Lancaster University

Faculty of Science and Technology

Math 492: Statistics Dissertation

MoneyBowl: How to win the Indian Premier League

A simulation based approach to player evaluation and team optimisation in Twenty20 Cricket

Author:
Matthew J. Triggs

Fylde College
Lancaster University
triggs52@gmail.com

Supervisor:
Dr. Andrew Titman

Dep. of Mathematics and Statistics
Lancaster University
a.titman@lancaster.ac.uk

May 2, 2015

Acknowledgements

In making this dissertation a reality, I would like to thank the following for their efforts and contributions:

- First and foremost, Dr. Andrew Titman for valuable advice and guidance on this dissertation. This certainly would not have been possible without his assistance, and I could not express how thankful I am in this small space.
- ESPN for making the player data available on their Cricinfo website and the IPL for publishing the auction results and player salaries.
- Daniel Musson, England and Wales Cricket Board, and James Buttler, Editor of Cricket Badger and former Media Manager of Yorkshire CCC, for taking the time to speak with me.
- My family. Especially my Mother for my love of mathematics, and my Father for my love of sports. I do not say it enough, but I appreciate all that you do for me - whether personal, pastoral, financial, or just being good role models.
- The friends I have neglected whilst I was working on this dissertation and were willing to listen to me drone on about writing it. In particular, Michael Gladstone, for knowing far more about cricket strategy than me and filling in the gaps in my knowledge, and George Bedford, for proofreading and correcting my many grammatical mistakes.
- Finally, I should like to thank Caroline Arnold for her support. "Not everything that can be counted counts, and not everything that counts can be counted".

Declaration

I, Matthew J. Triggs, declare that this submission is my own work. This dissertation is submitted to fulfil the requirements of a Master of Science (MSci) degree in Mathematics and Statistics at Lancaster University. I have not submitted it in substantially the same form towards the award of another degree or other qualification. It has not been written or composed by any other person and all sources have been appropriately referenced or acknowledged.

Matthew J. Triggs

Abstract

This dissertation uses simulation-based methods to evaluate cricket players for the Indian Premier League T20 competition prior to the 2015 season. Using techniques such as Expectation-Maximisation, Simulation and a Cumulative Probit Model for ordinal data, we consider the batting and bowling ability of a player independently, and then combine them into a metric called a WAVE (Wins Above Average) score, which is similar in concept to existing work in baseball. Data is used from the first six seasons of the tournament, and the seventh season acts as a test. We finally consider the problem of assembling a team within constraints ahead of the 2015 season and produce optimal hypothetical solutions.

Keywords: Cricket, Indian Premier League, T20, Simulation, Team Optimisation, EM Algorithm, Culmulative Probit Model for Ordinal Data, Sports

Contents

Chapter 1

Introduction

In all aspects of life, business or personal, it is economically rational to try to obtain the highest value product for the lowest possible costs. This especially holds true in professional sports, where stars are paid exorbitant sums of money to entertain millions, inspire countries and, most importantly, win. In order to promote parity between teams who could not necessarily compete financially, many leagues apply a "salary cap", setting a limit on what a team can pay players. This is common practice in American sporting leagues (the NFL, NBA and MLB all have salary caps of various amounts). The world's foremost Twenty20 (T20) cricket league, the Indian Premier League, also operates in this way. With a salary cap, it is important that teams optimise their on field performance with a limited budget - it is not possible to outspend the competition in order to attract the best players.

Sabermetrics, defined as "the search for objective knowledge about baseball" (SABR, 2014), is perhaps the best known field of sports statistics, with many years of research on how to optimise a baseball team within budget constraints/a salary cap. I believe that much of this theory can be applied to the game of cricket. Cricket has a wealth of statistical information (strike rate, economy, run rate - to name a few), but is suffering from what baseball has done and continues to do so - not being able to find the signal from the noise. This is not to say that Batting Average in cricket deserves the same disdain as Bill James suggested it does in baseball (Lewis, 2003, p.67), but by its definition[1], it can artificially inflate the ability of a cricketer who scores slowly and does not get out often. Criticism of Batting Average existed at least as early as 1945, when Sir William Elderton said in his paper *Cricket Scores and Some Skew Correlation Distributions* "I had to desert the old idea that a 'not out' innings had not been completed, which must, I think, be regarded as a pleasant fiction" (Elderton, 1945). This slow rate of scoring is no good in the Twenty20 format of the game. In addition, a statistic such as the number of half centuries which is highly important to test and one day cricket is, compared to Strike Rate and Batting Average, irrelevant in T20 cricket. A batsman that faces 6 bowls and is out for 30 is much more valuable (usually) than one who scores a half-century but faces 60 balls.

For all the money a player may generate in ticket sales, merchandise and sponsors, this dissertation will look solely at on-field success. For this purpose, the most important attribute of a player is how much they improve a team's chances of winning. In speaking to James Buttler (2015), Editor of Cricket Badger, he referenced a quote by Martyn Moxon, Yorkshire CC Director of Cricket on the signing of Herschelle Gibbs: "It doesn't matter how many runs he gets, how many games he plays - just how many wins he gets us".

The aim of this dissertation is to apply some of the same ideas from various other sports to T20 cricket to develop a statistic to measure the efforts of an individual player within a team

[1]Batting Avergae $= \frac{\text{Total Number of Runs}}{\text{Total of Number of Times Out}}$

performance, determine their influence on the number of wins that they contribute and provide a framework for player valuation. From this, it is hoped that teams can assemble the best possible team within budget constraints and so, this dissertation is aptly called "MoneyBowl: How to win the Indian Premier League".

1.1 The Game of Cricket

Cricket is an outdoor game played between two teams of 11 players on an oval/circular field, marked by a boundary. In the middle of the field is a rectangular pitch 22 yards long and 10 feet wide. The game is split into two innings of a designated number of overs, where the batting team will attempt to score a large number of runs and the fielding team of 11 players will attempt to restrict this. A team is made up of specialist batsmen (all players may bat in a game), bowlers (players who specialise in bowling), all-rounders (players that can both bat and bowl to high standard) and a wicketkeeper (specialist position that stands opposite the bowler, behind the wickets). The batting team bat in pairs, with one player "on strike" facing the bowler and one player at the bowler's end. A bowler will bowl six balls (excluding no balls), known as an over, from one end of the pitch, at which point another bowler will bowl from the opposite end. A batsman can be out in various ways, and the innings will continue until a designated number of overs has concluded, a total for victory has been reached or all of the batting team are out. More information, including the full laws of the game, is available in Marylebone Cricket Club (2013). A diagram of the cricket field is shown in the Appendix in Figure A.1.1.

Originally, cricket took place over a period of several days and each team had two innings each. However, even the five days allocated for full test matches frequently resulted in draws and spectator interest waned. As a result, a one day form of the game was formed, where each team is restricted to one innings of 50 overs. Twenty20 cricket is a further restricted format, whereby each team only has a maximum of 20 overs batting. This format prioritises high rates of scoring compared to One Day and First Class cricket (Norman and Clarke, 2010), which attracts record crowd numbers which translate into financial security. In England, Yorkshire CC director of cricket Martyn Moxon has said "T20 has been the saviour of cricket in some ways. I don't think there's any doubt in my mind that T20 is funding the county game." (Ramprakash, 2012). The money available in the short form of the game has deterred some of the top players (e.g. Chris Gayle, West Indies/Royal Chargers Bangalore) from playing test cricket in order to play T20 (Ramprakash, 2012). The most popular and most lucrative cricket league in the world is a T20 competition - the Indian Premier League.

In addition, in T20 cricket the first six overs of an innings are "Powerplay" overs - where the fielding team are only permitted to have two fielders outside of a certain distance from the playing wicket. This is to entice the opening (and usually best batsmen) to score runs quickly to make for an entertaining start to the game.

1.2 The Indian Premier League

The inaugural Indian Premier League took place in 2008 with 8 teams competing. Since then, the number of franchises (teams) increased to 11 in 2011 before returning to 8 by the start of the 2014 season. At the original franchise auction, the franchises were sold for $723.59 Million to Indian businesses and celebrities (Cricinfo Staff, 2008). The 2014 format featured 8 teams competing in a double round-robin tournament and took place in venues across India and the UAE between April and June. Each team plays every other team at their respective venues, with some teams sacrificing home games to allow matches to be played in the UAE. The competition concludes with

Players Retained	Cost
Capped	
1st	Rs. 12.5 crore
2nd	Rs. 9.5 crore
3rd	Rs. 7.5 crore
4th	Rs. 5.5 crore
5th	Rs. 4 crore
Uncapped	
1st	Rs. 4 crore

Players Retained	"Rights to Match"
5	1
4	1
3	1
2	2
1	2
0	3

Table 1.1: Players Retained, Salary & "Rights to Match"

a Page-McIntyre playoff system - a knockout system which gives preference to higher seeds.

All teams have a salary cap of 60 crore[2] Indian Rupees (Rs), which is approximately equal to \$9.66 Million USD. Each team is required to spend Rs. 36 crore per year. Both of these figures will increase by 5% yearly in 2015 and 2016. Each squad is allowed between 16 and 27 players, of which up to 9 may be overseas players (despite only four being allowed to appear in a single match). The main IPL auction takes place every three years with teams allowed to retain up to five players from the previous year at a set cost, one of which may be an uncapped player. Players are given one-year contracts with teams allowed to renew contracts for up to three years. Any player released from the contract at the end of the year enters into that years supplementary auction. The supplementary IPL Auction takes place yearly. A player may only play in the IPL if they have entered into the auction at the start of the year. Teams are also allowed a certain number of "Rights to Match", whereby a franchise is allowed to match the final bid for a player who played for them in the preceding season (IPL Desk in Mumbai, 2013). Table 1.1 details the number of Rights to Match a team may have, depending on the number of retained players.

1.3 Aims and Method Outline

Our aim is to find a way of valuing players based on on-field performance which can be used to build a winning team. A team can only recruit new players through the auction or via trades, so our aim should also be to build the best team possible with budget constraints. The most valuable attribute to a team is the number of wins a player can contribute as part of the team, so the valuation of the player will be tied to their Wins Above Average (or $WAVE$) score, which is explained in more detail below.

To evaluate the abilities of players, we will be making heavy use of simulated games, using an Expectation-Maximisation algorithm and also a Cumulative Probit Model for ordinal data to model how runs are scored/conceded by batsmen/bowlers respectively. We will then look into one method of optimisation using Microsoft Excel.

We will begin by finding the distribution of losing scores in every IPL match[3]. We will find the distributions for the number of balls faced in an innings for batsmen across the first six years of the IPL and use this, together with the runs scored data, to find a distribution for the "average" batsman. We will then find individual parameters for each cricketer to have played in the IPL. We will simulate how many games a team would win if their batting lineup consisted of 11 of that player (using the distribution of the losing scores), from which we can develop their Batting Score

[2]One crore is equal to ten million in the South Asian numbering system.

[3]We assume the margin of victory to be irrelevant. This may play a factor in determining tie breaks in the league, but we will just focus on winning matches in this regard. The winning score then simply needs to be treated as one more than the losing score.

($BATS$). We will then carry out a similar procedure for players who are deemed to be bowlers or all-rounders and find their Bowling Score ($BOWLS$). A player's $WAVE$ score will be a combination of their $BATS$ and $BOWLS$. The exact nature of this combination will be discussed in Chapter 6. Once we have a player's $WAVE$ score, we will examine the prices they sold for in the most recent (2015) auction, and provide examples of how $WAVE$ can be used to build a team within various constraints.

All of the data used in the modelling section of this dissertation is taken from the last seven years of data (since the IPL inaugural season) on the ESPN Cricinfo website (ESPN, 2015). Player salary data is taken from the IPL website (IPL, BCCI, 2015a).

Chapter 2

Literature Review

2.1 Inspiration

The main motivation for this dissertation comes from baseball; in particular *Moneyball* by Lewis (2003). The book is a non-fiction account of the 2001 and 2002 Major League Baseball seasons. Billy Beane, General Manager of the Oakland Athletics, is in charge of assembling their squad. He is highly critical of traditional scouting methods (he himself being projected to be a great player as an 18 year old, but never reaching that potential), and favours a statistical approach. He prioritises a statistic called On-Base Percentage (the percentage a batsman manages to get to at least first base out of every opportunity), which he thinks is neglected by other teams, who prioritise other statistics like Home Runs, Runs Batted In (RBI) or Batting Average. This enables him to find players that he believes are undervalued.

Oakland is not a popular market so, in addition to the salary cap, Beane must negotiate additional financial restrictions imposed by the owner. As a result, it is a priority for Oakland to sign undervalued talent. The book details Beane's trades, drafting strategy, unorthodox personnel management and the 2002 season where the Oakland A's set the American League record for consecutive wins with 20. Arguably more remarkably, Oakland managed to compete with, and almost defeat, the New York Yankees in the playoffs, who had spent over $125 Million compared to Oakland's $41 Million - the league's third lowest. They finished the season ranked 5th. It was turned into a film in 2011 starring Brad Pitt as Billy Beane, which was nominated for six Academy Awards.

The book also mentions the ground-breaking work by Bill James in his *Baseball Abstracts* in developing Sabermetrics. James pioneered the field, naming it after the Society for American Baseball Research. Many of the statistics used in modern day reporting and analysis were invented by him, including (and by no means limited to) Runs Created, Game Score and Pythagorean Winning Percentage. This book (and by default Billy Beane, Bill James and the Oakland A's) serve as an inspiration for this work, where I am asking the same questions (and hopefully finding some answers) but for a different sport. It is a tribute to this way of looking at sports that this dissertation is titled "MoneyBowl".

2.2 Statistical Modelling of Baseball

Sean Smith (founder of www.baseballprojection.com) developed a method of evaluating players called Wins Above Replacement, or WAR. WAR is an attempt to measure the value[1] of a player

[1] Value, in this sense, means a player's ability for hitting, pitching, fielding and base running - everything they could potentially do on the field to affect the outcome of the game.

(in wins) that they bring to a team above a player that could be signed off free agency or traded for cheaply. Alternatively, it could be thought of as the value a team would lose if a player got injured and a replacement had to be signed (Fangraphs, 2015). This statistic or idea was quickly adopted by much larger sabermetric sites, all of which have their own variations. At the core, however, they are all very similar. WAR goes into tremendous depth, for example, factoring in the size of certain grounds, whether a field is a "hitters park", etc. In the Indian Premier League, it does not make sense to think of "replacement" players, as the game lacks the depth (in terms of players, leagues, etc.) that baseball does. We will also be making assumptions about the field of play and wicket conditions, that will be discussed in Chapter 8. We hope, however, to find a similar metric to compare players to the average performances in the league.

The concept of Wins Above Average was introduced by Darowski (2012) as a potential alternative to WAR. He criticises WAR for measuring the very best "hall of fame" players, because the score couldn't match up to exactly how much better they were than everyone else in their era. Darowski does use Wins Above Average in his article to compare several players across the history of baseball, but it has seen little or no mainstream use (in player evaluation or in sports media).

2.3 Statistical Modelling of Cricket

Twenty20 cricket as a sport is, compared to the game of cricket itself, extremely young. As a result, much of the work done (including everything before 2003) is not directly applicable to T20 and the IPL, but can be used as a base for our work.

Elderton and Elderton (1912) first appear to suggest a Geometric Progression for modelling batting scores, but the idea is not formally stated until much later (Elderton, 1945). They also noticed that the frequency of a low total is greater than expected. It is also noted by Wood (1945) who, whilst agreeing that a geometric model is a reasonable fit, states "The series show discrepancies at each end, and particularly at the commencement". This is also commented on by Langdale (1945), who brought attention to the unusually high number of innings concluded before the fifth bowl is recorded. Langdale agrees with Elderton's criticism of the Batting Average statistic, and suggests that, with a large enough sample, "not out" batsmen's scores at the end of an innings could also follow a geometric progression.

Perera and Swartz (2013) considered the differences between T20 and One-Day cricket, and how the Duckworth-Lewis method may not be appropriate for the former. They consider overs and wickets as resources and investigate the change in resources over the course of the game. We will be extending this, and seeing how the change in resources affects the rate of runs scored and wickets.

After the first IPL season, Parker et al. (2008) investigated the original auction itself. The first auction had a lot of supplementary rules (and also no rights-to-match/player retention), so it was a complete auction for every player. Parker et al. found that there was a large premium on Indian players (as there was a maximum of four non-Indian players per team). Team valuations also increased for those with T20 experience, although this was as a time when T20 wasn't quite as prevalent, and the only real competitions were T20 internationals or in England [2]. The authors also note that batting and bowling strike rates were seen as highly significant.

Petersen et al. (2008) compared the magnitude of differences in indicators for key batting and bowling parameters between winning and losing teams from the first IPL season. Petersen et al. found that the best success indicators were taking more wickets in the game, taking more wickets in the last six overs, and having a higher run rate. They also note that winning bowlers captured

[2] Many English players did not take part in the original tournament as they were under contract with the ECB for an ODI Series in New Zealand

more wickets in the first and last six overs, and did a better job of limiting scoring in the middle eight. They conclude "Team tactics should focus on wicket-taking bowling and field placements in the first and last six overs and run restrictive field placing and bowling in the middle eight overs."

Ahmed et al. (2011) also developed a team optimisation model. However, their method of using batting average as the sole determinant of batting performance and the same for bowling average and bowling performance does not adequately meet the needs for T20 cricket, where as Petersen et al. (2008) pointed out, strike rate (scoring quickly) is so crucial. Also, fielding performance being measured by $\frac{\text{Total Catches Taken}}{\text{Number of games played}}$ unfairly inflates certain fielding positions (i.e. Slip-Fielders). It could also be argued that number of catches is actually an indicator on how good the captain is at setting a field as opposed to fielder skill.

Aparna et al. (2012) attempted an idea similar to what we are doing in terms of player evaluation, but with some noticeable differences. The authors simplify the league to three teams to make the problem easier to solve. Like Ahmed et al. (2011), their determination for fielding scores is arguable and, while their miscellaneous scores (factoring in marketability and potential) may be financially relevant, they are not statistically rigorous. However, the idea of varying the weighting of batting and bowling scores for an overall parameter (depending on the type of player[3]) should be acknowledged, although the use and application in this work was independently developed.

[3] Type, in this sense, means whether the player is defined as a batsman, bowler, all-rounder, batting all-rounder or bowling all-rounder

Chapter 3

Preliminary Data Analysis

3.1 Data Review

As mentioned in the introduction, all of the data used to model the batting and bowling distributions were obtained from the ESPN Cricinfo website. The website has individual pages for each season of the IPL summarising the results. From this, each match has an individual page with a variety of detailed match information. The number of balls faced by each batsmen and the runs scored can be taken from the Player vs Player tables of respective scorecards. This has been used to find the balls faced and distribution of scores for every batting innings. We are assuming that there is no home-field advantage, which may affect the scores. All of the results were obtained using free-to-use packages in R.

3.2 Losing Innings Total

3.2.1 Differences between first and second innings scores

As mentioned in Section 1.3, we wish to look at the distribution of the losing score for every IPL game. We will consider the losing scores for the first and second innings together and separately, and use a two-sample t-test to test to see whether it is appropriate to simulate from one combined distribution or two individual ones when we are calculating a player's batting score. Figure 3.1 shows the distributions of each of the two innings and a pooled distribution.

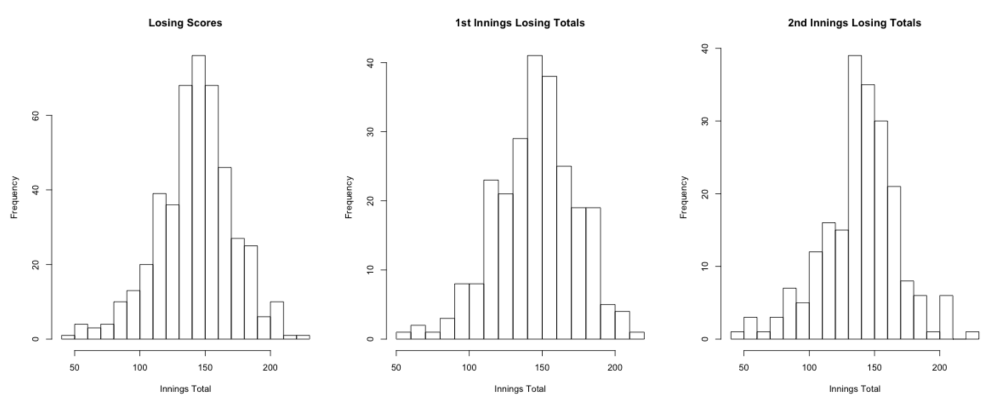


Figure 3.1: Histogram of losing scores (Left: total losing scores, Middle: first innings losing score, Right: second innings losing score)

It is not immediately obvious that there are any differences between the distributions, but we can calculate the mean and standard deviation of each, which is shown in Table 3.2. The larger standard deviation for the second innings can potentially be explained by teams batting more aggressively to chase down a target which would either result in a higher score or potentially higher wicket loss and lower scores.

