467	0.984749	10.15147	10.55373	13.41487	17.07837	0.726535	1.042716	5.398906	8.526771
11	Schofield	19	107	169	0.735254	0.779142	0.914828	0.848279	0.68716	11.35253	14.11538	7.908825	10.5451	2.011278	0.63348	6.108831	6.141863
12	Tremlett	16	95	40	0.907743	0.826716	0.934062	1.1059	1.034089	10.38375	11.10911	11.89241	14.75309	1.03913	1.00094	5.448331	7.818941
13	Sidebottom	17	139	-	1.134719	1.128861	1.081636	1.027421	0.892767	10.34153	10.23873	11.95552	15.94069	1.235893	1.251328	5.203732	7.607825
14	Odhiambo	0	21	-10	0.445917	0.280294	0.57507	0.610809	0.536848	13.82897	15.39688	-	-	-	0.338178	6.929719	-
15	Ongondo	2	84	-27	1.045875	1.255214	1.144313	0.884764	0.75779	10.93318	11.42751	-	-	-	0.89548	5.667631	-
16	Otieno	0	1	-	-0.01082	-0.01082	-0.01082	-0.01082	-0.01082	-	-	-	-	-	-	-	-
17	Varaiya	2	20	-	0.732358	0.285985	1.239166	1.387489	1.353419	8.706063	8.165755	-	-	-	1.814814	4.335394	-
18	Bond	4	98	236	1.656361	1.294573	1.222701	1.084966	0.887749	9.679137	9.466289	11.43459	15.24613	0.809996	1.209541	5.07029	7.984717
19	Gillespie	4	61	295	0.967494	0.830947	1.064449	1.063	1.125427	11.65037	11.93368	12.26161	16.34882	0.595166	0.657529	6.056592	8.260249
20	Martin	11	93	160	0.966804	1.260972	0.961105	1.08229	1.038856	10.72452	10.89013	16.78604	22.38138	0.366716	0.934775	5.576269	10.27891
21	Patel	8	60	65	1.065527	1.300106	0.971965	0.906745	0.774277	11.58289	13.94032	10.56836	14.09115	1.823894	0.949002	5.837314	6.809035
22	Rehman	11	72	-	1.547291	1.71377	1.254251	1.253801	1.179534	9.768309	10.73427	11.73814	14.20701	1.636238	1.376536	4.933104	7.238005
23	Anjum	9	68	-	1.057958	1.331264	1.089759	1.104699	0.89557	10.95374	10.56777	13.95262	18.60349	0.621322	1.130582	5.525498	8.514992
24	Blain	0	82	60	1.094503	0.760579	0.97905	0.961665	0.802604	11.72072	12.89753	-	-	-	0.57127	6.233016	-
25	Drummond	0	4	-	-	0.705053	0.717047	0.705053	0.546482	11.66826	9.225551	-	-	-	0.889074	5.828851	-
26	Lyons	0	29	-	0.658941	0.89774	0.595621	0.547397	0.541071	14.05773	17.91454	-	-	-	0.409678	7.007452	-
27	J_Nel	3	47	123	0.675264	0.318971	0.710981	0.698729	0.609833	12.83152	13.05925	15.68807	20.91743	0.533885	0.475941	6.581457	8.927104

Table B.3: Data used for the assessment of the bowlers performance measures (1 of 2)

		T20 Games	ListA Games	WC Fantasy Score	Fantasy MA 10	Fantasy MA 5	Fantasy EWA	Fantasy Mean	Fantasy Median	Lemmer CBR ListA	Lemmer FCBR ListA	Lemmer CBR T20	Lemmer FCBR T20	Gerber Sharp Bowling T20	Gerber Sharp Bowling ListA	Barr and Kantor Bowling ListA	Barr and Kantor Bowling T20
28	M_Morkel	20	21	351	1.301065	1.479952	1.14278	1.003925	0.930642	10.20876	10.97325	13.13728	15.43173	0.91221	1.078325	5.350455	8.143723
29	Tshabalala	23	45	-	0.782389	0.664667	0.76451	0.843768	0.665517	11.02786	13.25207	10.03913	13.3855	1.414638	0.889069	5.707675	7.047815
30	Collins	3	63	-	1.022611	0.796466	1.249872	1.304185	1.107945	9.91191	10.24052	8.050034	10.73338	3.18858	1.215421	5.118196	5.084608
31	Edwards	4	41	5	1.078185	1.217179	0.729905	0.706894	0.575296	10.77458	11.10513	13.34766	17.79688	1.3855	0.986658	5.520442	7.393635
32	Powell	3	66	6	1.25985	0.917656	1.136808	1.187143	1.132049	10.71299	10.96646	19.66281	26.21708	0.504446	0.899564	5.634774	10.17075
33	Rampaul	3	42	65	1.11419	1.202073	1.068735	1.035405	0.921978	10.28048	11.26735	14.40609	19.20812	0.434857	1.041106	5.364696	8.902929
34	Brent	1	107	65	1.579599	1.643222	1.327662	1.065363	1.022624	11.45763	12.25577	-	-	-	0.91411	5.81346	-
35	Mpofu	0	39	-	1.212576	1.010675	1.071393	1.036989	0.875296	11.14572	11.90294	-	-	-	0.778376	5.841039	-
36	Mupariwa	0	41	5	1.01447	0.938905	1.252513	1.174446	1.003672	11.71989	11.92589	-	-	-	0.687568	6.105563	-
37	Bracken	6	142	329	1.503175	1.745923	1.334687	1.270497	1.268676	9.942228	10.02588	15.53992	20.71989	0.411755	1.212722	5.119995	9.172165
38	Lee	3	181	364	1.249922	1.211409	1.426391	1.43194	1.371297	10.23228	10.69187	17.70492	23.60656	0.74769	1.020343	5.407999	9.064263
39	Razzak	3	92	252	0.729313	0.784562	0.922266	1.046186	0.917606	8.785259	9.221496	7.516484	8.290909	3.061659	1.595445	4.543716	5.090924
40	Agarkar	5	245	40	1.110745	0.830947	1.153719	1.291552	1.190402	11.18671	12.01503	9.688551	10.69062	1.538692	0.765726	5.918007	6.70599
41	H_Singh	14	197	248	0.909504	0.593849	0.766176	0.870291	0.797037	10.13258	11.55322	13.14826	17.53101	1.583164	1.391098	5.065955	7.36033
42	RP_Singh	3	67	426	1.134383	0.905639	0.984004	1.052891	0.941463	11.18894	11.5739	20.66087	27.54783	0.441259	0.837655	5.863982	10.57723
43	Sreesanth	4	47	275	0.403096	0.482613	0.499047	0.52617	0.405272	12.78262	12.93077	14.89091	19.85455	0.366412	0.373264	6.608327	9.4725
44	Vettori	1	267	551	1.179899	1.146054	1.262314	1.16552	0.974496	10.12357	11.70846	-	-	-	1.415477	5.047851	-
45	Asif	14	55	424	0.639897	0.642579	1.112809	1.175247	1.062026	11.29931	10.99756	10.20343	12.82342	1.578845	0.97678	5.733291	6.996311
46	Tanvir	10	13	221	0.846733	1.02587	0.767321	0.688786	0.651093	11.63555	13.61162	29.47826	39.30435	0.237052	0.647772	5.988472	-
47	Gul	9	63	565	1.058391	1.120299	1.129321	1.046949	0.938856	10.97671	11.3293	10.25995	10.41027	1.85414	0.865804	5.688512	6.780515
48	A_Nel	11	171	10	1.15825	1.412456	1.301372	1.281205	1.276927	9.784454	9.899643	10.89059	13.36556	2.274322	1.305431	5.026495	6.395962
49	Ntini	10	214	44	1.03635	1.015343	1.166567	1.208893	1.172896	10.04366	10.34458	15.36207	20.48276	0.69718	1.154063	5.204896	8.638172
50	D_Fernando	5	159	390	1.111042	1.729584	0.99885	0.955115	0.921978	11.26801	12.50585	15.09826	20.13102	0.315334	0.704622	5.959362	9.323017
51	Malinga	9	65	315	0.94568	0.602925	1.069196	1.059317	0.945312	10.51413	11.49638	14.48986	19.31981	0.902299	0.832624	5.615731	8.296405
52	Vaas	2	351	258	1.095047	1.395963	1.415259	1.413862	1.318779	9.832892	9.651908	-	-	-	1.334738	4.993675	-
53	Chawla	4	35	-	0.763086	0.781931	1.098239	1.174828	0.863625	10.90709	11.8188	11.46171	15.28228	1.253448	0.941566	5.690022	7.217597
54	Akhtar	12	177	-	1.258177	0.982636	1.231228	1.278159	1.206437	10.37715	10.69416	12.28905	14.98986	0.973813	0.931097	5.498767	7.972491
55	Muralitharan	8	374	-	1.418215	1.38381	1.336731	1.473392	1.382967	8.83504	9.407179	7.641892	9.36646	2.554837	1.71625	4.484226	5.67442

