Player Ratings in Continuous and Discrete Team Sports

Jonathan Sargent
B.Econ, M.Stat (AStat)

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy.

School of Mathematical and Geospatial Sciences
RMIT University
December 2013
Statement of Authorship

The candidate hereby declares that:

• except where due acknowledgement has been made, the work is that of the candidate alone;
• the work has not been submitted previously, in whole or in part, to qualify for any other academic award;
• the content of the thesis is the result of the work which has been carried out since the official commencement date of the approved research program;
• any editorial work, paid or unpaid, carried out by a third party is acknowledged.

Jonathan Sargent

November 2013
Acknowledgments

I wish to acknowledge the following people and organisations, without whose assistance, this dissertation would not have been possible.

Assoc Prof Anthony Bedford. Thank you for being a terrific supervisor. Your expertise, encouragement, attitude and humour have resulted in many satisfactory outcomes. We’ll always have the Wimbledon tickets story.

RMIT Sports Statistics Research Group. Thanks to everyone in the group over the years. I’ve loved chatting stats and sport with you. Special mention to Dr Adrian Schembri for being a great sounding board over the years. Thanks also to Cliff De Costa, the silent Supervisor.

Dr Peter Wright for his friendship over the years and astute Visual Basic tutelage. Your nested loop shaved at least six months off the research duration.

Prowess Sports for providing data used at various stages throughout this dissertation.

Ucinet for access to the network analysis software.

Examiners. Thank you for taking the time to review this dissertation. Your feedback has been greatly valued.

My Family. Thank you for constant support throughout my life.

Susan. Thank you for everything. You are everything.
Summary

In team sports, there is a mounting focus on the player membership of a team; how they perform as individuals, how they cooperate and how these actions contribute to the team’s success. Coaching staff, supporters, punters, and, ever increasingly, fantasy league coaches, are all key stakeholders in player-based information. This dissertation investigates player performance and player rating methodologies in continuous and discrete team sports, using the Australian Football League (AFL) and cricket as case studies. The one-on-one nature of discrete sports offers a less ambiguous measurement of a player’s contribution to their team’s success than continuous sports, as it is usually one player from each team involved in the immediate contest. Continuous sports, such as the AFL, are more convoluted as there are numerous contests occurring around the play at any one time, requiring that player rating models be expanded to account for as many of these contests as possible. After a discussion on the different types of player ratings, this dissertation progresses to the different types of AFL player data collected for analysis, how such data can be aggregated to produce a match “score”, $X$, for each player and how these scores can be approximated with a normal distribution. An Elo-influenced adjustable player rating system (APR) is discussed where simulation techniques pit the player being rated against an opponent player in the same position. The difference between the result of the observed contest ($Obs$) and the probability of him outscoring his opponent ($Exp$) dictated whether his rating increases or decreases after the match. The pre-match probabilities were found to be linearly related to his final score. Using Mahalanobis distance classification, we retrospectively reclassified the players with a series of game-related player data rather than assuming a player was confined to the same position for each match. The APR was rerun with the correct player positions resulting in an improved relationship between $Exp$ and $X$. Intra-position classification was also investigated. The main limitation of the APR was that the data and scoring was too player-centric, overlooking aspects of teamwork and a player’s contribution to the team. Player interaction was investigated which required transactional data to be analysed. A Visual Basic program, $LINK$, was written which looped through the transactional data to isolate link plays, or sets of relations involving two or more cooperating players from team $a$, where the ball’s movement effectively increased that
team’s scoring likelihood. Links continued until a score, turnover or dead play was realised. Player membership, transaction description, ground position, match period and link length were all recorded by the program, allowing a graphical output. The link plays contained send and receive data which fashioned an interaction matrix, revealing frequencies of player interaction. Network diagrams offered a graphic portrayal of the interactions between players in a match with node diameter indicating the prominence of a player in the network and connecting line width, the frequency of interactions between $i$ and $j$. The interaction matrix was then symmetrised so interactions between players could be simulated to estimate players’ influence on a team network. Interactions between pairs of players followed a negative binomial distribution with parameters estimated using a Pearson chi-squared approach. Player performance in a match was quantified using eigenvector centrality, an important network statistic, indicating a player’s level of interaction with other central players. Team strength was calculated by averaging each player’s centrality in the network. The team index for any match was adequately related to the score margin for that match making it possible to observe different players’ contribution to team performance (margin) when included and excluded from a simulated network. Data on the Geelong football club was used to conceptualise the methods in these chapters.

Apart from Schwartz (2006), in-play simulation methodology is yet to be fully investigated in cricket. For this section of the research, a Visual Basic program was written that called conditional probability distributions to simulate outcomes—runs and dismissals—while a limited overs cricket match was in-play. These likelihoods were conditional on the player’s order in the batting list, the delivery number—both discrete variables—and the type of batsman (fast, medium or slow scorer). The simulated batsman scores were then adjusted for team strength, innings and venue effects using multiple linear regression. This dissertation demonstrates the benefits of the model by fitting log-normal distributions to simulated innings ($n=500$) by Australia’s Ricky Ponting in the 2011 ODI World Cup quarter final. It was then possible to approximate how likely he was to achieve a certain score prior to the match, then at 10, 20 and 30 over intervals. It is anticipated that real-time information of a batter’s score expectations will add confidence to wagering in individual performance markets such as “highest score”, as well as live player-rating revisions.
Contents

1 Introduction .......................................................................................................................... 13
  1.1 Why the AFL and Cricket? .............................................................................................. 15
1.2 Literature Review ............................................................................................................. 19
  1.2.1 Statistical Analysis in the AFL ................................................................................. 20
  1.2.2 Continuous Sport Player Performance ................................................................. 21
  1.2.3 Team and Player Ratings ....................................................................................... 24
  1.2.4 AFL Player Interaction Simulation ........................................................................ 27
  1.2.5 AFL Player Performance Forecasts ...................................................................... 29
  1.2.6 Statistical Analysis of Batsmen in Cricket ............................................................ 30
  1.2.7 Conclusion ........................................................................................................... 31
  1.3 Research Questions and Publications ........................................................................... 32

2 The AFL and Cricket .......................................................................................................... 36
  2.1 The AFL ........................................................................................................................ 36
    2.1.1 History ................................................................................................................ 36
    2.1.2 Basic Rules and Mechanics ............................................................................. 37
    2.1.3 The Players ....................................................................................................... 39
    2.1.4 2012 Team and Player Rankings ..................................................................... 41
  2.2 Cricket .......................................................................................................................... 42
    2.2.1 Introduction ....................................................................................................... 42
    2.2.2 Basic Rules ...................................................................................................... 43
    2.2.3 Scoring ............................................................................................................ 44
    2.2.4 2012 Team and Player Rankings ..................................................................... 46
  2.3 Discussion .................................................................................................................... 46

3 Methods ............................................................................................................................. 48
  3.1 Linear Regression ........................................................................................................ 48
    3.1.1 Introduction ...................................................................................................... 49
    3.1.2 The Linear Regression Model .......................................................................... 49
    3.1.3 Assumptions ................................................................................................... 49
    3.1.4 Least Squares Regression ............................................................................... 50
    3.1.5 Ordinary Least Squares .................................................................................. 51
    3.1.6 Maximum Likelihood Estimation .................................................................... 51
    3.1.7 Analysis of Residuals ..................................................................................... 52
    3.1.8 Goodness of Fit ............................................................................................ 53
  3.2 Discriminant Analysis ................................................................................................ 54
    3.2.1 Introduction ...................................................................................................... 54
    3.2.2 The Model ...................................................................................................... 54
    3.2.3 Classification with Two Multivariate Normal Populations ......................... 55
    3.2.4 Fisher’s Classification with several Populations ........................................... 56
  3.3 Nonlinear Data Smoothers ......................................................................................... 57
  3.4 Optimisation ............................................................................................................... 59
    3.4.1 Introduction ...................................................................................................... 59

6
List of Figures

Figure 2.1: The Melbourne Cricket Ground with goal posts at the left and right............................................................39
Figure 4.1: Jimmy Bartel fantasy performance (Score) and value (Price) for the 2012 season........................................81
Figure 5.1: Linear relationship between average AFL fantasy player scores and margin for Geelong’s 2012 matches ..........................................................................................................................90
Figure 5.2: Linear relationship between re-optimised average player scores and margin for Geelong’s 2012 matches ........................................................................................................................................90
Figure 5.3: Boxplots showing the distribution of Geelong’s scores by player position (i. Equation (5.6), ii. Equation (5.7)) .........................................................................................................................................91
Figure 5.4: The adjustive AFL player rating system .........................................................................................................92
Figure 5.5: Joel Selwood’s scores fitted with a normal distribution..................................................................................94
Figure 5.6: Collingwood (left) and Sydney (right) midfielder scores fitted with normal distributions ........................95
Figure 5.7: Relationship between Exp and Xi for all Geelong midfield matches (2010–2011) .........................................98
Figure 6.1: Canonical discriminant function graph of n = 200 players ........................................................................111
Figure 6.2: Reclassification of forwards using Squared Euclidean distance .................................................................117
Figure 7.1: AFL coaching software (courtesy of Prowess Sports) .................................................................................124
Figure 7.2: Transactional data from round 8, 2011 ........................................................................................................127
Figure 7.3: LINK Production Menu .............................................................................................................................127
Figure 7.4: LINK output CSV file ..................................................................................................................................130
Figure 7.5: LINK output diagram of Geelong link, L_{00} ...............................................................................................130
Figure 7.6: Spatial-temporal matrix of Brisbane’s events (score proportions shaded) ....................................................132
Figure 7.7: Geelong’s interaction matrix from 2011’s grand final ...............................................................................134
Figure 7.8: Send/Receive event averages by position for Geelong’s grand final match .............................................135
Figure 7.9: Network diagram of Geelong’s 2011 grand final match ............................................................................137
Figure 7.10: Network diagram of Geelong’s loss in round 23, 2011 ............................................................................139
Figure 7.11: 2011 Geelong’s symmetrised interaction matrix from 2011’s grand final .................................................140
Figure 7.12: Frequency distribution of [Bartel, Mackey] interactions for 2011 ..........................................................140
Figure 7.13: Eigenvector centrality for Geelong’s 2011 grand final team .....................................................................144
Figure 7.14: Relationship between Geelong’s mean network rating and final score margin ........................................146
Figure 7.15: Relationship between Geelong’s mean individual rating and final score margin .....................................146
Figure 7.16: Interaction simulation module in LINK ...................................................................................................147
Figure 7.17: Simulated and observed interaction totals for Geelong matches ..............................................................148
Figure 7.18: Margin distributions with and without Bartel ..........................................................................................149
Figure 7.19: Eigenvector centrality approximately normal distribution for Geelong’s midfielder (2011) ....................151
Figure 8.1: Frequency distributions of kicks (KCK) and handballs (HBL) for all players in 2011 .................................154
Figure 8.2: Events leading to the optimal Tukey-smoothed AFL player performance forecast ..................................155
Figure 8.3: The T-EXP optimisation model in Excel ....................................................................................................158
Figure 8.4: Evolution of Mayne’s optimised smoother from match seven .................................................................160
Figure 8.5: Joel Corey’s match 18 score prediction after a peak in round 17 ..............................................................161
Figure 8.6-7: Histogram of sample players’ residuals / Distribution of Mayne’s simulated score forecasts ...............163
Figure 9.1: ODI batsman scores (2008–2011) fitted with an exponential distribution ..................................................172
Figure 9.2: ODI batsman contributions (2008–2011) fitted with an exponential distribution ....................................172
Figure 9.3: ODI run simulation flowchart ...................................................................................................................173
Figure 9.4: Dot charts displaying dismissal frequencies (a) per over, (b) per over by batting order ............................175
Figure 9.5: Mean runs scored per batting order ..........................................................................................................176
Figure 9.6: Boxplots of mean strike rate by batsman type ............................................................................................178
Figure 9.7: Polynomial smooth, f(X) of f(x) = P(x = 4| F_i,b_{1,...,300}) ........................................................................180
Figure 9.8: SimScore screen shot ..................................................................................................................................182
Figure 9.9: TR v G_{00} mean absolute error rates in pre- and intra-match states ..........................................................186
Figure 9.10: TR projection for Australia v South Africa (23/10/2011) ........................................................................187
Figure 9.11: Intraclass correlation of Sim\_mean and Sim\_S\_mean with holdout mean ........................................ 190
Figure 9.12: Ricky Ponting’s simulated runs fitted with log-normal distributions after (a) 0 overs, (b) 10 overs, (c) 20 overs, d) 30 overs ........................................................................................................ 191
Figure 11.1: Link play isolation code ................................................................................................................. 219
Figure 11.2: Eigenvector centrality code ............................................................................................................ 220
Figure 11.3: “Tukey” smoothing code ................................................................................................................ 221
Figure 11.4: SimScore code ................................................................................................................................ 222
List of Tables

Table 2.1: Final ladder of 2012 AFL regular season ................................................................. 41
Table 2.2: Brownlow Medal votes for 2012 ................................................................................. 42
Table 2.3: Highest win/loss ratio for ODI teams in 2012 ........................................................... 47
Table 2.4: Highest run scorers for 2012 in ODI matches .......................................................... 47
Table 4.1: Greater Western Sydney and Collingwood team list statistics for the 2012 season .... 64
Table 4.2: Brownlow Medal winners and respective team ranks .................................................. 65
Table 4.3: Send-receive performance indicators in an AFL match ............................................ 70
Table 4.4: Performance indicators for the Geelong Football Club in the 2011 grand final compared to its season average ........................................................................................................... 71
Table 4.5: Performance indicator totals for all clubs in 2011 for wins and losses, with p-values (α = 0.05) .......................................................................................... 71
Table 4.6: Standardised performance indicators for Geelong in the 2011 grand final ............. 72
Table 4.7: Multiple frequency player interactions for Geelong in the 2011 grand final .......... 75
Table 4.8: Correlation of 2012 batsmen average and contribution ranking with ICC ranking ... 78
Table 4.9: Jimmy Bartel performance indicator frequencies ....................................................... 81
Table 5.1: Joel Selwood’s ratings between 2010 and 2011 ......................................................... 96
Table 5.2: Final ratings results for Geelong midfielders in 2011 ............................................... 98
Table 5.3: Comparison of exponentially smoothed ratings and the APR model ....................... 100
Table 5.4: Consistency measures for Geelong midfielders between 2010 and 2011 ............... 104
Table 6.1: Classification function coefficients by position ....................................................... 110
Table 6.2: Tom Lonergan’s pre and post 2010 season classification ........................................ 112
Table 6.3: Prior probabilities of AFL-classified positions ....................................................... 114
Table 6.4: Classification posterior probabilities for Jonathan Brown in 2009 ......................... 115
Table 6.5: Performance covariance matrix (PCM) for Jonathan Brown ................................. 117
Table 6.6: Dissimilarity matrix of Squared Euclidean distance between forwards’ cov(Xi,Xj) ...... 118
Table 6.7: Dissimilarity matrix of Squared Euclidean distance: Brown and Akermanis [MRKGLS, HBRGLS] ................................................................. 119
Table 7.1: Probabilities and χ² values for [Bartel, Mackey] .......................................................... 142
Table 8.1: T-EXP player performance matrix (Chris Mayne example) ..................................... 159
Table 8.2: Comparison of forecast error from a Mean, EXP and T-EXP model ....................... 162
Table 8.3: Comparison of forecast error of 2011 Geelong midfield ........................................ 164
Table 9.1: First two overs of data from South Africa v India ODI match (12.01.2011) ............... 174
Table 9.2: Mean strike rates for batsmen types at early-, middle- and latter-innings stages ...... 178
Table 9.3: Probabilities for α=M1,5,11 at early- and late-innings stages ................................. 181
Table 9.4: Australia ODI team winning percentage and mean first innings runs by venue ........ 183
Table 9.5: Simulated match scenario showing run redistribution ............................................. 188
Table 9.6: Consistency coefficient (CV1) for 30 ODI batsmen (2008–2011) ............................ 192
Table 9.7: Consistency coefficient (CV2) for 30 ODI batsmen (2008–2011) ............................ 193
Table 11.1: APR ratings for top 100 midfielders in the AFL (2011) ........................................... 218
Table 11.2: ICC ranked batsmen as at November, 2012 ......................................................... 224
Table 11.3: Top 50 rated players in the MLS ................................................................. 225
Chapter 1

Introduction

“The way a team plays as a whole determines its success. You may have the greatest bunch of individual stars in the world, but if they don't play together, the club won't be worth a dime.” – Babe Ruth.

The principal reward of supporting a team or individual involved in a sporting contest—whether it is as a coach on the sidelines, a supporter in the stands or as a punter who has just laid a wager—is witnessing the team or individual win. In individual sporting contests such as singles tennis, victory is attributed to a player outplaying his or her direct opponent. With team sports such as football, victory is usually the result of the members of a team interacting and cooperating more effectively throughout the match than the members of the opposition team (Oliver, 2004). In these sports, the players cooperate by executing a series of game-specific skills between each other, ideally resulting in a score to the team. It is simple to either recall or research the person who won a particular individual contest—for example, Roger Federer winning the 2012 Wimbledon Championship. Similarly for team sports, the name of the victorious team is memorable—for example, Spain winning the 2010 World Cup. For a player in a team sport to be remembered (for the right reasons) in conjunction with the name of the winning team, an exceptional performance is required—for example, Andrés Iniesta, the scorer of the winning goal in the 2010 World Cup final. The fact that Iniesta scored the only goal in the match does not necessarily imply that he was the best or most
important player for Spain but, rather, that he was the player who was talented and/or lucky enough to be able to perform the ultimate act in a cooperative series of events leading up to the goal; this series consisted of seven passes of the football executed by five different Spanish players. Prior to scoring, Iniesta also had one effective pass in the series, illustrating that he was working with his teammates to help create his own victorious opportunity. The part that Sergio Ramos played to force the turnover that began the successful passing series in the Dutch forward half may not have been as spectacular and celebrated as the goal itself but was, arguably, just as important as Iniesta’s score because the goal may never have occurred without Ramos’s selfless act. The goal is the perfect result of Spain’s teamwork, exhibiting why teamwork is the crucial component of all non-individual sports.

While we can confirm Federer was the best male player at Wimbledon in 2012, how do we confirm who the best player was in the 2010 World Cup final? Was it Iniesta, Ramos or another Spanish player—one even from the opposition team? An answer may be sought from their singular acts—for example, the player with the most number of effective passes—but more realistically from a measure of each player’s contribution to Spain’s teamwork in the match. So, how is teamwork measured? In his chapter, “Teamwork”, Oliver (2004) discusses the difficulty in quantifying teamwork in fluid, “continuous” team sports such as basketball. Imagine this hypothetical situation for any continuous sport match: a team works cooperatively during a play, passing the ball effectively between its members 10 times and eventually to a designated scorer, but the scorer misses the shot he takes. While the players involved in the ball’s movement are effectively working as a team to increase the likelihood of a score, the scorer’s miss lets down the team. If the score is not increasing due to the shooter’s inaccuracy, how can we measure the other players’ contributions to teamwork? A “Score Assist” is an important record for players in basketball, credited to the player who passes the ball to a teammate who scores a field goal. To be fair to the final-pass player in this hypothetical situation, assuming the pass was not difficult to recover by the shooter, “Miss Assists” would need to be recorded, which sounds absurd. But what of the players involved in the passing chain prior to the score assist? Should the player who starts the series of 10 passes be credited more if the series results in a score, not a miss? Oliver (2004) argues that no tools have been established in basketball to
measure a particular player’s contribution to the likelihood of a team scoring. He does, however, mention “discrete” sports such as baseball being able to be more efficiently measured at the player level because the likelihood of victory is incrementally present in the form of batters running between the bases. A pitcher’s and the fielders’ prevention of runs is the converse measurement to a batter. The Australian Football League (AFL), Australia’s national and premier winter sport, provides another example of a complex, continuous sport with very little in the way of a statistical framework to quantify and award player performance and contributions to team performance. The premier summer sport in Australia is cricket, which can be likened to baseball; it is a discrete bat-and-ball sport with incremental additions to a total score called “runs”. Although official cricket player ratings are published on websites (ICC Rankings), there is very little focus on batting and bowling performances measured while the match is in progress. This dissertation is concerned with the discussion and development of player ratings and performance measurement tools for both the AFL and cricket, with both pre-match and in-play applications.

1.1 Why the AFL and Cricket?

There is a commonly held belief in the community that the field of statistics is bland—the awful subject you only just passed in first-year psychology or engineering that you would never have to worry about ever again. I completed a quantitative analysis course as part of an undergraduate economics degree at the University of Queensland and was immediately drawn to the field, to the point that I later pursued a Master’s degree in statistics at RMIT University with a view to moving into business-related employment. In certain assessment within the Master’s course, I was able to apply statistical methodology to real-world scenarios of my choosing and, so, discovered the beauty of statistical procedures within the world of sport. I quickly realised that mathematics and statistics were exponentially more enjoyable when practically applied to subjects in which a person is truly passionate—in my case, the AFL and cricket. The final piece of assessment for the Master’s degree was a minor thesis which, without any hesitation, I dedicated myself to the application of statistical procedures on AFL data. I emailed Associate Professor Anthony Bedford, who became known to me through his media release on Pythagorean projections in the AFL, requesting him as my supervisor,
and we decided an appropriate topic would be the development of AFL player ratings. We arrived at this topic quickly, citing the lack of objective information describing and predicting AFL player performance, as well as my interest in incorporating a team membership variable in match prediction models. At the completion of the Master’s degree, Assoc Prof Bedford and I agreed the research should continue at the PhD level.

Living in Victoria, it is easy to be swept up in the AFL “maelstrom”, even if you are new to the sport. The popularity of the AFL around Australia is confirmed by the potential for supporters to attend live games in five states or watch every game, live, across three television channels or on the official AFL website. Player and team statistics have evolved to become an integral part of any AFL match, displayed on screens at the match, on television and on numerous websites and mobile applications covering the AFL. With easy access to this match data, I originally began using statistical techniques, such as linear regression, to predict the outcome of AFL matches, using relatively simple inputs—team form, opponent form and venue. It did not take long for me to realise that a key component was absent in the predictions: the ability to take into account the impact on a match of the inclusion or exclusion of certain players. The extension of my prediction models to include a player component—namely, to predict how particular combinations of players would perform together and their contribution to victory—was the foremost motivation for this dissertation.

The numerical face of the AFL ensures a market for quantitative measurement tools will be ever-present, for coaches and punters alike. During a live telecast, frequent vision inside the coach’s box reveals numerous laptops with continuous live data feeds guiding the coaching staff’s decisions on substitutions and on-field player movements. While coaches rely on this data, it is common to hear them in media conferences admit that their techniques and decisions are mostly inherent and that “stats” do not tell the whole story. This is a reflection on the underuse of statistical methodologies in the AFL. The AFL, as a continuous sport, poses the same player performance measurement dilemmas as basketball, as outlined by Oliver (2004). The opportunity to provide scientific information that rates player performance is a tremendous challenge due to the frenetic nature of the game, exacerbated by the fact that, in each match, twenty-two players compete on each AFL side, the most
out of any recognised team sport. Thirty-six players are on the field at any one time—each side has eighteen players and four reserves—which must be of some encouragement for Oliver’s basketball quandary, given he only has ten players to contend with at any match moment. Another complexity in the AFL is the potential for any player on the field to score a goal, not just the forwards. A defender’s key role is to prevent opposition forwards from scoring, but it is not uncommon in a game to see the defender run up the ground and kick a goal. Quantifying the defender’s contribution to the team’s performance then becomes twofold: any scoring he prevents and any scoring he executes. The AFL’s current player performance measurement and ranking tools are scarce and include subjective awards as well as other basic cumulative ratings derived from on-field individual exploits, ignoring how effectively the match participants play off each other’s talents to increase the probability of a win. The most recognised player award is the Brownlow Medal, which is decided through a subjective method whereby umpires gather after each match and vote for the best players on the ground on a 3-2-1 basis (see Section 2.1.4 for 2012 results). The absence of an official, objective AFL player rating framework provided much encouragement throughout this research period. With a pre-match estimate of each player’s perceived contribution, or by means of live evidence during a match, a coach would have the luxury of making informed team list changes or live substitutions. This would be of outstanding benefit considering a rule change in 2010 that forced teams to designate one of their four reserves as a substitute who could be introduced into the game at any time but whose replacement could not return to the field. This rule was introduced to curb the dramatic rise in interchanges which were giving certain teams small physical advantages. A statistical teamwork contribution model would enable coaches to field an optimal team at any stage in the match; for example, a player could be substituted off if his contribution was negligible; another could be moved to the midfield if the data proved that his historical output was more valuable in that position.