Data Set	Mean	Standard Deviation
Losing Score	142.9	29.0
First Innings Losing	146.1	27.9
Second Innings Losing	139.1	29.9

Table 3.2: Innings Total Datasets

The difference in means between first and second innings is definitely of interest, and we will test their equivalence. We take as our null hypothesis (H0) that the distribution of first and second innings scores are from the same underlying distribution. Our alternate is the converse, that they have separate distributions. We will test at the 5% significance level using the `t.test()` function in R. The results of the test are shown in Table 3.3. With a value of $p = 0.01$, we reject the null hypothesis at the 5% significance level. In addition, the 95% confidence interval does not include zero, so we conclude that there is a difference between the two distributions.

Welch Two Sample t-test		
T Statistic	Degrees of Freedom	p-value
2.583	431.737	0.01012
95% Confidence Interval		
(1.681, 12.379)		

Table 3.3: Two Sample t-test on Losing.First and Losing.Second

As a result, when we undergo our simulations for the batting and bowling statistics, we will have to compare the scores against two different distributions, by simulating which innings a player's team batted in.

3.2.2 Change in innings scores over time

We also want to consider if there has been an increase or decrease in the scoring in an innings over the first six years in the IPL. We will do this by using a one-way ANOVA, which has the `oneway.test()` command in R. We will again test at the 95% significance level. Our null hypothesis is that all of the seasons have the same mean, and our alternative hypothesis is that there is greater variation in the means than what can be explained by chance. We obtain the results shown in Table 3.4.

One-way analysis of means (not assuming equal variances)			
F-Statistic	Num. DF	Denom. DF	p-value
2.706	6.00	195.44	0.015

Table 3.4: Output from R of a one-way ANOVA

A p-value of $p = 0.015 < 0.05$ implies that we should reject the null hypothesis at the 95% significance level and conclude that there is greater variation between the season means than what can be explained by chance. We now, therefore, wish to see if there is a trend over time that we can factor into our simulations. Figure 3.5 is a plot of the various means across the six test seasons.

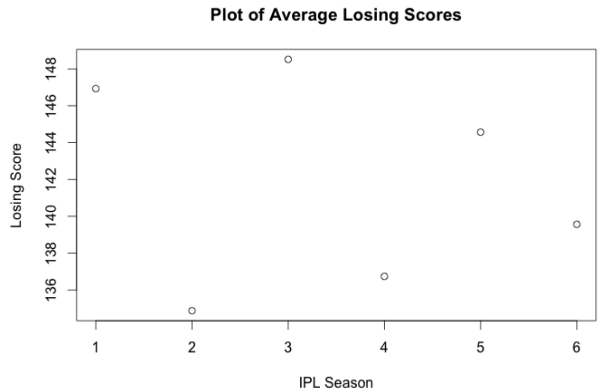


Figure 3.5: Plot of mean losing scores across six seasons

Parameter	Estimate	Std. Error	t-value	p-value
(Intercept)	146.931	3.767	39.003	≤2e-16 ***
Year2	-12.054	5.351	-2.253	0.0248 *
Year3	1.586	5.283	0.300	0.7642
Year4	-10.191	5.046	-2.019	0.0440 *
Year5	-2.363	5.031	-0.470	0.6388
Year6	-7.365	5.002	-1.472	0.1416
Year7	3.752	5.283	0.710	0.4779

Table 3.6: Linear Regression model of the mean losing scores and the season

Although ANOVA suggests more than random variation between the season means, we can see from the plot that there does not seem to be any sort of trend over time[1]. We can verify this (if needed) by creating a linear regression model with the various seasons treated as a factor and looking at the appropriate estimates. This output is shown in Table 3.6. There is no clear trend in the estimate scores for the years, and other than for the 2nd and 4th season, none of the p-values have any statistical significance. We are now in the situation where ANOVA concludes there is more difference in the means than what can be expected by random variation, but we have no way of modelling what this could be for future seasons. We will therefore make the decision not to account for year-on-year change in losing scores in our predictive/evaluative model.

3.3 Balls Faced by a Batsman

3.3.1 Geometric Distribution

We will begin by considering the work of Elderton (1945) and Langdale (1945) and examine a possible Geometric distribution for the number of balls faced. In order to assist in our calculations we will define “balls faced” as the number of balls faced where the batsman could have potentially scored a run. This has the effect of taking one from every ball faced. This allows us to model the balls faced as a geometric distribution as we can now have a zero value[2]. As we can see in Figure 3.7a), there is an inflated number at the lower end (we would expect a geometric distribution curve

[1] Indeed, we may expect that season 7 (which we will use for cross-validation) to be around the 142 mark if the pattern keeps converging. In actuality, the mean losing score for season 7 was 150.7.

[2] If a batsman was out on the first ball, they will now be recorded as a zero for balls faced, which is consistent with a geometric distribution.

to be smooth). This agrees with the observations of Wood (1945) about the discrepancies "at the commencement".

3.3.2 Zero-inflated Geometric Distribution

To account for this discrepancy, we will consider a Zero-inflated Geometric distribution (ZIG). We assume that the number of balls faced by an individual batsman are independent and identically distributed. A zero-inflated geometric distribution has the density function

$$X \sim \begin{cases} \text{Geom}(\theta) & \text{with prob. } p \\ 0 & \text{with prob. } (1-p). \end{cases}$$

Expectation-Maximisation Algorithm

We will find values for p and θ by using data augmentation and an Expectation-Maximisation (EM) algorithm. An EM Algorithm requires the Maximum Likelihood Estimate (MLE) and the expectation of the distribution. The algorithm then iteratively uses the expectation and the MLE to calculate the necessary parameters. We will first find the MLEs. We introduce the variable z_i with the following properties:

$$z_i = \begin{cases} 1 & \text{if } x_i \text{ is drawn from Geom}(\theta) \\ 0 & \text{otherwise.} \end{cases}$$

This enables us to express the Likelihood and the Log-Likelihood as

$$\begin{aligned} L(p,\theta;\mathbf{x},\mathbf{z}) &= \prod_{i=1}^{n} \left[(1-p)^{1-z_i} \left(p(1-\theta)^{x_i}\theta\right)^{z_i}\right] \\ \ell(p,\theta;\mathbf{x},\mathbf{z}) &= (n - \sum_{i=1}^{n} z_i)\log(1-p) + \log p \sum_{i=1}^{n} z_i + \log(1-p)\sum_{i=1}^{n} z_i + \log\theta \sum_{i=1}^{n} z_i \end{aligned}$$

We will make use of the fact that $\sum_{i=1}^{n}(x_i z_i) = \sum_{i=1}^{n} x_i$, and get the following MLEs.

$$\begin{aligned} \frac{d\ell}{dp} &= \frac{\sum_{i=1}^{n} z_i - n}{1-p} + \frac{\sum_{i=1}^{n} z_i}{p} \\ \Rightarrow \hat{p} &= \frac{\sum_{i=1}^{n} z_i}{n} \\ \frac{d\ell}{d\theta} &= \frac{-\sum_{i=1}^{n} x_i}{1-\theta} + \frac{\sum_{i=1}^{n} z_i}{\theta} \\ \Rightarrow \hat{\theta} &= \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} z_i}. \end{aligned}$$

We now calculate the expectation of z_i, to get

$$\begin{aligned} E[z_i|\mathbf{x},p,\theta] &= \frac{P(x_i=0|z_i=1)P(z_i=1)}{P(x_i=0|z_i=0)P(z_i=0) + P(x_i=0|z_i=0)P(z_i=1)} \\ &= \frac{p\theta}{p\theta + 1 - p}. \end{aligned}$$

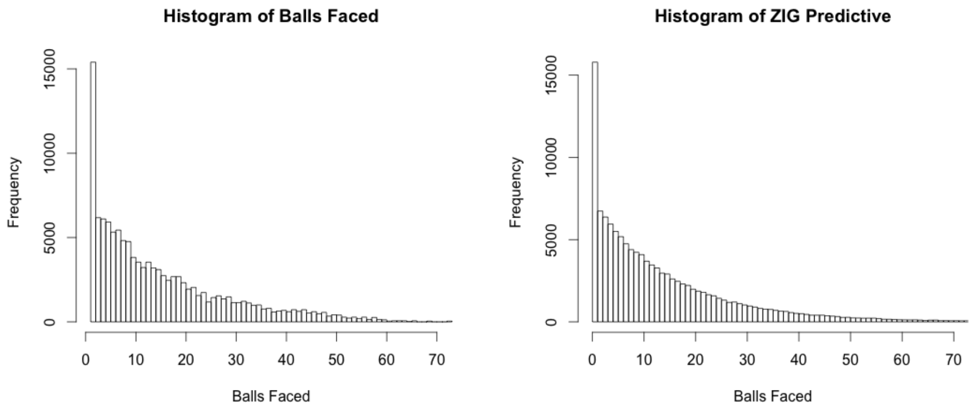


Figure 3.7: Comparison of Balls Faced (left) and fitted Zero-inflated Geometric Distribution (right).

After 100 iterations, we obtain the following estimates for $\hat{p}$ and $\hat{\theta}$

$$\hat{p} = 0.9949$$
$$\hat{\theta} = 0.0651.$$

We simulate data from this distribution using the above parameters. A comparison of both distributions[3] is shown in Figure 3.7.

3.4 Situational changes in Run Rates and Wickets

In order to create a realistic model for scoring in an innings, we need to factor in tactics throughout the game. If we make the assumption that the rate of scoring and rate of wicket loss will be the same throughout the game, the results we get (particularly in the final overs) will not reflect a real life T20 match[4]. As such, we will consider the following two things when building the simulation model:

- The distribution of wickets in the game, to identify the overs in which a higher proportion of wickets fell; and
- How the run rate changes throughout the game and investigate the dependence of the run rate on the game situation.

We will refer to the number of wickets lost and the current over as the **Game Situation**.

3.4.1 Distribution of Wickets

To account for the distribution of wickets throughout the match, we simply have to consider the over in which a wicket fell for every game played. We will then calculate a multiplicative constant for each over representing the relative frequency compared to the mean. Table 3.8 shows the frequency of wickets in each over throughout the first six years of the IPL tournament.

[3]The ZIG was simulated around 120,000 times, and the frequency of the actual balls faced was multiplied up for the sake of comparison of the graphs.

[4]Thanks to Michael Gladstone for pointing out that a score of 140/2 after 20 Overs would never happen. A team would bat more aggressively in the final overs as they do not have to worry about being bowled out.

Over	1	2	3	4	5	6	7	8	9	10
Frequency	161	196	210	194	196	201	170	173	188	174
Over	11	12	13	14	15	16	17	18	19	20
Frequency	214	199	204	217	235	268	285	362	361	478

Table 3.8: Frequency of Wickets per Over in the first six seasons of the IPL

Figure 3.9 illustrates these values on a graph. The horizontal line represents the mean number of wickets taken in a designated over. We can now clearly see why the original model was having trouble generating accurate scores towards the close of the innings. Table 3.10 shows the relative frequency of wickets in the over compared to the mean. For use in Chapter 6, we also calculate the mean of the wickets lost in an innings and get $\bar{\omega} = 5.856$.

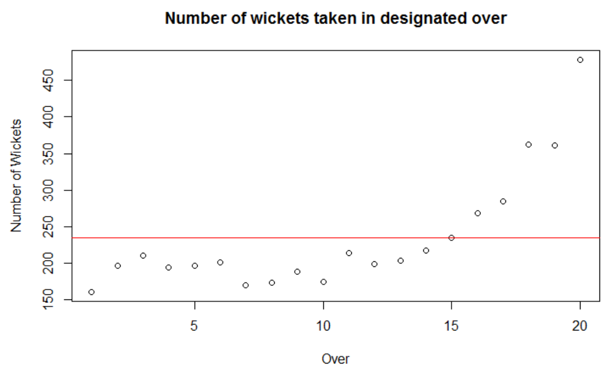


Figure 3.9: Histogram of Wickets Taken per Over

Over (i)	1	2	3	4	5	6	7	8	9	10
Proportion (Ψ_i)	0.687	0.837	0.896	0.828	0.837	0.858	0.726	0.738	0.802	0.713
Over (i)	11	12	13	14	15	16	17	18	19	20
Proportion (Ψ_i)	0.913	0.849	0.871	0.926	1.003	1.144	1.216	1.545	1.540	2.040

Table 3.10: Relative frequencies of the number of wickets fallen in a given over

3.4.2 Game Situation and Run Rate

Intuitively, we would expect the rate of scoring to vary throughout the game. Anyone who has much (if any) experience with cricket may say that at the start of an innings players “bat themselves in”, meaning they slowly accumulate runs and play conservatively to reduce the risk of getting out. They would also observe that, towards the end of an innings, especially when chasing a score (in the second innings) or with wickets to spare in the first innings, players would “have a bit of a slog”, meaning they will look to play aggressively and accumulate as many runs as possible in the limited balls left. Certain players may act contrary to these observations but, on the whole, they are reasonable expectations of the game and assumptions of basic strategy.

Table 3.11 gives the average modified run rate for an over, given the number of wickets that had fallen at the start of the over. We define the modified run rate to be

$$MRR = \frac{\text{Number of Runs Scored in an Over}}{\text{Number of Scoring Balls in an Over}}, \tag{3.1}$$

	Wickets Lost									
Over	0	1	2	3	4	5	6	7	8	9
1	5.973	-	-	-	-	-	-	-	-	-
2	7.439	6.74	-	-	-	-	-	-	-	-
3	8.109	7.224	6.077	-	-	-	-	-	-	-
4	8.462	7.832	7.066	-	-	-	-	-	-	-
5	8.857	8.261	7.455	7.292	-	-	-	-	-	-
6	9.568	8.354	7.801	6.594	-	-	-	-	-	-
7	7.655	6.868	6.48	5.798	5.653	-	-	-	-	-
8	8.302	7.642	6.879	7.005	7.3	-	-	-	-	-
9	8.007	8.172	7.666	6.603	5.823	-	-	-	-	-
10	7.931	7.829	7.45	6.662	6.08	-	-	-	-	-
11	9.002	9.124	7.71	7.156	6.74	7.17	-	-	-	-
12	9.056	8.629	7.678	8.198	7.557	6.508	-	-	-	-
13	-	8.822	8.106	7.755	7.601	7.198	5.489	-	-	-
14	-	9.697	8.509	8.458	7.76	7.016	6.191	-	-	-
15	-	11.13	9.246	8.602	8.177	7.802	7.056	-	-	-
16	-	10.31	9.789	9.386	8.531	8.425	7.329	6.57	-	-
17	-	12.73	10.65	10.41	9.49	9.308	9.11	7.142	-	-
18	-	11.8	12.3	11.23	10.57	9.426	10.08	7.848	-	-
19	-	-	12.46	11.19	11.42	10.89	10.59	8.86	7.826	-
20	-	-	-	13.12	12.18	12.44	11.5	10.52	8.615	6.865

Table 3.11: Mean number of runs scored in an over for a given game situation

where a scoring ball is defined to be a ball where a wicket did not fall. We use this as, while the traditional run rate provides an overall clearer picture of how a team is scoring across a match, we are only interested in the runs scored for an individual non-wicket ball for a particular game situation. This will become clearer when we begin the simulations in Chapters 4 and 5. In reality, there will be very little difference between these scores. We only include situations that have arisen at least 25 times. We can clearly see the run rate increasing towards the end of the game (as the overs progress as teams try to maximise their potential score) and the run rate trending downwards as the number of wickets lost increases. This is due to either batting conservatively to save losing more wickets, or the players being of lesser batting ability. This is especially the case once more than 8 wickets have fallen.

We can see in Figure 3.12 a visual representation of three of the columns of Table 3.11. This illustrates the points made above - the run rate is lower the more wickets that have fallen, and the run rate increases towards the end of the game. We can also see in the first graph the inflated run rate for the first six overs, where the Powerplay takes place[5].

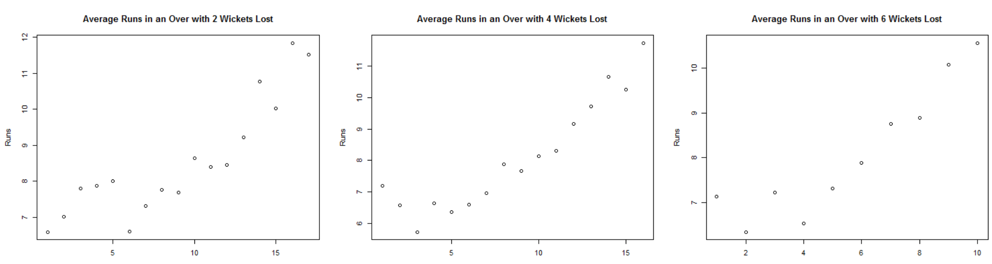


Figure 3.12: Expected/Average Run Rate for Game Situation with 2, 4, and 6 wickets lost

[5]See section 1.1

Chapter 4

Batting Score ($BATS$)

4.1 Original Simulation Model

We wish to assign each individual player a score that represents how many wins they contribute above an average player. We will do this by simulating games as if the entire batting line-up was made up of the given player. For a graphical illustration, Figure A.2.1 in the appendix is a representation of the original algorithm developed for this. We will use Zero-inflated Geometric distributions fitted using an Expectation-Maximisation algorithm to model the balls faced for each player. The probability of a certain number of runs in a ball can be calculated from the empirical data. We use the same formula and methods as described in detail in Section 3.3.2. An example of what the E-M Algorithm gives for some batting players is shown in Table 4.1. It should be pointed out that a player's ID number is for ease of coding use and has no effect on any player evaluation. It is simply the order in which a player has appeared as a batsman in the IPL. Players who have never batted but have bowled follow on after this. If a player has done neither, then they have not helped their team in a significant way and have not been included in the analysis. An implicit assumption in this chapter is that the individual batsman is batting against the average bowler. This may or may not be a reasonable assumption, and will be discussed in Chapter 8.

Remembering that p represents the probability the number of balls being from the Geometric Distribution, we can see that, for a number of batsmen displayed (SC Ganguly, BB McCullum, NM Coulter-Nile) and many others that are not, the number of balls faced in a typical innings is indistinguishable from a Geometric Distribution. In the case of Ganguly and McCullum, this is likely due to the fact they are excellent batsmen, as you would expect them to rarely get out on the first ball (certainly no more than any other ball, anyway). For Coulter-Nile, this value is more likely due to the limited number of innings played[1]. Given that we are drawing from a Geometric distribution, theta (θ) represents the probability of getting out on a given ball. To use the Geometric in this case, we assume that there is an equal chance on getting out every ball, and that a failure (Not Out, in this case) on a ball does not affect the probability of subsequent balls. Of course, in reality, both of these assumptions can be disputed. Different bowlers are of differing standards, so this probability is not constant. Also, many bowlers bowl different deliveries, and each may have a different probability of getting out. In addition, bowlers may "set-up" a batsman, so previous balls would have an effect on subsequent balls (Morgan-Mar, 2014).

For each ball in which a player was not out, we wish to simulate how many runs they scored on the given run. Using the data obtained from the ESPN website, we can calculate the empirical probabilities for a certain number of runs scored per batsman. We can see these probabilities for a

[1]Up to the end of IPL 6, Coulter-Nile had faced six balls.

ID	Batsman	θ	p
1	SC Ganguly	0.044	1.000
2	BB McCullum	0.048	1.000
3	RT Ponting	0.059	0.826
⋮	⋮	⋮	⋮
158	JD Ryder	0.063	0.994
159	KP Pietersen	0.047	0.999
⋮	⋮	⋮	⋮
372	NM Coulter-Nile	0.167	1.000

Table 4.1: Sample Expectation-Maximisation output for a ZIG Distribution for Batsmen (3 s.f.)

ID	Batsman	$P(Y=0)$	$P(Y=1)$	$P(Y=2)$	$P(Y=3)$	$P(Y=4)$	$P(Y=5)$	$P(Y=6)$
1	SC Ganguly	0.474	0.338	0.043	0.002	0.108	0.00	0.033
2	BB McCullum	0.443	0.327	0.054	0.002	0.126	0	0.048
3	RT Ponting	0.531	0.367	0.047	0	0.039	0	0.016
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
158	JD Ryder	0.419	0.317	0.066	0.007	0.151	0	0.041
159	KP Pietersen	0.365	0.374	0.067	0.007	0.118	0.002	0.067
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 4.2: Sample probabilities of runs scored by a batsman per ball

range of batsmen in Table 4.2. As our method of simulation involves generating a random variable, we will actually use the cumulative probabilities in the model.

The original idea was to simulate several thousand innings for each batsman, and compare to a direct, random sample from the existing scores. However, as mentioned in Section 3.4, the method of simulation outlined in Figure A.2.1 is not suitable, as it does not adequately reflect what would happen in an actual T20 game. This is in particular regard to how the tactics and aggressiveness of the batting team changes depending on the game situation. Instead, we need to consider a way of modelling this change in batting aggression and scoring rate over time.