Table B.4: Data used for the assessment of the bowlers performance measures (2 of 2)

		T20 Games	ListA Games	WC Fantasy Score	Fantasy MA 10	Fantasy MA 5	Fantasy EWA	Fantasy Mean	Fantasy Median	Lemmer Batting ListA	Lemmer CBR ListA	Lemmer FCBR ListA	Lemmer Batting T20	Lemmer CBR T20	Lemmer FCBR T20	Gerber Sharp All Ronder T20	Gerber Sharp All Ronder ListA	Barr and Kantor Batting ListA	Barr and Kantor Bowling ListA	Barr and Kantor Batting T20	Barr and Kantor Bowling T20
	TOTAL Correl	-	-	-	*0.479894*	*0.55657*	*0.45627*	*0.45761*	*0.483732*	*0.335773*	*-0.21581*	*-0.10758*	*0.109147*	*-0.09232*	*-0.09373*	*0.283272*	*0.358917*	*0.427573*	*-0.22538*	*0.338325*	*-0.21575*
1	Clarke	11	175	93	0.990035	0.786966	0.774988	0.764028	0.679142	27.87767	11.5591	14.91296	19.40956	18.3954	24.52721	0.927909	1.145283	50.06474	7.661444	22.20537	10.04187
2	Hogg	12	212	-	1.443654	1.699241	1.224169	1.07243	0.969282	9.219183	10.80517	12.93133	5.23715	10.64291	14.19055	0.981429	0.89481	32.19037	7.106144	19.46957	8.118338
3	Kapali	2	102	4	0.893353	1.107167	0.503702	0.408489	0.182731	16.56232	11.28792	12.67574	-	-	-	-	1.153958	39.50637	7.49455	-	-
4	Reza	3	45	22	0.708978	0.781265	0.702104	0.817844	0.855284	16.83303	10.6999	11.32588	-	13.75107	18.33476	0.524016	1.082576	38.33858	7.027711	10.98134	11.92348
5	Mahmuddullah	2	20	114	0.59908	0.768796	0.581956	0.574977	0.425143	26.86863	10.74458	12.21373	-	-	-	-	1.116571	37.80581	7.075226	-	-
6	Al_Hassan	1	42	304	0.958957	0.835248	1.480665	1.629029	1.640521	25.89456	10.0508	11.58706	-	-	-	-	1.367155	43.91707	6.585559	-	-
7	Collingwood	6	292	124	1.316076	1.818281	1.354943	1.060134	0.903818	36.80456	11.58882	13.91624	12.74199	10.16529	13.55372	1.099157	1.204628	45.84205	7.678722	32.10625	8.62976
8	Flintoff	10	265	339	1.490075	1.665233	1.572444	1.374635	1.244737	23.51746	9.566318	10.10752	22.20396	12.38951	16.51934	1.372707	1.306913	48.0637	6.221545	29.94107	7.753958
9	Maddy	43	310	279	0.753538	0.415542	1.185813	1.013764	0.785142	25.98824	11.50911	13.07869	18.09463	14.95443	19.93925	1.479061	1.068904	38.46506	7.63561	34.56503	8.950177
10	Snape	40	271	22	0.606186	0.880863	0.942858	0.929569	0.779142	12.06554	10.86046	12.81942	7.019269	11.42084	14.69336	1.160024	0.937545	32.33936	7.143608	17.10709	7.2817
11	Wright	29	71	57	0.808227	0.462487	0.959607	0.898028	0.768321	23.36263	12.56036	14.05435	114.0155	12.88235	17.17647	0.880066	0.689831	36.09512	8.393604	19.04787	8.836614
12	Bopara	30	86	-	0.853438	0.420142	0.981747	0.849828	0.663939	22.20141	11.84524	14.98499	7.062637	13.55745	18.0766	0.785471	0.86136	39.8715	7.931486	17.12238	8.744217
13	Tikolo	1	148	-6	1.204506	0.905838	1.366461	1.26594	1.028035	27.919	11.01787	13.10796	-	-	-	-	1.332007	49.18799	7.257068	-	-
14	Kamande	2	59	93	0.831718	0.612401	0.675587	0.549662	0.361789	6.506424	12.34787	14.35247	-	-	-	-	0.712183	23.69255	8.326697	-	-
15	C_Obuya	9	83	30	0.370207	0.493721	0.55604	0.7569	0.500235	7.858809	12.86359	14.42304	4.296103	11.86592	15.82123	0.838355	0.702764	28.3999	8.638707	15.93368	9.485195
16	Odoyo	2	143	88	1.025405	1.008864	1.326008	1.344034	1.248755	19.15471	10.69644	11.30537	-	-	-	-	1.028937	38.87951	7.025791	-	-
17	Onyango	2	34	-26	0.469232	0.598804	0.218455	0.113531	-0.06995	6.406422	14.27731	13.79183	-	-	-	-	0.320389	21.90499	9.68095	-	-
18	Suji	1	86	15	0.591929	0.376252	0.474446	0.601351	0.478495	4.595769	12.80174	13.06055	-	-	-	-	0.634514	21.59159	8.659033	-	-
19	N_McCullum	8	57	1	1.147621	1.330711	1.142951	1.02676	0.84407	-	10.80777	12.42572	12.12358	9.083613	12.11148	1.428148	0.961521	-	7.192451	22.02867	6.402441
20	Scott	8	44	-	1.096864	1.337954	1.074223	1.061956	1.070183	-	10.9629	11.74451	-	19.11134	25.48179	0.444544	0.794676	-	7.220822	-	10.94715
21	Alam	15	32	75	1.44698	0.811919	1.470411	1.346918	1.14707	33.274	11.19372	13.33376	6.292	8.775228	11.7003	1.37481	1.25028	54.42274	7.410885	20.31859	6.449739
22	Arafat	28	154	86	1.011229	0.898488	1.140394	1.279551	1.229614	13.42656	10.60762	11.44283	3.613213	13.14472	16.73092	0.449481	0.78744	32.87493	7.005616	10.94564	9.215758
23	R_Watson	0	106	2	0.817717	1.232878	0.874631	0.928921	0.703392	26.47448	12.5924	15.34168	-	-	-	-	0.95317	45.11112	8.413246	-	-
24	Brown	28	314	-19	0.732997	0.511338	0.968841	1.247576	1.125427	12.13208	10.64506	10.82712	5.64539	13.85728	15.74848	0.699901	0.925861	33.77992	6.995463	11.432	8.651421
25	Hamilton	3	163	17	0.589366	0.240415	0.617438	0.928118	0.719624	26.96505	10.57464	10.91028	-	-	-	-	0.90175	35.41554	6.963699	17.45329	-
26	Haq	0	51	84	1.153538	0.913277	1.014343	0.898	0.746919	10.91522	12.641	15.32273	-	-	-	-	0.597349	31.86642	8.452756	-	-
27	Maiden	1	41	-	0.477224	0.577864	0.608936	0.681084	0.465321	-	12.14951	13.76402	-	-	-	-	0.724721	-	8.142428	-	-
28	C_Wright	0	110	124	1.057669	0.711301	1.118736	1.25214	1.022624	10.16819	10.24127	10.05841	-	-	-	-	0.895687	31.43031	6.701692	-	-
29	Kemp	24	218	353	0.619709	0.858957	0.669123	1.090862	0.99557	24.31543	11.10146	12.45086	5.914128	11.77586	15.70114	0.921512	1.074353	44.17774	7.32152	20.25964	8.450587
30	Philander	16	35	148	0.829878	0.94949	0.672751	0.520245	0.34407	17.86184	11.41589	11.33193	9.797741	16.03526	21.38035	0.722909	0.824486	35.55047	7.552274	18.4566	10.00722
31	van_der_Wath	33	117	186	0.631709	0.688925	1.074229	1.150493	1.067097	22.51257	11.01754	11.9786	3.989224	13.52066	17.2635	0.602736	0.862254	47.66702	7.274188	10.9258	8.466647

Table B.5: Data used for the assessment of the all rounders performance measures (1 of 2)