For those of us not gifted enough to be directly involved with football as a player or coach, fantasy leagues are available online for our enjoyment. “Supercoach”, the most popular AFL fantasy football competition, attracted nearly 250,000 people in the state of Victoria alone in 2012, all competing for a first prize of $50,000. Supercoach offered this research a recreational opportunity for
the development of optimisation models that automatically selected the best possible players for a fantasy team given various constraints, for example salary cap. A fantasy coach could potentially run these models and enter a strong team without having to spend hours (or days) attempting to manually fit 30 players under a $10 million salary cap. Let linear programming do it for you in a matter of seconds! An important application resulting from this optimal fantasy team research was the ability to predict, with confidence, how certain players will perform against certain opponents. This model has real-world applications also, assisting coaches in their player trading and substitutions.

Australia’s other national sport is cricket, a discrete sport, analogous in many ways to baseball. A cricket player’s contribution to team performance can be measured more simply than that of an AFL’s player by observing player performance in three different roles (see Chapter 2):

1. Batsman—attempts to score (incremental) runs which are added to the team score when the ball is struck into space on the field.
2. Bowler—delivers the ball to the batsman in an over-arm style in an attempt to prevent the batsman from scoring runs, but ultimately to attempt to “dismiss” the batsman by various methods.
3. Fieldsman—on the same team as the bowler, attempts to prevent the batsman from scoring runs by retrieving a hit ball, but ultimately to “dismiss” the batsman.

A player’s contribution to his team’s performance can be simply demonstrated: a batsman who scores 100 of his team’s 300 runs in a match and a bowler who dismisses 5 of the 10 opposition batsmen could be said to have contributed $100/300 = 33.33\%$ and $5/10 = 50\%$ to the team’s batting and bowling performance, respectively. Although performance union makes this a naïve approach (some batsmen bowl and all bowlers may have to bat), it illustrates the point that the performance of participants in discrete sports like cricket and baseball can be more readily assessed as there is more of an onus on the individual. This onus is verified by the existence of the official ICC Cricket Rankings, where batters, bowlers and all-rounders (batters who regularly bowl) are rated as separate entities on their performances in a calendar year (see Chapter 9). These rankings are revised after each match for each batter, bowler and all-rounder. An exciting opportunity was realised in this research
where a player’s contribution could be measured and predicted while the game was in progress; a cricket player’s ranking and perceived contribution to team success could be updated in a real-time scenario. Such information would be highly advantageous for punters wagering on highest batsman score markets. The rules and mechanics of the AFL and cricket will be discussed in more detail in Chapter 2.

The research outlined in this dissertation is the result of a series of (mostly) productive ideas and their application, originating with a passion for sport and statistics, and driven by an opportunity to provide the research field and the AFL and cricket communities with information that can educate as well as guide decision-making in a professional and recreational capacity. Rather than developing a generic model for each sport, the complexity of the topic demanded a “modulated” approach, with symbiotic statistical modelling of the various factors that contribute to player performance, such as player skills (see Chapter 4), opponent strength (see Chapter 5), player positions (see Chapter 5, 6), player interaction (see Chapter 7) and in-play performance prediction (see Chapter 9). Throughout the research period, these modules have formed the topics of presentations at conferences in Manchester (the Institute of Mathematics and its Applications (IMA)), Boston (the New England Symposium on Statistics in Sports (NESSIS)), Leuven (MathSport International, Belgium) and at various cities in Australia (MathSport), radio interviews for Triple R in Melbourne and Pulse FM in Geelong and publications in the *Journal of Quantitative Analysis in Sport*, the *Journal of Sports Sciences and Medicine* and the *International Journal of Forecasting*. Prior to detailing these modules, a literature review and a comprehensive list of research questions and publications, relevant to the research, are to be examined.

### 1.2 Literature Review

This section explores previous research in sports statistics that has provided direction for and arguments within this dissertation. The majority of the literature is concerned with player performance and player ratings across a range of sports, including AFL, cricket, football, basketball and, even, Tiddlywinks. The review begins by discussing literature on statistical analysis in AFL and limited-
overs cricket before progressing logically to notational analysis in sport and an overview of papers on team and player ratings; literature that guided this research into simulating AFL player interactions, ratings and forecasts is then detailed. Although the scarcity of AFL player performance literature, at times, made it difficult to critically evaluate other publications concerning the sport, the challenge in adapting and augmenting techniques applied in other sports for analysing the AFL has been rewarding, adding a high degree of originality to this work. The literature outlined was either referred by colleagues or discovered through resource searches and in other papers’ references.

1.2.1 Statistical Analysis in the AFL

The complex fluidity of Australian Rules football (AFL) provides a rich source of statistical possibilities in such areas as performance analysis and prediction. As in any other competitive environment, numerous physical and mental conditions can affect performance, for example, weather, personal confidence, opponent strength and crowd size. Stefani and Clarke (1992) developed an AFL model to predict the winning margin of a given match, incorporating a home-ground advantage component and team strength difference. Home-ground advantage in the AFL is significant due to, among other things, the unbalanced draw. Clarke (2005) used linear regression to calculate individual home advantages for each club; this was necessary because of the different burdens each club faces with travel between five states in Australia and the characteristics unique to each ground. Clarke’s paper raised important venue implications which were apparent in player performance measurements in this research; for example, the Docklands venue in Melbourne is covered, thereby providing protection from the elements and a scoring advantage over teams playing without cover. Margin of victory in AFL was also estimated by Bailey (2000), who analysed with linear regression the predictive properties of the competing teams’ differences in age, weight, experience and number of kicks and handballs. We were reminded of the complexity of an AFL match by Forbes and Clarke (2004), who arrived at a minimum of seven states in a Markov process while an AFL match is in play: Team A in possession; Team B in possession; Ball in dispute; Team A goal (6 points); Team B goal; Team A behind (1 point); Team B behind. In comparison, Hirotsu (2003) was able to work with just four Markov states when analysing the characteristics of association football teams: Team A goal;
Team A in possession; Team B goal; Team B possession. The ball spends a lot more time in dispute in AFL, a by-product of being permitted to handle the ball.

The continuous linear scoring process in AFL offers punters a number of markets to bet on, including head-to-head, covering the line and total points scored in a match. Bailey (2000) used linear regression to arrive at a probability of victory for Team A, comparing this with head-to-head market prices offered and, finally, developing a fixed and Kelly wagering strategy; the fixed strategy proved more profitable than the Kelly approaches. Bailey and Clarke (2004) identified inefficiencies in AFL betting markets by predicting margins in matches with a multiple regression model where home advantage, travel, team quality and current form were all significant predictors. Team quality was measured at a team level by analysing past team scores and, also, at a player level where quality was an average of past match results associated with each player named in the team. The player-based approach produced lower margin errors than the team approach as well as a 15% return on investment, compared with the team’s 10%. The relative success of the player-based approach by Bailey and Clarke (2004) was very encouraging in the initial developmental stages of the AFL player ratings in this research.

1.2.2 Continuous Sport Player Performance

In Chapter 4, continuous sports are defined as having the (unrealistic) ability to fluidly progress without any dead or disputed phases for the entire in-play period. “Continuous” is a word to describe many “invasion” sports or sports where a team must enter the opponent’s zone to score. Bell and Hopper (2000) give a good classification of games played by individuals and teams. In team sports, notational analysis is primarily concerned with the interaction between players and the movements and behaviours of individual team members (Hughes and Bartlett, 2002). This concept underpins the team player performance measurement techniques pursued in this dissertation as well as by Oliver (2004). Notational analysis of individual performances in a team sport are the result of studying performance indicators relevant to the sport in question; such indicators include successful passes, tackles and shots on goal in football (see Section 4.2). Hughes and Bartlett (2002) state that the utility of performance indicators is reflected by their contribution to a successful performance or
outcome. However, coaches still focus on “negative” performance indicators, such as turnovers, to highlight areas of improvement for his or her players. Nevill et al (2002) described most performance indicators as discrete events, with comparisons drawn from the frequency distributions of factors such as field position and player roles. The sections of their paper relevant to this research detailed the effective application of discrete distributions (Poisson and Binomial) to individual performance indicators—effective kicks/handballs in AFL. A paper critical to the notational analysis in this research, by Reep and Benjamin (1968), dealt with passing between players in a football match, where a set of successful passes resulted in one of the following: a shot at goal, an infringement or an interception. They determined the probability of $x$ passes using a negative binomial distribution.

Further discussion on this approach takes place in the “AFL player interaction simulation” section of this review and in Chapter 7. Hughes and Franks (2005) augmented the work of Reep and Benjamin (1968) by normalising the same data—dividing the number of goals scored in each possession by the frequency of the sequence length—and proving that teams skilful enough to retain the ball and achieve longer passing sequences had a greater chance of scoring. The minute analytical detail within the flow of games was exhibited by Brillinger (2007), who analysed one particular sequence of passes ($n = 25$) by Argentina in the 2006 World Cup. Specifically, he used potential functions to simulate the motion of a soccer ball.

When analysing performance indicators at a team or player level, coaches may be interested in performance consistency or the degree of variability in match performances. Marquardt (2008) remarked that the more regularly investment managers beat their peers and their benchmarks, the greater the likelihood that skill, rather than luck, is driving their performance. Waldman (Waldman, M 2005) believes a (fantasy) team of players performing consistently at a desired level is more valuable than higher-scoring but more erratic players. We investigated a coefficient of variation which has been incorporated in the past to measure performance consistency—that is, points scored in basketball (Manley, 1988). Elderton (1909) used a coefficient of variation ($CV = 100 \times \frac{\text{standard deviation}}{\text{mean}}$) to measure a batsman’s scoring consistency where a coefficient closer to zero implied more consistent performances. Bracewell (2003) employed control charts to trigger an alarm when a rugby player’s
performance variability exceeded two standard deviations of his rating. While a parametric approach to consistency measurement is a relatively simple approach—that is, using a coefficient of variation—AFL performance measures are quite often right skewed, hence violating the correct use of moments from a normal performance distribution. Tukey (1957) and Box-Cox (1964) transformations were applied to AFL performance indicator distributions in sections of this work to achieve a near normal distribution. A closer approximation to the normal distribution was achieved by, firstly, transforming the performance data with respect to a player’s time-on-ground (James et al, 2005). We were able to prove that the consistency measures derived from transformed indicators were more highly correlated with player award votes than were untransformed ratings. Chinn (1996) offered other interesting transformation techniques, specifically power methods and maximum likelihood approaches. James et al (2005) overcame non-normal performance indicator distributions by using the median rather than the mean as an approximation of player performance. We used non-linear median smoothing to great effect to forecast player ratings.

If a coach was rating all of his players on frequency of kicking in each match for the season, he would benefit from a normalising procedure to account for players whose position exposes them to more play than others or, in other words, to rate each position as a separate entity. The importance of determining an individual player’s performance in a certain position in his team, for example, a defensive player, is reflected by the considerable body of work on the subject. James et al (2005) used medians with confidence limits in developing performance profiles in rugby which were dependent on player position. He even isolated intra-positional profiles: a “Prop”, “Hooker” and “Lock”, all forwards, each displayed differing frequencies of selected indicators, for example, successful and unsuccessful tackles. Discriminant analysis has been a common tool for classifying players into relevant positions. This approach greatly benefited this research because the data provided to us contained uniform player positions across all historical matches; that is, a forward was retrospectively classified as a forward for each match in a year even though he may have played one match as a defender. Chapter 6 demonstrates how we retrospectively reclassified each player into the relevant on-field positions by observing his performance indicators, thereby yielding a rating which more
accurately recognised his and his positional opponent’s position-specific indicators. Fratzke (1976) was able to determine basketball player ability and position using varying biographic data, while Sampaio et al (2006) employed discriminant analysis to maximise the average dissimilarities in game statistics between guards, forwards and centres in the National Basketball Association (NBA). Pyne et al (2006) concluded that AFL draft fitness assessments, involving statistical analysis on physical qualities such as height, mass and agility, were useful in determining future player position. A discriminant approach would certainly be useful for this purpose. A discriminant function was able to classify AFL players into one of four positions by maximising the Mahalanobis distance between the positional centroids as determined by a linear combination of player performance indicators. Chatterjee and Yilmaz (1999) also used a Mahalanobis distance measure, looking at skill variability between MVP basketball players. For the purposes of AFL player performance measurement, a player may be correctly assigned to a position, but how realistic is it to assume he played the entire match there? To mathematically ascertain this knowledge, posterior probabilities were calculated for each player’s defensive, forward, midfield and ruck roles in each match (James, 1985). Intra-position analysis was also of interest. James et al (2005) employed chi-squared testing to examine performance differences between rugby players in the same positions. We were able to investigate intra-position differences between AFL skills by examining two sets of covariates ([Mark, Goal], [Handball Receive, Goal]) in a two-dimensional space (Gordon, 1981). Players who received a handball before scoring were more likely to be smaller, mobile forwards, kicking a goal on the run rather than from a set shot (from a mark or free kick).

1.2.3 Team and Player Ratings

Ratings in sport offer an objective evaluation of the performance of a team or individual player, often calculated with prior performances in mind. A rating, say, on a numerical scale from 0 to 1000, where 1000 is the highest, yields a ranking of any competitor where the first ranked has achieved a rating closer to 1000, relative to the other competitors. Stefani (2010) offers a comprehensive survey of official rating systems published by recognised sports federations. Arpad Elo was a pioneer in the field of statistical ratings. His book, The Rating of Chess Players, Past and
Present (1978), detailed the Elo rating system with respect to the individual sport of chess. Elo (1978) arrived at a probability of victory in a match as the difference between observed performance—wins (1), losses (0) and draws (0.5)—and an expected performance, that is, a function of opponent strength. The Elo rating system has also been applied in team sports where performance can be quantitatively measured by score. The most notable application was to rate men’s national world football teams (World Football Elo Ratings 2007), not to be confused with the FIFA World Rankings (FIFA World Rankings 2007). Although parameters take on different interpretations depending on the sport being played, the Elo principle remains the same: a pre-match numerical rating with a post-match revision dependent on the quality of the opposition, the significance of the occasion and the extent (size) of the win when performance can be quantified. In the early stages of AFL player rating development, the approach was influenced by Elo, namely, an adjustive system where players are pitched in simulated “head-to-head” contests where player i’s score and opponent j’s score are randomly generated from two independent normal score distributions from that season, prior to the impending match (see Chapter 5). Opponent quality is, thus, accurately accounted for. Stefani and Clarke (1992) discussed accumulative and adaptive, or adjustive, rating systems, where adaptive systems cause ratings to rise or fall in line with good or bad performance, as is the case with Elo ratings. They used a least squares and an exponential smoothing rating system to help predict the outcome of AFL matches. Stefani (1997) offered a rating framework for individuals in professional competitions in golf, cross-country skiing, alpine skiing, men’s tennis and women’s tennis, as well as a team rating for football (soccer). He observed three phases for rating systems, adapted for these sports, over a season: weighting the observed results to produce points; aggregating these points to provide a seasonal value; and calculating a rating from the seasonal value. Like Elo, Stefani’s football ratings reflect a team’s performance relative to the opposing team. Rather than applying one rating system to many sports, Bedford (2004) applied three ratings systems to one sport, women’s handball, to determine which was the most predictive of World Cup matches from 2001 and 2003. He used an exponential smoothing model, a modified Elo model and a Pythagorean Projection model, the latter evaluating the expected win percentage for a team based on its accumulated and conceded scores. Bedford (2004) concluded the Pythagorean model was the best predictor for the tournaments, with Elo slow to react to certain
tourneyment results. Barrie (2003) was confronted with similar lag problems using Elo to rate participants in Tiddlywinks tournaments, prompting him to develop a tournament rating, with a player’s new rating being the weighted average of his original and tournament rating. The warning of lag effects offered by Bedford (2004) and Barrie (2003) was important prior to developing the adjustable player ratings in Chapter 5.

Objectively evaluating a player’s performance in a continuous team sport such as football or hockey using a statistical rating system is a more difficult task than rating a team. The scarcity of published literature on player ratings reflects a lack of understanding, a lack of demand or both. While it is trivial to observe the performance of a team using a set of final scores and margins, an equitable framework for player performance measurement is of continued debate. World football is one of the few sources of team sport player ratings, based on the actual performances of every player across Europe’s top five leagues (Castrol Football Rankings 2007). Rugby player ratings were developed by Bracewell (2003) using factor analysis, where each factor represented a core trait performance across nine positional clusters. Chapter 5 details the importance of position-specific ratings given the vastly different roles, say, a forward and a defender play in a single AFL match. Points Scored, Rebounds, Assists, Steals and Turnovers in basketball were analysed by Chatterjee and Yilmaz (1999) using a covariance matrix and Mahalanobis distances compared across a list of MVP winners to gauge the best overall performance. Oliver (2004) went to admirable efforts to establish a set of offensive and defensive ratings for each player in a basketball match; these were of an additive nature using frequencies of on-court skills such as effective passes and score assists executed by that player. More impressively, he developed, using relatively simple mathematics, a “difficulty theory for distributing credit to players in basketball”: the more difficult the contribution, the more credit it deserves. This concept of evaluating player contributions in team sports was also explored by Duch et al (2010) for the 2008 Euro Cup (football) tournament and was a key motivation for this work, especially because of the large number of players in an AFL side. Also of interest was the improvement in explained variation that Bailey (2000) achieved when calculating a probability of victory using individual

---

1 These rankings have since been discontinued across the European Leagues and isolated to the USA’s and Canada’s Major Soccer League (see Section 11.4 for the top 50 ranked players).
factors such as kicks and handballs, rather than team aggregates. This suggested that collective individual player ratings could be predictive of team success. The pragmatic approach adopted by Oliver (2004), particularly his “difficulty theory”, inspired certain developments for these AFL player ratings.

Discrete sports such as baseball and cricket offer a slightly improved source of literature on player ratings. Although not strictly classified as player ratings, Koop (2002) employed economic efficiency methods for comparing the batting performance between baseball players. This paper was interesting because it proved how different each batter’s roles are in baseball; in cricket, the designated batsmen’s roles are reasonably similar, especially in the longer format (test match cricket). The International Cricket Council (ICC) player ratings (ICC Rankings 2011) are internationally acclaimed and published for public consumption, calculated by a moving average of each batsman’s/bowler’s past performances and then weighted by various match conditions, such as quality of opponent, rate of scoring and match outcome. Croucher (2004) expanded on the naivety of the universally accepted rating, the batting/bowling “average”, by developing batsman/bowler ratings based on momentum, or batting/bowling average (mass) multiplied by batting/bowling strike rate (velocity). Kimber and Hansford (1993) addressed problems associated with traditional batting averages by investigating the hazard functions of batsmen.

1.2.4 AFL Player Interaction Simulation

A principle motivation for this research was evaluating an AFL player’s contribution to his team’s performance, in any match, by way of a player rating. While fantasy football ratings (match-by-match player scores) are generally a good approximation of performance, they are probably too player-centric, ignoring an important underlying concept that a team is supposed to be more than the sum of the individual players (Gould and Gatrell, 1979/80). Duch et al (2010) argued that the real measure of player performance is “hidden” in the team plays and not derived from strictly individual events associated with a certain player. Moreover, in their research on football-passing patterns from UEFA Euro 2004, Lee et al (2005) measured passing between players at a group level rather than at an individual level, demonstrating how a player’s passing patterns determined his location in the team.
network. Discussions about network analysis commonly refer to the use of relational data or the interactions that relate one agent (player) to another and, so, preclude the properties of the individual agents themselves (Scott, 2000). The chapter on distributing credit in Oliver (2004) and the paper by Duch et al (2010) are rare but engaging attempts at quantifying a player’s contribution to his team’s performance. The latter constructed a network of ball passing among the players on a football team to produce a probability of a shot at a network node, using network centrality measures. Centrality is one of the most widely studied concepts in network analysis and allows implicit assumptions about the prominence of an individual in a network (Lusher et al, 2010). Borgatti (2005) provides excellent definitions and applications of centrality in its various forms. In Chapter 7, symmetric interaction matrices are generated, in a similar approach to Gould and Gatrell (1979/80), for each match played by the Geelong Football Club in 2011 and negative binomial distributions (nbd) fitted to each player pair in the matrix so that their interaction frequency could be simulated. Pollard et al (1977) concluded that the nbd is a closer fit with events resulting from groups of players than from individual performances; for example, an improved fit is observed from batting partnerships in cricket rather than from individual batsman scores. Reep and Benjamin (1968) successfully modelled effective passes in world football with the nbd, while Pollard (1973) demonstrated how the number of touchdowns scored by a team in an American Football match closely followed the nbd. Nevill et al (2002) believed the difficulty involved in fitting a negative binomial distribution to frequency data lay mainly in the incapability of statistical programs. A Visual Basic looping module was written for this research that fitted a negative binomial distribution to each player pairing in the Geelong side, with parameters selected so as to minimise the Pearson chi-squared statistic. Pollard et al (1977) used a method of moments to estimate these parameters, but we found lower mean absolute errors for the large majority of the pairings using the chi-squared approach. Using eigenvector centrality (Borgatti, 2005), a measure of each node’s (player’s) influence on the network, we were able to calculate a player rating for each player in the simulated interaction matrix. Finally, we discovered a significant relationship between the team performance—an average of the player ratings (Duch et al, 2010)—and final score margins through the season. The network simulations’ effect on final margin meant we
could observe the chance of victory with different players in and out of the side and attach to each player an observed contribution.

### 1.2.5 AFL Player Performance Forecasts

Smoothing is a popular and effective forecasting technique when analysing sporting data. A predetermined number of past matches can be analysed and used to project how a player might perform in future matches. Clarke (1993) employed exponential smoothing to aid in the prediction of match results for the Australian Football League, while Bedford and Clarke (2000) investigated a ratings model for tennis players based on exponentially smoothing margins of victory. *Exploratory Data Analysis* by Tukey (1971) was extremely influential on this dissertation. Tukey (1971) and Velleman (1980) discuss nonlinear smoothing as a method of reducing the misleading effect of abrupt features in data sets prior to data exploration—for example, in sport, an unexpected poor performance because of injury or wet conditions. Chapter 8 demonstrates the positive effect of running a nonparametric 4253H smoother—the values are medians and the “H” is the Hanning weighted average—through a data set prior to forecasting. Sargent and Bedford (2007) employed a 4253H smoother to remove noise from player performance data when calculating simple AFL player ratings, while Shepherd and Bedford (2010) used the same smoother on the probabilities of winning a medal in pistol shooting. Gebski and McNeil (1984) examine the use of nonlinear smoothers as opposed to linear ones (where Gaussian assumptions should be met), identifying three important properties: resistance to outliers; retention of peaks and troughs (still resisting outliers); and repeatability until no further change occurs (still preserving peaks and troughs). Given the non-Gaussian nature of some of the AFL performance indicator data, we proposed nonlinear smoothers were appropriate for the analysis. With a minimum of seven games for each player (Velleman and Hoaglin, 1981, Janosky et.al., 1997), the objective was to forecast an AFL player’s performance scores for a subsequent round with minimised average season error. Goodman (1974) was able to reduce mean-squared error in exponentially smoothed data sets by *twicing* the smoothed residuals. Instead of exponentially smoothing residuals, we exponentially smoothed a brand-new nonlinear-smoothed series. Where Goodman (1974) applied Tukey (twicing) to exponential smoothing to reduce error, we exponentially
smoothed Tukey. Lower errors were achieved by smoothing the Tukey-smoothed series rather than using simple score means or exponential smoothing on its own. It is anticipated that this approach will translate to any sport where performance can be scored.