4.2 New Batting Innings Simulation Model

4.2.1 Cumulative (Ordered) Probit Model for Ordinal Response Data

To account for the change in run rates over time, we introduce a Cumulative Probit Model for Ordinal Response Data (Agresti, 2010). Very simply, this means that we use a latent variable to generate our random ball score value, with the latent variable adjustable according to the situation. We can use this method as the response variable (runs) are ordered and we use the cumulative probability values as thresholds for the model.

We begin by setting our thresholds for the cumulative distribution. For convenience, and to compensate for a lack of data for some batsmen, we will use the average cumulative probabilities as our thresholds and express them as π_i where $i = \{0, 1, 2, 3, 4, 5, 6\}$. We now need to define a latent variable which will vary according to the game situation. For convenience, we will use X_j where $X_j \sim N(0, 1)$. As X_j is a latent variable, we set the mean and variance to be arbitrary values - our modelling choice is simply using a Normal distribution rather than, say, a Logisitic.

In an ideal situation, we would like to have ball-by-ball data, as well as information pertaining to the point at which a batsman played in a match. We could then have a joint model that contains the effect of wickets and overs and also corrects for a batsman's quality. However, we only have a player's scoring distribution. The method we are using implicitly assumes that a player's innings are randomly distributed within a game. Whilst this could be the case for a middle-order batsman, it is obviously not true for an opening batsman. However, it should be a good approximation.

We would like to find the values of τ_i such that $P(X < \tau_i) = \pi_i$. We will call the vector of values τ our threshold values. The cumulative probabilities and threshold values are displayed in Table 4.3. Therefore, for each non-wicket ball, we have

$$\begin{aligned}
P(0\text{ Runs}) &= P(Y=0) = P(X_i < \tau_0) \\
P(1\text{ Run}) &= P(Y=1) = P(\tau_0 < X_i < \tau_1) \\
P(2\text{ Runs}) &= P(Y=2) = P(\tau_1 < X_i < \tau_2) \\
P(3\text{ Runs}) &= P(Y=3) = P(\tau_2 < X_i < \tau_3) \\
P(4\text{ Runs}) &= P(Y=4) = P(\tau_3 < X_i < \tau_4) \\
P(5\text{ Runs}) &= P(Y=5) = P(\tau_4 < X_i < \tau_5) \\
P(6\text{ Runs}) &= P(Y=6) = P(X_i > \tau_5).
\end{aligned}$$

i	0	1	2	3	4	5	6
π_i	0.397	0.774	0.839	0.843	0.959	0.959	1
τ_i	-0.260	0.753	0.991	1.006	1.735	1.739	∞

Table 4.3: Table of runs threshold values

The purpose in using this method to simulate runs is evident when we consider changing the distribution of X_i to be $X_i \sim N(\mu, 1)$. By adjusting the value of μ for different game situations, we can make small adjustments to the probabilities of certain runs from balls. Let S be the score on a ball, and let $u = \{0, 1, 2, 3, 4, 5, 6\}$. Also let $\mu_{\{Q,W\}}$ be the μ value for the game situation in the Q-th over, and where the team has lost W wickets. We have

$$\begin{aligned}
\mathbb{E}\left[S|\mu_{\{Q,W\}}\right] &= \sum_{i=1}^{6} u\mathbb{P}(Y = u) && (4.1) \\
&= \sum_{i=1}^{6} u\mathbb{P}\left(\tau_{u-1} - \mu_{\{Q,W\}} < X < \tau - \mu_{\{Q,W\}}\right) && (4.2) \\
&= \sum_{i=1}^{6} u\left[\Phi(\tau_u - \mu_{\{Q,W\}}) - \Phi(\tau_{u-1} - \mu_{\{Q,W\}})\right], && (4.3)
\end{aligned}$$

where Φ is the standard Normal CDF. We know the values of $\mathbb{E}\left[S|\mu_{\{Q,W\}}\right]$ from the run rates we calculated in Section 3.4.2, so this becomes a case of solving for the 200 different values for μ. Again, we exclude those situations that have arisen less than 25 times as in Table 3.11. For R, this is relatively straightforward using the `uniroot()` function[2]. We can see the different values for μ in Table 4.4.

A positive value for $\mu_{\{Q,W\}}$ indicates a higher scoring over than average, and vice versa. We can see an increase in μ values as the overs increase and also a decreasing value as the wickets lost increase - both of which correlate with the observations made in Section 3.4.2.

When simulating a non-wicket ball, we simulate one $X_i \sim N(\mu_{\{Q,W\}}, 1)$ random variable and compare this result to the individual batsman's threshold totals to find the run(s) scored from that

[2] `uniroot()` solves an equation for zero, so it is necessary to rearrange equation 4.3 to be in the form $0 = \sum_{i=1}^{6} u\left[\Phi(\tau_u - \mu_{\{Q,W\}}) - \Phi(\tau_{u-1} - \mu_{\{Q,W\}})\}\right] - \mathbb{E}\left[S|\mu_{\{Q,W\}}\right]$

	Wickets Lost (W)									
Over (Q)	0	1	2	3	4	5	6	7	8	9
1	-0.1848	-	-	-	-	-	-	-	-	-
2	0.0082	-0.0804	-	-	-	-	-	-	-	-
3	0.0885	-0.0184	-0.1701	-	-	-	-	-	-	-
4	0.1292	0.0559	-0.0384	-	-	-	-	-	-	-
5	0.1736	0.1062	0.0103	-0.01	-	-	-	-	-	-
6	0.2506	0.1168	0.0521	-0.0996	-	-	-	-	-	-
7	0.0346	-0.0638	-0.1149	-0.2098	-0.2309	-	-	-	-	-
8	0.1108	0.0330	-0.0623	-0.0462	-0.0090	-	-	-	-	-
9	0.0765	0.0959	0.0358	-0.0985	-0.2062	-	-	-	-	-
10	0.0676	0.0555	0.0096	-0.0906	-0.1698	-	-	-	-	-
11	0.1895	0.2029	0.0412	-0.027	-0.0804	-0.0252	-	-	-	-
12	0.1955	0.1481	0.0374	0.0989	0.0227	-0.1111	-	-	-	-
13	-	0.1696	0.0881	0.0465	0.0281	-0.0217	-0.2553	-	-	-
14	-	0.2643	0.1345	0.1287	0.0471	-0.0447	-0.1542	-	-	-
15	-	0.4100	0.2161	0.1450	0.0965	0.0523	-0.0396	-	-	-
16	-	0.3279	0.274	0.2312	0.1371	0.1250	-0.0054	-0.1028	-	-
17	-	0.5627	0.3624	0.3377	0.2423	0.2228	0.2014	-0.0288	-	-
18	-	0.4749	0.5219	0.4202	0.3542	0.2355	0.3038	0.0577	-	-
19	-	-	0.5368	0.4157	0.4387	0.3861	0.3559	0.1739	0.055	-
20	-	-	-	0.5981	0.5108	0.5355	0.4460	0.3497	0.1466	-0.0642

Table 4.4: Table showing the values of $\mu_{\{Q,W\}}$

ball. Should a situation arise where a value for $\mu_{\{Q,W\}}$ is not listed, we take the nearest horizontal value in the table.

4.2.2 Model Adjustments for Rate Change in Fall of Wickets

The Expectation-Maximisation method for determining the length of a batsman's innings should be seen as a reasonable approximation for a batsman where game situation is not significant. This may perhaps be true in longer formats of the game (such as test cricket) where we could perhaps assume that the current over does not matter towards scoring rates, but it is certainly not true for T20 Cricket. Instead, we will use the preliminary work done in Section 3.4.1. In particular, Table 3.10 (pg. 15) represents the simple proportional adjustments we will make to a batsman's probability of getting out in a particular over (Ψ_i). We will simply multiply the probability that a batsman gets out on a particular ball (θ) by the appropriate Ψ_i in Table 3.10. This will ensure that the fall of wickets more closely reflects a true game of T20.

4.2.3 Updated Batting Innings Simulation

We will illustrate the method we will use with an example. From Tables 4.1 and 4.2, we will use entry no. 159, Kevin Pietersen. The data for Pietersen is shown in Table 4.5 for convenience. We will still follow roughly the same process as outlined in Figure A.2.1, but with the new modifications for scoring and wicket rates.

As we now need to update the model continuously with the current game situation, we cannot simply simulate the number of balls faced and then simulate the runs for each ball as we did for the initial model. Instead, we will need to simulate one ball at a time and update the overs and wickets where appropriate (where there is a change). Using this approach we are, in effect, treating the Geometric part of the Zero-inflated Geometric as a sequence of Bernoulli Random Variables.

For example, say Pietersen enters a game in the 8th over, for the loss of 2 wickets. From Table 4.4, we know the run rate modifier $\mu_{\{8,2\}} - 0.0623$ and we know from Table 3.10 that $\Psi_8 = 0.738$. Assuming Pietersen does not get out due to the Zero-inflated part of the distribution[3], then there is a probability of $\theta \cdot \Psi_8 = 0.047 \cdot 0.738 = 0.0347$ that he will be out on any of the balls remaining in that over. If we assume (for the sake of argument) that Pietersen was not out, we need to find out how

[3]For Pietersen, $p = 0.999$, so there is a 99.9% chance that the balls faced will be drawn from the Geometric.

ID	Batsman	θ	p	τ_0	τ_1	τ_2	τ_3	τ_4	τ_5
159	KP Pietersen	0.047	0.999	-0.344	0.642	0.864	0.889	1.480	1.500

Table 4.5: KP Pietersen statistics for simulation

many runs he scored. We generate a random variable from the distribution $X_i \sim N(-0.0623, 1)$. Let us assume that $X_i = 0.5$. We see that $\tau_0 < 0.5 < \tau_1$ and conclude that, on that ball, Pietersen scored 1 run. This process is then repeated until 10 wickets have fallen, or there are no more overs left to play.

The algorithm depicted by Figure A.2.2 takes as input the batsman's ID number (in Pietersen's case - 159), looks up the other relevant data from various tables, and outputs a predicted innings total for that batsman. This is the major part in the larger simulation model.

4.3 Model Implementation and Full Simulation

We will now define our Batting Score ($BATS$) as the number of wins a player contributes to a team, based on his batting ability alone. Our method to do this is relatively simple: for a single iteration, we will first simulate whether a player's one-man team is batting first or second. Using the algorithm to generate an innings total mentioned in Section 4.2.3, we will compare their total with a random sample first innings score if they were chosen to bat second, and vice versa. We repeat this 2500 times for each batsman. The proportion of wins is a probability, which we can then use. We will multiply this probability by 14 to find out how many games they would win in an 14 game regular season of the IPL. We would expect the average batsman to win 7 out of 14 such games. Their $BATS$ is the number of wins above (or below) seven. Table 4.6 shows the win probability (φ) and the $BATS$ for a selection of batsmen. Their batting score should not be seen as a prediction of how well a team of that player will do in a year (as many of the top batsmen can not bowl). Instead, it should be treated as an independent assessment of batting ability.

Figure 4.7 shows the $BATS$ of all of the players for whom we have data, with certain key players highlighted. The orange circles on the graph represent the "Orange Caps", or highest run scorers for the first six years in the IPL (IPL, BCCI, 2015b)[4]. Player 163, CH Gayle, as well as winning the Orange Cap is noted as having the highest win proportion in the IPL at 0.955, which equates to a $BATS = 6.367$. This can be attributed to his extremely aggressive play style and tendency to accumulate a lot of runs, but also to his underlying p and θ values. A value of $p = 1$ means that he has no greater chance of getting out first ball than any other ball, and a $\theta = 0.037$ is the equivalent of facing an expected 27 balls per innings. With 25% of his balls faced going for four

[4]There are only five dots as no. 163, CH Gayle, won the award in back-to-back years in 2011 and 2012.

ID	Batsman	Win Probability (φ)	$BATS$
1	SC Ganguly	0.284	-0.504
2	BB McCullum	0.496	2.066
3	RT Ponting	0.032	-6.390
⋮	⋮	⋮	⋮
158	JD Ryder	0.504	1.378
159	KP Pietersen	0.721	4.435
⋮	⋮	⋮	⋮

Table 4.6: Batsmen, Batting Scores and Win Probabilities

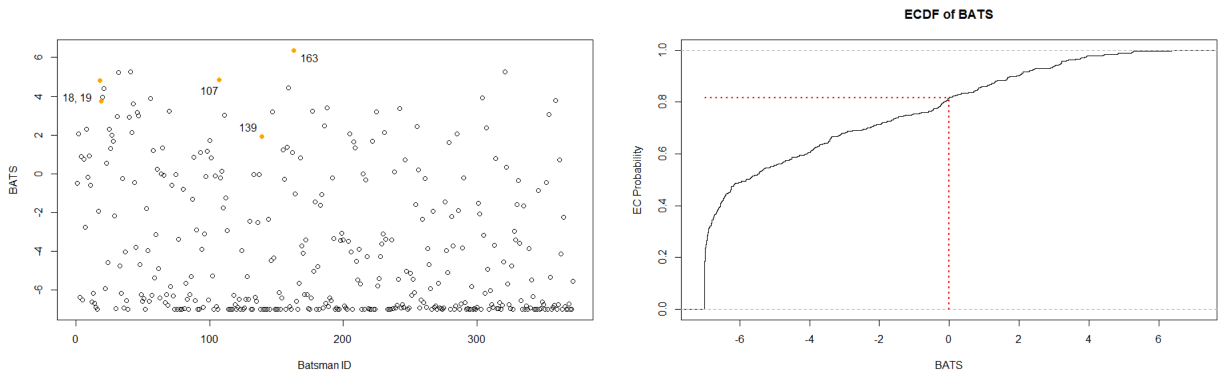


Figure 4.7: Plot of Batting Score ($BATS$)

runs or greater, it is easy to see why a team of Chris Gayles would perform so much better than the average.

There are a large number of players along the bottom of the graph with a $BATS$ equal or close to -7. These would almost entirely be bowlers, who have had to enter at the end of the innings. The limited number of balls that they face could skew the results downwards (as we treat an expired innings as completed (Elderton, 1945)), but this is unlikely to matter as these players are not judged on their batting and their BATS score will make up little or no part in their overall Wins Above Average (WAVE) score.

Figure 4.7 shows the empirical cumulative distribution of the batting score. As we can see, around 80% of the players for whom we have batting data do not contribute in a positive way to a team's chance of having a winning season when only batting score is taken into account. As mentioned in the previous paragraph, a large number of these would be bowlers (who contribute in another way). We can note, however, that in order to maximise the batting efficiency of a team, we want as many of the batsmen with a net positive $BATS$ as possible.

This does, however, suggest a methodological issue. If we are assuming that a $BATS = 0$ represents the average batsman, we would assume that there would be an equal number of players each side of 0. If we limit the plot to exclude players with a $BATS < -3.5$, we find the results shown in Figure 4.8. This may seem that we are ignoring a large number of batsmen, but it can be argued that if a player is getting a score this low, they should not really be considered a batsman at all. At the very least, they should not be picked in a T20 side based on their batting ability. It is likely that most of these players are bowlers (and so their $BATS$ will not count in the overall model) or "All-Rounders" which, in this case, might be a polite way of saying they are not exemplary with either the ball or the bat.

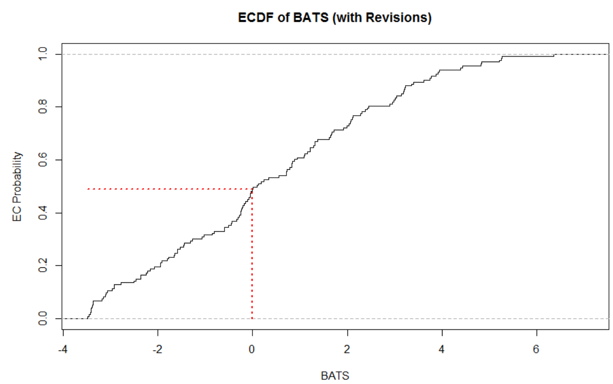


Figure 4.8: Empircal cumulative distribution of Batting Score, excluding values less than -3.5

Chapter 5

Bowling Score ($BOWLS$)

We now seek to obtain an objective value for the other side of the ball: the bowlers. In the Indian Premier League, bowlers may bowl a maximum of four overs, meaning that a minimum of five players must be proficient at bowling. The Bowling Score ($BOWLS$) not should represent how a bowling attack of five of the given player would perform against a random score. Just like we did for the Batting Score with the bowling ability of the players, we will not factor in the batting ability when we calculate the Bowling Score. We will bring the batting and bowling abilities of players together in the next chapter.

Data, as in the previous sections was obtained from the ESPN website. For each player, we have the scoring distribution against them and total wickets. From this, we can obtain the career economy and strike rate. The bowling strike rate is defined as the average number of balls bowled between wickets. The economy is the average number of runs per over. The economy does not play a part in the simulation model, but it is interesting nonetheless - a low economy, especially in T20, is a sign of a good bowler. The work of Petersen et al. (2008) suggests that a low economy is crucial for bowlers in the middle eight overs, whereas strike rate is more important for the first and last six.

An implicit assumption in this chapter is that the individual bowler is bowling against the average batsman. This may or may not be a reasonable assumption and will be discussed in Chapter 8.

5.1 Bowling Simulation Model

Generally speaking, we will follow the same simulation methods as in Chapter 5. An outline of the procedure used is in Figure A.2.2. For this, we need to calculate the probability a bowler gets a wicket on a ball for a given game situation and, subsequently, the cumulative probabilities for various run totals for a single ball (in the same game situation). The runs probability is quite simple to calculate and follows exactly the same method as the previous chapter (except now, obviously, we are looking at runs conceded rather than runs scored). We do, however, encounter difficulties for calculating the probability that a player gets a wicket on a given ball, which will be explored in the next section.

5.1.1 Modelling Wicket-taking Ability

We know from our work done in Section 3.3.2 that the number of balls a particular batsman faces in an innings follows a Zero-inflated Geometric distribution. Ideally, we would like to obtain similar data for a bowler, i.e. the distribution of the number of balls bowled between wickets. Regrettably,

	Wickets Lost (W)									
Over (Q)	0	1	2	3	4	5	6	7	8	9
1	0.0334	-	-	-	-	-	-	-	-	-
2	0.0408	0.0446	-	-	-	-	-	-	-	-
3	0.0486	0.0374	0.0357	-	-	-	-	-	-	-
4	0.0454	0.0385	0.0291	-	-	-	-	-	-	-
5	0.0390	0.04143	0.0473	0.0315	-	-	-	-	-	-
6	0.0465	0.0462	0.0340	0.0417	-	-	-	-	-	-
7	0.0465	0.0398	0.0297	0.0236	0.0245	-	-	-	-	-
8	0.0389	0.0377	0.0337	0.0369	0.0379	-	-	-	-	-
9	0.0380	0.0439	0.0389	0.0411	0.0269	-	-	-	-	-
10	0.0305	0.0368	0.0363	0.0376	0.0462	-	-	-	-	-
11	0.0667	0.0467	0.0481	0.0448	0.0250	0.0417	-	-	-	-
12	0.0598	0.0483	0.0447	0.0371	0.0315	0.0451	-	-	-	-
13	-	0.0511	0.0435	0.0434	0.0342	0.0489	0.0309	-	-	-
14	-	0.0465	0.0455	0.0421	0.0377	0.0680	0.0238	-	-	-
15	-	0.0473	0.0436	0.0531	0.0504	0.0513	0.0800	-	-	-
16	-	0.0490	0.0489	0.0530	0.0595	0.0655	0.0714	0.0720	-	-
17	-	0.0333	0.0726	0.0552	0.0676	0.0736	0.0610	0.0543	-	-
18	-	0.0256	0.0718	0.0742	0.0784	0.0988	0.0997	0.1070	-	-
19	-	-	0.0573	0.0646	0.0874	0.1040	0.0994	0.0964	0.1099	-
20	-	-	-	0.1288	0.0922	0.1370	0.1283	0.1683	0.1410	0.0901

Table 5.1: Table showing the values of $\omega_{\{Q,W\}}$

this data is not easily obtainable. Instead, we will use an alternate method for simulation using the Log-Odds.

We will use the values from Section 3.3.2 to represent the Zero-inflated Geometric distribution of balls faced for the average batsman. We have to make do with a bowler's strike rate to differentiate between them and their wicket taking ability. The strike rate represents the number of deliveries between wickets, so, assuming that $Y_i \sim \text{Geom}(\beta)$, we can say that

$$\beta_i = \frac{1}{\text{SR}_i}.$$

We also know from Section 3.4.1 that the rate of wickets lost varies between overs. This, however, was representative of the batsman versus an average bowler. We will instead find the probability of a wicket in an over for a given game situation, not just the over. It did not make sense to do this for the batsman as the batting ability (in the sense of not getting out) would not change for the number of wickets that have fallen if all batting players were the same (i.e. in our thought experiment/simulation). For our model to be effective, there should be some decrease in batting ability as the bowler takes wickets. We exclude game situations that have arisen less than 25 times[1], and we can see the probability of a wicket from a given game situation in Table 5.1. As we would expect, we see the probabilities increase towards the end of the game and as the number of wickets increase (i.e. as weaker batsman enter the game).