		T20 Games	ListA Games	WC Fantasy Score	Fantasy MA 10	Fantasy MA 5	Fantasy EWA	Fantasy Mean	Fantasy Median	Lemmer Batting ListA	Lemmer CBR ListA	Lemmer FCBR ListA	Lemmer Batting T20	Lemmer CBR T20	Lemmer FCBR T20	Gerber Sharp All Ronder T20	Gerber Sharp All Ronder ListA	Barr and Kantor Batting ListA	Barr and Kantor Bowling ListA	Barr and Kantor Batting T20	Barr and Kantor Bowling T20
32	H_Fernando	16	44	-	1.43922	1.551939	1.454419	1.374495	1.334513	23.77606	10.36719	10.72064	-	12.25272	16.33696	1.145298	1.240082	44.08241	6.786731	-	8.548253
33	Lokuarachchi	8	88	-	1.047287	0.655778	1.341399	1.356509	1.303981	19.18204	9.757316	11.26355	-	8.808578	11.74477	1.461908	1.029745	34.73988	6.36681	-	6.102302
34	Perera	10	55	-	0.846235	0.439349	0.944316	1.064202	0.90004	13.17895	9.48588	10.36888	-	13.45208	17.93611	0.858433	1.238973	32.68968	6.173317	-	7.917814
35	Wijekoon	8	98	36	1.327658	0.890799	1.181425	1.134179	1.044463	-	9.074418	9.413002	4.053111	11.55892	15.41189	0.772998	1.272934	-	5.877683	17.36201	8.052122
36	Sammy	4	50	-	0.673585	0.744028	1.053597	1.118801	1.071516	12.96824	10.18152	10.86566	-	11.08852	14.78469	0.854097	1.144114	38.60629	6.655116	-	7.52558
37	Utseya	1	84	73	0.203738	0.176895	0.498994	0.536303	0.363625	3.373441	10.47667	12.00582	-	-	-	-	0.98386	19.35717	7.154635	-	-
38	Chibhabha	2	46	85	0.181928	0.169537	0.315397	0.394475	0.24437	17.14448	13.9939	16.55274	-	-	-	-	0.882335	37.21347	9.547825	-	-
39	Chigumbura	2	89	274	1.381	1.001706	0.858488	0.663841	0.541884	17.89738	14.68936	16.00281	-	-	-	-	0.672838	35.66276	9.994755	-	-
40	Dabengwa	1	37	-30	0.353303	0.175798	0.257455	0.26006	-0.06104	9.121055	10.93179	12.86268	-	-	-	-	0.836362	28.78501	7.27912	-	-
41	Maruma	1	11	-	1.152796	0.886128	1.321243	1.416064	1.347266	1.982595	7.683885	7.816301	-	-	-	-	1.31099	16.34723	4.90598	-	-
42	Williams	1	35	-	0.892802	0.802938	0.957804	0.943983	0.898178	20.73566	11.17001	14.08449	-	-	-	-	1.105227	37.52414	7.452711	-	-
43	Symonds	17	388	181	0.341084	0.575663	0.730143	0.753532	0.55699	36.3051	11.36336	13.25962	136.3057	13.1553	16.23048	1.680201	1.180029	51.33732	7.508574	39.98705	8.17374
44	S_Watson	8	124	35	1.080698	0.884632	1.105502	1.034865	0.833249	24.13493	12.10969	14.52241	8.585799	12.19659	16.26212	0.866323	0.915944	43.12923	8.057336	20.37117	9.246501
45	Mascarenhas	24	199	139	1.084174	1.454098	1.244918	1.315559	1.220213	33.02579	9.784108	9.867591	15.32952	11.69554	12.87482	1.077791	1.096523	51.54283	6.370536	20.49072	8.151234
46	J_Sharma	8	42	105	0.880337	0.520874	0.981473	1.135523	1.06769	9.215421	10.85317	11.7632	39.64057	13.68729	18.24972	0.780088	0.775929	34.11563	7.156695	11.9263	8.088751
47	IK_Pathan	12	109	433	1.293345	1.561319	1.324577	1.355905	1.331034	16.56134	10.9669	11.78172	5.688481	11.80463	14.32985	0.70442	0.743732	32.89255	7.258252	9.33519	7.914414
48	YK_Pathan	4	38	32	1.378522	0.441514	1.284799	1.277928	0.968517	40.59255	11.88022	13.76845	-	8.071942	10.76259	1.39915	1.276203	54.03428	8.010951	25.11641	7.883903
49	Oram	5	173	142	1.446356	1.417179	1.310734	1.13977	0.995	23.75275	10.58038	10.78194	26.88863	17.72379	23.63171	1.209436	1.088245	41.3461	6.947767	27.26687	9.544967
50	Styris	29	260	185	0.881607	0.589134	1.232342	1.345715	1.190402	35.34558	10.78326	12.09756	13.95176	13.5949	17.27117	1.149794	1.164226	45.2029	7.087341	30.09739	8.532233
51	Malik	19	205	390	0.440651	0.587701	0.634132	0.91992	0.760426	31.50321	10.40355	12.03546	10.23995	10.59274	14.12365	1.72181	1.274929	46.49002	6.814688	39.40766	7.608856
52	Hafeez	25	112	355	1.523332	1.314061	1.457729	1.542042	1.436234	25.20908	10.32292	12.11136	24.77552	12.24583	16.32777	1.646535	1.479646	43.55558	6.790947	34.36202	7.295073
53	Afridi	12	317	709	1.908878	1.815649	1.392065	1.259294	1.082142	21.97957	11.09613	12.99556	45.39877	11.86495	15.81994	1.14975	1.117813	53.18419	7.322657	19.50824	7.38693
54	JA_Morkel	26	98	379	1.058278	1.02969	0.955885	0.903661	0.892963	18.24099	10.86448	11.97123	5.047812	14.93526	18.60599	0.668052	0.854704	37.30468	7.146349	15.59319	8.890657
55	Pollock	14	420	268	1.607043	1.437819	1.779996	1.667957	1.552191	20.68525	8.708931	8.7038	18.2238	13.03448	15.75	1.448316	1.261121	35.07051	5.608958	23.8451	7.082184
56	Dilshan	4	191	183	0.724745	0.816943	0.825377	0.810363	0.595201	20.66338	11.58308	13.95391	-	8.678405	11.57121	1.417487	1.255326	44.7162	7.70195	2.045361	5.623916
57	Jayasuriya	9	470	443	1.425122	2.232262	1.203611	1.047368	0.908708	36.38587	11.52256	14.33018	22.86322	9.026645	12.03553	1.177391	1.184263	54.02673	7.635485	25.70922	7.386963
58	Maharoof	6	97	5	0.84789	1.062629	0.48721	0.321887	0.133649	13.53401	10.77263	11.50194	6.457429	17.90576	23.87435	0.783743	0.80688	30.96205	7.088212	11.02625	9.228833
59	Bravo	5	101	-58	0.530208	0.327769	0.617185	0.540121	0.387161	15.94995	11.59854	13.23394	0.730794	11.02041	14.69388	1.126404	0.825555	35.81808	7.721814	8.451636	6.598478
60	Gayle	8	226	327	1.194583	1.337452	1.321814	1.507421	1.400473	32.8369	10.65586	12.38189	32.20173	9.299075	12.39877	1.590567	1.466706	55.29744	6.996886	25.11299	6.426915
61	DR_Smith	4	98	48	0.463031	0.479964	0.528494	0.50942	0.265225	6.316785	11.8823	12.94989	8.075609	11.2717	15.02893	1.053861	0.895807	34.18002	7.908773	10.43711	7.003277
62	Bhudia	2	2	13	-0.07413	-0.07413	-0.06983	-0.07413	-0.11362	-	-	-	-	-	-	-	-	-	-	-	-

Table B.6: Data used for the assessment of the all rounders performance measures (2 of 2)