1.2.6 Statistical Analysis of Batsmen in Cricket

The “one-on-one” nature of cricket and baseball—that is, a specialised bowler/pitcher competing directly with a specialised batter—offers countless directions for the analysis and practical application of the players’ statistics. Lewis (2005) went so far as to describe the game of cricket as a “sporting statistician’s dream”. There is extensive literature on player performance in the traditional and modern forms of cricket, mainly concerning batsmen. In an attempt to predict batsmen’s scores, numerous efforts have been made over the last century to retrospectively fit models. Early work by Elderton (1945) proposed test match cricket batting scores followed a geometric distribution, but subsequent research revealed this distribution provided an inadequate fit for zero and extreme scores (Wood, 1945, Kimber and Hansford, 1993, Bailey and Clarke, 2004). Bailey and Clarke (2004) log-transformed batting scores to alleviate this problem in the tail of the distribution, while Bracewell and Ruggiero (2009) employed a beta distribution to model zero scores, separate to non-zero scores that were modeled with a geometric fit. Kimber and Hansford (1993) preferred a product-limit estimator because it did not depend on parametric assumptions. Apart from an excellent paper by Swartz et al (2006), who applied a log-linear approach to simulate runs scored during any stage in a match for a proposed batting order, there is a scarcity of models that can statistically describe a batsman’s scoring progress and expectation while a match is in play. Swartz et al (2006) made mention of how simulation techniques in limited-overs cricket have not yet been fully explored. Simulation became a critical element in this research so in-play batsman performances could be estimated (see Chapter 9). Where this research forward simulated to assess and then reassess the probability of live batting outcomes, Swartz et al (2009) estimated the finite outcomes in a match with a Bayesian latent variable model, taking into account the characteristics of the individual batsmen as well as the bowlers. The inclusion of bowling quality was unprecedented for in-play research and has been flagged for augmenting this research. The estimates were conditional on the batsman’s order in the line-up, his
type (fast, medium or slow scoring) and the score (runs and dismissals), which described the state of
the match; bowling quality would increase the value of the model. Swartz et al (2006) employed
simulated annealing to analyse the many permutations within a team’s batting line-up to arrive at
optimal batting orders. This was very interesting and important work and should be highly regarded
by coaching and selection staff.

A limited-overs cricket team’s performance is not always accurately communicated to the consumer. If the team batting second, for example, achieves the required runs to win the match, that team is said to have won by however many batting resources (wickets) it still possesses at the match completion. Clarke and Allsopp (2001) and de Silva et al (2001) made use of the Duckworth-Lewis rain interruption rules (Duckworth and Lewis, 1998) to project a second innings winning score, after the match’s completion, to calculate a true margin of victory with respect to runs, not just wickets. The in-play batting estimate adjustments in this work were influenced by this method; that is to say, a second innings batsman could go on to make \( x \) runs if the match continued after the target was achieved.

Limited-overs cricket offers attractive betting markets due to the finite outcomes and abbreviated match duration. Bailey and Clarke (2004) designed strategies to maximise profits derived from wagering on one batsman outscoring another (“head-to-head”) during the 2003 ODI World Cup. The in-play simulation methodology adds further appeal to wagering on markets such as “highest-scoring batsman” because there is a real-time knowledge of the match conditions with which to improve the likelihoods of a batsman outscoring all others in his team. Easton and Uylangco (2007) were even able to provide some evidence of the ability of market odds to predict the outcomes of impending deliveries in ODI matches. This will ensure ongoing research for this in-play model.

1.2.7 Conclusion

The scarcity of published research on player ratings in the AFL indicated an entry point in the market for this work. The majority of papers on ratings have dealt with team ratings and/or player ratings in individual sports (Elo, 1978; Stefani and Clarke, 1992; Stefani, 1997; Stefani, 2010; Barrie,
2003), but little exists on player ratings in team sports. Originally, the AFL player ratings resembled individual ratings developed by Elo (1978). But, as the research progressed, we found that adjustive systems incorporating network centrality measures—for example, those used by Duch et al (2010) and Borgatti (2005)—were more appropriate because we could calculate each player’s contribution to the team network. The seminal works of Reep and Benjamin (1968) and Pollard (1977) guided the final layer of the work, inspiring the use of the negative binomial distribution to simulate interactions between the players in the network and, hence, to calculate player ratings for each competitor. The cricket batsmen score estimations have continued the exploration of simulation in the limited-overs format, pioneered by Swartz et al (2006) and Swartz et al (2009). Finally, special mention is to be made of Oliver (2004) whose tireless efforts to rate basketball players’ performances was a constant inspiration.

1.3 Research Questions and Publications

An outline of research questions and corresponding chapters and publications is important in emphasising the motivations and direction of this research. The initial research questions for this research dealt with appropriate development of an Elo-influenced AFL player ratings system; these questions form the core of Chapter 5. The AFL player research culminated in questions about the validity of network simulations for player performance which are answered in Chapter 7. Research questions concerning player performance simulation in limited-overs cricket are investigated in Chapter 9.

Chapter 5: Ratings Systems

1. Can the traditional Elo rating be reasonably adapted to reflect the performance of a player in a continuous team sport such as Australian Rules football?


2. What performance indicators best describe an AFL player’s performance in any match?
3. What distribution do such variables follow, and how does this affect their inclusion in the ratings system?
4. Can one rating formula apply to all players on the field, or is a separate formula required for each position?
5. What are the other factors that need to be considered when developing an AFL player-rating model?


6. What is the best measure of player performance consistency, and how can we validate the findings?


Chapter 6: AFL Positional Analysis

7. The AFL player data provided by Prowess Sports contained a player position variable (D = Defender, F = Forward, M = Midfielder, R = Ruck), but the assignations were uniform for each game. Can we retrospectively reclassify each player based on his accumulated performance indicators?

8. Is it possible to classify intra-position performance? Does this improve the analysis?

Chapter 7: Link Plays and Player Interaction Simulation

9. How does this research improve by analysing “transactional” performance data in an AFL match, rather than post-match aggregates?

10. Can we apply network analyses to the transactional data to measure the contribution of a player to his team’s performance, and how can we validate the findings?


11. Can we simulate player interactions? Which method is the most appropriate and why?

12. To what extent is the probability of team success dependent on these simulated interactions?

13. Do ratings based on network centrality reflect a player’s contribution to team success?


Chapter 8: Player Performance Forecasting

14. Can we generate accurate player performance predictions (expected values) using non-linear (Tukey) smoothing and validate the findings?

Chapter 9: In-play Simulation in Cricket

15. Can we simulate outcomes in discrete sports, specifically, in limited-overs cricket? What is the most appropriate method?

16. Are we able to develop player ratings from these simulations?

17. What other factors need to be considered in the cricket simulations for them to become more accurate?

Chapter 2

The AFL and Cricket

This chapter provides a summary on the history, rules and mechanics of the Australian Football League (AFL) and limited-overs (ODI) cricket, the primary sports of interrogation throughout this dissertation. The players’ roles in each sport are also introduced to provide context for the foremost subject of the research, player ratings. At the end of the AFL and cricket sections, there are examples of how current teams and players are able to be ranked in each sport with respect to awards and metrics. These ranking examples provide early context for the subsequent chapters of this dissertation.

2.1 The AFL

2.1.1 History

The Australian Football League (AFL) is the highest level of Australia’s national sport, Australian Rules football. It is played between 18 teams from five different states, with a season running annually between late March and late September. Each team plays 22 matches in the year, including a mid-season bye round. Originally founded as the Victorian Football League (VFL) in 1897 and consisting of only eight Victorian teams, the league grew to 12 teams in 1925, which would remain unchanged until 1987 when the West Coast Eagles (Western Australia) and Brisbane Bears (Queensland) were added. In 1982, the league expanded outside of Victoria and into New South Wales when the South Melbourne Football Club became the Sydney Swans; from this time, the
league was known as the Australian Football League. The addition of the Adelaide Crows (South Australia) in 1991 and Port Adelaide Power (South Australia) in 1997 ensured all but one (Tasmania) Australian states were represented. The Gold Coast Suns (Queensland) and Greater Western Sydney Giants (New South Wales) were added in 2011 and 2012 respectively, a reflection of the AFL’s push for recognition in areas of Australia where Australian Rules football is not the primary sport. While the bid to establish an AFL team in Tasmania is ongoing, the fluctuating levels of support have proven to be an obstacle. Nevertheless, the Hawthorn Hawks and North Melbourne Kangaroos each play a selection of “home” matches at grounds in the Tasmanian cities of Launceston and Hobart respectively.

The popularity of a sport is readily measured by match attendances. The AFL averaged crowds of 34,893 in regular 2011 season matches and 68,309 for finals matches (AFL Crowds 2011). In comparison, Rugby League, a football code most popular in New South Wales and Queensland, averaged only 16,273 spectators for the regular 2011 season and 37,937 for finals matches. The Melbourne Cricket Ground in Victoria was the most popular venue for watching AFL in 2011, averaging 53,972 spectators for the regular season and 76,111 for the finals. Subiaco Oval in Western Australia was the next most popular, averaging 35,915 spectators, while Melbourne’s Docklands Stadium averaged 32,825. The game’s nationwide appeal is evident, with an estimated 2.6 million Australians watching the grand final match on television in 2011. In 2010, 2.8 million people watched the drawn grand final and 2.68 million watched the rematch the following week (Knox D, 2011).

2.1.2 Basic Rules and Mechanics

AFL is an invasion game of similar fluidity and dynamics to football (soccer), except players are permitted to use their hands to punch (handball) the ball to the advantage of a team member. Kicking is the chief form of scoring or delivering the ball to a teammate. If player $i$ kicks the ball and it travels over 15 metres and is caught (marked) by player $j$ without the ball touching the ground, player $j$ may stop and quickly assess where his next kick will be directed without being confronted by an opposition player. If the ball is picked up off the turf by player $j$ in general play or the kick from

---

2 Rugby League is more commonly played in these areas.
player $i$ did not travel an estimated 15 metres, then he must deliver the ball to player $k$ (or back to player $i$) as effectively as possible or run the risk of being tackled by an opposition player. If he does not dispose of the ball by foot or by hand and is tackled by an opposition player, then a turnover results and the opponent is awarded a free kick, allowing him a few seconds to assess his next move.

The ultimate objective is to score a goal, worth six points, by kicking a ball through two upright posts at opposing ends of the ground, set roughly 6.4 metres apart on the boundary line (see Figure 2.1). If the ball is kicked between one of these goal posts and the smaller posts either side, one point is registered. The team with the most points at the completion of the match is declared the winner. Most goals are kicked within an arc that extends from boundary to boundary, 50 metres out from the goal face (see Figure 2.1). Powerful kickers are able to kick goals from beyond the arc, particularly when they are on the run. Goals are kicked from one of two scenarios:

1. Static play: a mark or a free kick that stops the play and allows the player to shoot for goal from a realistic distance. The kicker can stop, walk back, run in and kick the ball from where he marked the ball or was awarded the free kick, without opposition hindrance, similar to a punter in American Football.

2. General play: a few possibilities for goal scoring exist:
   
   i. a player runs within a realistic distance of the goal, having received the ball from a teammate/opposition player, and kicks it through the posts on the run
   
   ii. a player recovers a loose ball from the ground within scoring distance and kicks

   iii. a player kicks the ball off the ground and through the goal, as in football, without any use of the hands.

Opposition players are permitted to run at and tackle the player with the ball in Scenario 2. If players are tackled by the opposition and they did not have an opportunity to release the ball (the ball is said to be in dispute), the play stops and the umpire bounces the ball again where the player was tackled.

In the case where a ball bounces on and then out of the field of play, the umpire throws the ball back
into the field of play from the boundary (a throw-in) for the ruckmen to compete. If the ball is kicked out of the field of play without landing inside first, a free kick is awarded against the offending kicker.

Figure 2.1: The Melbourne Cricket Ground with goal posts at the left and right

2.1.3 The Players

Twenty-two players compete on each AFL side in each match, the most number out of any commonly recognised team sport. Players compete in four different positions—defenders, midfielders, forwards and rucks—and can rotate through these positions at any point in the game if the player is physically suitable. It is a forward’s role to kick goals or create goals (goal assists). As in any other invasion sport, forwards predominantly play in their forward half of the ground. Defenders play in their defensive half and must prevent opposition players from scoring goals. During the game, it is usual to see “one-on-one” contests—that is, one defender guarding one forward in an attempt to stop him marking or gathering the ball and kicking/creating a goal. A ruckman, usually the tallest player on the team, is analogous to a player involved in a centre jump in basketball. The umpire bounces the oval-shaped ball into the turf\(^3\) in the centre of the ground and the two competing ruckmen must simultaneously jump and “tap” down the ball with their hand to their respective midfielders waiting on the ground. A hypothetical passage of play might comprise the following transactions between players from team \(a\):

\(^3\) Australian Rules football is one of the few games in the world whose start requires an umpire to possess a unique physical skill.
1. The umpire bounces the ball and team a’s ruckman jumps and taps the ball down and it is gathered by midfielder $i$ from team $a$

2. Midfielder $i$ handballs to midfielder $j$ to avoid being tackled with the ball by a midfielder from team $b$

3. Midfielder $j$ runs away from the centre of the ground towards the forward 50-metre arc and kicks the ball 20 metres to a forward who is running towards him (away from the goal)

4. The forward marks the ball 30 metres out from the goal face at a 30-degree angle

5. The forward walks back 10 metres and jogs to the point where he marked the ball and kicks the ball for a goal

6. The players jog back to the centre of the ground and the umpire bounces the ball to recommence play.

This sequence of play represents an extremely efficient movement of the ball with a successful outcome. Most passages of play involve several more transactions between the players, including turnovers to the opposition, before a goal is scored. These transactions are essential in latter parts of this dissertation (see Chapter 7).

It is not uncommon to see players moved around in positions: for example, a ruckman who is tiring in the ruck contests can be designated a forward role where his height is an advantage in marking the ball and kicking goals. The ruckman has the ability to rest in this position because the ruck position requires him to follow the ball around the ground to compete each bounce/throw-in, whereas the forward position is more static, isolated mainly within the 50-metre arc.

It is preferable for players to possess physical strength, speed and endurance because a match lasts a minimum of 120 minutes, with some players running well in excess of 15 kilometres in one match. The dimensions of the ground also demand that players run long distances; playing fields are between 135 and 185 metres long and between 110 and 155 metres wide. The Melbourne Cricket Ground, the foremost AFL ground, is visualised in Figure 2.1.
2.1.4 2012 Team and Player Rankings

AFL teams are ranked after each match in a season according to the number of wins and their winning percentage (points for/points conceded x 100) (see Equation (5.3)); the top 8 ranked teams at the completion of the regular season progress to the finals series. Hawthorn finished the regular season on top of the 2012 season ladder, with 17 wins and a percentage of 154.6% (see Table 2.1). Somewhat unexpectedly, the team went on to lose to Sydney in the grand final. St Kilda required two more wins in order to make the finals. Greater Western Sydney, in its debut year, finished bottom of the ladder, with only two wins.

<table>
<thead>
<tr>
<th>Place</th>
<th>Team</th>
<th>P</th>
<th>W</th>
<th>D</th>
<th>L</th>
<th>%</th>
<th>Pts</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hawthorn</td>
<td>22</td>
<td>17</td>
<td>5</td>
<td>5</td>
<td>154.6</td>
<td>68</td>
</tr>
<tr>
<td>2</td>
<td>Adelaide</td>
<td>22</td>
<td>17</td>
<td>5</td>
<td>5</td>
<td>132.5</td>
<td>68</td>
</tr>
<tr>
<td>3</td>
<td>Sydney</td>
<td>22</td>
<td>16</td>
<td>6</td>
<td>6</td>
<td>140.6</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>Collingwood</td>
<td>22</td>
<td>16</td>
<td>6</td>
<td>6</td>
<td>116.5</td>
<td>64</td>
</tr>
<tr>
<td>5</td>
<td>West Coast</td>
<td>22</td>
<td>15</td>
<td>7</td>
<td>7</td>
<td>124.2</td>
<td>60</td>
</tr>
<tr>
<td>6</td>
<td>Geelong</td>
<td>22</td>
<td>15</td>
<td>7</td>
<td>7</td>
<td>117.1</td>
<td>60</td>
</tr>
<tr>
<td>7</td>
<td>Fremantle</td>
<td>22</td>
<td>14</td>
<td>8</td>
<td>8</td>
<td>115.7</td>
<td>56</td>
</tr>
<tr>
<td>8</td>
<td>North Melbourne</td>
<td>22</td>
<td>14</td>
<td>8</td>
<td>8</td>
<td>112.5</td>
<td>56</td>
</tr>
<tr>
<td>9</td>
<td>St Kilda</td>
<td>22</td>
<td>12</td>
<td>10</td>
<td>10</td>
<td>123.3</td>
<td>48</td>
</tr>
<tr>
<td>10</td>
<td>Carlton</td>
<td>22</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>108.0</td>
<td>44</td>
</tr>
<tr>
<td>11</td>
<td>Essendon</td>
<td>22</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>100.0</td>
<td>44</td>
</tr>
<tr>
<td>12</td>
<td>Richmond</td>
<td>22</td>
<td>10</td>
<td>1</td>
<td>11</td>
<td>111.6</td>
<td>42</td>
</tr>
<tr>
<td>13</td>
<td>Brisbane Lions</td>
<td>22</td>
<td>10</td>
<td>1</td>
<td>12</td>
<td>91.0</td>
<td>40</td>
</tr>
<tr>
<td>14</td>
<td>Port Adelaide</td>
<td>22</td>
<td>5</td>
<td>1</td>
<td>16</td>
<td>78.9</td>
<td>22</td>
</tr>
<tr>
<td>15</td>
<td>Western Bulldogs</td>
<td>22</td>
<td>5</td>
<td>1</td>
<td>17</td>
<td>67.0</td>
<td>20</td>
</tr>
<tr>
<td>16</td>
<td>Melbourne</td>
<td>22</td>
<td>4</td>
<td>18</td>
<td>18</td>
<td>67.5</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>Gold Coast</td>
<td>22</td>
<td>3</td>
<td>19</td>
<td>19</td>
<td>60.8</td>
<td>12</td>
</tr>
<tr>
<td>18</td>
<td>Greater Western Sydney</td>
<td>22</td>
<td>2</td>
<td>20</td>
<td>20</td>
<td>46.2</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 2.1: Final ladder of 2012 AFL regular season

Table 2.2 displays how AFL players may be ranked according to perceived performance. The Brownlow Medal (see Section 4.1) is the most prestigious individual award in the AFL with umpires deciding the best three players in a match with a 3, 2, 1 voting system. The best and fairest player for 2012 was Jobe Watson from Essendon. He was awarded 30 votes, with six 3-vote games (best on ground). It is worth noting that the top ten players are all midfielders, justifying the award’s nickname, the “Midfielder’s Medal”. Matthew Pavlich, mostly playing as a forward, is the only player in the top 20 who is not considered a midfielder. The most common explanation for this phenomenon is that midfielders are constantly following the ball and, hence, in the sight of umpires more than
forwards or defenders who are mostly in designated positions at either ends of the ground. This award
is discussed in more detail in subsequent chapters.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Player</th>
<th>Team</th>
<th>Matches</th>
<th>Votes</th>
<th>3-votes</th>
<th>2-votes</th>
<th>1-votes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Watson, Jobe</td>
<td>ESS</td>
<td>22</td>
<td>30</td>
<td>6</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>Cotchin, Trent</td>
<td>RIC</td>
<td>22</td>
<td>26</td>
<td>6</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Mitchell, Sam</td>
<td>HAW</td>
<td>21</td>
<td>26</td>
<td>4</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Swan, Dane</td>
<td>COL</td>
<td>18</td>
<td>25</td>
<td>6</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Thompson, Scott</td>
<td>ADE</td>
<td>22</td>
<td>25</td>
<td>4</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Ablett, Gary</td>
<td>GST</td>
<td>20</td>
<td>24</td>
<td>6</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>Dangerfield, Patrick</td>
<td>ADE</td>
<td>22</td>
<td>23</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>Kennedy, Josh</td>
<td>SYD</td>
<td>22</td>
<td>19</td>
<td>5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Beams, Dayne</td>
<td>COL</td>
<td>21</td>
<td>19</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>Hayes, Lenny</td>
<td>STK</td>
<td>22</td>
<td>19</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>Pavlich, Matthew</td>
<td>FRE</td>
<td>21</td>
<td>15</td>
<td>4</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>Selwood, Scott</td>
<td>WCE</td>
<td>22</td>
<td>15</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>Jack, Kieren</td>
<td>SYD</td>
<td>22</td>
<td>15</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>Pendlebury, Scott</td>
<td>COL</td>
<td>18</td>
<td>15</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>Stanton, Brent</td>
<td>ESS</td>
<td>20</td>
<td>14</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>Fyfe, Nathan</td>
<td>FRE</td>
<td>9</td>
<td>14</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>Selwood, Joel</td>
<td>GEE</td>
<td>20</td>
<td>14</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Jones, Nathan</td>
<td>MEL</td>
<td>21</td>
<td>14</td>
<td>1</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>Sewell, Brad</td>
<td>HAW</td>
<td>21</td>
<td>13</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>Deledio, Brett</td>
<td>RIC</td>
<td>21</td>
<td>13</td>
<td>2</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 2.2: Brownlow Medal votes for 2012

2.2 Cricket

2.2.1 Introduction

Cricket is an international sport played professionally between two teams from sixteen nations. There are three formats: test matches, “limited overs”, or One-Day Internationals (ODI), and T20, all of differing match durations. T20 is the shortest form of the game, internationally introduced in 2005 in an attempt to promote a more exciting brand of cricket due to the urgency with which runs must be scored. While attendance numbers are not as large as AFL, an average attendance of 24,174 for test matches in Australia since 2003 and 25,207 for ODI matches (Crowd Figures 2012) suggests a healthy national following. Recent market research found that nearly 7.8 million Australians (42% of the population) aged 14 years or above, watched cricket on television between October 2010 and September 2011 (Roy Morgan Research 2012).
2.2.2 Basic Rules

In this dissertation, the focus is on limited overs cricket (ODI) where each team needs to accumulate as many runs as possible for a maximum of 50 overs\(^4\) (one innings) or until 10 of the 11 batsmen in the batting team are dismissed (or “out”). The winning team is that which has accumulated the most runs when one of these terminal points is reached in the second innings. Each match consists of a maximum of 300 legitimate independent trials (deliveries)\(^5\) per innings, where designated bowlers from team \(a\) bowl consecutive deliveries to the “in” batsmen from team \(b\). It is the batsmen’s objective to score as many runs as possible through various means (see Section 2.2.3) to contribute to their team’s score. A batsman’s score is the aggregate of runs accumulated during his innings prior to:

a) himself being dismissed

b) all remaining teammates being dismissed or

c) the bowling team reaching its 50 overs without having dismissed 10 opponent batsmen.

For cases b) and c), the batsman would be recorded as “not out”.

While the avenues for scoring runs are nearly identical in ODI and traditional “test match” cricket, the durations differ, with a test match played over a maximum of five days. Test match cricket deliveries are of a continuous nature: prior to the match, it is unknown how many overs will be bowled in one of a possible four innings (two innings per team during the match); wickets are the only discrete resource. Given a batsman has only 50 overs to accumulate his score in ODI matches, his scoring rates tend to be more frenetic, yet more predictable, than in test matches\(^6\). In developing optimal scoring rates for ODI matches, Clarke (1988) worked under the commonly shared principle that batsmen are more cautious in early and middle stages of their team’s innings in an attempt to preserve team resources, while the latter stages of the innings are prone to aggressive batting to maximise team runs and to increase the probability of victory. A batsman, however, adopts the latter

\(^4\) An over consists of six deliveries.
\(^5\) Matches usually extend beyond 300 deliveries because illegitimate deliveries, such as ones deemed by the umpire to be too wide of the batsman, are required to be bowled again.
\(^6\) Test match batsmen are able to patiently bat for up to and exceeding 90 overs in a full day’s play.
tactic at his own peril because sustained aggression increases the risk of dismissal. The most common forms of batsman dismissals are:

1. Bowled: the bowler delivers the ball and the batsman is unable to prevent the ball from striking the wickets
2. Caught: a batsman strikes the ball and is caught by a fielder without the ball touching the ground
3. Leg Before Wicket (LBW): a delivery strikes the batter, most commonly on the leg, and the ball was deemed by the umpire to be going on to hit the stumps. This is a very basic explanation of LBW; more technical ones are available (Lords LBW 2012)
4. Run Out: a fielder or bowler breaks the wickets and the batsman is short of a painted line (crease) at the end to which he is running
5. Stumped: the batsman attempts to hit a delivery but misses, the ball is taken in the gloves of the wicketkeeper (a wicketkeeper is similar to a catcher in baseball) and he breaks the stumps with ball in hand and the batsman does not have his foot or bat within the crease.