We will also incorporate this into the log-odds. Instead of using the relative frequency for the change in wickets over time, we will now use the probability of a wicket for an individual ball in a given over.

We define the log-odds for a probability p to be $\log(\frac{p}{1-p})$. Let $\omega_{\{Q,W\}}$ represent the probability of a wicket in the Q-th over with W wickets currently down, with log-odds $\Omega_{\{Q,W\}}$. Also let β_i be the probability a bowler gets a wicket on a random ball, which, as mentioned above is the reciprocal of the Strike Rate. For a regular (non-first ball) bowl, the log-odds are expressed as

$$R_i = \log\left(\frac{\beta_i}{1-\beta_i}\right) + \Omega_{\{Q,W\}},$$

[1]In the simulation, for situations that have occurred less than 25 times, we take their nearest vertical value, with an exception for the 20th over, where we take the next horizontal value.

ID	Bowler	τ_0	τ_1	τ_2	τ_3	τ_4	τ_5
1	SC Ganguly	-0.536	0.821	1.061	1.077	1.640	1.640
4	DJ Hussey	-0.553	0.696	0.964	0.967	1.451	1.451
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
391	BW Hilfenhaus	0.085	0.811	0.994	1.014	1.709	1.709

Table 5.2: Bowling threshold values for a sample of players

with probability

$$P(\text{Wicket}|Q,i) = \frac{\exp(R_i)}{1+\exp(R_i)}.$$

For the first ball, we have to factor in the zero-inflation effect. From Section 3.3.2, we know that, for an average batsman, the probability of them being automatically drawn from the zero-inflation is $z = 1 - 0.9949 = 0.0051$. Let Z be the log-odds of this probability. Then for the first ball bowled to the average batsman by bowler i, we have $\beta_{i,0} = \beta_i(1-Z) + Z$ and

$$\begin{aligned} F_i &= \log\left(\frac{\beta_{i,0}}{1-\beta_{i,0}}\right) + \Omega_{\{Q,W\}} \\ \Rightarrow P(\text{Wicket}|\Omega_{\{Q,W\}}, i, Z) &= \frac{\exp(F_i)}{1+\exp(F_i)} \end{aligned}$$

5.1.2 Distribution of conceded runs

For every bowler, we have the runs that they have conceded in their career and the number of balls bowled. The threshold values for calculating the $\mu_{\{Q,W\}}$ will be the same as used in the batting simulation, but the thresholds can be calculated for the individual bowlers just as the individual batsmen (see Section 4.2.3). We will also use Table 4.4 to determine the run rate modification for the game situation. A comprehensive explanation of this method is given in Chapter 4. Table 5.2 gives the threshold values for certain batsmen. We can interpret higher threshold values for a bowler as a better economy. For example, BW Hilfenhaus has higher threshold values across the board than DJ Hussey, which is akin to Hussey's career economy of 8.62 being greater (and so, worse) than Hilfenhaus' 6.47[2].

5.2 Implementation and Full Simulation

We proceed as in the previous chapter and simulate on a ball-by-ball basis. We are assuming that a batsman faces bowls until he is out/the game is over, and the strike (batsman facing the bowler) does not change. This assumption is discussed in section 8.1.3. We simulate 2500 games for each bowler, randomly allocating each iteration to be batting first (bowling second) or vice versa. The proportion of wins gives us a probability, which we use to determine how many games a bowling attack exclusively of one player would win above the average number of wins in a season (7). This number above the average is the Bowling Score ($BOWLS$).

Table 5.3 gives the $BOWLS$ for some of the bowlers in the IPL. The highest Bowling Score calculated was for AM Rahane, who has only bowled one over, and took one wicket. This is an

[2]This data is not shown, but can be obtained easily from ESPN Cricinfo.

ID	Player	Win Probability	*BOWLS*
1	SC Ganguly	0.227	-3.819
4	DJ Hussey	0.047	-6.339
5	Mohammad Hafeez	0.593	1.327
⋮	⋮	⋮	⋮
316	SP Narine	0.867	5.1400
⋮	⋮	⋮	⋮
391	BW Hilfenhaus	0.741	3.3712
(93)	(AM Rahane)	(0.995)	(6.933)
(99)	(Sohail Tanvir)	(0.9036)	(5.6504)

Table 5.3: Bowlers, *BOWLS* and Win Probabilities

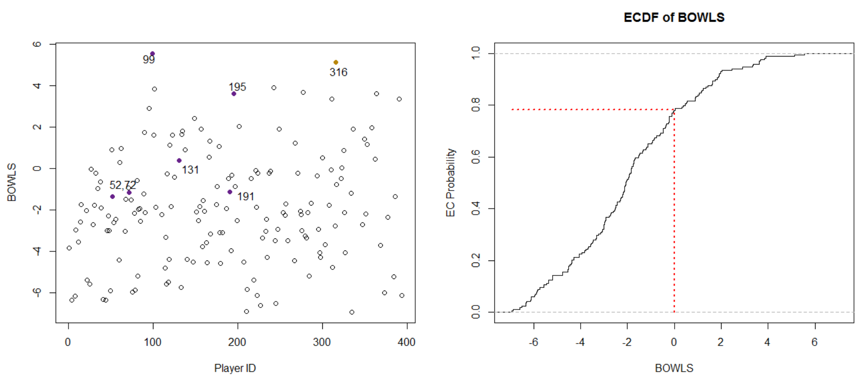


Figure 5.4: Plot of Bowling Score (left) and Cumulative Distribution (right)

impressive strike rate, but not enough to include in the Bowling Scores [3]. We will therefore exclude all bowlers who have bowled less than 100 balls. From 100 balls (which is approximately 17 overs, or 4 innings of play), we should have a good idea about a bowler's ability. We do this for bowlers, and not batsman, as a bowler is far more likely to "fluke" a good over than a batsman is. We would assume that players who have bowled less than this are not highly rated by their team mates and coaches. Excluding these players, the best record, statistically, belongs to Sohail Tanvir. Pakistani players, like Tanvir, played in the first season of the IPL but have not since due to political reasons between India and Pakistan (India Today, 2015). The best eligible bowler, therefore, is SP Narine. Narine was the 2012 Player of the Tournament, with an economy of 5.47 and 24 wickets.

Figure 5.4 shows the Batting Scores and the matching empirical cumulative distribution. The purple dots represent the winners of the "Purple Cap", which is awarded to the player with the most wickets in a year. It seems puzzling as to why some leading bowlers from previous years would have as low of a bowling score as they do. This has several possible explanations. The simplest is that the season in which they won the award was a fluke. For instance, this could be argued as the case for Morne Morkel (No. 72) who has struggled to get the same statistics in the IPL since the season in which he won the purple cap. As our data is based on a career average, this explanation has definite merit.

Another possible explanation is that our model favours bowlers with low average economies as opposed to wickets, given that the objective is to limit the runs scored, not to take all of the wickets. This would explain scores for other renowned bowlers, such as Shane Warne (54th, -0.6216) and

[3] The best record actually belongs to Adam Gilchrist, who bowled one ball and took a wicket. This broke the simulation, so he was excluded earlier on.

Ranking	ID	Player	Economy	Strike Rate	*BOWLS*
1	316	SP Narine	5.399	16.07	5.140
2	243	CK Langeveldt	6.803	12.08	3.903
3	277	Shakib Al Hasan	6.450	14.48	3.685
4	391	MM Sharma	6.296	15.20	3.623
5	195	SL Malinga	6.123	16.54	3.612
⋮	⋮	⋮	⋮	⋮	⋮
23	120	M Muralitharan	6.340	23.83	1.1424
47	38	SK Warne	7.075	20.95	-0.6216

Table 5.5: Top five-eligible BOWLS scores and comparisons

Muttiah Muralitharan (26th, 1.1424). It might be expected by cricket purists that these should score higher as they are both excellent spin bowlers capable of taking many wickets, but also susceptible to concede large amounts of runs. They would be welcome into any test or one-day side but, if anything, this highlights the difference between the various forms of the game.

Another way to consider this problem is to identify which bowlers achieved the highest scores and identify what they all have in common. The top five currently eligible bowlers (prior to the 2014 season), according to $BOWLS$, are shown in Table 5.5

We can immediately see why Narine's statistics are so good. His miserly economy rate, paired with a very good strike rate obviously has a large impact on the model. We can see why these noted players are not rated very highly. Muralitharan's economy is more than good enough to be in amongst the very top, but his strike rate is poor. This is undoubtedly the effects of the game situation modifier on the run rate.

Chapter 6

Wins Above Average ($WAVE$)

As outlined in the introduction, the purpose of this dissertation is to determine a way of evaluating player ability in the Indian Premier League. Wins Above Average, or $WAVE$, is the way we will do this. We want the WAVE score to represent how much better (or worse) a team will fare over the course of the season as a result of having that player, compared to our 'average' player.

6.1 Defining and Developing $WAVE$

For every player, we now have at least one of a batting or bowling score. If we do not have a score for a player (e.g. a batsman who has never bowled, or a bowler who is so far down the order that they have never batted), we will assume that score is the minimum of -7. Equally, if we have excluded a bowler due to the lack of balls bowled, they will receive a -7 score.

Using a simple additive formula to calculate the WAVE will not suffice. SP Narine, the best bowler from our data, would end up with a negative rating due his lack of batting ability. We will therefore have to define the WAVE on a role by role basis. First, we need to modify the Batting and Bowling Scores.

$BATS$ and $BOWLS$ represent a player's expected league performance if the entire batting or bowling attack was made up of the selected player. However, to determine the $WAVE$ score, we need to consider what effect one solitary version of a player will have on a team. Our function for determining $WAVE$ for a player i from category j is

$$WAVE_{i,j} = \phi\alpha_j BATS_i + \psi(1 - \alpha_j)BOWLS_i \tag{6.1}$$

where ϕ and ψ are standard modifications of the Batting and Bowling scores, and α and β are the weightings depending on the category of player.

We have calculated that the mean number of wickets lost in any innings is $\bar{\omega} = 5.856$. Therefore, hypothetically, 7.856 batsmen will appear in the average game. We set $\phi = \frac{1}{7.856}$. The value of ψ is even easier to calculate. For the simulation, we assume that the bowler bowls 20 overs. Each player can only bowl a maximum of five in a real game, so we just have to divide through by four. So, we set $\psi = \frac{1}{5}$. We will define the score weightings as shown in Table 6.1. This follows a similar weighting to Aparna et al. (2012) but, as we have different scores, we will weight things slightly differently. This is discussed further in Section 8.1.4.

Rather than go through ESPN Cricinfo to determine how we should define each player as, we will use our model to define the criteria for each position. This is shown in Table 6.2.

An issue does arise for the case of All-Rounders (AR). A player will maximise his WAVE score by being either exclusively a batsman or a bowler (depending whether his $BATS$ or $BOWLS$ is

Category	Batsman	Batting All-Rounder	True All-Rounder	Bowling All-Rounder	Bowler
j	1	2	3	4	5
α_j	1	0.65	0.5	0.35	0.05
$1-\alpha_j$	0	0.35	0.5	0.65	0.95

Table 6.1: Score weighting for player categories

	Criteria	
Category	*BATS*	*BOWLS*
Batsman	$BATS > -3.5$	-
Any All-Rounder	$BATS > -3.5$	$BOWLS > -3.5$
Bowler	-	$BOWLS > -3.5$

Table 6.2: Definition of Categories

higher). In addition to this, an AR will maximise their $WAVE$ by being a Batting All-Rounder (BAR) or a Bowling All-Rounder (BOAR), as opposed to a True All-Rounder (TAR). Whilst this is true, we are not looking to maximise the $WAVE$ of a player, but of the team. For instance, SK Raina (ID. 21) maximises his individual $WAVE$ score as a batsman, with $WAVE_{21,1} = 0.560$, compared to his Batting All-Rounder score of $WAVE_{21,2} = 0.223$. However, if his team needs him to bowl, Raina will still improve their chances of winning. There are multiple All-Rounder categories to attempt to portray a better picture of a player's ability. It also serves another benefit. The categories of All-Rounder could be seen as a players position on the batting order. A BAR may have to bat 4th, and so has a very high chance of batting in the game, whereas a BOAR may only be listed as 8th or 9th on the order. Any IPL team should have a view of the specific contributions they are looking for from an AR, and select from the appropriate category(/ies) accordingly.

6.1.1 $WAVE$ Scores

We will now highlight some interesting and notable results from the previous work. We will list, by category, the top ten players in each category. This is shown in the Appendix in Table 1.3. We many of the same names across many of the All-Rounders lists. If anything, this highlights the rarity of a player that can both bat and bowl at a sufficient level to be considered to have a net positive effect on their team. Bowling All-Rounders are seen as especially rare, with only the top six contributing extra wins to their teams. In addition four of these six are better suited to be played higher up the batting order (i.e. considered to be a TAR or BAR).

6.2 Results Evaluation

In order to evaluate the results from Chapter 6, we wish to see how teams perform given their respective team lists. We would expect the teams that have the largest combined $WAVE$ score to win the most games. To calculate the $WAVE$ scores, we used data from 2008-2013, which is IPL 1 to IPL 6 respectively. We will use the 2014 version of the tournament, known as IPL 7, to assess the validity of our model.

The 7th season of the IPL took place in 2014 and was the first year where the league dropped back down to just 8 franchises. The first 20 games were played in the UAE, and the remaining 36 regular season games (plus the 4 playoff matches) were played in stadia across India.

6.2.1 WAVE Scores for the 2014 Season

We will be using the results from the previous chapter as the $WAVE$ score for the players. We will omit the scores for players for whom we do not have data prior to the 2014 season (any ID greater than 397). For each team, we will select their typical line-up of 11 players[1]. The category of player (bowler, batsman, etc.) was inferred from the data and manually assigned to each player where there were any discrepancies. We can see the teams, players, $WAVE$ scores and totals in Table 6.3. We can also see the $WAVE$ score that would be assigned to an individual player for their performance solely for that season.

Our results, on the whole, seem to be reasonable. For 6 of the 8 teams, we were within 2.2 games of the actual total - this is an error per player of less than 0.2. Larger errors occurred for Kings XI Punjab and Delhi Daredevils, for which we will consider an explanation in more detail. Kings XI is the most glaring error, as the estimate is completely incorrect. The estimate for Delhi suggest more of systematic error within the model concerning extremes. We also correctly identified 3 of the 4 playoff teams, with the only exception again being Kings XI Punjab. For 6 of the 8 teams, their respective team $WAVE$ scores correctly indicate if the team will win 7 or more games (if positive) and 7 or less games (if negative).

6.2.2 Hypothesis Test of Model Validity

We have noted in the previous section that our model looks promising as an indicator for the performance of a majority of the teams. We will examine this more closely with a simple chi-squared hypothesis test. Our null hypothesis is that the difference between predicted wins and observed wins can be attributed to random variation, whilst our alternate hypothesis is that there is another underlying factor. We will use a 5% significance level, with $n = 8$ and 7 degrees of freedom.. We define the expected number of wins (the $WAVE$) with ϵ, and the observed number of wins by ω. For our test statistic, t, we have

$$\begin{aligned} t &= \sum_{i=1}^{8} \left(\frac{(\omega_i - \epsilon_i)^2}{\epsilon_i} \right) \\ &= 7.376 \\ P(T > t) &= 0.3909 \\ &> 0.05, \end{aligned}$$

so there is insufficient evidence to reject the null hypothesis and conclude that the variation from the expected (calculated) value for teams could be explained by random chance. We will now investigate the reasons why certain teams may have performed differently to expected.

6.2.3 Explanatory factors for poor estimates

Kings XI Punjab: Lack of previous data

We can see from Table 6.3 that KXI greatly outperformed their expected win total for the year. A key factor in that was the performance of GJ Maxwell. In the 2013 season, Maxwell had a batting average of just 18 runs, had taken no wickets as a bowler and had hit only one four in the tournament - certainly not worth the $1 Million USD that Mumbai paid for him at the start of the

[1]Having gone through every game, the players listed for each team are the ones with the most appearances

year. This is reflected in his $WAVE$ score of -0.788. His 2014 season could not be more different. His strike rate of 187 and average total of 34.5, with 48 fours and 36 sixes established him as the third highest run scorer in the tournament[2]. He was named as the 2014 Man of the Series. His ability is reflected in his season score of 0.807, which is a 1.5 game improvement. He was helped by improved performances from M Vohra ($\Delta = +0.7$), GJ Bailey (+0.2), R Dhawan (+0.7) and a good first year from AR Patel. One thing that all of these players have in common is that they have had very limited prior exposure to IPL (and so we do not have much available data). Dhawan had previously only played 6 games, Vohra had played 12, Maxwell 5, and Bailey 4. The more data we have on players in a team, the more reliable we would expect our data to be. For example, Mumbai had only two inexperienced players in CJ Anderson (rookie) and JJ Bumrah (2 games). It is no surprise the Mumbai team had the best estimate. Similarly, the Kolkata Knight Riders team has only one player (SA Yadav) who played less than a full season prior to the tournament. The next least-experienced player is Shakib Al Hasan with 15 games. KKR had the narrowest margin between their pre-2014 $WAVE$ score and 2014 season-long $WAVE$. Summing over the 2014 $WAVE$ score for Kings XI, we are still over one game away from the true score, but this could be put down to random variation [3].

Sunrisers Hyderabad: Individual player performance swings

We expected Hyderabad to win more than they lost, and we were wrong by two wins. Unlike Kings XI, where a variety of players are needed to explain this variation, we can attribute most of this change to the performance of DW Steyn and A Mishra. Steyn was projected to be a key net contributor to the team, with a $WAVE \simeq 0.5$. This reflected his 2013 performance of an excellent economy of 5.66 and bowling strike rate of 21.42. In 2014, his economy rose to 7.69 and strike rate rose to 30.54. He also took 8 less wickets in the season. This resulted in $\Delta WAVE = -1.236$. In addition, A Mishra's score sank as his economy increased from 6.35 to 9.06 and strike rate from 17.7 to 32.1. Mishra contributed a net change of -1.44. Cricket Country listed these two players in their "5 flops of the season" (Varma, 2014). From these players, we can see why it seemed that Sunrisers underperformed.

Delhi Daredevils - Model Flaw or Random Variation?

On first inspection, we can see that we predicted Delhi to win around 5 games in the 2014 season. Assessing the player performances for the season-long $WAVE$ score, we expect Delhi to win 4 games. In actual fact, Delhi won just 2. We were correct in thinking the Delhi team would struggle (they have the league lowest $WAVE$ score), but we seem to have underestimated by how much. We investigate if this is a problem with our model by simple hypothesis testing. We model the number of wins in a season to be a Binomial$(14, p)$ distribution, where

$$p = \frac{7 + \sum WAVE}{14}.$$

Our null hypothesis is that the $WAVE$ score for the team is a good estimate, and the difference for Delhi can be explained by random variation. Our alternative hypothesis is that random variation does not explain the score of Delhi Daredevils and there may be a systematic error in the model. We will test at the 5% significance level. We have $p_{\text{DDV}} = 0.337$ and so,

[2] Data taken from IPL, BCCI (2014)

[3] We can carry out a hypothesis test for this using the same method as in Section 6.2.3. We would get $P(X_i \geq 11|p) = 0.154$, which, again, we reject at the 5% level.

$$
\begin{aligned}
P(X_i \leq 2 | p = 0.337) &= 0.101 \\
&> 0.05,
\end{aligned}
$$

so there is insufficient evidence to reject the null hypothesis. It is worth noting, however, that there may be some issues with measuring extremely high or low performances, and it is definitely worth consideration in future years.