		T20 Games	ListA Games	WC Fantasy Score	Fantasy MA 10	Fantasy MA 5	Fantasy EWA	Fantasy Mean	Fantasy Median	Lemmer Batting ListA	Lemmer Batting T20	Gerber Sharp Keeper T20	Gerber Sharp Keeper ListA	Barr and Kantor Batting ListA	Barr and Kantor Batting T20	Barr and Kantor Combin
	TOTAL Correl	-	-	-	*0.406054*	*0.521207*	*0.567595*	*0.61999*	*0.636487*	*0.470664*	*0.272577*	*-0.4536*	*0.469398*	*0.57816*	*0.286051*	*0.462992*
1	Haddin	7	112	11	0.783784	0.823295	0.852959	0.756457	0.551892	37.72204	13.26962	1.229074	1.390081	74.88911	18.69716	74.88911
2	Rahim	3	36	139	0.376911	0.279995	0.362529	0.384374	0.221742	18.81756	-	0.780184	0.79269	44.94216	1.456137	44.94216
3	Prior	26	147	145	0.670789	0.613874	0.755822	0.61846	0.443678	29.11228	17.68693	1.101817	0.928204	61.82523	27.37597	61.82523
4	D_Obuya	2	60	23	0.401053	0.364266	0.493953	0.45471	0.309271	19.74451	-	-	0.760922	49.7305	-	-
5	Ouma	0	37	-10	0.270763	0.288706	0.237961	0.193425	0.056753	17.04895	-	-	0.600319	47.11666	-	-
6	Hopkins	10	113	-	0.788884	0.544298	0.565387	0.464045	0.411214	19.07862	-	0.668454	0.879184	46.05278	-	-
7	C_Smith	0	92	26	0.184178	0.065355	0.401839	0.470271	0.373339	19.16434	-	-	0.907928	53.08409	-	-
8	Ramdin	8	59	23	0.555372	0.394467	0.528929	0.543682	0.37346	19.99406	15.41283	0.818993	1.013902	61.13797	14.3672	61.13797
9	B_Taylor	2	87	260	0.510266	0.667926	0.545116	0.558328	0.448519	24.30035	-	-	0.904497	53.57295	-	-
10	Gilchrist	4	334	438	1.028181	1.122713	0.977005	0.888826	0.717743	39.38346	-	0.613591	1.443503	80.44618	20.08793	80.44618
11	Dhoni	5	136	285	0.585843	0.788933	0.780857	0.820217	0.555344	49.87349	-	1.084572	1.312789	76.60183	30.03179	76.60183
12	Karthik	11	66	93	0.360622	0.289431	0.548151	0.563519	0.443483	26.42248	6.796675	1.211806	0.949345	55.80878	20.28394	55.80878
13	BB_McCullum	16	150	265	0.442316	0.324443	0.555284	0.522982	0.408471	22.82852	33.22139	0.93121	0.849205	66.56426	31.25726	66.56426
14	Akmal	17	138	138	0.279165	0.368533	0.332447	0.362442	0.17346	24.94851	12.27691	1.020612	1.030103	65.16834	20.96352	65.16834
15	Boucher	9	309	213	0.594034	0.602203	0.692329	0.568026	0.466073	31.39334	8.810131	1.114329	1.002336	64.11819	27.90938	64.11819
16	de_Villiers	11	70	142	0.783465	0.849988	0.7693	0.708813	0.506872	42.74749	19.00535	1.128451	1.066042	71.71295	27.24034	71.71295
17	Sangakkara	6	267	166	0.637388	0.532414	0.748306	0.787773	0.638464	37.62298	103.0886	1.345673	1.215145	64.22173	40.20032	64.22173
18	Taibu	0	133	-28	0.747809	0.679292	0.421907	0.222726	0.078495	29.64534	-	-	0.930573	52.71349	-	-

Table B.7: Data used for the assessment of the wicket keepers performance measures

	List A			Twenty20		
Name	$\overline{BT}$	SR^t_{BAT}	BK^{LtA}_{BAT}	$\overline{BT}$	SR^t_{BAT}	BK^{Tw20}_{BAT}
Ashraful	17.4328	72.5300	44.0784	21.5000	179.1667	137.1281
Ahmed	23.3614	82.3200	53.0171	16.6667	161.2903	121.1400
Chowdhury	14.0870	72.2300	40.8059	10.0000	142.8571	102.1526
Nazimuddin	22.6111	76.0900	49.7995	62.0000	142.5287	128.3246
Iqbal	18.8800	74.2800	46.0322	6.0000	75.0000	54.5409
Pietersen	33.6946	87.5000	62.6940	22.0625	147.0833	115.7857
Shah	27.6776	76.1200	53.4577	26.9231	145.5301	117.6335
Solanki	25.8994	72.9300	50.7970	29.9500	142.2803	116.8949
Mishra	24.0938	73.0800	49.5970	25.5000	76.1194	66.3125
Obanda	15.3333	66.5100	39.8356	5.0000	29.4118	23.5217
Fulton	28.7471	73.7900	53.0857	16.8000	112.0000	88.1677
McMillan	25.9783	75.9400	52.2070	24.1429	146.9565	117.0208
Vincent	26.5967	69.8800	49.8651	23.4286	124.7148	101.0033
Nazir	26.8421	80.2100	54.7204	25.3889	162.6335	128.6733
Poonia	18.1250	67.6000	42.6817	10.6667	152.3810	108.9628
Duminy	23.5094	73.0000	49.1384	22.4400	112.4248	91.7489
Mubarak	24.2752	66.1200	46.5921	32.7500	130.3483	109.5087
Tharanga	29.4157	69.4300	51.4346	13.3333	105.2632	81.1161
Samuels	25.0067	75.6200	51.3750	30.0000	189.8734	150.4532
DS_Smith	25.0000	101.5400	62.2295	61.0000	179.4118	156.5895
Masakadza	24.1970	64.4000	45.7482	18.5000	123.3333	97.0894
Matsikenyeri	18.7672	68.7000	43.6602	7.5000	93.7500	68.1762
Sibanda	21.0521	62.4100	42.6954	10.0000	62.5000	49.6028
Hayden	39.2096	78.9600	61.8306	13.0000	216.6667	151.9486
Hodge	34.7286	87.5100	63.3642	41.9091	142.8719	122.3974
M_Hussey	32.0153	76.1700	56.2709	33.0833	126.0317	106.4688
Ponting	36.1404	80.4300	60.8204	33.0000	152.3077	125.5897
Y_Singh	28.7236	86.8100	58.9892	25.0769	124.4275	101.6680
Gambhir	32.2195	81.7300	59.0415	21.6000	103.8462	85.1896
Sehwag	28.8417	99.1400	64.4057	21.4545	136.4162	108.0315
R_Sharma	27.7037	73.8700	52.4415	31.5000	146.5116	120.6928
Uthappa	34.4792	91.9200	65.2595	21.2222	132.6389	105.2682
R_Taylor	33.5902	84.7500	61.3383	27.7500	176.1905	139.5552
ul_Haq	33.7551	88.1300	63.0268	25.2000	126.0000	102.8534
Butt	38.2353	78.3200	60.9658	17.3636	109.7701	86.9930
Y_Khan	26.3632	77.0700	52.9826	20.6923	128.0952	101.7849
G_Smith	38.2130	81.8200	62.7122	36.0526	140.6571	118.4673
Gibbs	30.2368	83.1400	58.3922	12.2857	126.4706	94.2508
Jayawardene	27.7428	76.5800	53.7118	5.0000	108.6957	73.7163
Silva	24.2276	70.5000	48.5444	23.6000	90.0763	76.0762
Sarwan	32.2663	76.5500	56.6077	20.0000	101.2658	82.5320
Chanderpaul	31.3240	70.6500	53.1758	33.5000	126.4151	106.9204

Table B.8: Batsmen Barr and Kantor weighting data (1 of 2)

	ListA			**Twenty20**		
Name	$\overline{BT}$	SR^t_{BAT}	BK^{LtA}_{BAT}	$\overline{BT}$	SR^t_{BAT}	BK^{Tw20}_{BAT}
Siddique	22.7333	72.4100	48.3097	NaN	-	-
McCallum	20.8750	63.8400	43.2018	NaN	-	-
Sheikh	4.4000	44.2300	19.7510	2.9000	0	-
Watts	19.0698	58.8800	39.7123	NaN	-	-
Bodi	24.4935	62.8700	45.2301	16.4074	0	-
Deonarine	18.7451	50.5100	35.7266	17.0000	0	-
Morton	30.1809	65.9300	50.1805	8.2500	0	-
Marumisa	19.5455	60.6800	40.8481	NaN	0	-
Bosman	26.2051	86.4400	56.9696	29.5217	0	-
	$\alpha = 0.6507$			$\alpha = 0.8739$		

Table B.9: Batsmen Barr and Kantor weighting data (2 of 2)