There are five other dismissals, but they occur too rarely for consideration in this thesis.

2.2.3 Scoring

In limited overs cricket, runs per delivery are scored more frequently than runs per pitch are scored in baseball. From this research the average total team score in a limited overs match innings was estimated at 235 runs, whereas baseball, on average, realises between four and five runs per innings (Studeman D, 2012). The following section attempts to explain the various scoring actions in limited overs cricket. The batsman facing the bowler (defined as “on strike”) is responsible for contributing to the team’s runs through a number of scoring actions. The most common scoring action is the on-strike batsman hitting the ball into a space on the field away from the fielders, allowing the two batsmen to run to their opposite end of the rectangular pitch (approximately 20 metres or 22 yards long). A run is scored each time the batsmen cross and safely reach their opposite end, which is analogous to a base earned in baseball. A fielder, from the bowling team, must retrieve a hit ball and throw it to either end of the pitch where there are “wickets” (three upright wooden poles aligned in the turf) that he or the bowler must attempt to break to dismiss the batsman. Batsmen can cross up to four
times from a single delivery should the ball be hit into a large enough space on the field. The second score action is for the batsman to hit the ball with sufficient force so it reaches the boundary of the field after touching the ground at least once; four runs are instantly awarded to the batsman, and the batsmen do not have to cross. The third score action sees the batsman forcibly strike the ball so that it clears the boundary without touching the ground; six runs are instantly awarded to the batsman, and the batsmen do not have to cross. If a batsman is dismissed, or he was unable to be dismissed by the bowling side before its 50 overs were bowled, his total runs scored are aggregated to produce his final score.

It is trivial to simulate an ODI cricket scoring scenario, where batsman \( i \) is facing the first delivery of the first over in the match. As an example: First delivery: hit into a space to the batsman’s left, allowing batsman \( i \) and batsman \( j \) to cross the length of the pitch once (that is, one run) before the ball is thrown by an opposition fielder back to the wicket keeper to end the play. Batsman \( j \) is subsequently on strike after the batsmen crossed for an odd number (one) of runs. Should they have crossed twice (two runs) or four times (four runs), batsman \( i \) would retain the strike position. Second delivery: batsman \( j \) chooses not to hit and allows the ball to go through to the wicket keeper; no runs are scored. Third delivery: batsman \( j \) hits firmly and the ball lands, bounces once and then reaches the field boundary (four runs). Fourth delivery: batsman \( j \) hits into space behind the pitch, resulting in a safe single cross of the batsmen (one run). Fifth delivery: batsman \( i \) hits the ball along the ground directly to a fielder on his right, but the batsmen do not attempt to cross over because the fielder quickly picks up and returns the ball to the wicket keeper. Final delivery: batsman \( i \) hits the ball into space for the batsmen to safely cross twice (two runs). Batsman \( i \) has scored 1 + 0 + 2 = 3 runs and batsman \( j \) has scored 0 + 4 + 1 = 5 runs. The end of the over is realised and the new bowler bowls from the opposite end to the previous over, meaning batsman \( i \) is on strike (he was at the non-striker’s end because a single run was scored on the last ball of the first over). The scoring process continues. In the case of dismissal, batsman \( i \) would be replaced by batsman \( k \) who would assume the scoring process with batsman \( j \). This sequence continues in the first innings until all 10 batting resources, or 300 delivery resources, are exhausted. No runs are awarded to a batsman on dismissal.
Runs can also be conceded by the bowling team where the batsman did not make contact with the ball. In the following events, the runs are added to the batting team’s total, not to the total of the batsman on strike. Should the ball hit the batter on the body, without hitting his bat, and rebound into enough space so that the batsmen can complete a run(s) or the ball reaches the boundary, a “leg bye” is recorded for the batting team. If the ball does not hit any part of the batsman and is unable to be retrieved by a member of the bowling team before the batsmen score a run(s) or the ball reaches the boundary, a “bye” is recorded. If the bowler delivers the ball outside a designated lateral range to the batsman, a “wide” is recorded; this is similar to a wild pitch in baseball. If the bowler does not have a part of his foot behind the bowling crease—a white line running across the pitch approximately four feet in front of the wickets—as he delivers the ball, a “no ball” is recorded. A batsman cannot be dismissed from a no ball except in the case of a run out.

2.2.4 2012 Team and Player Rankings

In Table 2.3, all ODI teams that competed in 2012 are displayed and arranged in descending order of win/loss ratio. Of teams playing over five matches, England had the best ratio (12 W, 2 L) while New Zealand had the lowest (4 W, 10 L). New Zealand did, however, achieve the highest score of the year: 373 against Zimbabwe. Teams that played fewer than six matches, such as Canada, were competing at the 2012 World Cup. The World Cup is held every four years; Sri Lanka hosted last year’s tournament.

The 20 highest run scorers in ODI cricket for the 2012 calendar year are displayed in Table 2.4. Kumar Sangakkara from Sri Lanka scored the most runs—1184 at an average of 43.85 runs per innings (Batting average = total runs scored / number of innings dismissed in). Sri Lanka’s heavy ODI schedule (33 matches) explains seven out of the 20 players being Sri Lankan.

2.3 Discussion

In Table 2.2, AFL player performance was ranked according to an additive award system where the player with the most “votes” at the completion was deemed the best and fairest player in the competition. In Table 2.4, ODI batsmen were ranked in descending order of batting average, a single,
<table>
<thead>
<tr>
<th>Team</th>
<th>Mat</th>
<th>Won</th>
<th>Lost</th>
<th>W/L</th>
<th>Tied</th>
<th>NR</th>
<th>Highest</th>
<th>Lowest</th>
</tr>
</thead>
<tbody>
<tr>
<td>England</td>
<td>15</td>
<td>12</td>
<td>2</td>
<td>6.00</td>
<td>0</td>
<td>1</td>
<td>288</td>
<td>182</td>
</tr>
<tr>
<td>South Africa</td>
<td>13</td>
<td>8</td>
<td>4</td>
<td>2.00</td>
<td>0</td>
<td>1</td>
<td>312</td>
<td>211</td>
</tr>
<tr>
<td>Ireland</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>2.00</td>
<td>0</td>
<td>1</td>
<td>237</td>
<td>163</td>
</tr>
<tr>
<td>India</td>
<td>16</td>
<td>9</td>
<td>6</td>
<td>1.50</td>
<td>1</td>
<td>0</td>
<td>330</td>
<td>138</td>
</tr>
<tr>
<td>West Indies</td>
<td>12</td>
<td>6</td>
<td>5</td>
<td>1.20</td>
<td>1</td>
<td>0</td>
<td>315</td>
<td>140</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>1.00</td>
<td>0</td>
<td>0</td>
<td>293</td>
<td>241</td>
</tr>
<tr>
<td>Kenya</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1.00</td>
<td>0</td>
<td>0</td>
<td>201</td>
<td>120</td>
</tr>
<tr>
<td>Netherlands</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1.00</td>
<td>0</td>
<td>0</td>
<td>256</td>
<td>-</td>
</tr>
<tr>
<td>Australia</td>
<td>25</td>
<td>11</td>
<td>12</td>
<td>0.92</td>
<td>1</td>
<td>1</td>
<td>321</td>
<td>158</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>33</td>
<td>14</td>
<td>16</td>
<td>0.88</td>
<td>1</td>
<td>2</td>
<td>320</td>
<td>43</td>
</tr>
<tr>
<td>Pakistan</td>
<td>17</td>
<td>6</td>
<td>10</td>
<td>0.60</td>
<td>0</td>
<td>1</td>
<td>329</td>
<td>130</td>
</tr>
<tr>
<td>New Zealand</td>
<td>15</td>
<td>4</td>
<td>10</td>
<td>0.40</td>
<td>0</td>
<td>1</td>
<td>373</td>
<td>206</td>
</tr>
<tr>
<td>Afghanistan</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>0.25</td>
<td>0</td>
<td>0</td>
<td>259</td>
<td>104</td>
</tr>
<tr>
<td>Canada</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.00</td>
<td>0</td>
<td>0</td>
<td>176</td>
<td>176</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0.00</td>
<td>0</td>
<td>0</td>
<td>231</td>
<td>158</td>
</tr>
<tr>
<td>Scotland</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>177</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2.3: Highest win/loss ratio for ODI teams in 2012

<table>
<thead>
<tr>
<th>Player</th>
<th>Matches</th>
<th>Innings</th>
<th>Not Out</th>
<th>Runs</th>
<th>Highest</th>
<th>Ave</th>
<th>100s</th>
<th>50s</th>
</tr>
</thead>
<tbody>
<tr>
<td>KC Sangakkara (SL)</td>
<td>31</td>
<td>29</td>
<td>2</td>
<td>1184</td>
<td>133</td>
<td>43.85</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>TM Dilshan (SL)</td>
<td>31</td>
<td>30</td>
<td>3</td>
<td>1119</td>
<td>160*</td>
<td>41.44</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>V Kohli (India)</td>
<td>16</td>
<td>16</td>
<td>2</td>
<td>1026</td>
<td>183</td>
<td>73.28</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>LD Chandimal (SL)</td>
<td>30</td>
<td>29</td>
<td>5</td>
<td>845</td>
<td>92*</td>
<td>35.20</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>DA Warner (Aus)</td>
<td>25</td>
<td>24</td>
<td>0</td>
<td>840</td>
<td>163</td>
<td>35.00</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>DPMD Jayawardene (SL)</td>
<td>30</td>
<td>27</td>
<td>2</td>
<td>785</td>
<td>85</td>
<td>31.40</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>WU Tharanga (SL)</td>
<td>28</td>
<td>26</td>
<td>1</td>
<td>752</td>
<td>71</td>
<td>30.08</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>DJ Hussey (Aus)</td>
<td>25</td>
<td>24</td>
<td>3</td>
<td>728</td>
<td>74</td>
<td>34.66</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>HM Amla (SA)</td>
<td>10</td>
<td>9</td>
<td>1</td>
<td>678</td>
<td>150</td>
<td>84.75</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>G Gambhir (India)</td>
<td>15</td>
<td>15</td>
<td>0</td>
<td>677</td>
<td>102</td>
<td>45.13</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>AN Cook (Eng)</td>
<td>15</td>
<td>15</td>
<td>1</td>
<td>663</td>
<td>137</td>
<td>47.35</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>MJ Clarke (Aus)</td>
<td>15</td>
<td>14</td>
<td>0</td>
<td>656</td>
<td>117</td>
<td>46.85</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>AB de Villiers (SA)</td>
<td>13</td>
<td>12</td>
<td>6</td>
<td>645</td>
<td>125*</td>
<td>107.50</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>MEK Hussey (Aus)</td>
<td>19</td>
<td>19</td>
<td>1</td>
<td>580</td>
<td>67</td>
<td>32.22</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>HDRL Thirimanne (SL)</td>
<td>28</td>
<td>20</td>
<td>1</td>
<td>571</td>
<td>77</td>
<td>30.05</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>IR Bell (Eng)</td>
<td>11</td>
<td>11</td>
<td>1</td>
<td>549</td>
<td>126</td>
<td>54.90</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>MS Wade (Aus)</td>
<td>25</td>
<td>24</td>
<td>0</td>
<td>546</td>
<td>75</td>
<td>22.75</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>AD Mathews (SL)</td>
<td>27</td>
<td>23</td>
<td>9</td>
<td>534</td>
<td>80*</td>
<td>38.14</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Misbah-ul-Haq (Pak)</td>
<td>17</td>
<td>17</td>
<td>5</td>
<td>477</td>
<td>72*</td>
<td>39.75</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Mohammad Hafeez (Pak)</td>
<td>17</td>
<td>17</td>
<td>0</td>
<td>476</td>
<td>105</td>
<td>28.00</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2.4: Highest run scorers for 2012 in ODI matches

Objective match metric which rates their performance as a batsman. These two sports differ in that the AFL does not possess a single match metric which can accurately assess the skill set of a player. As this dissertation progresses, it will become clearer why this is the case, and how the research assumed the necessary direction to overcome such an anomaly. Chapter 4 further investigates player performance measurement in continuous (for example, the AFL) and discrete (for example, ODI cricket) team sports.
Chapter 3

Methods

This chapter details the statistical methodology employed in subsequent chapters to help answer the research questions summarised in Chapter 1. Linear regression, employed in Chapters 5 and 7, is discussed as well as discriminant analysis—the cornerstone of the AFL positional analysis in Chapter 6. It was necessary to document nonlinear smoothing and optimisation as the methods are introduced as concurrent performance forecasting tools in Chapter 8.

3.1 Linear Regression

This section covers the linear regression component of this dissertation. To begin, the mathematical formulation of the linear regression model is described. This is followed by the various assumptions that must be satisfied in order to make inferences about the coefficients of the regression model. Then the estimation of parameters using two methods namely Ordinary Least Squares (OLS) and Maximum Likelihood Estimation (MLE) is showcased. After that, the calculation of the residuals (errors) are specified which are required to satisfy the various assumptions of linear regression. Finally, the coefficient of determination is explained which is used to assess the adequacy of the fitted model. It should be noted that the methodology described in this section has been extracted from Greene (2002) and Johnson and Wichern (2007).
3.1.1 Introduction

Linear regression remains one of the most widely used statistical techniques across many disciplines including econometrics, environmental sciences and biostatistics to name a few. Linear regression models are extremely powerful since they have the power to empirically isolate very complicated relationships between variables. Its use, however, is only appropriate under certain assumptions (discussed later) and is often misused.

3.1.2 The Linear Regression Model

In statistics, the multiple linear regression model is used to model the relationship between one or more independent variables and a dependent variable. The generic form of the model is given by:

\[
y = f(1, x_1, x_2, ..., x_K) + \epsilon.
\]

where \( y \) is the explained (or dependent) variable, \( x = (1, x_1, \ldots, x_K)' \) is a column vector of explanatory (or independent) variables, \( \beta = (\beta_0, \beta_1, \ldots, \beta_K)' \) is a column vector of coefficients and \( \epsilon \) allows for random variation in \( y \) for a fixed value of \( x \). The following \( n \times p \) matrix denotes the observed values for \( x \).

\[
X = \begin{bmatrix}
x_{12} & x_{22} & \cdots & x_{1p} \\
x_{21} & x_{22} & \cdots & x_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
x_{n1} & x_{n2} & \cdots & x_{np}
\end{bmatrix}
\]

(3.2)

The objective of the linear regression model is to estimate the unknown parameters \( \beta_0, \beta_1, \ldots, \beta_K \) which provide a “best fit” to a series of data points.

3.1.3 Assumptions

There are several assumptions of the linear regression model which must be satisfied in order to make inferences about the coefficients derived from the model. These assumptions are as follows:
A1. **Linearity:** \( y_i = x_i \beta_1 + x_i \beta_2 + \cdots + x_i \beta_K + \epsilon_i \). The model specifies a linear relationship between \( y \) and \( x_1, \ldots, x_K \).

A2. **Full Rank:** None of the independent variables are a perfect linear combination of the other independent variables.

A3. **Exogeneity of the independent variables:** \( E[\epsilon_i|x_1, x_2, \ldots, x_K] = 0 \). This states that the expected value of the disturbance at observation \( i \) in the sample is not a function of the independent variables observed at any observation, including this one. This means that independent variables will not carry useful information for prediction of \( \epsilon_i \).

A4. **Homoscedasticity and non-autocorrelation:** Each disturbance \( \epsilon_i \) has the same finite variance \( \sigma^2 \) and is uncorrelated with every other disturbance \( \epsilon_j \).

A5. **Exogenously generated data:** The data in \((x_{j1}, x_{j2}, \ldots, x_{jK})\) may be any mixture of constants and random variables. The process generating the data operates outside the assumptions of the model, that is, independently of the process that generates \( \epsilon_i \). Note that this extends A3. Analysis is done conditionally on the observed \( X \).

A6. **Normal Distribution:** The disturbances are normally distributed.

### 3.1.4 Least Squares Regression

Linear regression approximates the unknown parameters of the stochastic relation \( y_i = \mathbf{x}_i' \beta + \epsilon_i \). Firstly, it is important to distinguish between population quantities \( \beta \) and \( \epsilon_i \) and sample estimates, denoted \( \mathbf{b} \) and \( e_i \). Here the population regression is \( E[y_i|x_i] = \mathbf{x} \) while the estimate of \( E[y_i|x_i] \) is given by:

\[
\hat{y}_i = \mathbf{x}_i' \beta
\]  (3.3)

Here the disturbance associated with the \( i \)th data point is denoted:

\[
\epsilon_i = y_i - \mathbf{x}_i' \beta
\]  (3.4)

for a given value of \( b \), the estimate of \( \epsilon_i \) is given by the residual:

\[
e_i = y_i - \mathbf{x}_i' \mathbf{b}.
\]  (3.5)
Therefore, based on these definitions,

\[ y_i = x'_i \beta + e_i = x'_ib + e_i \]  \hspace{1cm} (3.6)

There are numerous methods for estimating the unknown vector, \( \beta \), of population quantities; most commonly used methods include *Ordinary Least Squares* (OLS) and *Maximum Likelihood Estimation* (MLE).

### 3.1.5 Ordinary Least Squares

Ordinary Least Squares (OLS) is perhaps the most frequently used method for estimating the unknown vector. The least squares coefficient vector minimizes the sum of the squared errors, which is given by:

\[
\sum_{i=1}^{n} e_{ii}^2 = \sum_{i=1}^{n} (y_i - x'_i \beta)^2
= (y - X\beta)'(y - X\beta)
= y'y - 2y'X\beta + \beta'X'X\beta \]  \hspace{1cm} (3.7)

where \( \beta \) denotes the choice of the coefficient vector.

To calculate the minimum, the partial derivative of (3.7) with respect to \( \beta \) is set to zero and solved.

\[
\frac{\partial}{\partial \beta} = -2X'y + 2X'X\beta = 0 \]  \hspace{1cm} (3.8)

Let \( b \) be the solution, therefore \( b \) satisfies the least squares normal equations given by:

\[
X'Xb = X'y \]  \hspace{1cm} (3.9)

Since the inverse of \( X'X \) exists based on the full rank assumption (A2.), then the solution is given by:

\[
b = (X'X)^+X'y \]  \hspace{1cm} (3.10)

### 3.1.6 Maximum Likelihood Estimation

Another method of estimating the unknown vector is that of *Maximum Likelihood Estimation* (MLE). The probability density function (pdf) of a random variable, \( y \), conditional on a given set of
parameters \( \theta \) is denoted \( f(y|\theta) \). The joint density (or likelihood function) of \( n \) independent identically distributed (iid) observations from given pdf is denoted:

\[
f(y_1, \ldots, y_n | \theta) = \prod_{i=1}^{n} f(y_i | \theta) = L(\theta | y)
\]  
(3.11)

It is more convenient to write the likelihood function after a log transformation. Also, the likelihood is written more conveniently as \( L \). Therefore:

\[
L = \ln L(\theta | y) = \sum_{i=1}^{n} \ln f(y_i | \theta)
\]  
(3.12)

In linear regression, the likelihood function for a sample of \( n \) independent, identically and normally distributed disturbances is given by:

\[
L = \left(2\pi \sigma^2 \right)^{-n/2} e^{-\frac{1}{2\sigma^2}(y-X\beta)'(y-X\beta)}
\]  
(3.13)

The transformation from \( \epsilon_i \) to \( y_i \) is \( \epsilon_i = y_i - x_i \), such that the Jacobian for each observation \( [\partial \epsilon_i / \partial y_i] \) equals one. Therefore, the likelihood function can now be written:

\[
L = \left(2\pi \sigma^2 \right)^{-n/2} e^{-\frac{1}{2\sigma^2}(y-X\beta)'(y-X\beta)}
\]  
(3.14)

and the log-likelihood is given by:

\[
\ln L = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{(y-X\beta)'(y-X\beta)}{2\sigma^2}
\]  
(3.15)

To maximise the log-likelihood the partial derivative of (3.15) is taken with respect to \( \beta \) and \( \sigma^2 \) which is given by:

\[
\begin{bmatrix}
\frac{\partial \ln L}{\partial \beta}
\frac{\partial \ln L}{\partial \sigma^2}
\end{bmatrix}
= \begin{bmatrix}
\frac{X'(y-X\beta)}{\sigma^2}
\frac{-n}{2\sigma^2} + \frac{(y-X\beta)'(y-X\beta)}{2\sigma^2}
\end{bmatrix}
= \begin{bmatrix}
0
0
\end{bmatrix}
\]  
(3.16)

The values which satisfy these equations are:

\[
\hat{\beta}_{ML} = (X'X)^{-1}X'y \quad \text{and} \quad \hat{\sigma}^2_{ML} = \frac{e'e}{n}
\]  
(3.17)

### 3.1.7 Analysis of Residuals

One such method to test the assumptions defined in Section 3.2.3, and thus the adequacy of the linear regression model, is plotting the residuals. The residual for the \( i \)th case is given by:

\[
\hat{z}_i = y_i - \hat{y}_i
\]  
(3.18)
where \( y_i \) is the observed outcome and \( \hat{y}_i \) is the predicted outcome. If a relationship exists between the residuals \( \hat{z}_i \) and any variable then there is an effect from that variable which has not yet been accounted for. A common plot includes the residuals \( \hat{z}_i \) against the fitted values \( \hat{y}_i \) which reveals outliers and whether the assumption of constant variance and linearity are appropriate. Additional measures used to detect outliers include Mahalanobis Distance and Cook’s Distance. Another common plot is the residuals \( \hat{z}_i \) against a time dependent predictor variable or the order number of the experiment, a smooth plot will show that the assumption of independence is not valid. Residual independence can also be checked using the Durbin-Watson statistic.

### 3.1.8 Goodness of Fit

The coefficient of determination, commonly denoted by \( R^2 \) is used to assess the goodness-of-fit of the linear regression model. \( R^2 \) is described as the amount of variation that can be explained by the regressors where \( R^2 \in [0, 1] \). If \( R^2 = 1 \) the values of \( x \) and \( y \) all lie on the same hyperplane such that all the residuals are zero. On the contrary, if \( R^2 = 0 \) the fitted values correspond to a horizontal line such that all the elements of \( b \) except the constant term are zero. The “variability” of the data is measured through different sum of squares where:

\[
SS_T = \sum_{i=1}^{n} (y_i - \bar{y})^2, \text{ total sum of squares}
\]

\[
SS_R = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2, \text{ regression sum of squares}
\]

\[
SS_E = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2, \text{ residual sum of squares}
\]

where:

\[
R^2 = 1 - \frac{SS_E}{SS_T} = \frac{SS_R}{SS_T}
\]  

(3.19)

Additional measures of goodness of fit which account for the complexity of the model include the adjusted \( R^2 \) and Akaike Information Criterion (AIC).
3.2 Discriminant Analysis

This section covers the discriminant analysis component of this dissertation (Chapter 6). To begin, the mathematical formulation of the model is described. This is followed by the various assumptions that must be satisfied in order to make inferences about the coefficients of the model. It should be noted that the methodology described in this section has been extracted from Johnson and Wichern (2007).

3.2.1 Introduction

Discriminant analysis is a multivariate statistical technique used to model the value of a dependent categorical variable based on its relationship to one or more predictors. It separates distinct sets of objects and also allocates new objects to previously defined groups. This process is popular in such areas as banking—to classify customers into predefined credit risk categories—and marketing—for allocating survey respondents that have already been segmented using a clustering technique.