Mumbai Indians (4th)				Chennai Super Kings (3rd)				Kolkata Knight Riders (2nd/W)				Kings XI Punjab (1st/RU)			
ID	Player	WAVE	True WAVE	ID	Player	WAVE	True WAVE	ID	Player	WAVE	True WAVE	ID	Player	WAVE	True WAVE
19	MEK Hussey	0.478	0.212	147	DR Smith	0.406	0.601	40	G Gambhir	0.369	0.021	204	CA Pujara	-0.443	-0.273
208	AP Tare	0.210	-0.602	2	BB McCullum	0.263	0.290	9	JH Kallis	-0.022	-0.634	41	V Sehwag	0.669	0.481
209	AT Rayudu	0.168	0.248	21	SK Raina	0.560	0.674	97	MK Pandey	-0.020	0.356	306	GJ Maxwell	-0.788	0.807
47	RG Sharma	0.379	0.335	20	MS Dhoni	0.502	0.490	58	RV Uthappa	0.154	0.668	399	AR Patel	-	0.454
222	KA Pollard	0.083	0.153	360	IC Pandey	-0.891	-0.848	31	YK Pathan	0.376	0.627	321	DA Miller	0.670	0.560
400	CJ Anderson	-	-1.239	304	F du Plessis	0.499	0.282	277	Shakib Al Hasan	0.675	0.433	198	GJ Bailey	-0.441	-0.234
62	Harbhajan Singh	0.157	0.135	35	RA Jadeja	-0.033	-0.242	309	SA Yadav	-0.891	-0.135	53	WP Saha	-0.230	0.757
392	JJ Bumrah	-0.891	-0.857	202	R Ashwin	0.343	-0.050	68	PP Chawla	-0.322	-0.043	358	MG Johnson	0.332	-0.344
195	SL Malinga	0.642	0.768	172	M Manhas	-0.435	-0.843	85	R Vinay Kumar	-0.520	-1.065	357	R Dhawan	-0.878	-0.167
15	Z Khan	-0.372	0.100	4	DJ Hussey	0.113	0.303	316	SP Narine	0.932	0.527	346	M Vohra	-0.109	0.675
131	PP Ojha	0.029	-0.029	364	MM Sharma	0.644	-0.022	191	M Morkel	-0.256	-0.195	368	Sandeep Sharma	-0.891	0.031
$\Sigma WAVE$		0.808	-0.718	$\Sigma WAVE$		1.971	0.628	$\Sigma WAVE$		0.277	1.273	$\Sigma WAVE$		-2.108	2.749
Actual Wins		7		Actual Wins		9		Actual Wins		9		Actual Wins		11	
Difference		0.808	-0.718	Difference		-0.029	-1.372	Difference		-1.723	-0.723	Difference		-6.108	-1.251
Rajasthan Royals (5th)				Sunrisers Hyderabad (6th)				Royal Challengers Bangalore (7th)				Delhi Daredevils (8th)			
ID	Player	WAVE	True WAVE	ID	Player	WAVE	True WAVE	ID	Player	WAVE	True WAVE	ID	Player	WAVE	True WAVE
307	SPD Smith	0.302	0.185	255	AJ Finch	0.308	0.077	17	PA Patel	-0.247	-0.298	347	JD Unadkat	-0.558	-0.828
93	AM Rahane	0.140	0.102	42	S Dhawan	0.269	0.208	120	M Muralitharan	0.172	-0.763	186	M Vijay	0.312	-0.292
352	SV Samson	-0.058	0.258	188	DA Warner	0.433	0.722	8	V Kohli	0.291	0.244	88	KD Karthik	0.107	0.180
32	SR Watson	0.664	-0.041	355	KL Rahul	-0.681	-0.347	27	Yuvraj Singh	0.163	0.014	87	MK Tiwary	-0.168	0.030
256	STR Binny	0.024	0.165	362	DJG Sammy	0.095	-0.535	111	AB de Villiers	0.384	0.690	155	JP Duminy	-0.104	-0.090
340	KK Nair	-0.883	0.586	29	IK Pathan	-0.275	-0.885	163	CH Gayle	0.810	-0.082	159	KP Pietersen	0.565	0.276
89	R Bhatia	-0.274	-0.020	167	KV Sharma	0.220	-0.121	302	S Rana	-0.265	-0.665	356	Q de Kock	-0.891	0.025
311	JP Faulkner	0.593	-0.397	95	DW Steyn	0.507	-0.729	103	AB Dinda	-0.397	-1.049	268	WD Parnell	0.190	-0.470
363	KW Richardson	-0.527	0.002	134	A Mishra	0.271	-1.171	402	MA Starc	-	-0.228	294	S Nadeem	-0.111	-0.849
81	DS Kulkarni	-0.154	-0.340	300	B Kumar	0.055	0.502	395	YS Chahal	-0.891	-0.257	236	KM Jadhav	-0.436	-0.112
371	PV Tambe	-0.890	0.158	184	NV Ojha	-0.136	0.390	297	VR Aaron	-0.815	0.385	348	Mohammed Shami	-0.887	-0.970
$\Sigma WAVE$		-1.063	0.664	$\Sigma WAVE$		1.073	-1.817	$\Sigma WAVE$		-0.792	-2.009	$\Sigma WAVE$		-2.286	-3.098
Actual Wins		7		Actual Wins		6		Actual Wins		5		Actual Wins		2	
Difference		-1.603	0.664	Difference		2.073	-0.817	Difference		1.208	-0.009	Difference		2.714	1.902

Table 6.3: $WAVE$ scores for each team, giving pre-season values and 2014 season values.

Chapter 7

Optimisation

We now wish to consider the applications of the work developed in the previous chapters. We will explore ways of optimising a team, within the player constraints outlined in the IPL rules and regulations (IPL Desk in Mumbai, 2013). For this, we will use Microsoft Excel, and the add-on OpenSolver by Mason (2012) to perform integer programming subject to linear constraints.

We will be considering optimisation problems prior to the 2015 season, using data accumulated from the previous seven tournaments. Players who have not appeared in the IPL prior to 2015 are not considered. This limits the practical use of this approach in the IPL immediately but is discussed in more detail in Section 8.4, along with the necessary work that would need to be carried out to include other players.

For players who were up for auction in 2015, we have exact salary values which we can use as a cost. This, however, gets more difficult for those retained ahead of the season as their salary negotiations are private. As a result, we have used the values for which they sold in 2014 before IPL 7. Certain extremely valuable players were retained prior to the 2014 auction[1], and so we will use the retention order listed on Cricinfo as the salaries for retained players (shown in Table 1.1). This assumption may seem like a stretch, but this chapter merely serves as an illustration of the possibilities of developing a Wins Above Average model.

7.1 Development of Linear Model

7.1.1 Initial Set Up

Our goal is to optimise the sum $WAVE$ score for a team of players. Where a chosen individual's score is given as ω_i, this is $\sum_{i=1}^{11} \omega_i$. We have several constraints to consider that are listed in Table 7.1, where x_i is the indicator variable for if the i-th player is selected. We also define $j = \{1, 2, 3, 4, 5\}$ to be the possible roles that a player could play within a team (Batsman, Batting All-Rounder, All-Rounder, Bowling All-Rounder, Bowler). The player uniqueness constraint ensures that a player cannot be selected multiple times.

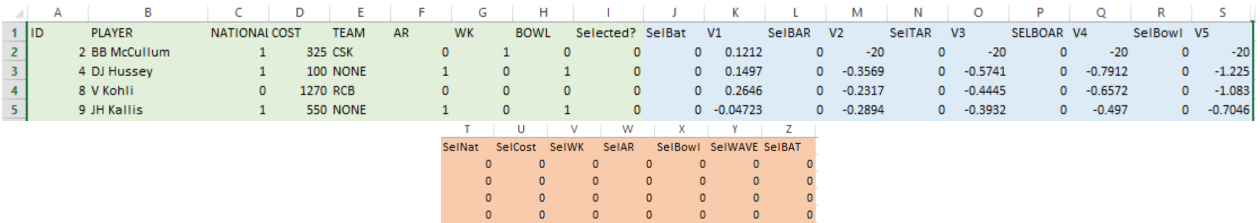
7.1.2 Methodology

As previously mentioned, we will use Microsoft Excel and OpenSolver for this problem. It is necessary to use OpenSolver as Excel is not able to consider the large number of variables which we will be using. We will set up a data frame for the players and their various attributes. This table needs indicators for the player's nationality, their respective roles in the team, their cost, and their

[1] More details about the auction set up and player retention are given in Section 1.2

Name	Coded Value	Constraint
Number of Players	Indicator	$\sum x_i = 11$
International Players	$t_i = 1$ for non-Indian players	$\sum x_i t_i \leq 4$
Wicket Keeper	$w_i = 1$ for Wicket Keepers	$\sum x_i w_i \geq 1$
Player Salary	c_i is player cost, β is team budget	$\sum x_i c_i \leq \beta$
Bowlers	$b_i = 1$ for Bowlers, All-Rounders	$\sum x_i b_i \geq 5$
Batsmen	$a_i = 1$ for Batsmen, All-Rounders	$\sum x_i a_i \geq 8$
Player Uniqueness	$\gamma_{i,j} = 1$ if the i-th player is selected for category j	$\sum_{j=1}^{5} \gamma_{i,j} \leq 1$
Non-Negativity	-	$x_i, \gamma_{i,j} \geq 0$
Integer Values	-	$x_i, \gamma_{i,j} \in \mathbb{Z}$

Table 7.1: Linear Optimisation Restrictions



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	ID	PLAYER	NATIONAl	COST	TEAM	AR	WK	BOWL	Selected?	SelBat	V1	SelBAR	V2	SelTAR	V3	SELBOAR	V4	SelBowl	V5
2	2	BB McCullum	1	325	CSK	0	1	0	0	0	0.1212	0	-20	0	-20	0	-20	0	-20
3	4	DJ Hussey	1	100	NONE	1	0	1	0	0	0.1497	0	-0.3569	0	-0.5741	0	-0.7912	0	-1.225
4	8	V Kohli	0	1270	RCB	0	0	0	0	0	0.2646	0	-0.2317	0	-0.4445	0	-0.6572	0	-1.083
5	9	JH Kallis	1	550	NONE	1	0	1	0	0	-0.04723	0	-0.2894	0	-0.3932	0	-0.497	0	-0.7046

T	U	V	W	X	Y	Z
SelNat	SelCost	SelWK	SelAR	SelBowl	SelWAVE	SelBAT
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

Figure 7.2: Screenshot of Excel Table

$WAVE$ scores for each of the five positional categories. We can see this table in Figure 7.2. We can see the constraints (and the OpenSolver UI) in Figure 7.3[2].

[2]Certain references reference the sum of various totals for the constraints (i.e. cell AC3 is the total number of international players).

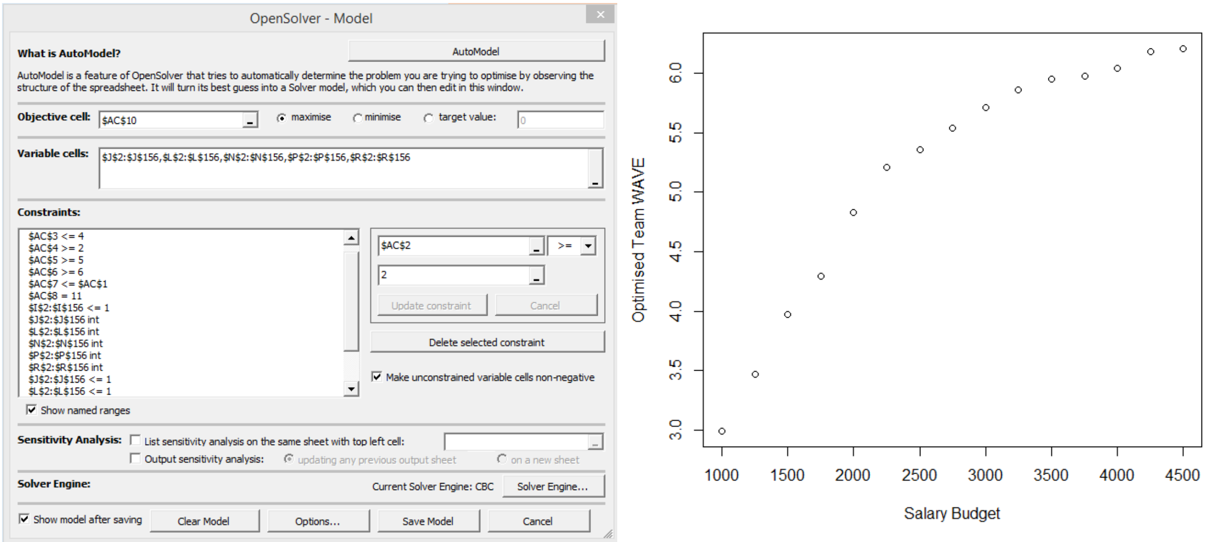


Figure 7.3: OpenSolver UI displaying constraints and optimum team scores

7.2 Implementation and Scenarios

We will now show how this process works for a range of situations. We will consider:

- the best team[3] on an unlimited (unconstrained) budget;
- the best team on a IPL budget (Rs. 60 crore);
- the best team on a realistic budget (Rs. 40 crore); and
- the hypothetical minimum spend possible to have a positive team $WAVE$ score.

The difference between what we have called a "realistic" budget and the IPL budget is that the IPL budget of Rs. 60 crore is for the entire squad. A team will not, therefore, spend all of the budget on the first eleven players. The "realistic" budget hopes to account for this by setting aside a certain amount of the cap for reserve players.

The way the optimisation process is set up means that it is best for calculating eleven players. This can be attributed to the extra rules in place regarding international players (i.e. there can be up to nine in the squad, but only four are allowed to play per game). We assume that a team would want at least two All-Rounders selected for a team. We can see in Figure 7.3 how the optimum team $WAVE$ score changes as the budget increases.

7.2.1 General Optimised Team Evaluation

Table 7.5 shows the optimised teams for the first three scenarios listed. Certain players[4] are included in every team we have created. This is likely due to the relationship between $WAVE$ score and cost. Each player costs less than Rs. 80 cr per $WAVE$ score[5]. Karanveer Singh is the best example

[3] By "team", we are considering eleven players

[4] V Sehwag, S Dhawan, WP Saha, SL Malinga, M Vohra and Karanveer Singh

[5] Not displayed, but easily calculated

ID	Player	Position	$WAVE$
23	S Badrinath	Batsman	0.0294
44	Y Venugopal Rao	Batsman	-0.0330
110	JA Morkel (Int)	Bat. AR	-0.2959
113	YV Takawale	Batsman (WK)	-0.2326
172	M Manhas	All-Rounder	-0.392
184	NV Ojha	Batsman	-0.0107
290	AC Blizzard (Int)	Batsman	-0.0250
340	KK Nair	Batsman	0.4758
341	GH Vihari	Bowler	-0.1232
411	S Gopal	Bowler	0.0645
413	Karanveer Singh	Bowler	0.5526
	Total Cost		Rs. 2.95 cr
	Total $WAVE$		0.01

Table 7.4: Least expensive side with a positive Team $WAVE$

of this, costing only Rs. 0.1 crore but contributing 0.55 wins above the average player. This rating could be inflated due to limited sample size, but it is nevertheless impressive. Notably, none of these players cost more than Rs. 10 crore, which elite players (MS Dhoni, Yuvraj Singh, etc.) demand.

We now consider the cost of the least expensive team to finish with an above-average record. We change the linear programming in Excel slightly to now minimise the cost, but ensure that the Team $WAVE > 0$, in addition to all other constraints. The team is shown in Table 7.4. Assuming our statistics are valid and correct, a (marginally) winning team can be assembled for less than Rs. 3 crore. This is much below the IPL minimum spend of Rs. 36 crore, and lacks the big-name attraction of even the 2014 Delhi Daredevils. $WAVE$, however, assures us that they would perform better. We should perhaps be sceptical, as this team relies heavily on the bowling success of Karanveer Singh (previously discussed) and the batting ability of KK Nair. Both have had excellent seasons since they were up for auction (Nair in 2013, Singh in 2014) so it is very unlikely that they will be obtainable now for the price which they were sold.

Unlimited Budget					Rs. 60 crore Budget					Rs. 37.5 crore Budget				
ID	Player	Position	$WAVE$	Cost	ID	Player	Position	$WAVE$	Cost	ID	Player	Position	$WAVE$	Cost (Rs. Lakh)
20	MS Dhoni	Batsman (WK)	0.4910	1300	20	MS Dhoni	Batsman (WK)	0.4910	1300	41	V Sehwag	Batsman	0.6139	320
21	SK Raina	Batsman	0.5685		950 41	V Sehwag	Batsman	0.6139	320	42	S Dhawan	All-Rounder	0.3029	950
32	SR Watson (Int)	Bat. AR	0.3843	1250	42	S Dhawan	All-Rounder	0.3029	950	53	WP Saha	Batsman (WK)	0.3234	220
41	V Sehwag	Batsman	0.6139	320	53	WP Saha	Batsman (WK)	0.3234	220	101	BJ Hodge (Int)	All-Rounder	0.2312	100
42	S Dhawan	All-Rounder	0.3302	950	159	KP Pietersen (Int)	Bat. AR	0.2692	200	107	SE Marsh (Int)	Batsman	0.6184	220
53	WP Saha	Batsman (WK)	0.3234	220	163	CH Gayle (Int)	Batsman	0.7788	950	111	AB De Villiers (Int)	Batsman (WK)	0.4865	750
163	CH Gayle (Int)	Batsman	0.7788	950	195	SL Malinga (Int)	Bowler	0.6913	550	195	SL Malinga (Int)	Bowler	0.6913	550
195	SL Malinga (Int)	Bowler	0.6913	550	316	SP Narine (Int)	Bowler	0.8014	950	340	KK Nair	Batsman	0.4758	75
316	SP Narine (Int)	Bowler	0.8014	950	340	KK Nair	Batsman	0.4758	75	346	M Vohra	Batsman	0.5409	400
346	M Vohra	Batsman	0.5409	400	346	M Vohra	Batsman	0.5409	400	399	AR Patel	Bowler	0.3744	75
413	Karanveer Singh	Bowler	0.5526	10	413	Karanveer Singh	Bowler	0.5526	10	413	Karanveer Singh	Bowler	0.5526	10
Total Cost				Rs. 78.50 Cr	Total Cost				Rs. 59.25 cr.	Total Cost				Rs. 36.70 cr.
Total $WAVE$				6.0490	Total $WAVE$				5.8412	Total $WAVE$				5.2113

Table 7.5: Optimised (best possible teams) under constraints

Chapter 8

Discussion

In this chapter, we will discuss the methodology we have used to get our results, issues arising from said methodology and alternative methods. We will also discuss further steps that may be taken and possible areas for improvement. We will discuss the Batting Scores, Bowling Scores and WAVE Scores in turn.

8.1 Assumptions

8.1.1 General Assumptions and Comments

Firstly, we need to make some assumptions about the field of play. We assume that all players play on the same sized field. A larger outfield may reduce the probability of fours and sixes. In addition, we assume that altitude has no effect (higher altitude means less air resistance so the ball could potentially travel further/faster through the air). Equally, pitch conditions may vary between stadia. It could be argued that these may balance out over the course of the year, but if one ground provides a significant advantage, the home side would have the benefit of that for more games. We do not know if these assumptions are appropriate without more investigation, but they are necessary to proceed.

One thing that would greatly improve this work is if ball-by-ball data was freely available. There is obviously a proprietary issue to this, as the data compilation would be relatively time expensive. A sports statistics supplier like Opta would have this data, but their business model relies on selling it to teams or to the gambling industry. Ball-by-ball data would enable us to expand the batting and bowling model to specific situations (e.g. to see how a particular bowler does against a left or right-handed batsman, examine a bowlers effectiveness against openers vs the tail-end, etc.).

In addition, we are assuming that all players are of equal fielding ability, so this does not contribute to player evaluation. Whilst this is obviously not the case in real life, it can be argued that it is a good (and necessary) assumption to make, as a player would not be playing at this standard of cricket if they could not catch the ball. We also assume that Wicket Keeping skill is equal amongst players who ever play in that position.

A shrewd reader may look at the Team $WAVE$ table (6.3) and comment on the sizeable variation between players, even within teams whose performance we seem to have predicted reasonably accurately. Michael Lewis, author of Moneyball, also noticed this in the methods used by Beane et. al at the Oakland Athletics, and expressed his doubts: “My problem can be simply put: every player is different. Every player must be viewed as a special case. The sample size is always one. [Beane’s] answer is equally simple: baseball players follow similar patterns, and these patterns are etched into the record books. Of course, every so often a player may fail to embrace his statistical destiny, but on a team of twenty-five players the statistical aberrations will tend to cancel each other out"[1] (Lewis, 2003). Beane's opinion on baseball statistics is relevant to cricket in this regard, and provides a sufficient explanation for the variation.

We also assume for both $BATS$ and $BOWLS$ that wides and no balls are not bowled. This is an unrealistic assumption, but it seemed improper to have an arbitrary simulation for it for the batsmen, and between bowlers, there did not seem to be too much difference in the numbers of wides and no balls bowled. This was made as a simplification and further research could investigate their effect on the game and whether they should be included. Musson (2015) argued that no balls (in particular ones leading to a "free hit" - where the bowler's foot is in front of the crease) should be factored in. These, however, should factor into the bowler's economy and score distribution.

Crucially, we also assume that past performance from a player is indicative of future performance. This is not always the case, and for a small sample size could present a large error. In addition, we assume that there is no regression to the mean. This is unlikely to be a good assumption, and needs further research. Schall and Smith (2000) found that it was an important factor to consider in baseball, so it is likely to also be important in cricket.

Buttler (2015) commented on something briefly touched on in the introduction. A player's salary and worth to a team is often measured in far more than their ability on the field. "Icon" players, as they were known at the start of the IPL, are as much of a marketing and brand asset as they are a cricketer. Merchandise alone may pay a player's salary, and large auction acquisitions could be seen as statements of intent for the upcoming year. We cannot calculate the effect of this based on their on-field statistics, so the value of a player's mass appear is to be determined by the highest bidder.

8.1.2 Batting Score

We assume that a batsman faces balls until he is bowled out, and there is no swapping ends of batsmen. This is obviously not how the game of cricket works, but the assumption does simplify the coding of the problem. In addition, the only difference (as we are assuming all batsmen are the same) is that the zero-inflation for the second batsman will happen earlier in the innings than what is modelled. We also assume that a batsman does not play significantly better or worse against different types of bowler, and that the handedness of the bowler has no impact against said batsmen.