	List A			Twenty20		
Name	AVE^t_{BWL}	$ECON^t_{BWL}$	BK^{LtA}_{BWL}	AVE^t_{BWL}	$ECON^t_{BWL}$	BK^{Tw20}_{BWL}
Clark	27.1333	4.3565	5.2308	60.0000	7.5000	9.2336
Johnson	31.4107	5.0546	6.0677	27.0000	6.0000	6.9738
Anderson	25.9494	4.6933	5.5686	34.7647	8.9320	10.2322
Broad	32.6829	5.2618	6.3161	15.0000	5.8696	6.4470
Kirtley	22.5396	4.6062	5.3989	24.6897	7.5767	8.5268
Schofield	25.9143	5.2027	6.1088	9.4762	5.8529	6.1419
Tremlett	24.5839	4.6084	5.4483	19.6842	7.0566	7.8189
Ongondo	28.1798	4.7425	5.6676	10.6667	4.5714	4.9756
Bond	22.1195	4.3048	5.0703	17.6667	7.3103	7.9847
Gillespie	30.9000	5.0535	6.0566	20.2500	7.4769	8.2602
Martin	26.9259	4.6813	5.5763	37.4000	8.9048	10.2789
Patel	36.6923	4.7589	5.8373	16.6667	6.1644	6.8090
J_Nel	39.1000	5.3993	6.5815	38.0000	7.6000	8.9271
M_Morkel	24.0313	4.5280	5.3505	24.5789	7.2031	8.1437
Edwards	29.1800	4.5881	5.5204	31.0000	6.3051	7.3936
Powell	25.6040	4.7624	5.6348	82.0000	8.0656	10.1708
Rampaul	24.7455	4.5266	5.3647	28.6667	7.8182	8.9029
Brent	34.3304	4.7724	5.8135	20.0000	10.0000	10.7177
Bracken	24.9020	4.2948	5.1200	34.8000	7.9091	9.1722
Lee	23.3463	4.5969	5.4080	72.0000	7.2000	9.0643
Razzak	20.0400	3.8530	4.5437	9.5000	4.7500	5.0909
Agarkar	26.8583	5.0025	5.9180	13.7143	6.1935	6.7060
H_Singh	30.8584	4.1445	5.0660	29.4444	6.3095	7.3603
RP_Singh	27.8876	4.9311	5.8640	99.0000	8.2500	10.5772
Sreesanth	37.1754	5.4543	6.6083	29.2500	8.3571	9.4725
Asif	33.5079	4.7121	5.7333	14.9583	6.4299	6.9963
Gul	28.4048	4.7577	5.6885	15.5385	6.1837	6.7805
A_Nel	24.5597	4.2142	5.0265	19.7500	5.6429	6.3960
Ntini	24.5714	4.3805	5.2049	37.7143	7.3333	8.6382
D_Fernando	27.2056	5.0342	5.9594	31.0000	8.1579	9.3230
Malinga	23.4851	4.7903	5.6157	32.8333	7.1205	8.2964
Vaas	26.3571	4.1509	4.9937	28.0000	7.0000	8.0409
Hilfenhaus	33.3750	4.0523	5.0035	13.5000	6.3529	6.8503
Tait	22.3434	5.0064	5.8142	12.8889	6.6923	7.1456
Mortaza	28.1238	4.5027	5.4079	37.0000	9.2500	10.6255
Rasel	26.2692	4.0790	4.9141	Inf	5.3750	-
Rahman	27.5000	4.6262	5.5288	NaN	NaN	-
Sidebottom	30.4539	4.2762	5.2037	20.5000	6.8144	7.6078
Odhiambo	53.6667	5.5200	6.9297	NaN	NaN	-
Otieno	Inf	7.7500	-	NaN	NaN	-
Varaiya	24.6087	3.5747	4.3354	48.0000	8.0000	9.5698
Rehman	26.9897	4.0842	4.9331	20.6667	6.4416	7.2380
Anjum	33.0395	4.5298	5.5255	27.6667	7.4700	8.5150

Table B.10: Bowler Barr and Kantor weighting data (1 of 2)

	List A			**Twenty20**		
Name	AVE^{t}_{BWL}	$ECON^{t}_{BWL}$	BK^{LtA}_{BWL}	AVE^{t}_{BWL}	$ECON^{t}_{BWL}$	BK^{Tw20}_{BWL}
Blain	28.4245	5.2659	6.2330	NaN	NaN	-
Drummond	40.0000	4.7059	5.8289	NaN	NaN	-
Lyons	58.9474	5.5309	7.0075	NaN	NaN	-
Tshabalala	28.8000	4.7682	5.7077	14.3548	6.5122	7.0478
Collins	24.4583	4.3017	5.1182	11.2500	4.6552	5.0846
Mpofu	27.7170	4.9130	5.8410	NaN	NaN	-
Mupariwa	30.9074	5.0988	6.1056	NaN	NaN	-
Vettori	31.5000	4.1186	5.0479	Inf	4.3333	-
Tanvir	32.4118	4.9640	5.9885	Inf	9.8261	-
Chawla	27.2766	4.7806	5.6900	19.4000	6.4667	7.2176
Akhtar	23.6498	4.6759	5.4988	21.0000	7.1591	7.9725
Muralitharan	22.2234	3.7536	4.4842	9.1765	5.3793	5.6744
	$\beta = 0.9$			$\beta = 0.9$		

Table B.11: Bowler Barr and Kantor weighting data (2 of 2)

	List A			**Twenty20**		
Name	$\overline{BT}$	SR^t_{BAT}	BK^{LtA}_{BAT}	$\overline{BT}$	SR^t_{BAT}	BK^{Tw20}_{BAT}
Clarke	29.3657	79.6200	50.0647	18.1818	134.2282	22.2054
Reza	20.9333	64.8900	38.3386	8.3333	131.5789	10.9813
Al_Hassan	26.4524	68.2500	43.9171	26.0000	92.8571	29.5296
Collingwood	25.5137	76.3100	45.8420	27.3333	136.6667	32.1063
Flintoff	23.7434	88.7500	48.0637	24.8000	163.1579	29.9411
Maddy	25.5774	54.8500	38.4651	29.7209	134.5263	34.5650
Snape	13.7122	68.2000	32.3394	13.7250	124.2081	17.1071
Wright	11.4648	97.8600	36.0951	14.9310	170.4724	19.0479
Kamande	8.7627	56.2700	23.6926	3.0000	42.8571	3.9139
C_Obuya	10.5181	67.3700	28.3999	12.7778	116.1616	15.9337
Odoyo	19.4685	70.9500	38.8795	13.5000	135.0000	16.9955
Onyango	5.5000	72.8600	21.9050	4.0000	66.6667	5.2996
Alam	30.6563	89.6500	54.4227	16.8000	112.5000	20.3186
Arafat	12.0519	78.6800	32.8749	8.2857	134.1040	10.9456
Brown	15.5510	66.3200	33.7799	8.6786	136.5169	11.4320
Kemp	21.3578	83.1200	44.1777	16.5417	125.6329	20.2596
Philander	15.4286	73.4700	35.5505	14.8750	128.6486	18.4566
van_der_Wath	15.4274	127.1400	47.6670	8.3636	121.0526	10.9258
Utseya	5.6310	56.6500	19.3572	7.0000	87.5000	9.0113
Chigumbura	20.9888	56.5500	35.6628	2.5000	50.0000	3.3732
Dabengwa	10.5405	68.9600	28.7850	7.0000	53.8462	8.5843
Symonds	25.9948	92.7800	51.3373	33.4118	201.4184	39.9871
S_Watson	21.5887	78.7300	43.1292	16.7500	118.5841	20.3712
Mascarenhas	16.6633	137.6100	51.5428	16.8333	120.2381	20.4907
J_Sharma	11.7381	86.2800	34.1156	9.1250	132.7273	11.9263
IK_Pathan	12.1743	78.0700	32.8926	6.8333	154.7170	9.3352
YK_Pathan	23.2895	112.3400	54.0343	20.0000	195.1220	25.1164
Oram	18.5202	83.1300	41.3461	22.4000	160.0000	27.2669
Styris	23.7192	79.2000	45.2029	25.3448	141.3462	30.0974
Malik	25.1512	79.3200	46.4900	34.6316	126.0536	39.4077
Hafeez	31.2411	58.1500	43.5556	28.9600	160.1770	34.3620
Afridi	22.7697	111.2200	53.1842	15.4167	162.2807	19.5082
JA_Morkel	14.6633	84.0300	37.3047	12.3077	131.1475	15.5932
Pollock	12.3881	86.6900	35.0705	19.3571	155.7471	23.8451
Dilshan	22.8220	80.2600	44.7162	1.5000	33.3333	2.0454
Jayasuriya	29.6532	91.0300	54.0267	21.0000	158.8235	25.7092
Maharoof	10.7835	77.4800	30.9621	8.6667	96.2963	11.0263
Bravo	14.0792	80.6800	35.8181	6.4000	103.2258	8.4516
Gayle	35.7478	80.8100	55.2974	20.8750	132.5397	25.1130
DR_Smith	15.5000	67.9900	34.1800	8.0000	114.2857	10.4371

Table B.12: All rounder Barr and Kantor batting weighting data (1 of 2)