3.2.2 The Model

Given a set of independent variables, discriminant analysis attempts to find linear combinations of those variables that best separate the groups of cases. These combinations are called discriminant functions and take the form:

\[ d_{ik} = b_{0k} + b_{1k}x_{i1} + \ldots + b_{pk}x_{ip} \]  

(3.21)

where \( d_{ik} \) is the value for the \( k^{th} \) discriminant function for the \( i^{th} \) case, \( p \) is the number of predictors, \( b_{jk} \) is the value of the \( j^{th} \) coefficient of the \( k^{th} \) function and \( x_{ij} \) is the value of the \( i^{th} \) case of the \( j^{th} \) predictor. The number of functions is \( \min(m - 1, p) \) where \( m \) is the number of groups (four for this dissertation—one for each player position). The procedure automatically chooses a first function that will separate the groups as much as possible. It then chooses a second function that is both uncorrelated with the first function and provides as much further separation as possible. The procedure continues adding functions in this way until reaching the maximum number of functions as determined by the number of predictors and categories in the dependent variable.
In the case of two populations \((\pi_1, \pi_2)\), a new observation, \(x_0\), is allocated to the population with the largest posterior probability, \(P(\pi_1 \mid x_0)\). By Bayes’s rule, the posterior probabilities are:

\[
P(\pi_1 \mid x_0) = \frac{P(\pi_1 \text{ occurs and we observe } x_0)}{P(\text{we observe } x_0)}
\]

\[
= \frac{P(\text{we observe } x_0 \mid \pi_1)P(\pi_1)}{P(\text{we observe } x_0 \mid \pi_1)P(\pi_1) + P(\text{we observe } x_0 \mid \pi_2)P(\pi_2)}
\]

\[
= \frac{p_1 f_1(x_0)}{p_1 f_1(x_0) + p_2 f_2(x_0)}
\]

\[
P(\pi_2 \mid x_0) = 1 - P(\pi_1 \mid x_0) = \frac{p_2 f_2(x_0)}{p_1 f_1(x_0) + p_2 f_2(x_0)}
\]

(3.22)

3.2.3 Classification with Two Multivariate Normal Populations

If the case of two populations, say we had \(n_1\) observations of the multivariate random variable:

\[
X' = [X_1, X_2, \ldots, X_p]
\]

(3.23)

from \(\pi_1\) and \(n_2\) measurements of this quantity from \(\pi_2\), with \(n_1 + n_2 - 2 \geq p\). The respective sample mean vectors and covariance matrices are:

\[
\bar{x}_1 = \frac{1}{n_1} \sum_{j=1}^{n_1} x_{1j}, \quad S_1 = \frac{1}{n_1 - 1} \sum_{j=1}^{n_1} (x_{1j} - \bar{x}_1)(x_{1j} - \bar{x}_1)'
\]

(3.24)

\[
\bar{x}_2 = \frac{1}{n_2} \sum_{j=1}^{n_2} x_{2j}, \quad S_2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)(x_{2j} - \bar{x}_2)'
\]

(3.25)

Since it is assumed that the parent populations have the same covariance matrix, \(\Sigma\), the sample covariance matrices, \(S_1\) and \(S_2\), are pooled to derive a single unbiased estimate of \(\Sigma\). In particular, the weighted average

\[
S_{\text{pooled}} = \left[ \frac{n_1 - 1}{(n_1 - 1) + (n_2 - 1)} \right] S_1 + \left[ \frac{n_2 - 1}{(n_1 - 1) + (n_2 - 1)} \right] S_2
\]

(3.26)

is an unbiased estimator of \(\Sigma\) if the data matrices \(X_1\) and \(X_2\) contain \textit{random} samples from the populations \(\pi_1\) and \(\pi_2\), respectively. The classification rule is:
Allocate $x_0$ to $\pi_1$ if
\[
(x_1 - x_2)\mathbf{S}_{\text{pooled}}^{-1}x_0 - \frac{1}{2} (x_1 - x_2)\mathbf{S}_{\text{pooled}}^{-1}(x_1 + x_2) \geq \ln \left( \frac{c(1 | 2)}{c(2 | 1)} \right) \frac{p_2}{p_1} \]
(3.27)
where $p_1$ and $p_2$ are prior probabilities of $\pi_1$ and $\pi_2$, respectively, and $c(1 | 2)$ is the cost of misclassifying an observation from $\pi_2$ as $\pi_1$.

Allocate $x_0$ to $\pi_2$ otherwise.

Fishers’s linear discriminant function maximally separates the two populations, using:
\[
D^2 = (x_1 - x_2)' \mathbf{S}_{\text{pooled}}^{-1} (x_1 - x_2) \]
(3.28)
where $D^2$ is the sample squared distance between the two means.

Fisher’s solution allocation rules are as follows:

Allocate $x_0$ to $\pi_1$ if
\[
y_0 = (x_1 - x_2)' \mathbf{S}_{\text{pooled}}^{-1}x_0 \\
\geq \hat{m} = \frac{1}{2} (x_1 - x_2)' \mathbf{S}_{\text{pooled}}^{-1}(x_1 + x_2) \]
(3.29)

or
\[
y_0 - \hat{m} \geq 0
\]

Allocate $x_0$ to $\pi_2$ if
\[
y_0 - \hat{m} < 0
\]

3.2.4 Fisher’s Classification with several Populations

The motivation behind the Fisher discriminant analysis is the need to obtain a reasonable representation of the populations that involves only a few linear combinations of the observations, such as $a'_1x, a'_2x$ and $a'_3x$. The primary purpose is to separate populations and to classify. It is not necessary to assume that the $g$ populations are multivariate normal, but we do assume that the $p \times p$ population covariance matrices are equal.

Let $\hat{\lambda}_1, \hat{\lambda}_2, ..., \hat{\lambda}_s > 0$ denote the $s \leq \min(g - 1, p)$ nonzero eigenvalues of $W^{-1}B$ and $\hat{e}_1, ..., \hat{e}_s$ be the corresponding eigenvectors (scaled so that $\hat{e}_i' \hat{S}_{\text{pooled}} \hat{e}_i = 1$). Then the vector of coefficients $\hat{a}$ that maximises the ratio
\[
\frac{\hat{a}' \hat{B} \hat{a}}{\hat{a}' \hat{W} \hat{a}} = \frac{\hat{a}' \left( \sum_{i=1}^{g} (x_i - \bar{x})(x_i - \bar{x})' \right) \hat{a}}{\hat{a}' \left( \sum_{i=1}^{g} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i) \right) \hat{a}}
\]
(3.30)
is given by $\hat{a}_1 = \hat{e}_1$. The linear combination of $\hat{a}_1^T x$ is called the sample first discriminant. The choice $\hat{a}_2 = \hat{e}_2$ produces the sample second discriminant, $\hat{a}_2^T x$, continuing until we obtain $\hat{a}_k^T x = \hat{e}_k^T x$, the sample $k^{th}$ discriminant, $k \leq s$.

A reasonable classification rule is one that assigns $y$ to population $\pi_k$ if the square of the distance from $y$ to $\mu_{kY}$ is smaller than the square distance from $y$ to $\mu_{iY}$ for $i \neq k$. If only $r$ of the discriminants are used for allocation, the rule is:

Allocate $x$ to $\pi_k$ if

$$\sum_{j=1}^{r} (y_j - \mu_{kY})^2 = \sum_{j=1}^{r} [a'_j(x - \mu_k)]^2 \leq \sum_{j=1}^{r} [a'_j(x - \mu_i)]^2, \text{ for all } i \neq k$$

(3.31)

### 3.3 Nonlinear Data Smoothers

This section investigates nonlinear data smoothing, or Tukey smoothing, as introduced by Tukey (1971). This methodology applies median smoothing to contiguous values in a dataset and is applicable when dealing with non-normal datasets, as outlined in Chapter 8. This outline begins by describing how median combinations can smooth numerous adjacent values, then explains the concepts of repeating, splitting, twicing and Hanning, and how these features can be merged to create the best smooth. Chapter 8 provides a thorough theoretical and practical examination of nonlinear smoothing.

Any smoother separates the original sequence into a smooth sequence and a residual sequence, or:

$$x_t = f(x_t) + r_t$$

(3.32)

where $f$ is reasonably smooth and $r_t$ represents a mild departure from the true $f(x_t)$ except for abrupt changes or spikes in data. Velleman (1980) defines an odd-span running median of $v$ as:

$$y_t = \text{med}\{x_{t-u}, ..., x_t, ..., x_{t+u}\} \quad \text{where } v = 2u + 1$$

(3.33)
Given a data set, \( x_1 = 5, x_2 = 6, x_3 = 4, x_4 = 7, x_5 = 4 \), from Equation (3.33), a running median of span-3 produces the smooth, \( y_1 = 5, y_2 = 5, y_3 = 6, y_4 = 4, y_5 = 4 \), where \( x_1 \) and \( x_5 \) form the end points of the smoothed series. An odd-span running median tracks upward movement in progressive data points but removes spikes of \( u \) consecutive points or fewer. An even-span running median is defined as:

\[
y_{t+1/2} = \text{med}\{x_{t+1}, \ldots, x_{t+u}\} \quad \text{where } v = 2u \tag{3.34}
\]

A span-4 median, for example, sees the two middle \( x \) values averaged then housed at \( t = 2.5 \). Applying Equation (3.34) to the example data set gives \( y_1 = 5, y_2 = 5.5, y_3 = 5, y_4 = 7, y_5 = 4 \), where \( x_1 \) and \( x_5 \) again form the end points of the smoothed series. Such basic smooths prove even-span median smoothers generally show less resistance to spikes than odd spans of similar length.

A short discussion is necessary of some of the elegant features of nonlinear smoothing. An important attribute is the capacity to combine medians of the same or varying span in a single smooth. Extensive running median combinations to smooth one-dimensional data are illustrated in Tukey (1971). Resmoothing the span-3 smooth with a span-3 smooth (denoted \( 33 \)), again using Equation (3.33), gives \( y^*_1 = 5, y^*_2 = 5, y^*_3 = 5, y^*_4 = 4, y^*_5 = 4 \); Equation (3.34) could then be applied to the \( 33 \) smooth to offer a \( 334 \) smooth, and so on. As a rule, the fewer running medians involved, the less resistant the smooth will be to abrupt changes in the data. Repeated running medians of the same span will result in no further change within a short space. For example, a \( 333 \) smooth on a given data set provides the same smooth as \( 3333, 33333 \), et cetera. Tukey (1971) simplified this with the symbol, \( R \) (repeat smooth on previous median in sequence until no further change occurs), so \( 3R = 3333.. \).

Re-roughing or twicing (\( T \)) a data set improves the fit of a model by fitting the smooth \( f \) in Equation (3.32) to the residuals, \( r_n \), to obtain \( f(r) \). The smoothed residual is then added to the original smooth to obtain Equation (3.35). The re-roughing process can be extended to thricing and beyond.

\[
y_i^{(\text{twice})} = f(r) + y_i \tag{3.35}
\]

Hanning (\( H \)) entails a weighted moving average (Equation (3.36)), normally applied in a smooth sequence after nonlinear smoothing (in the same sequence) has removed any spikes.

\[
z_t = \frac{1}{4} y_{t-1} + \frac{1}{2} y_t + \frac{1}{4} y_{t+1} \tag{3.36}
\]

58
Tukey (1971, p.527) observed Hanning “pulls tops down and bottoms up”, giving a more curvaceous appearance to the smooth. He also explores the concept of splitting ($S$). This process dissects the sequence into shorter subsequences at all places where two successive values are identical, runs a $3R$ smooth on each subsequence, reassembles them, and polishes the result with a $3$ smooth. An $S$ is likely to be included in a smooth sequence where plateaus or valleys, two points wide, are corrupting the data. Each of the smoothing sequence components outlined here will be active in Chapter 8 where smoothing sequences are optimised on a player-by-player basis to minimise average forecast error.

3.4 Optimisation

In Chapter 8, an optimisation technique is employed to arrive at the best possible fit for a forecasting model. This section summarises the optimisation process, describing some applications and then the mathematical formulation of the optimisation technique employed. The appropriate methodology has been summarised from Nocedal and Wright (1999).

3.4.1 Introduction

Optimisation techniques are in common use across a wide range of fields, for example, finance—how to minimise risk or maximise gain from investments, and transport—how to optimally allocate loads for shipment. To utilise optimisation techniques, the objective must first be formulated, that is, a quantitative measure of performance of the system, for example, maximising profits. This measure is commonly referred to as the objective function, which is typically either minimised or maximised. The objective function is dependent on certain characteristics of the system known as variables. The goal of any optimisation problem is to find values of the unknown variables which optimise the objective function. Typically these unknown variables have constraints placed on them, for example, the amount of capital available for investment.

3.4.2 Mathematical Formulation

The mathematical formulation of an optimisation problem may use the following notation:
Let

\( \cdot x \) be a vector of unknown parameters

\( \cdot f(x) \) be the objective function, which is to be maximised or minimised

\( \cdot c \) be a vector of constraints which the unknown parameters must satisfy

Now the optimisation problem can be written as:

\[
\min_{x \in \mathbb{R}} f(x) \text{ or } \max_{x \in \mathbb{R}} f(x)
\]

subject to:

\[c_i(x) \leq 0, \ i = 1, 2, \ldots, m.\]

For example, a sporting club might wish to maximise the profit after membership sales from a particular advertisement. Let \( x_1 \) denote the number of memberships needed to be sold, \( x_2 \) denote the advertising costs and \( x_3 \) denote the labour costs. Let \( z \) denote the profit measured in thousands of dollars, so:

\[
z = x_1 - x_2 - x_3
\]

If the cost of advertising and labour per ticket sold is $10 and the cost of labour per ticket is $10 and the advertising budget was $50,000, to maximise profits, the objective function, with constraints becomes:

\[
\max z = x_1 - x_2 - x_3
\]

subject to:

\[x_2 + x_3 \leq 50,000\]

\[x_2, x_3 \geq 0.\]

### 3.4.3 Important Considerations

There are many important considerations that should be taken into account when defining an optimisation problem. Let us consider four general issues which may arise.

(i) Is the optimisation problem discrete or continuous or a combination of the two? Discrete optimisation usually refers to problems in which the optimal solution is derived from a finite set of feasible solutions, that is, a vector of integers. However, continuous optimisation problems refer to
problems in which the optimal solution is derived from an infinite set of feasible solutions, that is, a vector of real numbers. Typically speaking, continuous optimisation models are easier to solve since the behaviour of the function at all points close to $x$ are similar due to the smoothness of the function.

However, the same cannot be said about discrete optimisation models due to their discrete nature. Optimisation models that have both discrete and continuous variables are referred to as *mixed integer programming* problems.

(ii) Is the optimisation problem *stochastic* or *deterministic*? Stochastic optimisation problems arise when the model is not fully specified, that is, there is some unknown quantity at time of formulation. For example, in economics and finance an important characteristic of companies is future cash flow which is always unknown but can be estimated. Deterministic models on the other hand, are models that are fully specified, that is, there is no unknown quantity at time of formulation.

(iii) Is the optimisation problem *constrained* or *unconstrained*? A constrained optimisation model has explicit constraints on the unknown parameters which must be met in order for the objective function to be feasible. A constraint could simply be a bound place on a variable $a \leq x_1 \leq b$; declaring a variable must take integer values $x_2 \in \mathbb{Z}$; a more general linear constraint $\Sigma x_i \leq c$; or a nonlinear inequality which is a complex function comprising several variables. For unconstrained optimisation models every possible solution is feasible.

(iv) Is the local solution also the global solution? Many computer algorithms seek only a local solution, that is, the objective function is smaller than all other values within its vicinity. Furthermore, many computer algorithms have no built-in functions to check for local/global solutions. However, many non-linear functions have several local minimums in which case one would be interested in which one of these local minimums is also the global minimum, that is, the best solution of all such minima.

### 3.4.4 Optimisation Algorithms

An optimisation algorithm is an iterative numerical procedure for finding the values of the vector $x$ that maximises (or minimises) the objective function $f(x)$ subject to the constraints $c$. The
algorithm begins with an initial estimate of the unknown parameters $x_0$ then a sequence of improved estimates $(x_i), i=1,\ldots,\infty$ are generated until no more improvements can be made or a solution is approximated with sufficient accuracy. The strategy of going from one iteration to the next is what separates the algorithms from one another. Some of the most common optimisation algorithms include Monte Carlo Sampling and Latin Hypercube Sampling.

3.4.4.1 Monte Carlo Sampling

Monte Carlo methods are a class of computational algorithms that utilise repeated random or pseudo-random numbers. These methods are typically used when computing an exact solution is unfeasible or impossible. Although there is not one definitive Monte Carlo method, the approach of many Monte Carlo methods are similar. Typically, a domain of possible inputs is defined of which inputs are generated randomly, then a deterministic computation is performed using these inputs and finally the results of the individual computations are aggregated into a final result.

3.4.4.2 Latin Hypercube Sampling

To understand the statistical method of Latin Hypercube Sampling it is crucial to comprehend the Latin Hypercube. Firstly, a Latin square is $n \times n$ square filled with $n$ different colours such that each colour is represented only once in each row and each column. Similarly, a Latin Hypercube is the generalization of this concept to an arbitrary number of dimensions. Latin Hypercube sampling uses a statistical technique known as “stratified sampling without replacement”, whereby sampling is undertaken from a function of $N$ variables with each variable being split into $M$ equally probable intervals. The $M$ sample points are then placed such that the Latin Hypercube is satisfied.
Chapter 4

Player performance

“Talent wins games, but teamwork and intelligence wins championships.” – Michael Jordan

Team and player ratings are typically a function of some quantitative estimate of performance in the sporting arena. This chapter investigates player performance in continuous and discrete team sports, establishing important metrics from which the AFL individual player ratings (Chapter 5), the AFL cooperative ratings (Chapter 7) and in-play cricket batsman simulations (Chapter 9) are derived.

4.1 Introduction

An ever-present conundrum in the sporting arena is whether it is best to have a team of champions or a champion team. Ideally, a coach would have both; however, salary cap constraints prevent this from being a common occurrence. While teams go to great lengths to secure the services of champion players, it is generally accepted that a team needs to be more than the sum of the individual players (Gould and Gatrell, 1979/80). A professional team will nearly always contain at least one champion, but for that team to develop, each champion must sacrifice his or her ego for the good of the team. The newest team in the AFL, the Greater Western Sydney Giants (GWS), offers an interesting case study: prior to the team’s debut match in 2012, recruiters were set the task of putting together a team that could at least endure the physical demands of the game and could develop as quickly as possible into a competitive side. To achieve this, experienced senior players and inexperienced young players were recruited. The former were in the twilight of their careers but had all achieved success at previous clubs; they were drafted in the hope that their on-field skills and
leadership would educate the younger players. The make-up of the team differed considerably to
other, more established teams such as Collingwood, the 2010 champions. Table 4.1 shows notable age
dispersion; compared with Collingwood, GWS has around a third of the number of players over 25
years of age. The youth element in the GWS side is evident, with Collingwood having around a third
of the number of players under 20 years old. The current Collingwood team structure is what the
GWS is striving for: a balance of youth and experience. GWS would be striving for as many as
possible of its younger players to develop into champion players, moving together as a unit to produce
a champion team. Although the GWS only won two games in 2012, the players have displayed a
cohesive style of play which would probably not have been possible without the selfless influence of
the senior players.

<table>
<thead>
<tr>
<th></th>
<th>Ave. Age</th>
<th>Ave. Games</th>
<th>No. over 25 years</th>
<th>No. under 20 years</th>
<th>Ave. games over 25 years</th>
<th>Ave. games under 25 years</th>
<th>Ave. games under 20 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>GW Sydney</td>
<td>21.5</td>
<td>30.5</td>
<td>6</td>
<td>27</td>
<td>177.7</td>
<td>10.0</td>
<td>5.9</td>
</tr>
<tr>
<td>Collingwood</td>
<td>23.5</td>
<td>67.8</td>
<td>17</td>
<td>10</td>
<td>145.4</td>
<td>25.2</td>
<td>2.0</td>
</tr>
</tbody>
</table>

Table 4.1: Greater Western Sydney and Collingwood team list statistics for the 2012 season

Describing an AFL player as a champion is resorting mostly to subjectivity and opinion, but it
is confirmed if the player wins the highest individual award, the Brownlow Medal. The umpires
involved in each regular season match gather at the match completion and award votes of 3, 2 and 1 to
the best three players on the ground, with 3 being the best. The Monday before the grand final, the
votes are retrospectively (round-by-round) made public and the medal awarded to the player with the
highest tally after the final round’s votes are revealed. A noteworthy trend is that the award winner is
from a team that performed strongly in the season, because players from the victorious team tend to
attract the umpires’ attention more frequently than players from the losing side. From Table 4.2, nine
of the fourteen winners\(^7\) since the year 2000 have played for teams who finished the season ranked in
the top 3 (out of 16 teams). The lowest any winner’s team finished after the final regular round was 8\(^{th}\)
(Chris Judd for the Carlton Blues in 2010). Interestingly, the next lowest was Judd in 2007 with his
former club, the West Coast Eagles, finishing 7\(^{th}\). Moreover, nine of the 14 winners went on to play in

\(^{7}\) In 2003, three players shared the award, all tallying an equal number of votes.
the grand final match (*) that week. This is strong evidence that the champion player and champion team are not mutually exclusive in the AFL.

<table>
<thead>
<tr>
<th>Year</th>
<th>Player</th>
<th>Team</th>
<th>Team Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>Dane Swan</td>
<td>Collingwood *</td>
<td>1</td>
</tr>
<tr>
<td>2010</td>
<td>Chris Judd</td>
<td>Carlton</td>
<td>8</td>
</tr>
<tr>
<td>2009</td>
<td>Gary Ablett</td>
<td>Geelong *</td>
<td>2</td>
</tr>
<tr>
<td>2008</td>
<td>Adam Cooney</td>
<td>Western Bulldogs</td>
<td>3</td>
</tr>
<tr>
<td>2007</td>
<td>Jimmy Bartel</td>
<td>Geelong *</td>
<td>1</td>
</tr>
<tr>
<td>2006</td>
<td>Adam Goodes</td>
<td>Sydney *</td>
<td>4</td>
</tr>
<tr>
<td>2005</td>
<td>Ben Cousins</td>
<td>West Coast *</td>
<td>2</td>
</tr>
<tr>
<td>2004</td>
<td>Chris Judd</td>
<td>West Coast</td>
<td>7</td>
</tr>
<tr>
<td>2003</td>
<td>Mark Ricciuto</td>
<td>Adelaide</td>
<td>6</td>
</tr>
<tr>
<td>2003</td>
<td>Nathan Buckley</td>
<td>Collingwood *</td>
<td>2</td>
</tr>
<tr>
<td>2003</td>
<td>Adam Goodes</td>
<td>Sydney</td>
<td>4</td>
</tr>
<tr>
<td>2002</td>
<td>Simon Black</td>
<td>Brisbane Lions *</td>
<td>2</td>
</tr>
<tr>
<td>2001</td>
<td>Jason Akermanis</td>
<td>Brisbane Lions *</td>
<td>2</td>
</tr>
<tr>
<td>2000</td>
<td>Shane Woewodin</td>
<td>Melbourne *</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 4.2: Brownlow Medal winners and respective team ranks

In certain cases, it is possible to assess the contribution a champion makes to a side without having to resort to complex quantitative analysis. Chris Judd, commonly regarded as the greatest footballer of the decade, was drafted by the Carlton Football Club from the West Coast Eagles Football Club at the end of 2007, to debut in 2008. In the three years prior to Judd arriving at the club, Carlton finished 16th (2005), 16th (2006) and 15th (2007). In the three seasons after his debut, Carlton finished 7th (2009), 8th (2010) and 5th (2011). One could attribute the team’s improvement in the team standings to factors independent of Judd’s signing—for example, the development of the younger talent and/or changes in coaching staff—but it is widely accepted that Judd’s presence in the side has transformed the team into a premiership contender. Judd is not only an extremely skillful player, but also has leadership qualities and an on-field presence that lifts those around him to higher performance levels. While the Judd effect makes an interesting case for the champion player, it is difficult to isolate other instances in the AFL.
In 2007, the Carlton Football Club was financially in a position to offer Judd an attractive salary in the courting procedure. Some clubs, however, are financially unable to make attractive offers to lure star players, mostly due to salary cap constraints or recent poor performance which reduces revenues earned from gate and membership sales. In Michael Lewis’s 2004 book *Moneyball*, he recounted the pioneering recruitment techniques of Billy Beane, General Manager for the Oakland A’s Major League baseball team. Beane was forced to cope with one of the lowest payrolls in baseball—the New York Yankees payroll in 2002 was $126 million, while the Oakland A’s payroll was around $40 million—by recruiting players who were undervalued by the market. Beane’s recruiting was driven by statistical analyses that looked beyond the conventional wisdom of attaching higher prices to players with, for example, higher batting averages, runs batted in and strike-outs. He hired a statistically minded Harvard graduate, Paul De Podesta, who had run analyses to test which of baseball’s performance indicators drove winning percentage; on-base percentage and slugging percentage became an Oakland A’s batter’s key attributes. He sought batters from other clubs and the minor leagues who excelled at such indicators, highly valuing anyone who could regularly draw a walk. The Oakland A’s were deprived of recognised champion players due to salary constraints, but by using statistical analysis to take advantage of undervalued players, Oakland came close to becoming a champion team. In 2002, the A’s finished first in the American League West (103 wins and 59 losses), but went on to lose to the Minnesota Twins (2-3) in the post-season. This was an extraordinary result for players who were being paid far less than their market value during that period. Barry Zito, pitcher for the A’s, was paid the very low sums of $200,000 in 2000 and $500,000 in 2002, the year he would win the Cy Young award for best pitcher in the American league. Interestingly, Zito was a vital component of the San Francisco’s pitching line-up in the 2012 World Series, outpitching Justin Verlander, the 2011 American league Cy Young award winner, in Game 1 (8-3).