8.1.3 Bowling Score

For our model, we have not factored in the change of batsmen (facing the bowl) after odd numbered runs and at the end of the over. This may have a slight effect on the probability of wickets at the beginning of the simulated innings. The assumption, therefore, is that a batsman bats continuously until they are out. For example, at the start of an innings, we inflate the zero probability of one batsman for their first ball. If they then go on to score one run, Batsman 2 will then be facing his first ball, and so that ball probability should be from the zero-inflated distribution. Under this version of the model, this does not happen. If a wicket falls early in the game (this is only an issue for the first batting pair), this may affect the run rate as the game situation will have changed. However, as this will be the same for all bowlers, this should affect them all in the same way - a case of two wrongs actually making a right. It is believed that this assumption is minor, and as long as the number of iterations in the simulation is large enough, will not affect the results.

[1] This is cited in Singh's book on mathematics in The Simpsons, specifically referencing the time when Lisa became the manager of Bart's baseball team and used sabermetric principles (Singh, 2013).

We also assume that the bowler is bowling against the average batsman. The effects of this may be most significant towards the end of an innings, where the batsmen are of lesser ability. We know from section 4.3 that there are a large number of very poor batsmen (who we speculate to mainly be bowlers). In this case, the wicket loss adjustments for the over and the run rate adjustments for the game situation may not adequately reflect a true game. There is no current alternative without more data. However, this would be the same for all bowlers, so any affect should offset with enough simulations.

8.1.4 $WAVE$

Our weightings for the WAVE score have been assigned arbitrarily, based on the work of Aparna et al. (2012). They have been chosen to reflect and balance All-Rounder ability across the three AR categories, the work done by Aparna, and also personal choice and opinion. The weightings selected are certainly up for debate, and should be tailored on a case-by-case basis to exactly what a team might want from an All-Rounder: do they want a Pietersen-type AR, who is chiefly a batter who can bowl if needed, or vice versa?

When we calculate the $WAVE$ score for a team in Chapter 6.2, we assume that the same eleven players play in every game across the year. This is not actually the case due to injuries, fatigue, international commitments, etc. The smallest number of players that appeared for a side in 2014 was 17 for Sunrisers Hyderabad. It may, therefore, be inappropriate to just select 11 players to evaluate a team's $WAVE$ score. A better alternative in practice may be to maximise the $WAVE$ score for the entire squad. The issue with this is that the notion of "Wins Above Average" depends on eleven players playing. It would not make sense to use a $WAVE$ score to project wins using a squad score (unless we are breaking the squad down into what proportion of games people would play, which is impossible to know in advance).

8.2 Potential Biases in Results

It is important to acknowledge any biases that may have arisen throughout the process. Opening batsmen typically do not make it to the end of an innings, where the run rate will accelerate. As a result, their scores may reflect the position in which they are utilised by the team. A player who consistently ends up batting in the closing overs would likely end up with a higher strike rate (and potentially higher $BATS$ and $WAVE$) than the opening batsmen, even if the opener is the more gifted player. The same argument is true for bowlers. A team usually uses their best bowlers in the opening six and closing four overs, where a batting team will be batting most aggressively. It may then appear that they concede more runs (lowering their $BOWLS$) than another bowler who usually bowls the middle ten overs, despite being a superior player.

As mentioned above, we assumed that ground sizes were all the same. Daniel Musson highlighted that this is not the case, and the size of the ground should have a significant effect on the runs scored and the wickets lost at a particular venue - especially for a team that play half of their games there. It has not been possible to find pitch size for various stadia to standardise for this.

8.3 Alternative Steps & Potential Improvements

Ball-by-ball data would enable us, when modelling bowlers, to obtain the exact number of deliveries bowled between wickets. It may be the case that a bowler is particularly effective early on against a batsman (before a batsman can get used to the bowler's style, for instance), but not so effective once a batsman is familiar to the style of bowling, or as the bowler gets tired. In any case, ball-by-ball data would enable us to find a better way of simulating the number of balls bowled between wickets than the simple inverse strike rate formula used.

Currently, the $WAVE$ score for a player is difficult to calculate and only after extensive simulation can we be certain of any results. Calculation of $BATS$ and $BOWLS$ for all players took well in excess of 10 hours. This limits its usefulness for a broadcaster, casual observer or spectator, who wishes to know how well a player is doing over the course of the season (or even extrapolate the performance of a game). This could be improved by streamlining the existing code, or by using a language which may produce quicker results.

A player's $WAVE$ score is currently liable to take on extreme values due to a limited data set. Approaching this problem using a Bayesian model could help limit this. A possible technique would be to assume all players are "average" batsmen and bowlers until proven otherwise.

In my discussion with Daniel Musson (2015), he highlighted that a player's fielding score should be considered in their $WAVE$, and a statistic like "Runs Saved" could be developed to determine their contribution. He suggested that whilst fielding is not as important to evaluating a player as batting and bowling, a player may not be selected based on their fielding skill. This requires much more in-depth data than it would be possible to source for this dissertation, but his suggestion is good one. In addition, he was sceptical of how we have defined "All-Rounders", and instead suggested that future work should look at the number of overs bowled per game, and use this to assign the categories.

8.4 Future Research

To account for a fielding score, we would have to introduce and develop something similar to what is described by Winston (2009) in "7. Evaluating Fielders". This details the error in measuring fielding ability as the number of errors a player makes, and instead suggests alternatives. This could potentially be applicable to cricket, but it is not something that has been considered. The effect of an individual player's fielding will have on the overall result of a game or season is arguably minimal compared to their batting and bowling ability. Runs are scored much more frequently in cricket than in baseball, so one mistake is unlikely to cost a team. However, it may perhaps be worth investigating, for academic reasons if nothing else.

One thing which may be of interest to the reader is calculating exactly how much a player is worth, based on their $WAVE$ score. This was intended to be one of the things to consider in this work, but we have opted for a different route, using the values given by the market. The value of a player could potentially be calculated from the scores we have assigned, but this price would need to factor in nationality, player category, commercial appeal and other benefits a player may bring to a club. This work has been developed in Kansal et al. (2014), but it would be interesting to see if the $WAVE$ score can be used as a determinant of player salary. In the Appendix, Figure A.1.2 shows a plot of $WAVE$ against salary, which could be of interest.

For optimisation, we use the career statistics of players up until the start of the 2015 season. This involves up to seven seasons of data for certain players. One thing we have not accounted for is the regression (or improvement) of players over time. Swiroff et al. (1989) and James (1982) independently concluded that baseball players reach their peak at around 26 or 27 years of age. If we assume that something similar happens in cricket, we would need to factor in player depreciation over the course of their career. We could also improve the optimisation process with the exact salaries for each player.

A flaw in this work is that we can only account for, evaluate and consider players that have appeared in the IPL. However, T20 leagues exist all over the planet - most notably the Australian "Big Bash" League, the English "T20 Bash" and the T20 World Cup. There are numerous smaller regional competitions between lower level teams. All of these players could potentially be targets in the IPL Auction (should the players declare themselves eligible), but we cannot currently optimise them. This is best seen with a player like Trent Boult (NZ). Boult, a bowler, has accumulated a strike rate of 21.6 and an economy of 7.7 (both extremely respectable[2]) and went in the 2015 auction for Rs. 3.8 crore - the second highest non-Indian player that year. However, his performances in domestic cricket cannot be accounted for as we only look at IPL data. There is currently no way to predict how well a player's skill will translate from one league to another, and similar work to Major League Equivalents (James, 1985)[3], which predicts MLB ability from their statistics in lower baseball leagues, would need to be done before it can be accounted for.

[2] Quickly considering Table 5.5, he would probably have a score close to M. Muralitharan.
[3] Referenced in Winston (2009).

Chapter 9

Conclusion

The principal objective of this dissertation was to develop an objective way of measuring player performance across positions and teams. In the course of this work, we have demonstrated a method for the evaluation of a player's batting and bowling ability and how to evaluate a player's impact on a team as an All-Rounder. This is through the development of $WAVE$, or Wins Above Average. We can now compare exactly how much a player has contributed to his team's on field success.

We have used a range of statistical techniques including, but not limited to, hypothesis testing, expectation-maximisation, Monte Carlo methods, optimisation and log-odds adjustments. We have also explored methods in optimising a team under constraints, either imposed by league rules or by a team themselves, and suggested alternative and future work that can be undertaken to enhance the practicality, applicability and reliability of results.

As previously mentioned, our optimisation model selects eleven players to form a team. The methods and constraints used can be adapted to meet the needs of an individual team, to see how they could best optimise their team given their current line up. For instance, for the Chennai Super Kings, having retained MS Dhoni, they do not need to purchase another Wicket Keeper. Equally, some of their salary cap will be reduced so that needs to changed (and so on and so forth for every player so retained). We have not carried this out, as this is very much dependent on having a large number of players available for auction (i.e. at the three-yearly super auction). The only way this could be used effectively for any other auction is by developing the work done mentioned in Section 8.4 concerning projecting IPL statistics from other leagues.

It is my hope that $WAVE$, or a method similar, can be used to evaluate players in T20 cricket. Player comparison across positions is not only useful for teams, but also of interest to fans. For teams, however, in choosing whether or not to purchase a player, it is unwise to rely totally on statistics, just as it is unwise to rely on the "eye-test". Musson (2015) said that to coaches, statistics aren't as important as the player's skill and mental state in determining a line-up, but also said that cricket analysis is "massively data-driven" and it's use depends upon the coach making the decision[1]. However, as Dayn Perry, noted columnist for Baseball Prospectus, said "having to choose between statistical methods and regular scouting is like choosing between beer and tacos". Both are good individually, but it is much better to have both. If nothing else, a player's $WAVE$ score can be used to settle friendly arguments in cricket clubs across the country.

This study outlines a new way of thinking about player evaluation and, hopefully, provides a useful platform for further research to build on in the future.

[1]Peter Moores, current head coach of England is known for a very statistics-driven approach.

Bibliography

Agresti, A. (2010). *Analysis of Ordinal Categorical Data.* John Wiley & Sons, New Jersey, 2 edition.

Ahmed, F., Deb, K., and Jindal, A. (2011). Evolutionary multi-objective optimization and decision making approaches to cricket team selection. Technical Report 2011016, KanGAL, Indian Institute of Technology, Kanpur, 208016, India.

Aparna, T. C., Athul, U., and Vivek, R. S. (2012). Dynamically determining the optimal set of players and their bid prices. Revenue Management, Bharathidasan Institute of Technology.

Brettenny, W. (2010). Integer optimisation for the selection of a fantasy league cricket team. Master's thesis, Port Elizabeth, South Africa.

Buttler, J. (2015). Personal communication.

Cricinfo Staff (2008). Big business and Bollywood grab stakes in IPL. `http://www.espncricinfo.com/ipl/content/story/333193.html`.

Darowski, A. (2012). Wins above average: Taking the "replacement" out of war. `http://www.beyondtheboxscore.com/2012/4/24/2971859/wins-above-average-taking-the-replacement-out-of-war`.

Elderton, W. P. (1945). Cricket scores and some skew correlation distributions: (an arithmetical study). *Journal of the Royal Statistical Society*, 108(1/2):1–11.

Elderton, W. P. and Elderton, E. M. (1912). *Primer of statistics.* London, A. and C. Black.

ESPN (2015). Cricinfo. `www.espncricinfo.com`.

Fangraphs (2015). WAR. `www.fangraphs.com/library/misc/war`.

India Today (2015). No Pakistan player in IPL till ties are resolved: PCB. `http://indiatoday.intoday.in/story/indian-premier-league-pakistan-cricket-board-cricket-bilateral-ties/1/417448.html`.

IPL, BCCI (2014). Glenn Maxwell. `http://www.iplt20.com/teams/kings-xi-punjab/squad/282/Glenn-Maxwell/`.

IPL, BCCI (2015a). IPL Player Auction. `www.iplt20.com/playerauction`.

IPL, BCCI (2015b). IPL Statistics. `www.iplt20.com/stats`.

IPL Desk in Mumbai (2013). Pepsi ipl 2014 player regulations. `http://www.iplt20.com/news/2013/announcements/4248/pepsi-ipl-2014-player-regulations`.

James, B. (1982). *The Bill James Baseball Abstract*. Balantine, New York.

Kansal, P., Kumar, P., Arya, H., and Methaila, A. (2014). Player valuation in indian premier league auction using data mining technique. In *Contemporary Computing and Informatics (IC3I), 2014 International Conference on*, pages 197–203.

Langdale, G. R. (1945). Discussion on "Cricket Scores". *Journal of the Royal Statistical Society*, 108(1/2):26–27.

Lewis, M. M. (2003). *Moneyball: the art of winning an unfair game*. W. W. Norton.

Marylebone Cricket Club (2013). *The Laws of Cricket*. Marylebone Cricket Club, 2000 code 5th edition.

Mason, A. J. (2012). OpenSolver - an open source add-in to solver linear and integer programmes in Excel. In Klatter, D., Lathi, H.-J., and Schmedders, K., editors, *Operations Research Proceedings 2011*, Operations Research Proceedings, pages 401–406. Springer Berlin Heidelberg.

Morgan-Mar, D. (2014). Bowling strategy and tactics. `http://www.dangermouse.net/cricket/bowlstrategy.html`.

Musson, D. (2015). Personal communication.

Norman, J. and Clarke, S. (2010). Optimal batting orders in cricket. *Journal of the Operational Research Society*, 61(6):980–986.

Parker, D., Burns, P., and Natarajan, H. (2008). Player valuations in the indian premier league. Technical report, Frontier Economics, London.

Perera, H. and Swartz, T. B. (2013). Resource estimation in T20 cricket. *IMA J Management Math*, 24(3):337–347.

Petersen, C., Pyne, D. B., Portus, M. J., and Dawson, B. (2008). Analysis of Twenty/20 Cricket performance during the 2008 Indian Premier League. *International Journal of Performance Analysis in Sport*, 8(3):63–69.

Ramprakash, M. (2012). Has twenty20 changed cricket? `http://www.bbc.co.uk/sport/0/cricket/19624160`.

SABR (2014). A guide to sabermetric research. `http://sabr.org/sabermetrics`.

Schall, E. M. and Smith, G. (2000). Do baseball players regress towards the mean? *The American Statistician*, 54:231–235.

Singh, S. (2013). *The Simpsons and their Mathematical Secrets*. Bloomsbury.

Swiroff, S., Hirdt, S., and Hirdt, P. (1989). *The 1989 Elias Baseball Analyst*. Collier Books, New York.

Varma, D. (2014). IPL 2014: Chris Gayle, Dale Steyn among flops of the season. `http://www.cricketcountry.com/articles/ipl-2014-chris-gayle-dale-steyn-among-flops-of-the-season-144216`.

Winston, W. L. (2009). *Mathletics: How Gamblers, Managers, and Sports Enthusiasts Use Mathematics in Baseball, Basketball, and Football*. Princeton University Press.

Wood, G. H. (1945). Cricket scores and geometric progression. *Journal of the Royal Statistical Society*, 108(1/2):12–40.

Appendix A

Supplementary Materials

A.1 Graphs and Tables (not included in main body)

Figure A.1.1: A Diagram of the Cricket Field with some field positions (from the points of view of a right-handed batsmen. Source: Wikipedia.)

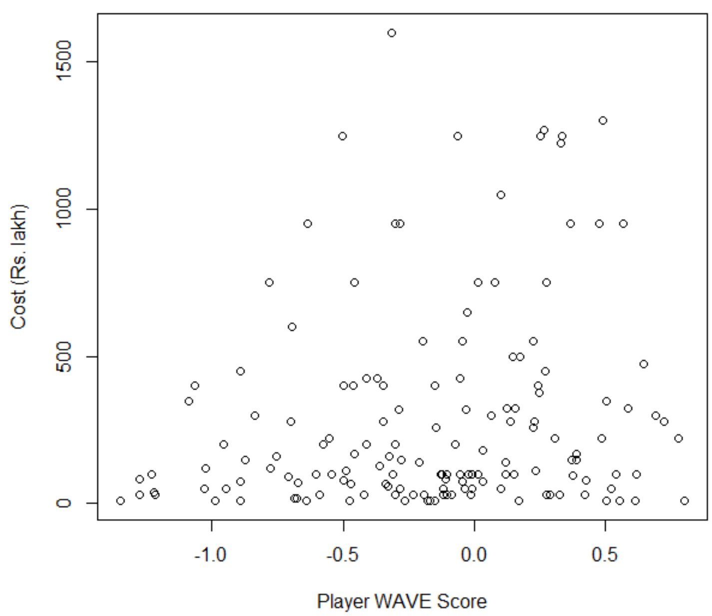


Figure A.1.2: Plot of Salary against calculated $WAVE$ score

	Category														
	Batsman			Batting All-Rounder			True All-Rounder			Bowling All-Rounder			Bowler		
Rank	ID	Player	$WAVE$	ID	Player	$WAVE$	ID	Player	$WAVE$	ID	Player	$WAVE$	ID	Player	$WAVE$
1	163	CH Gayle	0.811	32	SR Watson	0.416	32	SR Watson	0.310	101	BJ Hodge	0.248	99	Sohail Tanvir	1.015
2	321	DA Miller	0.670	159	KP Pietersen	0.260	101	BJ Hodge	0.215	32	SR Watson	0.204	316	SP Narine	0.932
3	41	V Sehwag	0.669	21	SK Raina	0.223	159	KP Pietersen	0.129	90	MF Maharoof	0.097	102	Umar Gul	0.699
4	32	SR Watson	0.664	101	BJ Hodge	0.182	27	Yuvraj Singh	0.125	362	DJG Sammy	0.095	243	CK Langeveldt	0.698
5	107	SE Marsh	0.617	27	Yuvraj Singh	0.163	362	DJG Sammy	0.093	27	Yuvraj Singh	0.086	277	Shakib Al Hasan	0.675
6	18	ML Hayden	0.613	56	ST Jayasuriya	0.150	21	SK Raina	0.079	61	SM Pollock	0.051	364	MM Sharma	0.644
7	159	KP Pietersen	0.565	31	YK Pathan	0.121	61	SM Pollock	0.046	159	KP Pietersen	-0.002	195	SL Malinga	0.642
8	21	SK Raina	0.560	362	DJG Sammy	0.092	31	YK Pathan	0.012	322	Azhar Mahmood	-0.047	391	BW Hilfenhaus	0.596
9	20	MS Dhoni	0.502	47	RG Sharma	0.087	56	ST Jayasuriya	0.004	21	SK Raina	-0.066	311	JP Faulkner	0.593
10	304	F du Plessis	0.499	46	A Symonds	0.051	90	MF Maharoof	-0.011	31	YK Pathan	-0.098	95	DW Steyn	0.507

Table 1.3: Top Ranked Player in each position

A.2 Flowcharts of Simulation Methods

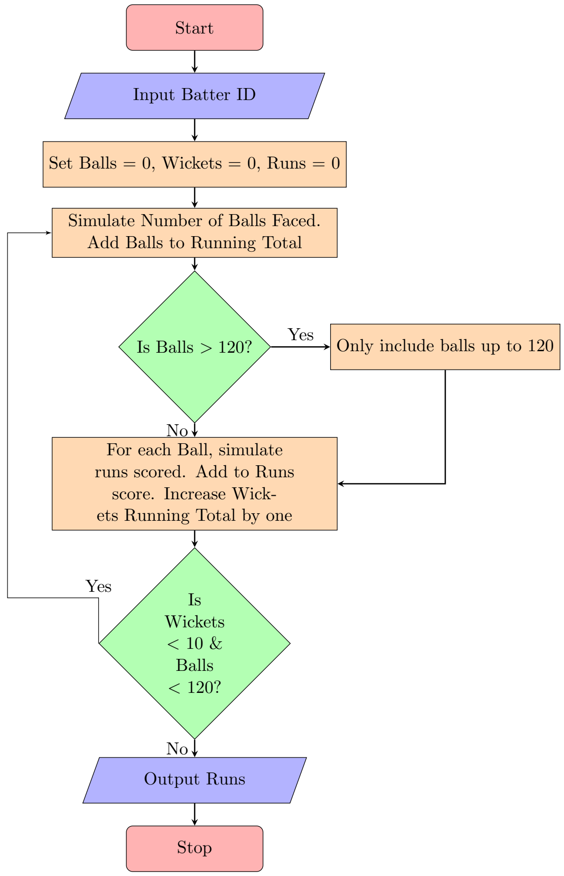


Figure A.2.1: One iteration of the original algorithm for a simulated batting innings