	ListA			Twenty20		
Name	$\overline{BT}$	SR^t_{BAT}	BK^{LtA}_{BAT}	$\overline{BT}$	SR^t_{BAT}	BK^{Tw20}_{BAT}
Hogg	11.5189	78.6800	32.1904	15.9167	119.3700	19.4696
Kapali	20.8824	68.7800	39.5064	18.0000	120.0000	21.7602
Mahmuddullah	21.3500	62.1400	37.8058	5.5000	47.8261	6.8280
Bopara	22.0233	66.8100	39.8715	13.7667	121.9500	17.1224
Tikolo	30.2365	75.1000	49.1880	24.0000	114.2857	28.0536
Suji	7.8953	51.7900	21.5916	4.0000	51.8500	5.1681
Scott	7.7045	-	-	6.8750	0.0000	-
R_Watson	25.1698	74.9300	45.1111	NaN	-	-
Hamilton	16.5337	68.6900	35.4155	13.6667	157.6923	17.4533
Haq	13.8627	65.7200	31.8664	NaN	-	-
Maiden	8.8293	-	-	2.0000	0.0000	-
C_Wright	11.6545	74.4800	31.4303	NaN	-	-
H_Fernando	23.0000	77.6200	44.0824	21.4375	0.0000	-
Lokuarachchi	13.2841	80.1500	34.7399	11.0000	0.0000	-
Perera	17.2000	57.1400	32.6897	9.2000	0.0000	-
Sammy	14.4800	90.5800	38.6063	7.5000	0.0000	-
Maruma	3.2727	66.2100	16.3472	1.0000	0.0000	-
Williams	17.2857	73.6300	37.5241	38.0000	0.0000	-
Bhudia	0.0000	0.0000	-	4.5000	81.8182	6.0142
N_McCullum	9.8772	-	-	17.6250	163.9535	22.0287
Wijekoon	18.6531	-	-	13.8750	130.5882	17.3620
Chibhabha	23.0000	56.5500	37.2135	1.0000	40.0000	1.4461
	$\alpha = 0.5349$			$\alpha = 0.1$		

Table B.13: All rounder Barr and Kantor batting weighting data (2 of 2)

	List A			Twenty20		
Name	AVE^t_{BWL}	$ECON^t_{BWL}$	BK^{LtA}_{BWL}	AVE^t_{BWL}	$ECON^t_{BWL}$	BK^{Tw20}_{BWL}
Clarke	30.4211	5.0212	7.6614	56.0000	8.2963	10.0419
Reza	28.4524	4.5785	7.0277	22.2500	11.1250	11.9235
Al_Hassan	31.2381	4.0872	6.5856	31.0000	7.7500	8.9024
Collingwood	35.2674	4.8130	7.6787	13.6667	8.2000	8.6298
Flintoff	23.0344	4.1659	6.2215	22.1111	6.9017	7.7540
Maddy	29.2444	5.0599	7.6356	31.9091	7.7712	8.9502
Snape	29.1982	4.6405	7.1436	19.0278	6.5446	7.2817
Wright	39.5000	5.2220	8.3936	21.6296	8.0000	8.8366
Kamande	47.5217	4.8831	8.3267	21.0000	7.0000	7.8129
C_Obuya	43.9565	5.2474	8.6387	17.7000	8.8500	9.4852
Odoyo	28.2739	4.5857	7.0258	28.0000	7.0000	8.0409
Onyango	48.0000	5.9274	9.6810	40.0000	8.0000	9.3970
Alam	27.6429	4.9510	7.4109	11.3684	6.0561	6.4497
Arafat	24.4500	4.7766	7.0056	22.0286	8.3653	9.2158
Brown	26.9703	4.6263	6.9955	26.5652	7.6375	8.6514
Kemp	29.4400	4.7800	7.3215	18.2174	7.7593	8.4506
Philander	34.5000	4.7416	7.5523	33.8000	8.7414	10.0072
van_der_Wath	27.5676	4.8360	7.2742	25.4333	7.4926	8.4666
Utseya	49.5172	3.9550	7.1546	8.3333	6.2500	6.4324
Chigumbura	44.9070	6.3070	9.9948	22.0000	11.0000	11.7895
Dabengwa	24.2941	5.0315	7.2791	1.0000	1.0000	1.0000
Symonds	32.6912	4.7841	7.5086	24.5625	7.2331	8.1737
S_Watson	36.5900	5.0679	8.0573	18.8000	8.5455	9.2465
Mascarenhas	24.7764	4.2018	6.3705	18.3333	7.4492	8.1512
J_Sharma	26.8036	4.7752	7.1567	28.1667	7.0417	8.0888
IK_Pathan	26.1071	4.9033	7.2583	19.1333	7.1750	7.9144
YK_Pathan	47.1154	4.6548	8.0110	9.4444	7.7273	7.8839
Oram	31.2752	4.3818	6.9478	55.0000	7.8571	9.5450
Styris	29.0266	4.6011	7.0873	25.5926	7.5519	8.5322
Malik	29.4010	4.3541	6.8147	15.4000	7.0355	7.6089
Hafeez	33.3267	4.1710	6.7909	23.2273	6.4142	7.2951
Afridi	33.6907	4.5874	7.3227	20.7692	6.5854	7.3869
JA_Morkel	29.2768	4.6390	7.1463	32.1176	7.7082	8.8907
Pollock	22.5062	3.6643	5.6090	31.5000	6.0000	7.0822
Dilshan	38.7193	4.6957	7.7019	12.2500	5.1579	5.6239
Jayasuriya	35.8640	4.7531	7.6355	11.4211	7.0378	7.3870
Maharoof	27.4054	4.6836	7.0882	70.0000	7.3684	9.2288
Bravo	28.3578	5.1833	7.7218	19.5000	5.8500	6.5985
Gayle	30.5053	4.4562	6.9969	12.8889	5.9487	6.4269
DR_Smith	38.7667	4.8593	7.9088	19.2000	6.2609	7.0033

Table B.14: All rounder Barr and Kantor bowling weighting data (1 of 2)

	List A			Twenty20		
Name	AVE^t_{BWL}	$ECON^t_{BWL}$	BK^{LtA}_{BWL}	AVE^t_{BWL}	$ECON^t_{BWL}$	BK^{Tw20}_{BWL}
Hogg	28.3333	4.6514	7.1061	15.0952	7.5777	8.1183
Kapali	38.1042	4.5535	7.4946	NaN	NaN	-
Mahmuddullah	33.1818	4.4064	7.0752	NaN	NaN	-
Bopara	28.1250	5.3815	7.9315	24.7368	7.7901	8.7442
Tikolo	31.2301	4.6404	7.2571	NaN	NaN	-
Suji	49.5676	5.0733	8.6590	NaN	NaN	-
Scott	32.1887	4.5676	7.2208	52.5000	9.1971	10.9471
R_Watson	36.6615	5.3591	8.4132	NaN	NaN	-
Hamilton	25.2109	4.6950	6.9637	NaN	NaN	-
Haq	35.8824	5.4276	8.4528	NaN	NaN	-
Maiden	43.6087	4.8689	8.1424	14.5000	7.2500	7.7704
C_Wright	25.7969	4.4342	6.7017	NaN	NaN	-
H_Fernando	28.5714	4.3689	6.7867	19.7333	7.7895	8.5483
Lokuarachchi	22.9127	4.3004	6.3668	11.8462	5.6687	6.1023
Perera	27.9545	3.8863	6.1733	27.5714	6.8929	7.9178
Sammy	26.2203	4.3721	6.6551	17.1667	6.8667	7.5256
Maruma	20.5000	3.1654	4.9060	31.0000	7.7500	8.9024
Williams	40.9524	4.4216	7.4527	28.0000	7.0000	8.0409
Bhudia	Inf	14.0000	-	Inf	9.1500	-
N_McCullum	39.4091	4.2709	7.1925	12.2727	5.9559	6.4024
Wijekoon	25.8316	3.7342	5.8777	18.0000	7.3636	8.0521
Chibhabha	55.6364	5.5636	9.5478	14.0000	7.6364	8.1135
	$\beta = 0.7655$			$\beta = 0.9$		

Table B.15: All rounder Barr and Kantor bowling weighting data (2 of 2)

	List A			Twenty20		
Name	$\overline{BT}$	SR^t_{BAT}	BK^{LtA}_{BAT}	$\overline{BT}$	SR^t_{BAT}	BK^{Tw20}_{BAT}
Ouma	15.8649	58.69	47.1167	NaN	-	-
Hopkins	16.0885	56.94	46.0528	12.7000	-	-
C_Smith	17.9457	66.07	53.0841	NaN	-	-
Taibu	19.6917	64.3	52.7135	NaN	-	-
Haddin	27.8393	91.44	74.8891	15.1429	124.7059	18.6972
Rahim	15.0278	56.06	44.9422	1.0000	42.8571	1.4561
Prior	22.1837	76.03	61.8252	22.8462	139.4366	27.3760
D_Obuya	17.3833	61.48	49.7305	13.5000	122.7273	16.8343
Ramdin	16.8983	79.25	61.1380	10.7500	195.4545	14.3672
B_Taylor	24.0115	62.99	53.5730	15.0000	107.1429	18.2591
Gilchrist	31.9222	96.94	80.4462	16.2500	135.4167	20.0879
Dhoni	32.0956	91.3	76.6018	24.6000	180.8824	30.0318
Karthik	18.4697	69.76	55.8088	16.7273	115.0000	20.2839
BB_McCullum	14.4133	90.64	66.5643	25.8125	175.0000	31.2573
Akmal	19.0217	83.55	65.1683	17.1176	129.9107	20.9635
Boucher	16.8835	83.93	64.1182	24.0000	108.5427	27.9094
de_Villiers	33.6143	83.56	71.7129	22.3636	160.7843	27.2403
Sangakkara	30.9363	74.42	64.2217	34.1667	173.7288	40.2003

Table B.16: Wicket keeper Barr and Kantor weighting data

Appendix C

Team Selection Data

The data provided in Tables C.1 and C.2 are the data on which the various integer programs were run.