A champion player may be described as one whose performance is consistently of a level that is required to win a championship. But how does one track the rise of a champion or any player, for

---

As part of the Judd deal, Carlton also traded a young forward to West Coast named Josh Kennedy, who has blossomed into a fine player. Consequently, Carlton has, at times since 2008, badly been in need of a forward of Kennedy’s ilk.
that matter? What are the necessary measurement tools required for anyone to observe a player’s performance in a match or a season? Answering these questions begins with *notational analysis*.

### 4.2 Notational Analysis

The type of sport being played dictates the level of responsibility a competitor assumes in deciding the outcome of the contest. The overall performance of a competitor in an individual sporting contest—for example, a golf tournament—is directly related to his or her final placing in the contest; if the golfer plays well over the four rounds, he or she is more likely to finish in the top placings and vice versa. In team sports such as football, an individual’s performance is not necessarily correlated with the result: a competitor can perform well below expectation and the team still realise victory in a given match; or a competitor can perform extremely well in a side that has realised a large defeat. Effective description of player performance in a team sport requires clarification of the type of game in which the player is involved. Hopper and Bell (2000) provide the following classifications of games: target games (golf, lawn bowls); net/wall games (tennis, volleyball); batting and fielding games, also known as bat and ball games (cricket, baseball); territory games, also known as invasion games (Australian Rules football, football and basketball). This research defined certain invasion games, such as the ones just mentioned, as “continuous” or having the ability to fluidly progress without any dead or disputed phases for the entire match duration. Of course, this is an unrealistic concept; events such as a goal or a penalty occur frequently in these sports, allowing both teams to become involved in a momentary “ceasefire”. While continuous play is possible, we made this distinction to further classify invasion games. It could be argued that American football (gridiron) is a “discrete” invasion game, where the match is a collection of many consecutive “downs” or advances by a team in an attempt to carry the ball at least 10 yards so as to be awarded another four downs to continue the advance to the scoring zone. Downs are executed until a touchdown, turnover or period end occurs, but a “ceasefire” is observed between each down, where players receive a brief respite. For the purposes of this research, when the term “discrete” is mentioned, we are referring to a bat and ball game, namely, cricket or baseball. Australian Rules football is clearly a continuous game in comparison to gridiron. Because of the rules, the expanse of the field and the number of players in the
match, AFL is prone to relatively long periods of continuous, open play without any ceasefire, some passages of play even exceeding 30 “transactions” (see Chapter 7) between 10 or more players over multiple minutes. Discrete sports, however, are governed by “one-on-one” contests between, for example, a bowler and a batsman, with a match made up of several finite contests: at least 300 individual deliveries per team in limited-overs cricket and at least 81 pitches per team in baseball (3 strikes x 3 batters x 9 innings). Play is “dead” after each offensive (run) or defensive (dismissal/strike out) action associated with each event. The aforementioned “one-on-one” contest in discrete sports makes player performance measurement a less ambiguous task than that of continuous sports because one can aggregate the result of each of these contests and observe the players involved at each phase; the more a player is involved in successful offensive and/or defensive outcomes, the more easily he can be credited with a good performance. We were reminded of the complexity of an AFL match by Forbes and Clarke (2004), who arrived at a minimum of seven states in a Markov process while an AFL match is in play: Team A in possession; Team B in possession; Ball in dispute; Team A goal (6 points); Team B goal; Team A behind (1 point); Team B behind. In comparison, Hirotsu (2003) was able to work with just four Markov states when analysing the characteristics of association football teams: Team A goal; Team A in possession; Team B goal; Team B possession. The ball spends a lot more time in dispute in AFL, a by-product of being permitted to handle the ball.

4.2.1 AFL Performance Indicators

Notational analysis of performance in sport involves the identification, collection and/or analysis of performance indicators, or open skills, by analysts, coaches and punters to assess the performance of an individual, team or elements of a team (Hughes and Bartlett, 2002). Hughes and Bartlett (2002) define a performance indicator as a “selection, or combination, of action variables that aim to define some or all aspects of performance”, for example, the number of goals in a football match scored by team \( a \) / player \( i \) or the number of successful passes executed by team \( a \) / player \( i \). Target games, as defined in Hopper and Bell (2000), offer a different viewpoint with respect to

---

9 A pitching team may throw fewer than 81 pitches—for example, an out on a first pitch, the team turning double or triple plays (retiring two or three batters with one pitch) or not being required to pitch in the bottom of the ninth if a lead is maintained by that team—but for this dissertation, we will assume 81 is the minimum.
performance indicators. Games like golf and archery, where an individual is competing indirectly against others, involve predominantly closed skills which are developed and measured by biomechanical methods, for example, a golfer’s swing. A player’s performance in these games, although a result of how effectively these closed skills are executed, is usually reflected in a cumulative score which is ranked with the other competitors’. Hughes and Bartlett (2002) suggest sport biomechanics are not as relevant to team sports—where training, psychology, fitness and teamwork are favoured—but certainly remain important. For example, Ball et al (2002), researched the major factors impacting accuracy when kicking for goal in AFL, research that should be widely greeted by coaching staff.

As outlined in Chapter 2, the ultimate objective in AFL is to score a goal by kicking a ball through two upright posts at either end of the ground. Like other invasion sports, scoring is the result of a series of critical events (performance indicators) executed between the individuals involved in the contest (Nevill et al, 2002). For this research, specifically, Chapter 7, any performance indicator that resulted in a i) score, ii) successful pass or iii) turnover was termed a send, while any successful performance indicator that resulted from of a send was termed a receive. Send-receive events are all discrete in nature (Nevill et al, 2002), whether they are the number of kicks by player $i$ or the number of times player $j$ marks (catches) a kick from player $i$. This framework is analogous to $r$-pass movement in football (Reep and Benjamin, 1968). Table 4.3 lists the send-receive performance indicators, with abbreviations, relevant to the AFL research in this dissertation. Basic definitions of each indicator can be found in Section 11.1. These indicators were selected because they represent the primary skills executed by the players to maintain a fluid direction of play with which to establish a goal, score a goal or prevent a goal. All performance indicators in Table 4.3, except Goal, can be described as transitional variables because they are contributing to the continuous flow of play in a match, even in the case of a turnover, for example, team $a$ kick followed by team $b$ mark [KCK$_a$, MRK$_b$]. Goals can be described as terminal variables because they produce a stop and restart of play where teams have to contest possession of the ball. The most frequently occurring send indicators in an AFL game are kicks (KCK) and handballs (HBL), while the most frequent receives are ball gets
BG. BG is defined as a player taking possession of the ball, normally off the ground, when it is in dispute. A coach would prefer that a KCK and HBL is followed by a MRK and HBR respectively, rather than a BG, because this indicates a more efficient movement of the ball between the players\(^\text{10}\). A BG becomes more valuable, however, if player \(i\) from team \(a\) retrieves the ball after a KCK or HBL from player \(h\) on team \(b\), that is, a turnover. While coaches and analysts highlight performance indicators that contribute to a successful performance or outcome, such as a goal, they also focus on “negative” performance indicators, such as free kicks against and turnovers, to highlight areas of improvement for his or her players. For example, in an AFL match, a high frequency of “goals kicked from turnovers” for team \(a\) is a positive performance indicator but a negative for team \(b\); increasing trends in “scores from turnovers” through the season may prompt team \(b\)’s coach to review passing accuracy in their defensive zone.

<table>
<thead>
<tr>
<th>Send</th>
<th>Receive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kick (KCK)</td>
<td>Goal (GLS)</td>
</tr>
<tr>
<td>Handball (HBL)</td>
<td>Behind (BHS)</td>
</tr>
<tr>
<td>Kick In (KIN)</td>
<td>Mark (MRK)</td>
</tr>
<tr>
<td>Hit Out (HIT)</td>
<td>Handball Receive (HBR)</td>
</tr>
<tr>
<td>Knock On (KNK)</td>
<td>Ball Get (BG)</td>
</tr>
<tr>
<td>Spoil (SPL)</td>
<td>Smother (SMT)</td>
</tr>
</tbody>
</table>

Table 4.3: Send-receive performance indicators in an AFL match

The performance indicators just discussed can be aggregated or averaged to any point in the season(s) and reviewed internally or compared with an opponent’s equivalent data for the ensuing match. Table 4.4 displays the results of a simple notational analysis of the Geelong Football Club’s key send-receive indicators from the 2011 grand final. Coaching staff can compare how Geelong performed in the match indicators (Frequency) with respect to the team’s season average (Average). Geelong was down on its season average in all selected indicators except BG, an interesting outcome because Geelong won the game comfortably; indicator frequencies are significantly higher for wins (see \(t\)-test results for all matches in the 2011 season in Table 4.5). This is analogous to the conclusion

\(^{10}\) Unlike a round football, AFL uses an oblong ball which can bounce unpredictably and, hence, out of a player’s control. Taking control of the ball in the air is preferred.
that football teams skilful enough to retain possession for longer periods than their opposition have a
greater chance of scoring and, therefore, winning (Hughes and Franks, 2005).

<table>
<thead>
<tr>
<th></th>
<th>KCK</th>
<th>MRK</th>
<th>HBL</th>
<th>HBR</th>
<th>BG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>211</td>
<td>65</td>
<td>127</td>
<td>75</td>
<td>235</td>
</tr>
<tr>
<td>Average</td>
<td>217</td>
<td>98</td>
<td>169</td>
<td>129</td>
<td>180</td>
</tr>
<tr>
<td>Diff</td>
<td>-6</td>
<td>-33</td>
<td>-42</td>
<td>-54</td>
<td>+55</td>
</tr>
</tbody>
</table>

Table 4.4: Performance indicators for the Geelong Football Club in the 2011 grand final compared to its season average

<table>
<thead>
<tr>
<th></th>
<th>KCK</th>
<th>MRK</th>
<th>HBL</th>
<th>HBR</th>
<th>BG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss</td>
<td>204</td>
<td>84</td>
<td>161</td>
<td>116</td>
<td>183</td>
</tr>
<tr>
<td>Win</td>
<td>235</td>
<td>101</td>
<td>170</td>
<td>126</td>
<td>192</td>
</tr>
<tr>
<td>Diff</td>
<td>+31</td>
<td>+17</td>
<td>+9</td>
<td>+10</td>
<td>+9</td>
</tr>
<tr>
<td>p</td>
<td>0.000</td>
<td>0.000</td>
<td>0.053</td>
<td>0.010</td>
<td>0.057</td>
</tr>
</tbody>
</table>

Table 4.5: Performance indicator totals for all clubs in 2011 for wins and losses, with $p$-values ($\alpha = 0.05$)

Assuming the necessary data was collected, the notational analyst may be asked by coaching staff to extend the analysis to a player level so they can review each player’s performance through the indicator frequencies. Each cell, $Y$, in Table 4.6 represents each Geelong player’s key indicator frequency after the 2011 Geelong grand final match, standardised with respect to his season average by:

$$Y_{im} = (X_{im} - \bar{X}_{im})$$

(4.1)

where $X$ is the frequency of each $m = \text{KCK, MRK, HBL, HBR, BG}$ in the final match for player $i$ and $\bar{X}$ is each indicator’s mean for all of each player’s matches up to but not including the final. The decision to standardise the statistics using Equation (4.1) allows the consumer a more contextual observation of the performance. For example, how would one interpret Bartel achieving 20 kicks? This is a satisfactory performance if he usually averages 10 kicks but unsatisfactory if he usually averages 30. From Table 4.6, no player exceeded his season average in all indicators. However, Hawkins surpassed his average in all but HBR, which is reasonable given his forward position role is more linked with marking (MRK); a smaller forward is more likely to achieve a higher frequency of HBR than a key forward (see Section 6.2.3). Interestingly, Hawkins (5 votes) was voted the third best
player on the ground (out of 44) in the grand final voting; the voting for this award, the Norm Smith Medal, is subjectively undertaken by a panel of AFL experts. The second best player in the grand final, Selwood (9 votes), was in deficit in HBL only, while the best on ground, Bartel (13 votes), was in deficit in MRK only\(^\text{11}\). In the absence of a voting system, would it be possible to scrutinise Table 4.6 and determine who the best player on the ground was? The answer is probably no. However, a simple approximation could be made by summing each player’s \(Y_m\) values in Table 4.6 and then ranking the results in descending order (see Total column).

<table>
<thead>
<tr>
<th>Player</th>
<th>KCK</th>
<th>MRK</th>
<th>HBL</th>
<th>HBR</th>
<th>BG</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Bartel</td>
<td>0</td>
<td>-1</td>
<td>3</td>
<td>1</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>26 Hawkins</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>-3</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>5 Varcoe</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>-4</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>11 Corey</td>
<td>-2</td>
<td>-2</td>
<td>5</td>
<td>-6</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>14 Selwood</td>
<td>3</td>
<td>0</td>
<td>-2</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>8 Hunt</td>
<td>0</td>
<td>-3</td>
<td>1</td>
<td>-1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>35 Chapman</td>
<td>-1</td>
<td>-2</td>
<td>0</td>
<td>-4</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>28 Christensen</td>
<td>2</td>
<td>-2</td>
<td>2</td>
<td>-1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4 Mackie</td>
<td>-4</td>
<td>0</td>
<td>1</td>
<td>-2</td>
<td>3</td>
<td>-2</td>
</tr>
<tr>
<td>45 Ling</td>
<td>6</td>
<td>-3</td>
<td>-8</td>
<td>-5</td>
<td>7</td>
<td>-3</td>
</tr>
<tr>
<td>40 Wojcinski</td>
<td>3</td>
<td>3</td>
<td>-4</td>
<td>-2</td>
<td>-6</td>
<td>-6</td>
</tr>
<tr>
<td>13 Lonergan</td>
<td>0</td>
<td>-2</td>
<td>-4</td>
<td>-1</td>
<td>-2</td>
<td>-9</td>
</tr>
<tr>
<td>30 Scarlett</td>
<td>-3</td>
<td>2</td>
<td>-2</td>
<td>-5</td>
<td>0</td>
<td>-9</td>
</tr>
<tr>
<td>6 Ottens</td>
<td>-2</td>
<td>-3</td>
<td>-3</td>
<td>-2</td>
<td>0</td>
<td>-10</td>
</tr>
<tr>
<td>7 Taylor</td>
<td>-4</td>
<td>-1</td>
<td>-1</td>
<td>-3</td>
<td>-1</td>
<td>-10</td>
</tr>
<tr>
<td>12 West</td>
<td>-1</td>
<td>-1</td>
<td>-3</td>
<td>-1</td>
<td>4</td>
<td>-10</td>
</tr>
<tr>
<td>22 Duncan</td>
<td>-1</td>
<td>-5</td>
<td>-5</td>
<td>-1</td>
<td>-1</td>
<td>-13</td>
</tr>
<tr>
<td>20 Johnson</td>
<td>0</td>
<td>-3</td>
<td>-7</td>
<td>-1</td>
<td>-2</td>
<td>-13</td>
</tr>
<tr>
<td>27 Stokes</td>
<td>-2</td>
<td>-5</td>
<td>-7</td>
<td>-4</td>
<td>4</td>
<td>-14</td>
</tr>
<tr>
<td>44 Enright</td>
<td>-7</td>
<td>-2</td>
<td>-4</td>
<td>-4</td>
<td>-1</td>
<td>-18</td>
</tr>
<tr>
<td>9 Kelly</td>
<td>-1</td>
<td>-2</td>
<td>-9</td>
<td>-9</td>
<td>1</td>
<td>-20</td>
</tr>
<tr>
<td>31 Podsiadly</td>
<td>-7</td>
<td>-6</td>
<td>-3</td>
<td>-2</td>
<td>-2</td>
<td>-20</td>
</tr>
</tbody>
</table>

Table 4.6: Standardised performance indicators for Geelong in the 2011 grand final

The top 5 players in Table 4.6 are as follows: Bartel: 11, Hawkins: 9, Varcoe: 9, Corey: 5 and Selwood: 3; interestingly, the top 3 candidates in the Norm Smith medal voting are contained in our top 5. For this example analysis, the indicators did not take into account critical aspects of the indicators which contribute more to team success, including levels of cooperation and the efficiency with which each was executed (Oliver, 2004). While Bartel was credited with three handballs more than his season average, there is no evidence in the data to suggest his handballs were effective and

\(^{11}\) Scott Pendlebury from Collingwood was awarded 2 votes.
not all turned over to the opposition. Moreover, by summing the raw frequencies of each indicator, rather than a weighted sum, the analysis assumed that no indicator was more valuable than another, that is, each indicator frequency was multiplied by a weight of 1.0. This is a flawed assumption; nevertheless, there is sufficient evidence that this simple, hypothetical notational system possesses some merit, given the correlation with the subjective voting results.

The procedure just outlined is a simple example of notational analysis which evolved into the development and validation of a very basic player performance measurement system. Specifically, this is an individual notational system as its components are the result of individual exploits, without any measurement of teamwork, strategy or a player’s perceived contribution to the final match result. How then does the notational analyst measure the performance of a player in a multi-team sport if each player is within their rights to say they made a valid contribution to Geelong’s victory, even if it is not necessarily evident in the statistical records? Part of the answer to this problem lies in the analysis of player interaction.

4.2.2 AFL Player Interaction

Hughes and Bartlett (2002) note that notational analysis of team and match-play sports focuses on the movements and behaviours of the individual players as well as the interactions between the players. Player interaction is defined in this research as the collection and linking of open skills executed by the players that contribute to a dynamic system of team plays in a match. Table 4.6 is a summary of the individual exploits of the players in the match, some of which generate player performance “scores” in Chapter 5. However, the summary does not provide information on the effectiveness of each statistic or linking information, for example, the percentage of “sends” resulting in a successful “receive”. Player interaction analysis is concerned with the player membership and open skill constitution of consecutive passages of play, where each play comprises at least two players [Send, Receive] from team a. This concept of player interaction underpins the team player performance measurement techniques pursued in Chapter 7, hence, the send-receive notation, where each receive is dependent on a send. Duch et al (2010) argue that the real measure of player performance is “hidden” in these team plays and is not strictly derived from individual events.
associated with player $i$, such as those represented in Table 4.6. Moreover, Pollard et al (1977) noticed
the fit of statistical models improved at a group level rather than by individual performances,
specifically, lower error rates when modelling batting partnerships in cricket rather than individual
batsman scores. These works were motivations to extend the AFL player performance measurement
in this dissertation to a group level, specifically, by investigating the effect that numerous interactions
between groups of players has on a team’s success. Oliver (2004) discusses the necessity of
teamwork, namely, “…a way of cooperating to increase your team’s probability of scoring or to
decrease the opponent’s probability of scoring”. He gives simple examples of interactions in
basketball from an offensive perspective—player $i$ passing to player $j$ if $j$ has a better chance of
scoring—and a defensive perspective—“double-teaming” a player with a higher likelihood of scoring.
Such cooperative actions are comparable to particular events within an AFL match:

i. Player $i$ should kick to player $j$ if $j$ is closer to the goal or able to take the shot from an improved
angle
ii. Defenders often “double-team” a forward who has a reputation for kicking several goals in a
match.

Turning our attention to Geelong’s 2011 grand final match, a simple notational system can be
established for player interaction, provided the transactions between all players involved in the match
have been recorded. Frequencies of each send ($s$) from player $i$ resulting in an effective receive ($r$) by
player $j$ have been tabulated in Table 4.7. Only multiple frequencies of each interaction have been
selected in the table due to the sheer size of the total player combinations: $n = 22 \times 21 = 462$ (22
players in the side with each interacting with, at most, 21 others on the side). The modal interaction in
the match, with the allotted indicators, was [Bartel$_{HBL}$, Duncan$_{HBR}$] with a frequency of 4. Only 3 BG
interactions exist in Table 4.7, opposing the finding that BG is the modal indicator in Geelong’s
individual statistics. This confirms the earlier discussion on MRK and HBR being the preferred
receive indicators for fluent ball movement; BG suggests the ball is in dispute prior to the “get”. With
4 sends and 5 receives, Bartel is the most cooperative player in the sample, which corresponds with
his “best on ground” award. Hawkins achieved 5 receives, while Scarlett achieved 4 sends, explained
by Scarlett’s defensive role, continually looking to deliver (send) the ball out of Geelong’s defensive zone to prevent the opposition from scoring (see Section 7.3.1). Hawkins, on the other hand, is a scoring forward, meaning teammates are continually looking to deliver the ball so as to be received by him. A successful defensive passage of play will, more times than not, begin with a defender, continue through a midfielder and end with a shot at goal by a forward.

<table>
<thead>
<tr>
<th>Players ((i, j))</th>
<th>Interaction ((s, r))</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Bartel, Duncan]</td>
<td>[HBL, HBR] 4</td>
</tr>
<tr>
<td>[Hunt, Bartel]</td>
<td>[KCK, MRK] 3</td>
</tr>
<tr>
<td>[Chapman, Hunt]</td>
<td>[KCK, BG] 2</td>
</tr>
<tr>
<td>[Otten, Selwood]</td>
<td></td>
</tr>
<tr>
<td>[Varcoe, West]</td>
<td></td>
</tr>
<tr>
<td>[Kelly, Bartel]</td>
<td></td>
</tr>
<tr>
<td>[Kelly, Hawkins]</td>
<td></td>
</tr>
<tr>
<td>[Ling, Hawkins]</td>
<td></td>
</tr>
<tr>
<td>[Scarlett, Chapman]</td>
<td></td>
</tr>
<tr>
<td>[Scarlett, Mackie]</td>
<td></td>
</tr>
<tr>
<td>[Christensen, Hawkins]</td>
<td>1, 2</td>
</tr>
</tbody>
</table>

Table 4.7: Multiple frequency player interactions for Geelong in the 2011 grand final

A limitation in player performance measurement is that notational analysts who are gathering statistics from a match, either at the ground or from a video recording, can only realistically focus on the actions immediately surrounding the ball. For example, point i. above might be denoted as \([\text{KCK}, \text{MRK}]\), using the abbreviations from Table 4.3, and the players involved would realise additions to their standardised indicators in Table 4.6. These statistics, however, fail to show that player \(j\) may have run 100 metres, at full pace, to put himself in the position to receive the mark and improve his team’s chance of scoring. In this instance, the interaction between players \(i\) and \(j\) is no ordinary transaction, highlighted by player \(j\)’s selfless effort to put himself in a position to continue the passage of play and to increase his team’s chance of scoring. In point ii., the two defenders may have combined to legally prevent the forward’s path to mark the ball and may not receive any statistical recognition because the forward was not close to the ball and, hence, outside the notational analyst’s vision. The defenders may each have been credited with a “shepherd”\(^{12}\), but this would depend on the proximity of their efforts to the immediate view of the notational analyst. Player \(j\) and the two

\(^{12}\) The act of legally blocking an opposition player to prevent him reaching the ball or a contest.
defenders have each made sacrifices for the good of the team but are either inadequately represented or completely absent from the notational analyst’s statistical records. The lack of statistical detail concerning cooperative “off-the-ball” acts—that is, those not directly involved with the immediate play—is a limitation in many team sport data collections. Extra personnel could be hired to attend games to record “off-the-ball” indicators, but then data collection becomes an increasingly costly task. Cameras with wide-angle lenses could be used to record these indicators but at the expense of close-up detail such as the names of the players involved in the passage of play. Instead, much team sport notational analysis makes do with the visible transactions executed by the players. Our research focuses on the recorded transactions executed by player $i$ that connect player $i$ with either player $j$ or a terminal point (GLS, BHS or “out of bounds”).