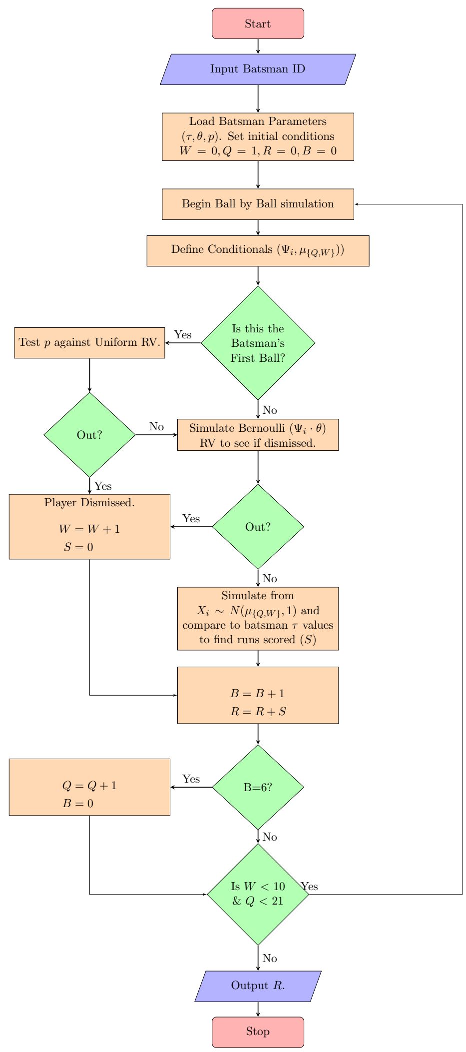


Figure A.2.2: Flowchart showing the updated batting simulation algorithm for one innings

Appendix B

Complete Player $WAVE$ scores

ID	NAME	BATS	BOWLS	Bat	Batsman.Score	BatAR	BattingAR.Score	AR	AR.Score	BowlAR	BowlingAr.Score	Bowl	Bowler.Score
1	SC Ganguly	-0.43	-2.27	1	-0.05	1	-0.19	1	-0.25	1	-0.31	1	-0.43
2	BB McCullum	0.95	-7.00	1	0.12	0	0.00	0	0.00	0	0.00	0	0.00
3	RT Ponting	-6.59	-7.00	1	-0.84	0	0.00	0	0.00	0	0.00	0	0.00
4	DJ Hussey	1.18	-6.49	1	0.15	1	-0.36	1	-0.57	1	-0.79	1	-1.23
5	Mohammad Hafeez	-6.57	1.27	0	0.00	0	0.00	0	0.00	0	0.00	1	0.20
6	R Dravid	0.96	-7.00	1	0.12	0	0.00	0	0.00	0	0.00	0	0.00
7	W Jaffer	-1.04	-7.00	1	-0.13	0	0.00	0	0.00	0	0.00	0	0.00
8	V Kohli	2.08	-5.77	1	0.26	1	-0.23	1	-0.44	1	-0.66	1	-1.08
9	JH Kallis	-0.37	-3.70	1	-0.05	1	-0.29	1	-0.39	1	-0.50	1	-0.70
10	CL White	1.16	-7.00	1	0.15	0	0.00	0	0.00	0	0.00	0	0.00
11	MV Boucher	-0.46	-7.00	1	-0.06	0	0.00	0	0.00	0	0.00	0	0.00
12	B Akhil	-6.41	-4.19	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.84
13	AA Noffke	-7.00	-5.18	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.03
14	P Kumar	-6.81	-2.69	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.56
15	Z Khan	-6.92	-1.59	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.35
16	SB Joshi	-7.00	-6.57	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.29
17	PA Patel	-1.64	-7.00	1	-0.21	0	0.00	0	0.00	0	0.00	0	0.00
18	ML Hayden	4.68	-7.00	1	0.60	0	0.00	0	0.00	0	0.00	0	0.00
19	MEK Hussey	3.04	-7.00	1	0.39	0	0.00	0	0.00	0	0.00	0	0.00
20	MS Dhoni	3.86	-7.00	1	0.49	0	0.00	0	0.00	0	0.00	0	0.00
21	SK Raina	4.47	-2.35	1	0.57	1	0.20	1	0.05	1	-0.11	1	-0.42
22	JDP Oram	-5.92	-5.49	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.08
23	S Badrinath	0.23	-7.00	1	0.03	0	0.00	0	0.00	0	0.00	0	0.00
24	K Goel	-4.64	-6.94	1	-0.59	0	0.00	0	0.00	0	0.00	1	-1.35
25	JR Hopes	2.07	-5.88	1	0.26	1	-0.24	1	-0.46	1	-0.67	1	-1.10
26	KC Sangakkara	1.47	-7.00	1	0.19	0	0.00	0	0.00	0	0.00	0	0.00
27	Yuvraj Singh	2.16	-0.83	1	0.28	1	0.12	1	0.06	1	-0.01	1	-0.14
28	SM Katich	1.41	-7.00	1	0.18	0	0.00	0	0.00	0	0.00	0	0.00
29	IK Pathan	-2.49	-2.22	1	-0.32	1	-0.36	1	-0.38	1	-0.40	1	-0.44
30	T Kohli	-6.99	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
31	YK Pathan	2.89	-2.23	1	0.37	1	0.08	1	-0.04	1	-0.16	1	-0.41
32	SR Watson	4.58	0.07	1	0.58	1	0.38	1	0.30	1	0.21	1	0.04
33	M Kaif	-5.02	-7.00	1	-0.64	0	0.00	0	0.00	0	0.00	0	0.00
34	DS Lehmann	-6.03	-7.00	1	-0.77	0	0.00	0	0.00	0	0.00	0	0.00
35	RA Jadeja	-0.50	-1.10	1	-0.06	1	-0.12	1	-0.14	1	-0.16	1	-0.21
36	M Rawat	-6.92	-7.00	1	-0.88	0	0.00	0	0.00	0	0.00	0	0.00
37	D Salunkhe	-3.92	-6.32	1	-0.50	1	-0.77	1	-0.88	1	-1.00	1	-1.23
38	SK Warne	-6.57	-0.41	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.12
39	SK Trivedi	-6.98	-1.67	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.36
40	G Gambhir	2.17	-7.00	1	0.28	0	0.00	0	0.00	0	0.00	0	0.00
41	V Sehwag	4.82	-6.47	1	0.61	1	-0.05	1	-0.34	1	-0.63	1	-1.20
42	S Dhawan	2.14	1.67	1	0.27	1	0.29	1	0.30	1	0.31	1	0.33
43	AC Gilchrist	3.65	-7.00	1	0.47	0	0.00	0	0.00	0	0.00	0	0.00
44	Y Venugopal Rao	-0.26	-6.55	1	-0.03	1	-0.48	1	-0.67	1	-0.86	1	-1.25
45	VVS Laxman	-3.68	-7.00	1	-0.47	0	0.00	0	0.00	0	0.00	0	0.00
46	A Symonds	2.98	-4.24	1	0.38	1	-0.05	1	-0.23	1	-0.42	1	-0.79
47	RG Sharma	2.87	-2.16	1	0.37	1	0.09	1	-0.03	1	-0.15	1	-0.39
48	SB Styris	-4.82	-3.82	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.76

ID	NAME	BATS	BOWLS	Bat	Batsman.Score	BatAR	BattingAR.Score	AR	AR.Score	BowlAR	BowlingAr.Score	Bowl	Bowler.Score
49	AS Yadav	-6.02	-7.00	1	-0.77	0	0.00	0	0.00	0	0.00	0	0.00
50	SB Bangar	-6.70	-5.89	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.16
51	WPUJC Vaas	-6.50	-0.45	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.13
52	RP Singh	-7.00	-1.34	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.30
53	WP Saha	2.54	-7.00	1	0.32	0	0.00	0	0.00	0	0.00	0	0.00
54	LR Shukla	-3.95	-2.88	1	-0.50	1	-0.53	1	-0.54	1	-0.55	1	-0.57
55	L Ronchi	-6.71	-7.00	1	-0.85	0	0.00	0	0.00	0	0.00	0	0.00
56	ST Jayasuriya	3.93	-2.62	1	0.50	1	0.14	1	-0.01	1	-0.17	1	-0.47
57	DJ Thornely	-6.33	5.80	0	0.00	0	0.00	0	0.00	0	0.00	1	1.06
58	RV Uthappa	2.39	-7.00	1	0.30	0	0.00	0	0.00	0	0.00	0	0.00
59	PR Shah	-5.57	-7.00	1	-0.71	0	0.00	0	0.00	0	0.00	0	0.00
60	AM Nayar	-3.31	-4.85	1	-0.42	1	-0.61	1	-0.70	1	-0.78	1	-0.94
61	SM Pollock	0.27	0.43	1	0.03	1	0.05	1	0.06	1	0.07	1	0.08
62	Harbhajan Singh	-5.33	1.10	0	0.00	0	0.00	0	0.00	0	0.00	1	0.17
63	S Chanderpaul	-6.45	-7.00	1	-0.82	0	0.00	0	0.00	0	0.00	0	0.00
64	LRPL Taylor	0.09	-7.00	1	0.01	0	0.00	0	0.00	0	0.00	0	0.00
65	DPMD Jayawardene	1.24	-7.00	1	0.16	0	0.00	0	0.00	0	0.00	0	0.00
66	S Sohal	-0.08	-7.00	1	-0.01	0	0.00	0	0.00	0	0.00	0	0.00
67	B Lee	-6.68	-2.93	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.60
68	PP Chawla	-6.27	-1.28	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.28
69	WA Mota	-6.72	-3.07	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.63
70	Kamran Akmal	3.54	-7.00	1	0.45	0	0.00	0	0.00	0	0.00	0	0.00
71	Shahid Afridi	-5.63	-0.80	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.19
72	DJ Bravo	-0.81	-0.69	1	-0.10	1	-0.12	1	-0.12	1	-0.13	1	-0.14
73	MA Khote	-6.34	-4.40	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.88
74	A Nehra	-6.99	-1.72	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.37
75	GC Smith	0.01	-7.00	1	0.00	0	0.00	0	0.00	0	0.00	0	0.00
76	Pankaj Singh	-7.00	-6.43	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.26
77	RR Sarwan	-3.92	-7.00	1	-0.50	0	0.00	0	0.00	0	0.00	0	0.00
78	S Sreesanth	-7.00	-1.55	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.34
79	VRV Singh	-7.00	-5.52	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.09
80	SS Tiwary	-0.97	-7.00	1	-0.12	0	0.00	0	0.00	0	0.00	0	0.00
81	DS Kulkarni	-6.65	-0.41	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.12
82	AB Agarkar	-5.48	-4.92	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.97
83	M Kartik	-6.29	-2.83	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.58
84	I Sharma	-6.99	-2.33	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.49
85	R Vinay Kumar	-6.24	-2.97	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.60
86	Shoaib Malik	-5.42	-5.68	1	-0.69	0	0.00	0	0.00	0	0.00	1	-1.11
87	MK Tiwary	-1.16	-6.95	1	-0.15	1	-0.58	1	-0.77	1	-0.96	1	-1.33
88	KD Karthik	1.08	-7.00	1	0.14	0	0.00	0	0.00	0	0.00	0	0.00
89	R Bhatia	-6.52	-1.60	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.35
90	MF Maharoof	-2.74	2.21	1	-0.35	1	-0.07	1	0.05	1	0.17	1	0.40
91	VY Mahesh	-6.99	-1.96	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.42
92	TM Srivastava	-6.99	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
93	AM Rahane	0.86	6.96	1	0.11	1	0.56	1	0.75	1	0.94	1	1.33
94	B Chipli	-3.49	-7.00	1	-0.44	0	0.00	0	0.00	0	0.00	0	0.00
95	DW Steyn	-6.86	2.28	0	0.00	0	0.00	0	0.00	0	0.00	1	0.39
96	DB Das	-3.29	-7.00	1	-0.42	0	0.00	0	0.00	0	0.00	0	0.00

ID	NAME	BATS	BOWLS	Bat	Batsman.Score	BatAR	BattingAR.Score	AR	AR.Score	BowlAR	BowlingAr.Score	Bowl	Bowler.Score
97	MK Pandey	0.60	-7.00	1	0.08	0	0.00	0	0.00	0	0.00	0	0.00
98	SA Asnodkar	1.50	-7.00	1	0.19	0	0.00	0	0.00	0	0.00	0	0.00
99	Sohail Tanvir	-6.38	5.67	0	0.00	0	0.00	0	0.00	0	0.00	1	1.04
100	Salman Butt	2.04	-7.00	1	0.26	0	0.00	0	0.00	0	0.00	0	0.00
101	BJ Hodge	0.94	1.72	1	0.12	1	0.20	1	0.23	1	0.26	1	0.33
102	Umar Gul	-5.29	4.23	0	0.00	0	0.00	0	0.00	0	0.00	1	0.77
103	AB Dinda	-7.00	-2.16	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.46
104	HH Gibbs	0.15	-7.00	1	0.02	0	0.00	0	0.00	0	0.00	0	0.00
105	DNT Zoysa	-6.83	-5.63	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.11
106	D Kalyankrishna	-7.00	-6.71	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.32
107	SE Marsh	4.86	-7.00	1	0.62	0	0.00	0	0.00	0	0.00	0	0.00
108	SP Fleming	-0.18	-7.00	1	-0.02	0	0.00	0	0.00	0	0.00	0	0.00
109	S Vidyut	0.27	-2.31	1	0.03	1	-0.14	1	-0.21	1	-0.29	1	-0.44
110	JA Morkel	-2.19	-1.64	1	-0.28	1	-0.30	1	-0.30	1	-0.31	1	-0.33
111	AB de Villiers	3.82	-7.00	1	0.49	0	0.00	0	0.00	0	0.00	0	0.00
112	Misba	-1.20	-7.00	1	-0.15	0	0.00	0	0.00	0	0.00	0	0.00
113	YV Takawale	-1.83	-7.00	1	-0.23	0	0.00	0	0.00	0	0.00	0	0.00
114	RR Raje	-6.99	-4.40	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.88
115	PJ Sangwan	-7.00	-3.53	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.71
116	Mohammad Asif	-7.00	-5.41	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.07
117	GD McGrath	-7.00	-0.46	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.13
118	Joginder Sharma	-6.25	-4.77	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.95
119	MS Gony	-6.66	-3.33	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.68
120	M Muralitharan	-7.00	0.29	0	0.00	0	0.00	0	0.00	0	0.00	1	0.01
121	M Ntini	-6.87	-1.29	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.29
122	DT Patil	-6.46	-7.00	1	-0.82	0	0.00	0	0.00	0	0.00	0	0.00
123	A Kumble	-7.00	1.41	0	0.00	0	0.00	0	0.00	0	0.00	1	0.22
124	S Anirudha	-4.63	-7.00	1	-0.59	0	0.00	0	0.00	0	0.00	0	0.00
125	MM Patel	-6.99	-0.65	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.17
126	CK Kapugedera	-6.90	-7.00	1	-0.88	0	0.00	0	0.00	0	0.00	0	0.00
127	A Chopra	-6.50	-7.00	1	-0.83	0	0.00	0	0.00	0	0.00	0	0.00
128	T Taibu	-5.26	-7.00	1	-0.67	0	0.00	0	0.00	0	0.00	0	0.00
129	J Arunkumar	-6.42	-7.00	1	-0.82	0	0.00	0	0.00	0	0.00	0	0.00
130	DB Ravi Teja	-2.45	-3.24	1	-0.31	1	-0.43	1	-0.48	1	-0.53	1	-0.63
131	PP Ojha	-7.00	-0.71	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.18
132	U Kaul	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
133	TM Dilshan	0.10	-5.35	1	0.01	1	-0.37	1	-0.53	1	-0.69	1	-1.01
134	A Mishra	-6.47	0.85	0	0.00	0	0.00	0	0.00	0	0.00	1	0.12
135	AD Mascarenhas	-6.64	2.95	0	0.00	0	0.00	0	0.00	0	0.00	1	0.52
136	NK Patel	-2.65	-7.00	1	-0.34	0	0.00	0	0.00	0	0.00	0	0.00
137	LA Pomersbach	-0.13	-7.00	1	-0.02	0	0.00	0	0.00	0	0.00	0	0.00
138	Iqbal Abdulla	-6.99	0.76	0	0.00	0	0.00	0	0.00	0	0.00	1	0.10
139	SR Tendulkar	2.23	-6.97	1	0.28	1	-0.30	1	-0.56	1	-0.81	1	-1.31
140	DP Vijaykumar	-7.00	-4.21	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.84
141	PM Sarvesh Kumar	-7.00	-4.89	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.97
142	Shoaib Akhtar	-7.00	5.88	0	0.00	0	0.00	0	0.00	0	0.00	1	1.07
143	Abdur Razzak	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
144	LPC Silva	-2.15	-7.00	1	-0.27	0	0.00	0	0.00	0	0.00	0	0.00

ID	NAME	BATS	BOWLS	Bat	Batsman.Score	BatAR	BattingAR.Score	AR	AR.Score	BowlAR	BowlingAr.Score	Bowl	Bowler.Score
145	H Das	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
146	SP Goswami	-4.56	-7.00	1	-0.58	0	0.00	0	0.00	0	0.00	0	0.00
147	DR Smith	3.95	-2.86	1	0.50	1	0.13	1	-0.03	1	-0.20	1	-0.52
148	SD Chitnis	-4.65	-4.00	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.79
149	CRD Fernando	-7.00	2.81	0	0.00	0	0.00	0	0.00	0	0.00	1	0.49
150	VS Yeligati	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
151	L Balaji	-7.00	-2.24	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.47
152	A Mukund	-6.17	-7.00	1	-0.78	0	0.00	0	0.00	0	0.00	0	0.00
153	Younis Khan	-6.99	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
154	RR Powar	-6.38	-2.51	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.52
155	JP Duminy	2.10	-2.69	1	0.27	1	-0.01	1	-0.14	1	-0.26	1	-0.50
156	A Flintoff	-0.37	-6.54	1	-0.05	1	-0.49	1	-0.68	1	-0.87	1	-1.25
157	T Thushara	-7.00	1.28	0	0.00	0	0.00	0	0.00	0	0.00	1	0.20
158	JD Ryder	1.32	-1.83	1	0.17	1	-0.02	1	-0.10	1	-0.18	1	-0.34
159	KP Pietersen	4.09	-0.99	1	0.52	1	0.27	1	0.16	1	0.05	1	-0.16
160	T Henderson	-6.79	-2.13	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.45
161	Kamran Khan	-7.00	-2.39	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.50
162	RS Bopara	1.13	-2.60	1	0.14	1	-0.09	1	-0.19	1	-0.29	1	-0.49
163	CH Gayle	6.12	-3.58	1	0.78	1	0.26	1	0.03	1	-0.19	1	-0.64
164	MC Henriques	-2.36	-4.45	1	-0.30	1	-0.51	1	-0.60	1	-0.68	1	-0.86
165	R Bishnoi	-6.56	-7.00	1	-0.83	0	0.00	0	0.00	0	0.00	0	0.00
166	FH Edwards	-7.00	1.08	0	0.00	0	0.00	0	0.00	0	0.00	1	0.16
167	KV Sharma	-6.27	-0.08	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.06
168	PC Valthaty	0.61	-2.80	1	0.08	1	-0.15	1	-0.24	1	-0.34	1	-0.53
169	RJ Quiney	-3.68	-7.00	1	-0.47	0	0.00	0	0.00	0	0.00	0	0.00
170	AS Raut	-4.29	-6.78	1	-0.55	0	0.00	0	0.00	0	0.00	1	-1.31
171	Yashpal Singh	-6.31	-7.00	1	-0.80	0	0.00	0	0.00	0	0.00	0	0.00
172	M Manhas	-3.85	-1.47	1	-0.49	1	-0.42	1	-0.39	1	-0.36	1	-0.30
173	AA Bilakhia	-6.21	-7.00	1	-0.79	0	0.00	0	0.00	0	0.00	0	0.00
174	AN Ghosh	-6.99	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
175	BAW Mendis	-7.00	-1.20	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.27
176	DL Vettori	-6.50	-0.75	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.18
177	MN van Wyk	3.21	-7.00	1	0.41	0	0.00	0	0.00	0	0.00	0	0.00
178	RE van der Merwe	-5.23	1.16	0	0.00	0	0.00	0	0.00	0	0.00	1	0.19
179	TL Suman	-1.57	-3.30	1	-0.20	1	-0.36	1	-0.43	1	-0.50	1	-0.64
180	Shoaib Ahmed	-7.00	-4.71	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.94
181	GR Napier	-5.03	0.03	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.03
182	KP Appanna	-7.00	-2.89	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.59
183	LA Carseldine	-1.74	6.97	1	-0.22	1	0.34	1	0.59	1	0.83	1	1.31
184	NV Ojha	-0.08	-7.00	1	-0.01	0	0.00	0	0.00	0	0.00	0	0.00
185	SM Harwood	-6.96	2.49	0	0.00	0	0.00	0	0.00	0	0.00	1	0.43
186	M Vijay	1.94	-7.00	1	0.25	0	0.00	0	0.00	0	0.00	0	0.00
187	SB Jakati	-6.68	-2.12	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.45
188	DA Warner	3.95	-7.00	1	0.50	0	0.00	0	0.00	0	0.00	0	0.00
189	RJ Harris	-6.83	-0.28	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.10
190	D du Preez	-6.57	3.91	0	0.00	0	0.00	0	0.00	0	0.00	1	0.70
191	M Morkel	-6.68	-1.27	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.28
192	AD Mathews	-0.08	-4.20	1	-0.01	1	-0.30	1	-0.42	1	-0.55	1	-0.80