The data given in Table C.3 to Table C.4, provide additional output of the integer optimisation procedure discussed in Section 4.4.1.

#	Name	Price	Type	Fantasy Score				Performance Measure				Adjusted Performance Measure			
				Stage 1	Stage 2	Stage 3	Stage 4	Stage 1	Stage 2	Stage 3	Stage 4	Stage 1	Stage 2	Stage 3	Stage 4
1	Bracken	85000	BWL	0	0	0	0	1.682902	1.682902	1.682902	1.682902	0	0	0	0
2	Lee	115000	BWL	214	0	0	0	1.37013	0.961993	0.961993	0.961993	1.37013	0	0	0
3	Noffke	75000	BWL	24	0	0	0	1.40005	1.230091	1.230091	1.230091	1.40005	0	0	0
4	Zoysa	75000	BWL	25	41	-10	0	1.770384	1.279612	1.135805	0.88757	1.770384	1.279612	1.135805	0
5	Ojha	50000	BWL	80	116	30	86	1.427098	0.719775	0.559174	0.329376	1.427098	0.719775	0.559174	0.329376
6	McGrath	85000	BWL	165	235	190	-16	1.380274	1.062229	0.995789	0.727692	1.380274	1.062229	0.995789	0.727692
7	RP_Singh	115000	BWL	187	209	125	28	1.84228	1.218316	1.260056	0.984299	1.84228	1.218316	1.260056	0.984299
8	Kartik	85000	BWL	153	42	-10	0	1.361334	0.622776	0.746842	0.593644	1.361334	0.622776	0.746842	0
9	Vettori	100000	BWL	25	45	0	0	1.694573	1.591921	1.479506	1.479506	1.694573	1.591921	0	0
10	Tanvir	75000	BWL	10	389	215	374	1.612676	0.958738	1.528153	2.072008	1.612676	0.958738	1.528153	2.072008
11	Akhtar	85000	BWL	0	0	207	0	1.358459	1.358459	1.358459	1.201273	0	0	1.358459	0
12	D_Fernando	75000	BWL	0	0	50	322	1.836954	1.673565	1.673565	1.535952	1.836954	0	1.673565	1.535952
13	Malinga	90000	BWL	0	0	0	0	1.465829	1.465829	1.465829	1.465829	0	0	0	0
14	Gony	50000	BWL	230	115	40	326	1.466941	1.372183	0.804291	0.532364	1.466941	1.372183	0.804291	0.532364
15	Amarnath	50000	BWL	135	-20	20	0	1.346627	0.789744	0.440445	0.402146	1.346627	0.789744	0.440445	0
16	Hayden	90000	BAT	331	0	0	0	1.009651	1.093769	1.093769	1.093769	1.009651	0	0	0
17	Hodge	75000	BAT	0	57	0	0	0.94072	0.94072	0.875399	0.875399	0	0.94072	0	0
18	M_Tiwary	95000	BAT	59	31	28	51	1.045437	0.956172	0.896159	0.839216	1.045437	0.956172	0.896159	0.839216
19	Y_Singh	120000	BAT	284	89	76	288	1.016843	1.07127	1.011679	0.952263	1.016843	1.07127	1.011679	0.952263
20	Sehwag	105000	BAT	254	364	222	68	0.916924	0.998154	1.108229	1.067212	0.916924	0.998154	1.108229	1.067212
21	Goel	50000	BAT	67	21	16	0	1.147537	0.980824	0.90125	0.870713	1.147537	0.980824	0.90125	0
22	Marsh	55000	BAT	0	324	261	678	0.930429	0.930429	1.076518	1.149172	0	0.930429	1.076518	1.149172
23	D_Hussey	90000	BAT	148	267	181	2	1.394518	1.290114	1.338489	1.211618	1.394518	1.290114	1.338489	1.211618
24	Ganguly	120000	BAT	110	210	391	148	0.917094	0.855342	0.912493	1.005249	0.917094	0.855342	0.912493	1.005249
25	Tendulkar	125000	BAT	0	0	87	205	1.097413	1.097413	1.097413	1.035189	0	0	1.097413	1.035189

Table C.1: Data used in the integer programming problem (1 of 2)

#	Name	Price	Type	Fantasy Score				Performance Measure				Adjusted Performance Measure			
				Stage 1	Stage 2	Stage 3	Stage 4	Stage 1	Stage 2	Stage 3	Stage 4	Stage 1	Stage 2	Stage 3	Stage 4
26	R_Taylor	75000	BAT	226	93	0	0	0.995316	1.04896	1.078682	1.078682	0.995316	1.04896	0	0
27	Lehmann	65000	BAT	18	0	0	0	1.433942	1.335278	1.335278	1.335278	1.433942	0	0	0
28	G_Smith	95000	BAT	179	97	217	184	1.046893	1.103191	1.007947	1.051667	1.046893	1.103191	1.007947	1.051667
29	Ravi_Teja	50000	BAT	0	6	89	89	0.922156	1.022156	0.928453	0.83961	0	1.022156	0.928453	0.83961
30	S_Tiwary	50000	BAT	19	24	0	0	1.352419	1.146374	0.95988	0.95988	1.352419	1.146374	0	0
31	S_Watson	75000	ALR	466	275	341	373	1.873803	1.992532	1.87668	1.895829	1.873803	1.992532	1.87668	1.895829
32	Kallis	120000	ALR	67	167	27	81	1.297772	0.704759	0.754502	0.743012	1.297772	0.704759	0.754502	0.743012
33	P_Kumar	75000	ALR	128	303	2	61	1.951165	1.022648	1.279206	0.425125	1.951165	1.022648	1.279206	0.425125
34	Mascarenhas	80000	ALR	0	0	71	0	1.437075	1.142939	1.142939	2.086285	0	0	1.142939	0
35	J_Sharma	75000	ALR	215	-17	25	0	1.423536	1.024719	0.337036	0.241287	1.423536	1.024719	0.337036	0
36	Hopes	80000	ALR	173	18	93	163	1.519513	1.652251	1.116652	0.425125	1.519513	1.652251	1.116652	0.425125
37	Nayar	50000	ALR	216	29	97	87	1.190356	0.827271	0.348526	0.482575	1.190356	0.827271	0.348526	0.482575
38	Thornely	55000	ALR	0	84	114	0	1.254747	1.074739	0.99048	0.758332	1.254747	1.074739	0.99048	0
39	Malik	95000	ALR	49	72	56	0	1.434838	1.121416	0.463425	0.409806	1.434838	1.121416	0.463425	0
40	Hafeez	75000	ALR	118	30	15	0	1.506341	0.750606	0.566834	0.180008	1.506341	0.750606	0.566834	0
41	Afridi	110000	ALR	169	151	152	26	1.619187	1.341888	0.808121	1.141328	1.619187	1.341888	0.808121	1.141328
42	Jadeja	50000	ALR	147	65	21	84	1.528625	0.811896	0.363846	0.04213	1.528625	0.811896	0.363846	0.04213
43	Pollock	95000	ALR	223	259	270	56	2.264098	0.854081	1.459214	2.294145	2.264098	0.854081	1.459214	2.294145
44	Bravo	75000	ALR	241	73	379	0	1.768118	1.086409	1.026429	1.731141	1.768118	1.086409	1.026429	0
45	Khote	50000	ALR	56	0	0	0	1.39419	0.627856	0.627856	0.627856	1.39419	0	0	0
46	Gilchrist	110000	WKR	384	237	111	167	0.983242	1.124971	1.131672	1.082106	0.983242	1.124971	1.131672	1.082106
47	Dhoni	150000	WKR	347	141	173	173	0.836072	0.969757	0.9276	0.944436	0.836072	0.969757	0.9276	0.944436
48	BB_McCullum	105000	WKR	423	0	0	0	1.037883	1.1718	1.1718	1.1718	1.037883	0	0	0
49	de_Villiers	85000	WKR	0	94	98	0	0.885068	0.947764	0.906756	0.839127	0	0.947764	0.906756	0
50	Takawale	50000	WKR	0	103	110	96	0.885773	0.885773	0.907686	0.859551	0	0.885773	0.907686	0.859551