The performance variables and methodologies outlined in this chapter become important in subsequent chapters; in Chapter 5 we aggregate individual player performance indicators to arrive at a player-rating system that incorporates player form and opponent strength on a match-by-match basis. Chapter 7 is concerned with player interaction within a match network; we simulate the interactions between any pair of players and observe how this contributes to a team’s success. The findings of the two systems will be critically discussed.

### 4.2.3 Cricket

Notational analysis is concerned with general match indicators, tactical indicators and technical indicators (Hughes and Bartlett, 2002). A critical point of focus regarding the notational analysis of invasion sports is the interaction between teammates and their contributions to scoring. The previous section revealed how interaction between the players can be quantified using a simple additive approach of the different open skills players execute in order to achieve one of three primary play outcomes: setting up a goal, scoring a goal or preventing a goal. Bat and ball sports offer a very different set of parameters with which to quantify player performance, namely, the outcomes of a series of one-on-one contests between a bowler/pitcher and a batsman/batter, with very isolated cases of interaction between the players on the same team. In cricket, the outcomes from each delivery are runs (including zero runs) and wickets, or dismissals. From a notational perspective, there is far less
attention paid to the different physical skills a cricketer executes during a match, unlike in the AFL. For example, the number of pull shots a batsman performs or the number of slower balls a bowler delivers is not recorded for public consumption; such actions are more likely to be analysed and practised at closed training sessions. The performance indicators of interest for each batsman are runs scored, while for the bowler, they are wickets taken and the number of runs conceded. This lack of notational interest in physical skills is most likely because runs (0,…,6) are the most likely outcome after any exchange (a delivery) between a bowler and a batsman (see Chapter 9); as long as runs are being scored, it does not really matter in what technical fashion (within reason) the batsmen are doing so. In sports where longer intervals exist between scoring opportunities, such as in the AFL, there is more emphasis on the players’ open skills because they can be scrutinised to investigate why a team is or is not scoring or to predict who is most likely to score next or, indeed, win the match. Moreover, every batsman has a chance to score runs in a cricket match, whereas in the AFL, it is rare to see more than 10 (out of 22) goal scorers; therefore, the other positions’ unique skills come under scrutiny as team/opponent goals are generated from the efficiency/inefficiency of the execution of these skills. Baseball and the AFL may be related in this way, also. The rarity of run scoring in baseball gives way to an abundance of descriptive match statistics—for example, batting average, on-base percentage and runs batted in (the list of collected statistics is extremely long)—to describe which batter or team is most likely to score or strike out next. A baseball pitcher’s physical skills are also quantified for public consumption, for example, the number of hits from his curve ball versus his fast ball. A cricket batsman’s average (total runs scored / number of innings dismissed in) and bowler’s average (total runs conceded / total wickets taken) are also publicised, but there is far more emphasis on a baseball player’s data, given the rarity of scoring. Croucher (2004) expanded on the cricket batting/bowling average by developing batsman/bowler ratings based on momentum or batting/bowling average (mass) multiplied by batting/bowling strike rate (velocity). However, public offerings are kept at simple averages. As with the AFL performance indicators discussed earlier in this chapter, cricket performance indicators can be expressed as individual exploits or as contributions to team totals (teamwork). Batsman $i$’s runs, $r_i$ in match $t$, can be expressed as a percentage of total team runs, $R$, simply denoted as, $C_i = r_i/R$, where $C$ is described as the batsman’s contribution to the team runs total.
<table>
<thead>
<tr>
<th>Batsman</th>
<th>Rank</th>
<th>Ave</th>
<th>Batsman</th>
<th>Rank</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>HM Amla</td>
<td>3</td>
<td>75.33</td>
<td>HM Amla</td>
<td>3</td>
<td>0.32</td>
</tr>
<tr>
<td>KP Pietersen</td>
<td>11</td>
<td>70.25</td>
<td>KP Pietersen</td>
<td>11</td>
<td>0.30</td>
</tr>
<tr>
<td>V Kohli</td>
<td>37</td>
<td>60.35</td>
<td>BJ Watling</td>
<td>87</td>
<td>0.27</td>
</tr>
<tr>
<td>Shakib Al Hasan</td>
<td>30</td>
<td>59.25</td>
<td>CKB Kulasekara</td>
<td>64</td>
<td>0.25</td>
</tr>
<tr>
<td>Nasir Jamshed</td>
<td>57.75</td>
<td>BRM Taylor</td>
<td>64</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>BJ Watling</td>
<td>87</td>
<td>56.33</td>
<td>Shakib Al Hasan</td>
<td>30</td>
<td>0.25</td>
</tr>
<tr>
<td>AB de Villiers</td>
<td>7</td>
<td>53.75</td>
<td>Nasir Jamshed</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>IR Bell</td>
<td>17</td>
<td>52.30</td>
<td>V Kohli</td>
<td>37</td>
<td>0.22</td>
</tr>
<tr>
<td>AN Cook</td>
<td>5</td>
<td>46.64</td>
<td>IR Bell</td>
<td>17</td>
<td>0.22</td>
</tr>
<tr>
<td>MJ Clarke</td>
<td>1</td>
<td>44.69</td>
<td>AB de Villiers</td>
<td>7</td>
<td>0.22</td>
</tr>
<tr>
<td>LRPL Taylor</td>
<td>44.60</td>
<td>AN Cook</td>
<td>5</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>MK Tiwary</td>
<td>43.00</td>
<td>LRPL Taylor</td>
<td>0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G Gambhir</td>
<td>35</td>
<td>42.81</td>
<td>CR Woakes</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>BRM Taylor</td>
<td>64</td>
<td>42.33</td>
<td>MJ Clarke</td>
<td>1</td>
<td>0.18</td>
</tr>
<tr>
<td>KC Sangakkara</td>
<td>4</td>
<td>42.11</td>
<td>GJ Maxwell</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>BB McCullum</td>
<td>28</td>
<td>41.60</td>
<td>KC Sangakkara</td>
<td>4</td>
<td>0.17</td>
</tr>
<tr>
<td>AB Barath</td>
<td>41.50</td>
<td>G Gambhir</td>
<td>35</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>GJ Maxwell</td>
<td>40.67</td>
<td>BB McCullum</td>
<td>28</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>Anamul Haque</td>
<td>39.00</td>
<td>Azhar Ali</td>
<td>12</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>Tamim Iqbal</td>
<td>32</td>
<td>38.56</td>
<td>GJ Bailey</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>BMJ Mendis</td>
<td>38.25</td>
<td>Tamim Iqbal</td>
<td>32</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>MS Dhoni</td>
<td>37.43</td>
<td>T Taibu</td>
<td>0.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TM Dilshan</td>
<td>37.30</td>
<td>MS Dhoni</td>
<td>39</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Azhar Ali</td>
<td>36.73</td>
<td>JM Bairstow</td>
<td>73</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>MJ Guptill</td>
<td>36.36</td>
<td>TM Dilshan</td>
<td>27</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>DA Warner</td>
<td>35.48</td>
<td>E Chigumbura</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GJ Bailey</td>
<td>34.83</td>
<td>Anamul Haque</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IJL Trott</td>
<td>34.17</td>
<td>AB Barath</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC Smith</td>
<td>33.80</td>
<td>MK Tiwary</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.8: Correlation of 2012 batsmen average and contribution ranking with ICC ranking

A score of 20 by a top batsman may be viewed as a low score on its own, but if the total team score was low, say 120, due to difficult batting conditions (for example, a cracked pitch or humidity), the batsman’s score may be deemed as satisfactory. In this instance, it is trivial to compare the individual average approach to the average contribution to team totals. In Table 4.8, batsmen on the left are ranked according to their 2012 batting average (Ave) (only the top 30 are shown) and on the right according to their 2012 average contribution (C) to the team’s score. Notice how there is quite a deal of change in the top 30 batsman membership in both columns, suggesting that the two performance indicators reveal different aspects of the batsmen. The “Rank” column is where the batsman was ranked according to the ICC cricket player ratings as at 31 December 2012 (see basic description of ICC rating system in Chapter 9). We can validate which is the preferred method of quantifying a batsman’s performance through a Spearman’s correlation analysis on the top 100 batsmen’s “Ave” and “C” rank and their ICC rank. There is marginal evidence that the batting contribution ($\rho = 0.172$) better explains a batsman’s performance than the batting average method ($\rho = 0.131$) with respect to
the ICC rankings, but both values are far from significant (2-tailed). In Chapter 9, simulation methods will be investigated to predict how each individual batsman will perform during a match, with respect to the state of the match resources and the type of batsman being assessed.

4.3 AFL Fantasy Sports

This century, the evolution of online media has increasingly focused on interaction with the media user. Websites, television and radio have all embraced technology that allows the audience to become involved with its preferred content, whether using an online forum to offer opinion on a newspaper article or joining a “community” using social media such as Twitter or Facebook. Sporting media is certainly no different, with sports fans able to access online information and “chat” with fellow punters across a seemingly endless range of sports, from basketball and football through to games such as chess and even Tiddlywinks. A typical sporting website will not only offer pre-game news, statistics and odds but also in-game content such as live scores and team/player performance indicator frequencies. At the centre of interactive sports media are “fantasy sports”. A fantasy sport enables members of the public to compete for prize money in an online version of their favourite sports; specifically, they assume the role of an online “coach” and build a team that competes against other fantasy coaches coexisting in cyberspace.

The fantasy sports explosion occurred concurrently with the internet boom of the mid-1990s as statistics and other relevant information were more readily uploaded and accessed by worldwide participants. The popularity is evident in the numbers alone; the Fantasy Sports Trade Association estimated that 32 million people over the age of 12 in the US and Canada played fantasy sports in 2010 (Fantasy Sport 2010). The Australian Football League’s official fantasy league, the “Dream Team”, had over 320,000 participants in 2010. To become involved, fantasy coaches register and then pick their side at the start of the season; the success of the side from week to week is dependent on the real-life performance of each selected player on the AFL field. The strategy may sound simple—pick the historically best-performing players—but the following constraints are in place to prevent this approach:
1. 30 players are to be selected in a side
2. 9 defenders, 8 midfielders, 4 rucks and 9 forwards are to be selected
3. Every coach must adhere to a “salary cap” or ceiling.

The final point above requires coaches to be prudent with their selection strategy because each player comes with a price which is commensurate with his talent. Given the salary cap, selecting a team of high performers is unfeasible, so a mix of these and cheaper “up-and-coming” players is preferable. This strategy is comparable to the strategy of the GWS coaches, outlined in Section 4.1. Moreover, coaches have a given number of “trades” allowing them to swap an injured or under-performing player for another player who is of no greater value and is not already in the side. After each game, a player’s value fluctuates, governed by a moving average of his past $k$ performances; that is, his price will increase in value following a cluster of good matches and vice versa. Player $i$’s performance is reflected by a weekly score:

$$SCORE_i = 3*KCK + 1*HBL + 3*MRK + 4*TKL + 6*GLS + 1*BHS + 1*FF - 3*FA$$ \hspace{1cm} (4.2)

where: KCK = kick, HBL = handball, MRK = mark (catch), TKL = tackle, GLS = goal, BHS = behind (one point), FF = free kick (for) and FA = free kick (against). These represent the frequency of on-field performance indicators collected for any player $i$. The coefficients attached to each performance indicator have been subjectively chosen to reflect the perceived worth of each collected variable to the outcome of the game. Every fantasy side’s score after each round of football is simply \(\Sigma SCORE_i (i = 1, \ldots, 22)\). Oliver (2004) developed similar linear equations for individual offensive and defensive performances in basketball, moving so far as to calculate an individual’s contribution to a game using indicators such as field goal attempts (FGA), field goals made (FGM), points scored (PTS) and assists (AST).

Jimmy Bartel from the Geelong Football Club provides a neat example in the AFL; in round 1, he achieved a score of 110, calculated by substituting his performance indicator frequencies for the match into Equation (4.2):

\[^{13}\text{Twenty-two players are scored, while 8 are “reserves”}\).
Table 4.9: Jimmy Bartel performance indicator frequencies

<table>
<thead>
<tr>
<th>KCK</th>
<th>MRK</th>
<th>HBL</th>
<th>TKL</th>
<th>GLS</th>
<th>BHS</th>
<th>HIT</th>
<th>FF</th>
<th>FA</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>8</td>
<td>13</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>110</td>
</tr>
</tbody>
</table>

Figure 4.2 gives his starting price at round 1 as $482,200, which remained relatively constant in proportion to an average score of 105 between rounds 1 and 4. His sharp decrease in value was in response to a drop in performance, averaging 81 between rounds 5 and 7. An average score of 98 between rounds 8 and 13 contributed to a steady improvement in value up to round 13.

Figure 4.1: Jimmy Bartel fantasy performance (Score) and value (Price) for the 2012 season

Equation (4.2) represents another very simple example of a player performance metric. By summing or averaging these scores across each of Bartel’s matches, a player rating and, therefore, ranking can be generated after each performance. As this thesis progresses, more complex ratings
methodologies will be explored, ones where any subjectivity involved in parameter and coefficient selection is removed and replaced with scientific approaches. The proceeding chapter will introduce different team and player rating methodologies and culminate in an objective AFL player rating system that responds to the original research questions from this research.
Chapter 5

Ratings Systems

5.1 Introduction

A simple but rigorous AFL player rating system is discussed in this chapter, which not only effectively rates each player in their allocated position, but offers a pre-match performance prediction, a powerful concept in team and player ratings. This chapter seeks to answer questions posed in the infancy of the research and is the origin of the different research strands in this dissertation. Chapter 6 augments the positional component of the ratings outlined in this chapter, the results of the network methodology in Chapter 7 are frequently compared to this model while Chapter 8 investigates a median smooth performance forecasting model that can be applied when Gaussian assumptions attached to this model are violated. A detailed discussion on team and player ratings is necessary prior to the description of player ratings systems originating from this research.

5.2 Team and Player Ratings

Ratings in sport are derived from evaluations of the performance of a team or individual player, most often with prior performances in mind. Stefani (2012) offers a succinct distinction between a rating and a ranking: “A rating is a numerical value assigned to a competitor, based on results and other factors while a ranking is the ordinal placement based on ratings.” Rating systems are beneficial to numerous parties; athletes and coaches can track form and progress and use the rating as a motivational tool, while sporting federations can publish the top-ranked (and bottom-ranked)
sides for public consumption. With ratings in place, a league and, indeed, world ranking can be produced: teams and/or players can be compared for improved team selection; and players and teams can be compared in hypothetical situations for betting purposes. Stefani and Clarke (1992) and Stefani (2012) define three types of rating systems across a range of internationally recognised sports: subjective, accumulative and adjustive (or adaptive). Subjective rating systems, as the name suggests, offer the least scientific approach to rating competitors, with a panel of experts ranking the competitors after each assessable round. Accumulative systems are the most widespread, converting performance to points which are added over a specified number of rounds to produce the rating. They are especially attractive for individual events like archery or diving because competitors’ final point totals are commensurate with their performance (accuracy, in an archer’s case). A limitation of accumulative ratings is the capacity to overlook absent competitors, particularly in a team sport. For example, if a rating system is concerned with the sum of performance-dependent points, allocated to football players over a season, players who aren’t able to participate in certain matches due to injury risk being “leapfrogged” by their teammates who play more games and have the opportunity to earn more points. The fact that an injured player has fewer points—and, hence, a lower rating—does not necessarily mean he is a lesser player than one who has acquired more points. Adjustive rating systems account for leapfrogging because ratings can increase, decrease or remain constant depending on above, below and met expectations respectively; the latter would apply to competitors who did not participate for reasons such as injury. Adjustive systems have interesting and unique properties that require discussion. The generalised form of the rating is:

$$R_n = R_o + k(Obs - Exp)$$  \hspace{1cm} (5.1)$$

where $R_n$ is the new rating for player/team $i$, $R_o$ is the old rating, $Obs$ is the observed result of a contest, $Exp$ is the expected result and $k$ is a multiplier that assumes different interpretations depending on the contest. A key attribute of these ratings is that $Exp$ is predictive in nature, providing a pre-match approximation of a player’s or team’s performance. $Exp$ is usually probabilistic, describing the likelihood of a player or team defeating the opponent, so is a function of opponent strength ($R_{oi} - R_{oj}$, where $j$ is the opposition player/team). The opposition strength parameter
introduces another attractive property of the adjustable system: a lesser participant can realise a heavy defeat against a stronger opponent without experiencing a heavy rating decrease—a powerful concept ignored by many ratings systems. Conversely, the stronger opponent may only realise a slight ratings increase having soundly beaten a lesser opponent.

5.2.1 Elo ratings

Arpad Elo was a pioneer in the field of statistical ratings and no sport statistics conference would be complete without at least one reference to the Elo ratings system. His book, *The Rating of Chess Players, Past and Present* (1978), details the Elo rating system with respect to the individual sport of chess. The Elo rating system is arguably the most recognised adjustable system and has since been used in research on other individual sports such as tennis (Bedford and Clarke, 2000) and, even, Tiddlywinks (Barrie, 2003). The general Elo ratings equation is the same as Equation (5.1), but the expected performance of player $i$ is calculated as a probability of victory before the match, under a logistic curve, using:

$$\text{Exp} = \frac{1}{10^{(R_i - R_j)/400} + 1}$$  \hspace{1cm} (5.2)

where $R_i$ and $R_j$ are the strengths of players $i$ and $j$ respectively. Elo’s assumption was that chess performance followed a normal distribution, but governing Chess bodies have since revised the assumption so that Equation (5.2) follows a logistic distribution. The $k$ multiplier from Equation (5.1) is set at different values, depending on the levels of chess players, to adjust for overperformance or underperformance relative to player $i$’s Exp. Rating movements are dependent on: the difference between observed performance in the match—wins (1), losses (0) and draws (0.5)—and expected performance; a probability of victory for player $i$; recognising opponent strength; and the level of the competitors in the game.

The Elo rating system also has been applied in team sports, where performance can be quantitatively measured by score or margin. The Elo ratings provide a pre-match probability of team $a$ beating team $b$ which is subtracted from the observed result to determine the direction of the rating
movement. The most notable application was to rate men’s national world football teams (World Football Elo Ratings 2007), not to be confused with the FIFA World Rankings (FIFA World Rankings 2007). Bedford (2004) used a modified Elo ratings model to predict outcomes of the 2003 Women’s World Cup Handball tournament. Although parameters take on different interpretations depending on the sport being played, the Elo principle remains the same: a pre-match numerical rating with a post-match revision dependent on the result; the quality of the opposition; and the extent (size) of the win when performance can be quantified. A segment of this research applied Elo ratings to AFL teams to support an argument that the current AFL ladder system was not the fairest representation of each team’s performance through the season. The current AFL team ranking system is based on aggregating team $i$’s premiership points, $w$ to round $n$ in a given season, where:

$$w = \begin{cases} 
4 & \text{win} \\
2 & \text{draw} \\
0 & \text{loss}
\end{cases}$$

If two or more teams are tied on aggregate points, $\sum w$, at match $n$, a percentage formula, $p$, is referred to (Equation (5.3)); the team with a higher $p$ is ranked higher on the ladder than the team they are point-tied with:

$$p = \frac{\sum_{t=1}^{n} SF_t}{\sum_{t=1}^{n} SA_t} \times 100$$  \hspace{1cm} (5.3)

where $SF$ are points scored in match $t$ by team $a$, $SA$ are points conceded in match $t$ by team $a$ and $n$ is the most recent round. Our Elo ratings model for AFL teams takes the form of Equation (5.1) with some notable variations: $k$ becomes a linear representation of the margin of victory or:

$$k = 25 + 2.5 \left( \frac{|SF_i - SA_i|}{6} \right)$$  \hspace{1cm} (5.4)

where the denominator represents the value of a goal in Australian Rules football, thereby establishing the number of goals in the margin. Thus, $k$ becomes a linear function of score margin with intercept and slope optimised so as to retrospectively maximise the correct match result predictions (using
Equation (5.5)). In world football Elo ratings, \( k \) reflects the importance of the tournament (for example, \( k = 60 \) for World Cup finals) and another multiplier, \( g \), is the goal difference index. Because our ratings were only calculated with one competition in mind, a match importance index was redundant, and so ignored. Another variation in team sport Elo ratings is that home-ground advantage is incorporated in pre-match probability calculations:

\[
Exp = \frac{1}{10^{(R_d - R_b + h)/400} + 1}
\]

(5.5)

where \( h \) signals a team’s strength at its home ground. This is an important inclusion because 55% of AFL matches are won by the home team\(^{14}\), with an average score of 96 compared with the visitor average score of 88 \(^{15}\). The Elo model is probably a fairer representation of the teams in the AFL than the current ranking system, given strong teams will not receive as large a rating increase after defeating a much weaker side. Equation (5.3) will increase the strong team’s percentage at the expense of the weaker team, whereas Elo handicaps the strong team with respect to the weaker team. The current system has serious implications for a team that is hoping to play in the finals but could miss out due to a difficult draw in the lead-up to the finals. Moreover, a team that may not be as worthy of playing finals may progress due to an easier draw leading up to the finals. Clarke (2005) discusses the repercussions of an unbalanced draw, employing linear regression to calculate individual home advantages for each club. An Elo approach may be another solution to the problems associated with an unbalanced draw.

### 5.2.2 Player Ratings in Continuous Team Sports

This chapter has focussed on the applications of adjustive ratings models for individual contests as well as for continuous sporting teams, but what of the ratings’ application to competitors involved in continuous team sports? How to generate player ratings for competitors in team sports presents the core research question for this dissertation and has provided countless hours of discussion and puzzlement. A renowned AFL journalist, Mike Sheahan, publishes an annual list of whom he

---

\(^{14}\) The majority of AFL teams share a home ground with another team so in some matches there is no home advantage.

\(^{15}\) Calculated using data from the 2012 AFL season.
considers to be the 50 best players (Mike’s Top 50 2012). The list receives a mixture of feedback—mostly negative—from the AFL community and general public. A Geelong supporter, for example, will always argue that Jimmy Bartel is a far better player than Sheahan gives him credit for each year. People mostly agree on whom the best footballer is in Australia (Lance Franklin or Gary Ablett Jr.), but how does one distinguish between the fourth and fifth or 49th and 50th best player? Is it possible for opinion and science to operate concurrently or will arguments between different parties prevent player ratings from ever being truly accepted as a meaningful measurement in team sports? The scarcity of published research on player ratings in continuous team sports reflects a lack of consensus, understanding, demand or all of the above. Rugby player ratings were developed by Bracewell (2003) using factor analysis, where each factor represented a core trait performance across nine positional clusters. Oliver (2004) went to admirable efforts to establish a set of offensive and defensive ratings for each player in a basketball match; this was an additive system using frequencies of on-court skills such as effective passes and score assists executed by that player. More impressively, he developed, using relatively simple mathematics, a “difficulty theory for distributing credit to players in basketball”: the more difficult the contribution, the more credit it deserves. This will be discussed further in Chapter 7 when cooperative ratings are investigated through network analysis.