ID	NAME	BATS	BOWLS	Bat	Batsman.Score	BatAR	BattingAR.Score	AR	AR.Score	BowlAR	BowlingAr.Score	Bowl	Bowler.Score
193	J Botha	-3.20	-0.41	1	-0.41	1	-0.29	1	-0.24	1	-0.20	1	-0.10
194	C Nanda	-7.00	-0.33	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.11
195	SL Malinga	-6.94	3.87	0	0.00	0	0.00	0	0.00	0	0.00	1	0.69
196	Mashrafe Mortaza	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
197	A Singh	-7.00	-0.80	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.20
198	GJ Bailey	-1.56	-7.00	1	-0.20	0	0.00	0	0.00	0	0.00	0	0.00
199	AB McDonald	-2.86	-2.27	1	-0.36	1	-0.40	1	-0.41	1	-0.42	1	-0.45
200	Y Nagar	-3.25	-5.14	1	-0.41	1	-0.63	1	-0.72	1	-0.81	1	-1.00
201	SS Shaikh	-6.84	-7.00	1	-0.87	0	0.00	0	0.00	0	0.00	0	0.00
202	R Ashwin	-6.88	1.42	0	0.00	0	0.00	0	0.00	0	0.00	1	0.23
203	Mohammad Ashraful	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
204	CA Pujara	-3.61	-7.00	1	-0.46	0	0.00	0	0.00	0	0.00	0	0.00
205	OA Shah	2.05	-7.00	1	0.26	0	0.00	0	0.00	0	0.00	0	0.00
206	Anirudh Singh	-4.14	-7.00	1	-0.53	0	0.00	0	0.00	0	0.00	0	0.00
207	Jaskaran Singh	-7.00	-3.88	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.78
208	AP Tare	-1.55	-7.00	1	-0.20	0	0.00	0	0.00	0	0.00	0	0.00
209	AT Rayudu	1.23	-7.00	1	0.16	0	0.00	0	0.00	0	0.00	0	0.00
210	R Sathish	-4.48	-6.82	1	-0.57	0	0.00	0	0.00	0	0.00	1	-1.32
211	R McLaren	-5.54	-5.87	1	-0.70	0	0.00	0	0.00	0	0.00	1	-1.15
212	AA Jhunjhunwala	-4.19	-6.79	1	-0.53	0	0.00	0	0.00	0	0.00	1	-1.32
213	P Dogra	-5.82	-7.00	1	-0.74	0	0.00	0	0.00	0	0.00	0	0.00
214	A Uniyal	-7.00	-5.80	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.15
215	MS Bisla	-0.73	-7.00	1	-0.09	0	0.00	0	0.00	0	0.00	0	0.00
216	YA Abdulla	-7.00	-0.55	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.15
217	EJG Morgan	-0.27	-7.00	1	-0.03	0	0.00	0	0.00	0	0.00	0	0.00
218	JM Kemp	-4.30	2.72	0	0.00	0	0.00	0	0.00	0	0.00	1	0.49
219	S Tyagi	-7.00	-5.18	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.03
220	RS Gavaskar	-7.00	-6.37	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.25
221	SE Bond	-7.00	-0.17	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.08
222	KA Pollard	1.75	-2.95	1	0.22	1	-0.06	1	-0.18	1	-0.31	1	-0.55
223	S Ladda	-7.00	-5.94	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.17
224	DP Nannes	-7.00	-0.15	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.07
225	MJ Lumb	3.09	-7.00	1	0.39	0	0.00	0	0.00	0	0.00	0	0.00
226	DR Martyn	-5.77	-7.00	1	-0.73	0	0.00	0	0.00	0	0.00	0	0.00
227	S Narwal	-5.59	-6.42	1	-0.71	0	0.00	0	0.00	0	0.00	1	-1.25
228	AB Barath	-4.25	-7.00	1	-0.54	0	0.00	0	0.00	0	0.00	0	0.00
229	Bipul Sharma	-3.65	-3.39	1	-0.46	1	-0.54	1	-0.57	1	-0.60	1	-0.67
230	FY Fazal	-3.16	-6.98	1	-0.40	1	-0.75	1	-0.90	1	-1.05	1	-1.35
231	AC Voges	2.02	-6.44	1	0.26	1	-0.28	1	-0.52	1	-0.75	1	-1.21
232	MD Mishra	-3.16	-7.00	1	-0.40	0	0.00	0	0.00	0	0.00	0	0.00
233	UT Yadav	-7.00	-2.94	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.60
234	J Theron	-7.00	-2.64	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.55
235	SJ Srivastava	-7.00	-4.11	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.83
236	KM Jadhav	-2.55	-7.00	1	-0.32	0	0.00	0	0.00	0	0.00	0	0.00
237	R Sharma	-6.98	-0.57	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.15
238	Mandeep Singh	-0.56	-7.00	1	-0.07	0	0.00	0	0.00	0	0.00	0	0.00
239	SW Tait	-6.95	-0.21	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.08
240	KB Arun Karthik	-6.72	-7.00	1	-0.86	0	0.00	0	0.00	0	0.00	0	0.00

ID	NAME	BATS	BOWLS	Bat	Batsman.Score	BatAR	BattingAR.Score	AR	AR.Score	BowlAR	BowlingAr.Score	Bowl	Bowler.Score
241	KAJ Roach	-5.68	-6.96	1	-0.72	0	0.00	0	0.00	0	0.00	1	-1.36
242	PD Collingwood	3.30	2.02	1	0.42	1	0.41	1	0.41	1	0.41	1	0.41
243	CK Langeveldt	-6.94	4.04	0	0.00	0	0.00	0	0.00	0	0.00	1	0.72
244	VS Malik	-6.99	-3.88	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.78
245	A Mithun	-6.92	-6.47	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.27
246	AP Dole	0.72	-2.07	1	0.09	1	-0.09	1	-0.16	1	-0.24	1	-0.39
247	AN Ahmed	-5.84	-3.18	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.64
248	RS Sodhi	-6.99	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
249	DE Bollinger	-6.92	2.25	0	0.00	0	0.00	0	0.00	0	0.00	1	0.38
250	S Sriram	-5.12	-7.00	1	-0.65	0	0.00	0	0.00	0	0.00	0	0.00
251	B Sumanth	-6.20	-7.00	1	-0.79	0	0.00	0	0.00	0	0.00	0	0.00
252	C Madan	-5.57	-7.00	1	-0.71	0	0.00	0	0.00	0	0.00	0	0.00
253	AG Paunikar	-6.02	-7.00	1	-0.77	0	0.00	0	0.00	0	0.00	0	0.00
254	MR Marsh	-1.76	-2.01	1	-0.22	1	-0.29	1	-0.31	1	-0.34	1	-0.39
255	AJ Finch	1.77	-6.43	1	0.23	1	-0.30	1	-0.53	1	-0.76	1	-1.21
256	STR Binny	-1.51	-1.81	1	-0.19	1	-0.25	1	-0.28	1	-0.30	1	-0.35
257	Harmeet Singh	-6.95	-1.79	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.38
258	IR Jaggi	-6.59	-7.00	1	-0.84	0	0.00	0	0.00	0	0.00	0	0.00
259	DT Christian	-2.42	-3.37	1	-0.31	1	-0.44	1	-0.49	1	-0.55	1	-0.66
260	RV Gomez	-6.66	-2.48	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.51
261	MA Agarwal	-0.57	-7.00	1	-0.07	0	0.00	0	0.00	0	0.00	0	0.00
262	AUK Pathan	-6.91	-6.80	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.33
263	UBT Chand	-4.76	-7.00	1	-0.61	0	0.00	0	0.00	0	0.00	0	0.00
264	DJ Jacobs	-4.54	-7.00	1	-0.58	0	0.00	0	0.00	0	0.00	0	0.00
265	Sunny Singh	-5.93	-7.00	1	-0.75	0	0.00	0	0.00	0	0.00	0	0.00
266	NJ Rimmington	-7.00	-3.04	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.62
267	AL Menaria	-1.83	-4.35	1	-0.23	1	-0.46	1	-0.55	1	-0.65	1	-0.84
268	WD Parnell	-6.91	0.15	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.02
269	JJ van der Wath	-6.80	-6.46	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.27
270	R Ninan	-7.00	-2.81	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.58
271	S Aravind	-5.80	-0.01	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.04
272	MS Wade	-6.93	-7.00	1	-0.88	0	0.00	0	0.00	0	0.00	0	0.00
273	TD Paine	-6.99	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
274	SB Wagh	-7.00	-2.13	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.45
275	AC Thomas	-6.92	-2.24	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.47
276	JEC Franklin	-1.53	-2.39	1	-0.20	1	-0.29	1	-0.34	1	-0.38	1	-0.46
277	Shakib Al Hasan	-0.32	2.25	1	-0.04	1	0.13	1	0.20	1	0.28	1	0.42
278	DH Yagnik	-5.38	-7.00	1	-0.69	0	0.00	0	0.00	0	0.00	0	0.00
279	BJ Haddin	1.64	-7.00	1	0.21	0	0.00	0	0.00	0	0.00	0	0.00
280	S Randiv	-7.00	-3.21	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.65
281	NLTC Perera	-2.21	-3.21	1	-0.28	1	-0.41	1	-0.46	1	-0.52	1	-0.62
282	NL McCullum	-3.70	-4.40	1	-0.47	1	-0.61	1	-0.68	1	-0.74	1	-0.86
283	JE Taylor	-7.00	-2.18	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.46
284	J Syed Mohammad	-6.79	-5.02	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.00
285	RN ten Doeschate	-0.83	-2.60	1	-0.11	1	-0.25	1	-0.31	1	-0.38	1	-0.50
286	TR Birt	-2.04	-7.00	1	-0.26	0	0.00	0	0.00	0	0.00	0	0.00
287	AG Murtaza	-6.99	-1.72	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.37
288	Harpreet Singh	-6.68	-7.00	1	-0.85	0	0.00	0	0.00	0	0.00	0	0.00

ID	NAME	BATS	BOWLS	Bat	Batsman.Score	BatAR	BattingAR.Score	AR	AR.Score	BowlAR	BowlingAr.Score	Bowl	Bowler.Score
289	M Klinger	-3.91	-7.00	1	-0.50	0	0.00	0	0.00	0	0.00	0	0.00
290	AC Blizzard	-0.20	-7.00	1	-0.02	0	0.00	0	0.00	0	0.00	0	0.00
291	I Malhotra	-6.55	-7.00	1	-0.83	0	0.00	0	0.00	0	0.00	0	0.00
292	L Ablish	-7.00	-3.62	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.73
293	CA Ingram	-6.99	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
294	S Nadeem	-7.00	-1.34	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.30
295	P Parameswaran	-7.00	-2.74	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.56
296	CJ Ferguson	-5.75	-7.00	1	-0.73	0	0.00	0	0.00	0	0.00	0	0.00
297	VR Aaron	-6.83	-1.29	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.29
298	AA Chavan	-6.99	-4.04	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.81
299	ND Doshi	-7.00	-5.00	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.99
300	B Kumar	-6.89	1.50	0	0.00	0	0.00	0	0.00	0	0.00	1	0.24
301	Y Gnaneswara Rao	-1.34	-5.14	1	-0.17	1	-0.47	1	-0.60	1	-0.73	1	-0.98
302	S Rana	-4.76	-3.82	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.76
303	BA Bhatt	-6.98	-3.39	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.69
304	F du Plessis	3.29	-7.00	1	0.42	0	0.00	0	0.00	0	0.00	0	0.00
305	RE Levi	-3.40	-7.00	1	-0.43	0	0.00	0	0.00	0	0.00	0	0.00
306	GJ Maxwell	5.68	-6.96	1	0.72	1	-0.02	1	-0.34	1	-0.65	1	-1.29
307	SPD Smith	2.26	-7.00	1	0.29	0	0.00	0	0.00	0	0.00	0	0.00
308	MN Samuels	-4.92	-1.83	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.38
309	SA Yadav	-2.46	-6.28	1	-0.31	1	-0.64	1	-0.78	1	-0.93	1	-1.21
310	KK Cooper	-5.38	-0.29	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.09
311	JP Faulkner	-4.68	-0.43	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.11
312	HV Patel	-6.95	-4.79	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.95
313	DAJ Bracewell	-3.63	5.92	1	-0.46	1	0.11	1	0.36	1	0.61	1	1.10
314	DJ Harris	0.41	-6.74	1	0.05	1	-0.44	1	-0.65	1	-0.86	1	-1.28
315	Ankit Sharma	-6.09	-0.08	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.05
316	SP Narine	-6.99	4.45	0	0.00	0	0.00	0	0.00	0	0.00	1	0.80
317	GB Hogg	-6.54	-0.59	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.15
318	RR Bhatkal	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
319	CJ McKay	-6.89	-6.08	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.20
320	N Saini	-4.60	-7.00	1	-0.59	0	0.00	0	0.00	0	0.00	0	0.00
321	DA Miller	5.06	-7.00	1	0.64	0	0.00	0	0.00	0	0.00	0	0.00
322	Azhar Mahmood	0.59	-0.47	1	0.08	1	0.02	1	-0.01	1	-0.03	1	-0.09
323	P Negi	-6.08	-3.45	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.69
324	RJ Peterson	-6.78	-2.61	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.54
325	KMDN Kulasekara	-6.93	0.77	0	0.00	0	0.00	0	0.00	0	0.00	1	0.10
326	A Ashish Reddy	-4.69	-1.34	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.29
327	V Pratap Singh	-7.00	-3.90	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.79
328	BB Samantray	-2.80	-7.00	1	-0.36	0	0.00	0	0.00	0	0.00	0	0.00
329	MJ Clarke	-3.32	0.60	1	-0.42	1	-0.23	1	-0.15	1	-0.07	1	0.09
330	Gurkeerat Singh	-1.60	-7.00	1	-0.20	0	0.00	0	0.00	0	0.00	0	0.00
331	AP Majumdar	-0.32	-7.00	1	-0.04	0	0.00	0	0.00	0	0.00	0	0.00
332	PA Reddy	-3.40	-7.00	1	-0.43	0	0.00	0	0.00	0	0.00	0	0.00
333	K Upadhyay	-6.57	-6.88	1	-0.84	0	0.00	0	0.00	0	0.00	1	-1.35
334	P Awana	-7.00	-2.39	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.50
335	AD Russell	-5.27	-6.89	1	-0.67	0	0.00	0	0.00	0	0.00	1	-1.34
336	A Chandila	-7.00	2.04	0	0.00	0	0.00	0	0.00	0	0.00	1	0.34

ID	NAME	BATS	BOWLS	Bat	Batsman.Score	BatAR	BattingAR.Score	AR	AR.Score	BowlAR	BowlingAr.Score	Bowl	Bowler.Score
337	CA Lynn	-0.94	-7.00	1	-0.12	0	0.00	0	0.00	0	0.00	0	0.00
338	Sunny Gupta	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
339	MC Juneja	-3.91	-7.00	1	-0.50	0	0.00	0	0.00	0	0.00	0	0.00
340	KK Nair	3.74	-7.00	1	0.48	0	0.00	0	0.00	0	0.00	0	0.00
341	GH Vihari	-5.36	-0.47	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.12
342	MDKJ Perera	-6.23	-7.00	1	-0.79	0	0.00	0	0.00	0	0.00	0	0.00
343	R Shukla	-6.95	-6.20	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.22
344	B Laughlin	-7.00	-6.77	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.33
345	AS Rajpoot	-7.00	-4.72	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.94
346	M Vohra	4.25	-7.00	1	0.54	0	0.00	0	0.00	0	0.00	0	0.00
347	JD Unadkat	-7.00	-3.12	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.64
348	Mohammed Shami	-7.00	-5.19	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.03
349	BMAJ Mendis	-6.64	-2.58	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.53
350	R Rampaul	-6.59	1.55	0	0.00	0	0.00	0	0.00	0	0.00	1	0.25
351	CH Morris	-7.00	-2.26	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.47
352	SV Samson	1.32	-7.00	1	0.17	0	0.00	0	0.00	0	0.00	0	0.00
353	SMSM Senanayake	-6.99	1.44	0	0.00	0	0.00	0	0.00	0	0.00	1	0.23
354	BJ Rohrer	2.95	-6.99	1	0.38	1	-0.25	1	-0.51	1	-0.78	1	-1.31
355	KL Rahul	-2.66	-7.00	1	-0.34	0	0.00	0	0.00	0	0.00	0	0.00
356	Q de Kock	-2.59	-7.00	1	-0.33	0	0.00	0	0.00	0	0.00	0	0.00
357	R Dhawan	-5.45	-1.72	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.36
358	MG Johnson	-6.65	0.38	0	0.00	0	0.00	0	0.00	0	0.00	1	0.03
359	LJ Wright	3.75	-6.83	1	0.48	1	-0.17	1	-0.44	1	-0.72	1	-1.27
360	IC Pandey	-7.00	-4.96	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.99
361	X Thalaivan Sargunam	-6.73	-7.00	1	-0.86	0	0.00	0	0.00	0	0.00	0	0.00
362	DJG Sammy	-1.34	-3.95	1	-0.17	1	-0.39	1	-0.48	1	-0.57	1	-0.76
363	KW Richardson	-5.59	-2.25	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.46
364	MM Sharma	-7.00	1.46	0	0.00	0	0.00	0	0.00	0	0.00	1	0.23
365	CM Gautam	-3.23	-7.00	1	-0.41	0	0.00	0	0.00	0	0.00	0	0.00
366	UA Birla	-6.87	-7.00	1	-0.88	0	0.00	0	0.00	0	0.00	0	0.00
367	Parvez Rasool	-7.00	-4.75	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.95
368	Sandeep Sharma	-7.00	0.86	0	0.00	0	0.00	0	0.00	0	0.00	1	0.12
369	S Kaul	-6.94	-5.37	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.06
370	Sachin Baby	-6.99	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
371	PV Tambe	-6.99	-0.43	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.13
372	NM Coulte	-5.86	-7.00	1	-0.75	0	0.00	0	0.00	0	0.00	0	0.00
373	P Amarnath	-7.00	-6.00	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.18
374	B Geeves	-7.00	-6.99	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.37
375	Gagandeep Singh	-7.00	-6.61	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.30
376	A Nel	-7.00	-5.64	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.12
377	AM Salvi	-7.00	-1.99	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.42
378	Anureet Singh	-7.00	-1.48	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.33
379	RR Bose	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
380	SS Sarkar	-7.00	-4.07	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.82
381	RA Shaikh	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
382	C Ganapathy	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
383	MB Parmar	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
384	TG Southee	-7.00	-5.50	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.09

ID	NAME	BATS	BOWLS	Bat	Batsman.Score	BatAR	BattingAR.Score	AR	AR.Score	BowlAR	BowlingAr.Score	Bowl	Bowler.Score
385	AA Kazi	-7.00	-6.99	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.37
386	Anand Rajan	-7.00	-0.76	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.19
387	P Prasanth	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
388	RW Price	-7.00	-6.99	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.37
389	M de Lange	-7.00	-4.16	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.84
390	TP Sudhindra	-7.00	-6.99	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.37
391	BW Hilfenhaus	-7.00	0.10	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.03
392	JJ Bumrah	-7.00	-4.37	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.87
393	S Badree	-7.00	-3.46	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.70
394	JO Holder	-7.00	-6.23	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.23
395	YS Chahal	-7.00	-1.92	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.41
396	P Suyal	-7.00	-3.34	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.68
397	MG Neser	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
398	NJ Maddinson	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
399	AR Patel	-7.00	2.21	0	0.00	0	0.00	0	0.00	0	0.00	1	0.37
400	CJ Anderson	-7.00	-6.49	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.28
401	JDS Neesham	-7.00	-6.87	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.35
402	MA Starc	-7.00	-1.16	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.26
403	BR Dunk	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
404	RR Rossouw	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
405	Shivam Sharma	-7.00	-3.85	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.78
406	LMP Simmons	-7.00	-6.17	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.22
407	VH Zol	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
408	BCJ Cutting	-7.00	-2.87	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.59
409	Imran Tahir	-7.00	0.85	0	0.00	0	0.00	0	0.00	0	0.00	1	0.12
410	BE Hendricks	-7.00	-5.16	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.02
411	S Gopal	-7.00	0.57	0	0.00	0	0.00	0	0.00	0	0.00	1	0.06
412	R Tewatia	-7.00	-0.01	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.05
413	Karanveer Singh	-7.00	3.14	0	0.00	0	0.00	0	0.00	0	0.00	1	0.55
414	V Shankar	-7.00	-7.00	1	-0.89	0	0.00	0	0.00	0	0.00	0	0.00
415	K Santokie	-7.00	-5.00	0	0.00	0	0.00	0	0.00	0	0.00	1	-1.00
416	PJ Cummins	-7.00	-1.75	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.38
417	JW Hastings	-7.00	-3.64	0	0.00	0	0.00	0	0.00	0	0.00	1	-0.74