Table C.2: Data used in the integer programming problem (2 of 2)

Stage 1			Stage 2			Stage 3			Stage 4		
Name	*Meas*	*Fan. Pts*	*Name*	*Meas*	*Fan. Pts*	*Name*	*Meas*	*Fan. Pts*	*Name*	*Meas*	*Fan. Pts*
Batsmen			**Batsmen**			**Batsmen**			**Batsmen**		
D Hussey	1.39452	148	D Hussey	1.29011	267	D Hussey	1.33849	181	D Hussey	1.21162	2
D Lehmann	1.43394	18	D Lehmann	1.33528	0	D Lehmann	1.33528	0	D Lehmann	1.33528	0
S Tiwary	1.35242	19	S Tiwary	1.14637	24	S Tiwary	0.95988	0	M Hayden	1.09377	0
K Goel	1.14754	67	K Goel	0.98082	21	V Sehwag	1.10823	222	S Marsh	1.14917	678
Bowlers			**Bowlers**			**Bowlers**			**Bowlers**		
D Zoysa	1.77038	25	D Zoysa	1.27961	41	S Tanvir	1.52815	215	S Tanvir	2.07201	374
RP Singh	1.84228	187	N Bracken	1.68290	0	N Bracken	1.68290	0	N Bracken	1.68290	0
D Fernando	1.83695	0	D Fernando	1.67357	0	D Fernando	1.67357	50	D Fernando	1.53595	322
			D Vettori	1.59192	45	D Vettori	1.47951	0			
All Rounders			**All Rounders**			**All Rounders**			**All Rounders**		
S Watson	1.87380	466	S Watson	1.99253	275	S Watson	1.87668	341	S Watson	1.89583	373
S Pollock	2.26410	223	J Hopes	1.65225	18	S Pollock	1.45921	270	S Pollock	2.29415	56
P Kumar	1.94117	128							D Mascarenhas	2.08629	0
Wicket Keeper			**Wicket Keeper**			**Wicket Keeper**			**Wicket Keeper**		
B McCullum	1.03788	423	B McCullum	1.1718	0	B McCullum	1.1718	0	B McCullum	1.1718	0
TOTAL	17.9050	1 704	**TOTAL**	15.7972	691	**TOTAL**	15.6137	1 279	**TOTAL**		1 805
Changes		-	*Changes*		3	*Changes*		3	*Changes*		3

Table C.3: Solution of integer programming problem: The prospective approach using $\underline{M}$, including optimised values of the performance measures

Stage 1			Stage 2			Stage 3			Stage 4		
Name	*Meas*	*Fan. Pts*	*Name*	*Meas*	*Fan. Pts*	*Name*	*Meas*	*Fan. Pts*	*Name*	*Meas*	*Fan. Pts*
Batsmen			**Batsmen**			**Batsmen**			**Batsmen**		
D Hussey	1.39452	148	D Hussey	1.29011	267	D Hussey	1.33849	181	D Hussey	1.21162	2
D Lehmann	1.43394	18	G Smith	1.10319	97	G Smith	1.00795	217	G Smith	1.05167	184
S Tiwary	1.35242	19	S Tiwary	1.14637	24	V Sehwag	1.10823	222	V Sehwag	1.06721	68
K Goel	1.14754	67	K Goel	0.98082	21	K Goel	0.90125	16	S Marsh	1.14917	678
Bowlers			**Bowlers**			**Bowlers**			**Bowlers**		
D Zoysa	1.77038	25	D Zoysa	1.27961	41	S Tanvir	1.52815	215	S Tanvir	2.07201	374
RP Singh	1.84228	187	RP Singh	1.21832	209	RP Singh	1.26006	125	RP Singh	0.98430	28
D Fernando	1.69457	0	D Vettori	1.59192	45	D Fernando	1.67357	50	D Fernando	1.53595	322
All Rounders			**All Rounders**			**All Rounders**			**All Rounders**		
S Watson	1.87380	466	S Watson	1.99253	275	S Watson	1.87668	341	S Watson	1.89583	373
S Pollock	2.26410	223	S Pollock	0.85408	259	S Pollock	1.45921	270	S Pollock	2.29415	56
P Kumar	1.95117	128	P Kumar	1.02265	303	P Kumar	1.27921	2	S Afridi	1.14133	26
Wicket Keeper			**Wicket Keeper**			**Wicket Keeper**			**Wicket Keeper**		
B McCullum	1.03788	423	A Gilchrist	1.12497	237	A Gilchrist	1.13167	111	A Gilchrist	1.08211	167
TOTAL	17.9050	1 704	**TOTAL**	13.6046	1 778	**TOTAL**	14.5345	1 750	**TOTAL**	15.4853	2 278
Changes		-	*Changes*		3	*Changes*		3	*Changes*		2

Table C.4: Solution of integer programming problem: The prospective approach using $\underline{R}$, including optimised values of the performance measures

Appendix D

Additional Discussion of Performance Measures

D.1 Batsmen

For BK_{BAT}^{LtA} an α value of 0.6507 is used and for BK_{BAT}^{Tw20} an α value of 0.8739 is used. Worth noting is that α represents the weight of the strike rate, and this weight is lower for the ListA format of the game than for the Twenty20 format. This value agrees with the popular belief (Sharp et. al., 2009) that the strike rate is the most important aspect to consider when selecting batsmen for Twenty20 cricket. Also this shows that the strike rate is an important statistic in the evaluation of batsmen in ListA cricket.

D.2 Bowlers

The Barr and Kantor bowling measures (BK_{BWL}^{LtA} and BK_{BWL}^{Tw20}) are calculated by selecting the value of β which would minimise value of the correlation coefficient. For BK_{BWL}^{LtA} the value of $\hat{\rho} = -0.1856$. This value is achieved by setting $\beta = 0.9$ according to the restrictions discussed in Section 3.3.1. The correlation for BK_{BWL}^{Tw20} is $\hat{\rho} = -0.01028$, and is reached by setting $\beta = 0.9$. This value of β provides support to the common belief that a bowler's most important measure is their economy rate.

D.3 All Rounders

The Barr and Kantor measures BK_{BAT} and BK_{BWL} for which the correlations are optimised, by changing the values of α and β respectively, provided relatively good results. The values of α and β are provided in Table D.1.

Measure	**Optimised $\hat{\rho}$**	α	β
BK_{BAT}^{LtA}	0.427573	0.5348598	-
BK_{BAT}^{Tw20}	0.338325	0.1	-
BK_{BWL}^{LtA}	-0.22538	-	0.765454
BK_{BWL}^{Tw20}	-0.21575	-	0.9

Table D.1: All Rounders: Values of α and β for measures BK and BK_{BWL}

The values given for β in the above table indicate that the economy rate is the more important measure for a bowler (which is similar to the results given in Section 4.1.2). The values of α, however, indicate that for all rounders the strike rate is only slightly more important than the mean batting score for ListA matches and of very little importance in the Twenty20 format of the game. This is contrary to the belief that the strike rate of a batsman is the main measure of concern for Twenty20 matches. A simple explanation for this discrepancy could be that the amount of data available is limited, especially for Twenty20 cricket, and thus this should be investigated again when more data is available.

D.4 Wicket Keepers

The Barr and Kantor BK_{BAT}^{LtA} measure resulted in a correlation estimate of $\hat{\rho} = 0.57816$. Since this is a ListA measure the BK_{BAT}^{Tw20} measure needs to be included into consideration. An interesting observation is that when the estimated correlation was optimised for BK_{BAT}^{LtA} an α value of 0.832098 resulted. When a similar optimisation was conducted for BK_{BAT}^{Tw20} it resulted in $\alpha = 0.1$. This value is inconsistent with the belief that strike rate is more important than the mean batting score in the Twenty20 game. Again, this can be explained by the lack of sufficient Twenty20 data. The α value for BK_{BAT}^{LtA} is what is expected, and indicates that once a format is established (as ListA cricket is) the results tend to agree with the popular belief.

In order to combine the BK_{BAT}^{LtA} and the BK_{BAT}^{Tw20} measures the methodology described for the batsmen category in Section 4.3.1 is implemented. That is,

$$BK_{BAT}^{com} = \left(BK_{BAT}^{Tw20}\right)^{\gamma} \left(BK_{BAT}^{LtA}\right)^{1-\gamma}$$

where $0 \leq \gamma \leq 1$. It is then desired to optimise the linear correlation between BK_{BAT}^{com} and the total fantasy league score in the 2007 ICC World Twenty20 by changing the value of γ. *Solver* is used to perform this optimisation and the value of γ which optimised the correlation is $\gamma = 0$ and the resulting correlation is $\hat{\rho} = 0.462992$. As such, this formulation was not considered as a performance measure for wicket keepers.