5.2.3 An AFL Player Performance Estimate

World football is one of the few continuous team sports that has adapted and published official player ratings; the Castrol rankings are based on the actual performances of every player across the Major Soccer League (USA and Canada) (MSL Castrol Football Rankings 2013). Specifically, after each match, every player is given index scores that reflect the outcomes of every touch of the ball from different parts of the pitch and are then converted in an adjustive points system16. In the AFL, the only publicly accessible player-rating system with a disclosed equation is that calculated in the AFL-sponsored fantasy football competition (see Section 4.3). Each player’s rating at any stage in the season is the average of his performance scores, \( X_n \), to the most recent match, \( n \), where:

16 The algorithm is not publicly accessible.
\[ X_i = 3*KCK + 1*HBL + 3*MRK + 4*TKL + 6*GLS + 1*BHS + 1*HIT + 1*FF - 3*FA \]  \hspace{1cm} (5.6)

This could be described as a subjective-adjustive system because:

i) The weightings attached to each performance indicator are inexactely chosen

ii) Player \( i \)’s rating (average) can rise, fall or remain constant in line with overall performance.

The key issue with averaging all performances \( X_{t-1,\ldots,n} \) is that the most recent performances, described as a player’s “form”, are not accurately portrayed. These ratings are only comparatively informative. An accumulative system is also in place for the AFL fantasy competition—summing each \( X_t \)—but this system is not as informative because it biases players who have played the majority of games in the season. If a fantasy coach has a selection option where, say, player \( i \) has a higher rating and lower price than player \( j \), statistically speaking, player \( i \) is the preferred choice\(^{17}\). A major drawback of the AFL fantasy ratings—and, indeed, any subjective rating system—is the lack of scientific validation for the final ratings. The number of kicks is an obvious inclusion in Equation (5.6), but the weighting of 3 attached to the kick frequency represents a perceived value of the performance indicator on player \( i \)’s team’s chances of victory, relative to the other indicators. The equation suggests that a tackle is more valued than a kick. The weight of -3 for each free kick player \( i \) concedes reflects the negative impact on the team, that is, losing possession of the ball. How do we know these are the optimal allocations? It is possible to assess the coefficient allocation by calculating a team rating—an average of the player scores from each match in a season(s)—then plotting them with respect to the score margin from each of those matches. Figure 5.1 shows a linear relationship (\( y = 5.52x - 332.03 \)) between Geelong’s team ratings average and each margin from matches in the 2011 season. The strong positive relationship (\( R^2 = 0.6705 \)) supports the use of the AFL rating system; the better the players perform to increase the team rating, the greater their chance of victory (higher margin). Furthermore, the coefficients in the fantasy equation can be optimised using linear programming to maximise the relationship (\( R^3 \)) between the average of each \( X_t \) and the final margins from Geelong’s matches (see Figure 5.2).

\(^{17}\) A fantasy coach may choose player \( j \) over \( i \) for subjective reasons; for example, the coach may know that \( i \) is more injury prone and could potentially miss upcoming games.
The new equation is:

\[ X_t = 1.9*KCK + 0.6*HBL + 0.5*MRK + 1*TKL + 13.7*GLS + 3.2*BHS + 0.7*HIT + 0.9*FF - 0.9*FA \]  \hspace{1cm} (5.7)

Figure 5.1: Linear relationship between average AFL fantasy player scores and margin for Geelong’s 2012 matches

Figure 5.2: Linear relationship between re-optimised average player scores and margin for Geelong’s 2012 matches

The optimisation naturally gave greater weighting to scoring (GLS, BHS) to increase the margin and the chance of victory. While the \( R^2 \) is greatly improved, the equation \( (y = 8.39x - 220.92) \) is flawed—it overvalues scorers (forwards, denoted as “F” in Figure 5.3) and undervalues players not prone to goal kicking, most notably, defenders (“D”). Note, the scale of performance scores has decreased in the second boxplot due to the different weightings, but the increase in a forward’s scores relative to
the other positions is still evident. This problem will be addressed in Chapter 6 using optimal positional allocation.

![Boxplots showing the distribution of Geelong’s scores by player position](image)

Figure 5.3: Boxplots showing the distribution of Geelong’s scores by player position (i. Equation (5.6), ii. Equation (5.7))

Oliver (2004) perhaps best summed up these attempts at player performance evaluation, with respect to basketball: “All these formulas are just approximate ways of representing someone’s opinion about the quality of players.” Equations (5.6) and (5.7) are naive descriptions of how an individual player performed in a given match; their value can be enhanced by including them as parameters in an adjustive rating system, as in Equation (5.1).

### 5.3 An Adjustive AFL Player Rating System

While it could be argued that in the infancy of this research the Elo model was adopted for producing AFL player ratings, it would be more correct to label it an adjustive rating system influenced by Elo. For ease of reference, we titled the system the adjustive AFL player rating system or APR. The complexities of the AFL—namely, the number of players on the field and the variety of players’ roles—made ambiguous and, at times, erroneous the application of the original Elo formula, particularly the expectation calculation. From a player perspective, the AFL is not a traditional head-to-head competition where the observed result is a win, loss or draw, as in chess; instead, players are
said to have played well, moderately or poorly\textsuperscript{18}. Although two players will “match up” on each other, the match-up rarely continues for the whole game, demanding that the data collection process account for each and every match-up on the field at any stage in the game, which is virtually impossible. Moreover, different performance appraisal methods may be required for the match-ups in order to declare the “winner” of the match-up. For example, a defender who was matched up on a forward for the entire game would need his goal prevention assessed, while the forward would need his goal kicking/assists assessed; two separate equations are required. We are now out of the bounds of a logical application of the classical Elo model, but it is possible to arrive at a logical rating system that is respectful of the Elo methodology.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig5.4.png}
\caption{The adjustive AFL player rating system}
\end{figure}

\subsection{The APR Model}

The APR model assumed the form of Equation (5.1) and, like other adjustive systems, added rating points for above-expectation performance and subtracted points for below-expectation performance. This ratings methodology was adopted and developed under the assumption that players whose observed performance consistently exceeded their expected performance were more likely to achieve higher rankings. For the sake of simplicity, each player’s observed performance after match $t$, was measured by Equation (5.6); a more rigorous approach to AFL player performance measurement is pursued in Chapter 7. While Equation (5.6) can itself be described as a rating—a rating for that player in that match—the APR model was designed to depict a player’s performance and form across one or multiple seasons. For this chapter, Equation (5.6) is a player’s performance in match $t$. The

\textsuperscript{18} Ruckmen, it can be argued, do contest head-to-head, being the most specialised position within the game.
motivation for the APR model was to be able to take advantage of the predictive element of the ratings \((Exp)\), that is, to arrive at a parametric estimate of how a player might perform in an upcoming match. These estimates would determine whether a player is a sensible pick in a fantasy league given a certain set of match conditions, for example, opposition strength. Such knowledge is also beneficial when predicting the winner of player awards. It was important to respect Elo’s “head-to-head” methodology, so the agreed approach involved calculating the probability of player \(i\) outscoring player \(j\) in an upcoming match, where player \(j\) was a randomly selected opponent from the same position as \(i\) (for example, midfield) in match \(t\). A player’s position (Defender, Forward, Midfielder, Ruck) was initially allocated as per the AFL fantasy league website and assumed to be uniform for each player across the entire season, which is not realistic because players can play in different positions from week to week and, even, during a match. Chapter 6 details a player position classification system which scientifically allocates players to one of the four positions after each match, with confidence levels, using a range of AFL performance indicators. For the purposes of establishing the APR model, the positions from the fantasy league site were considered adequate.

From Equation (5.1), the observed and expected values for each player \(i\) are denoted as:

\[
Obs = \begin{cases} 
1 & \text{if } X_i > \bar{X}_j \\
0.5 & \text{if } X_i = \bar{X}_j \\
0 & \text{if } X_i < \bar{X}_j 
\end{cases}, \quad \text{Exp} = P(X_i > X_j) \quad (5.8)
\]

where \(X\) is a random variable describing the performance of players \(I\) from team \(a\) and \(j\) from team \(b\), and \(\bar{X}_j\) is the average observed performance of the opponents in the same position as player \(i\) in match \(t\). Like Elo (1978), we modelled the players’ performance, \(X\), as a normally distributed random variable, so each \(X_i \sim N(\mu_i, \sigma_i)\). It was, hence, possible to arrive at estimates of each player’s score, \(\hat{X}_i\), from each player’s unique distribution in his allocated position. Figure 5.5 shows Geelong midfielder Joel Selwood’s performance score distribution between 2010 and 2011, fitted with a normal distribution. The estimated random opponent player score, \(\hat{X}_j\), was derived from the normal distribution of the player \(j\) scores—in the same position as \(i\)—in matches leading up to but not
including \( t \) in that year. By generating a unique normal performance distribution for each position in each team using matches prior to \( t \), the model was able to account for opponent strength—not just the team as a whole, but the relative strengths of each team’s positions. Figure 5.6 offers a comparison between the midfielder performance distribution of Geelong’s round 23 (Sydney) and round 24 (Collingwood) opponents in the 2011 season. Collingwood played Geelong in the 2011 grand final and was considered to have an exceptional midfield, realising a mean score of close to 89 and 90% of scores falling between 43 and 137. Sydney finished seventh on the ladder in 2011 and was considered to have a weaker midfield, with a mean of under 80 and 90% of scores falling between 31 and 121 (see Figure 5.6). A Geelong midfielder, for example, would have a greater chance of outscoring a Sydney midfielder than a Collingwood midfielder.

Figures 5.5: Joel Selwood’s scores fitted with a normal distribution

\[ \text{Exp} \] was calculated by simulating 1000 different \( \hat{x}_i \) and \( \hat{x}_j \) from the player and opponent normal distributions respectively and recording the percentage of \( \hat{x}_i > \hat{x}_j \). Simulations were run on every Geelong player’s matches between 2010 and 2011 so that the ratings could be adjusted on a match-by-match basis. Apart from the old and new ratings, the final value to be calculated in Equation (5.1) was the multiplier, \( k \), which determined the severity of the ratings fluctuation for each player. Because the result of any player match-up could be quantified, \( k \) measured the size of the match-up victory or defeat for player \( i \), using:

![Figure 5.5: Joel Selwood’s scores fitted with a normal distribution](image-url)
where $\bar{X}_j$ is the average score of like-positioned opponents to $i$ after match $t$.

$$k = |X_i - \bar{X}_j|$$ (5.9)

Figure 5.6: Collingwood (left) and Sydney (right) midfielder scores fitted with normal distributions

Supplementary APR system details are as follows:

1. Ratings were produced from data collected over the 2010 and 2011 seasons
2. Each player started with a rating of 500 as of round 1, 2010
3. Ratings were calculated after the player’s third match in the rating period
4. Only Geelong midfielders were rated for this stage of the research
5. For future research, forwards will be measured against opponent forwards and defenders against opponent defenders and ruckmen against opponent ruckmen
6. Round 1, 2011 was excluded due to limited knowledge of team form for that year.

5.3.2 Model Validation and Results

It is simple to demonstrate a ratings update for Joel Selwood, after the final match (round 24) of the 2011 regular season, by substituting this data into Equation (5.1) to obtain $R_o$:

$$R_o = 813.8 \quad \text{(round 23 rating)}$$

$$k = |118 - 81| = 37$$

$Obs = 1$

$Exp = 0.716$

$$R_n = 813.8 + 37(1 - 0.716) = 824.4$$
Table 5.1 shows Joel Selwood’s ratings fluctuations for each match in the 2010 and 2011 season. Selwood was in the top 3 midfielders at Geelong and was consistently achieving high scores, so rarely does he realise a drop in ratings. His lowest score was 27 in round 1 of 2011 which we did not record. In this game, Selwood was injured and was substituted out of the match, which poses a question about whether it is fair to penalise players who have sustained an injury in a match, preventing them from improving their current score. The next section details transformation techniques which interpolate scores with respect to “time on ground”. Selwood’s biggest ratings
decrease was -6.9 against Port Adelaide in round 4, 2010, only achieving 60. Although he fell 10 points behind Port Adelaide’s midfield average, his $Exp$ was 0.664, suggesting strong midfield opposition, and prevented a sharp decrease in rating. If his expectation to win a match-up in that game was 90%, his rating would have fallen a further 10 points; the system is harsher on players who are “supposed” to play better than they actually do. His sharpest increase was 27.9 against Melbourne in round 19, 2011. His expectation to win a match-up was 0.772, but his score was 178, which is outstanding. If his $Exp$ had been set at 0.664, like in the Port Adelaide match, his rating would have increased a further 14 points. A higher $Exp$ value suggested Selwood was expected to play well against the weaker opposition, so his rating increase was handicapped.

It was important that the APR model possessed a reliable predictive element. We ran an “internal” validation on the model by measuring the relationship between each player’s $Exp$ values and his resulting performance scores, $X_t$, for every Geelong match from 2010 to 2011. Figure 5.7 reveals a satisfactory linear relationship between the $Exp$ and $X_t$ values ($R^2 = 0.341$), suggesting performance can be predicted with modest confidence. The outlier sitting close to the x-axis near $Exp = 0.8$ is Jimmy Bartel who was concussed in round 13, 2011 in the first quarter, having only achieved $X_t = 2$; he was taken from the field and did not return. It is anticipated that the accuracy of the model can be enhanced by an additional variable, possibly ground advantage (travel effects), to improve the fit.

A problematic area for ratings models, particularly those concerning players in team sports, is validation of the final output. How does the notational analyst know that the ratings, calculated up to the prior match, are reflective of observed on-field performance? We decided to compare the final 2011 ratings for the Geelong midfield to the club’s “best and fairest” voting results. The award system is accumulative, with votes awarded by Geelong’s internal management after each game with each player given a rating out of 10 where 10 is the highest. Firstly, we summarised every player’s rating results after his last measured match. In Table 5.2, $f(Exp)$ is the average $Exp$ value for player $i$ through the rating period, $Rating$ is his final rating (the table is sorted in Descending order of $Rating$).
Figure 5.7: Relationship between $Exp$ and $X_i$ for all Geelong midfield matches (2010–2011)

*Movement* is his average rating movement per round and *Rank* is where each player finished in the best and fairest voting. The winner of the award for 2011 was Corey Enright—he achieved 150 points (ranked one)—who throughout the year was classified as a defender but more through the midfield than as a key defender who tends to stay closer to the opposition goals. The encouraging result was that our top-five-rated Geelong midfielders finished ranked between second and sixth place in the award voting. This gave us confidence to use the APR as a valid ratings model.

<table>
<thead>
<tr>
<th>Player</th>
<th>Matches</th>
<th>$f(Exp)$</th>
<th>Rating</th>
<th>Movement</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joel Selwood</td>
<td>38</td>
<td>0.661</td>
<td>823.10</td>
<td>8.54</td>
<td>6</td>
</tr>
<tr>
<td>Jimmy Bartel</td>
<td>41</td>
<td>0.726</td>
<td>678.95</td>
<td>4.36</td>
<td>3</td>
</tr>
<tr>
<td>Cameron Ling</td>
<td>36</td>
<td>0.487</td>
<td>670.37</td>
<td>4.73</td>
<td>4</td>
</tr>
<tr>
<td>Joel Corey</td>
<td>30</td>
<td>0.580</td>
<td>648.27</td>
<td>4.94</td>
<td>2</td>
</tr>
<tr>
<td>James Kelly</td>
<td>21</td>
<td>0.593</td>
<td>623.51</td>
<td>5.88</td>
<td>5</td>
</tr>
<tr>
<td>Paul Chapman</td>
<td>20</td>
<td>0.672</td>
<td>608.37</td>
<td>5.42</td>
<td>-</td>
</tr>
<tr>
<td>Allen Christensen</td>
<td>16</td>
<td>0.206</td>
<td>589.01</td>
<td>5.56</td>
<td>-</td>
</tr>
<tr>
<td>Gary Ablett *</td>
<td>21</td>
<td>0.824</td>
<td>560.09</td>
<td>2.86</td>
<td>-</td>
</tr>
<tr>
<td>Mitchell Duncan</td>
<td>18</td>
<td>0.427</td>
<td>499.75</td>
<td>-0.01</td>
<td>-</td>
</tr>
<tr>
<td>Josh Cowan</td>
<td>3</td>
<td>0.049</td>
<td>493.52</td>
<td>-2.16</td>
<td>-</td>
</tr>
<tr>
<td>Simon Hogan</td>
<td>12</td>
<td>0.223</td>
<td>465.92</td>
<td>-2.84</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 5.2: Final ratings results for Geelong midfielders in 2011

(* Gary Ablett was only rated for 2010 because he was drafted by another club for the 2011 season.*)
5.3.3 Exponentially Smoothed Ratings

Section 5.2 discussed an elegant feature of adjustive ratings: a performance prediction parameter, Exp, while Section 5.4.1 demonstrated how this parameter could be estimated by simulating pre-match performances from player and opponent (normal) distributions. There will be instances where a parametric approach to estimating Exp is not feasible due to violation of assumptions, for example, normality, in which case, an exponential smoothing approach could be considered. Stefani and Clarke (1992) used an exponential smoothing approach to develop adaptive AFL team ratings, adjusted for home advantage, while Bedford (2004) applied an exponentially smoothed ratings model to predict the outcome of women’s handball World Cup matches between 2001 and 2003. Bedford and Clarke (2000) also investigated a ratings model for tennis players based on exponentially smoothing margins of victory. For this research, the exponential smoothing procedure retained the predictive property of adjustive ratings by offering a $t + 1$ forecast of performance $x_t$ using:

$$\hat{x}_{t+1} = x_t + \theta_t (\hat{x}_t - x_t)$$

(5.10)

where $\hat{x}_{t+1}$ is the forecast for player $i$ after observation $t$, and $\theta_t$ is the smoothing parameter, optimised to minimise the average error ($E$) after $n$ matches:

$$\min \left[ E = \frac{1}{n} \sum_{t=1}^{n} |x_t - \hat{x}_t| \right]$$

subject to: $0 \leq \theta_t \leq 2$

(5.11)

Using a Visual Basic loop, Equation (5.11) was executed after each player’s match ($t \geq 2$ for each $i$) and the resulting estimate, $\hat{x}_{t+1}$ from Equation (5.10), substituted as Exp into Equation (5.1). The Obs parameter was simply $x_t$ while $k$ was set at 1 for all $i$.

Table 5.3 compares the exponential ratings (ranked by “Rating”) to the APR model: the correlation with club champion voting is not as evident as with APR where the top five ranked players

---

19 $k$ could assume various forms but is outside the bounds of this research.
all received votes. With exponential ratings, the bottom three rated midfielders received votes, suggesting the approach is not recognising the performances of the better players. Moreover, it is unrealistic that Allen Christensen and Mitch Duncan are higher rated players than Jimmy Bartel, and that James Kelly was the lowest rated midfielder over the ratings period. An explanation for these flaws is that the exponential smoothing approach heavily and unfairly penalised good performers who were expected to score well but failed to do so in a match. Similarly, average performers who were expected to score in the lower ranges but scored highly in the impending match were erroneously over-credited. For example, in round 1, 2011, Joel Selwood was expected to achieve 118 points (Equation (5.10)) but only scored 27 points due to his removal from the match from injury. The resulting ratings decrease of 91.39 was too excessive given the circumstances. Conversely, Mitch Duncan, a lesser player to Selwood, was only expected to achieve 28 points by Equation (5.10) but achieved 109 points, earning a rating increase of 81, likewise deemed too excessive given Duncan’s capabilities. Chapter 8 of this thesis demonstrates the value of exponentially smoothed forecasts from smoothed player performance scores, \( f(X_t) \) where normality assumptions may be violated.

<table>
<thead>
<tr>
<th>Player</th>
<th>Matches</th>
<th>Rating</th>
<th>Rank</th>
<th>APR</th>
<th>Rank</th>
<th>Votes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joel Corey</td>
<td>30</td>
<td>753.88</td>
<td>1</td>
<td>648.27</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Joel Selwood</td>
<td>38</td>
<td>583.91</td>
<td>2</td>
<td>823.10</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Paul Chapman</td>
<td>20</td>
<td>568.29</td>
<td>3</td>
<td>608.37</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Mitch Duncan</td>
<td>18</td>
<td>524.34</td>
<td>4</td>
<td>499.75</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Allen Christensen</td>
<td>16</td>
<td>505.57</td>
<td>5</td>
<td>589.01</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Josh Cowan</td>
<td>3</td>
<td>490.00</td>
<td>6</td>
<td>493.52</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Gary Ablett</td>
<td>21</td>
<td>435.00</td>
<td>7</td>
<td>560.09</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Simon Hogan</td>
<td>12</td>
<td>389.26</td>
<td>8</td>
<td>465.92</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Cameron Ling</td>
<td>36</td>
<td>378.73</td>
<td>9</td>
<td>670.37</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Jimmy Bartel</td>
<td>41</td>
<td>329.96</td>
<td>10</td>
<td>678.95</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>James Kelly</td>
<td>21</td>
<td>221.63</td>
<td>11</td>
<td>623.51</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 5.3: Comparison of exponentially smoothed ratings and the APR model

5.3.4 Injury Compensation

As discussed in the previous section, in round 1, 2011, Joel Selwood was injured and was only able to achieve 27 points because he was removed from the game. This was significantly less than his 2010 season average of 106 points and represents an unfair deduction in his APR rating.
because he may have gone on to at least achieve his average\(^2\). This was a case of bad luck rather than lack of skill and should consequently be accounted for. We investigated a data transformation that recognised the inability of an AFL player to achieve at least what is expected of him due to injury. Using “time-on-ground” data, we were able to prove that a competitor who plays partial matches can have a final quantified performance estimate based on the actual performance achieved to the moment of his removal from the contest. In AFL, players are removed from the ground, primarily, to be rested but also due to injury or for disciplinary reasons. Very rarely does a player spend every minute of the match on the ground. James et al (2005) outlined a data transformation to account for time-on-ground shortcomings in their rugby performance measurements. A simple formula was adapted for AFL players:

\[
\Delta_i = x_i \left(\frac{125}{y}\right)
\]

(5.12)

where \(x_i\) is the performance score for player \(i\) and \(y\) = time on ground (minutes).

The nature of the transformation (see Figure 5.8) rewarded players who had performed well in a partial match, but left relatively unchanged poor performances in partial matches. Assuming 125 minutes in a game of AFL (four 30-minute quarters, plus time on), each player’s performance

\(^2\) Round 1 scores did not contribute to the ratings, but it is important to account for such a case in subsequent rounds.
measurement was expressed as a function of time spent on the ground. This provides an acceptable transformation. However, the nature of the function (grey curve in Figure 5.8) dictates that one performance point in one minute of playing time will be extended to 125 points for the match. This is an overly generous forecast and is improved by the use of:

\[
\Delta_x = x \left( \frac{125}{y} \left( \log_{10} \frac{125}{y} + 1 \right) \right)
\]

(5.13)

Given that taking the square root and/or logarithm of cases in an asymmetric data set are basic but, sometimes, effective near-normal transformation methods (Chinn, 1996), it is logical that this method is preferred to Equation (5.12). Furthermore, Equation (5.13) yields a far more realistic estimation of final performance in a partial match; although a one-minute-one-performance-point match is rare, a forecast score of 35 by Equation (5.13) is prudent. While this is a simple compensation for player injury, it was impossible to retrospectively determine, from our dataset, which players had sustained an injury during a match. To effectively account for player injury, a full list of injuries, and the period at which they occurred, would be required, which was not deemed a priority for this stage of the research.

5.4 Performance Consistency

This section details a simple player ranking system that is a function of a player’s performance consistency, that is, the ability to play well on a consistent basis. Why is the issue of performance consistency worthy of discussion? Of fund management, Marquardt (2008) observed that the more regularly managers beat their peers and their benchmarks, the greater the likelihood that skill, rather than luck, is driving their performance. The same concept can be applied to sporting performances. In fantasy sport leagues, there is much interest in selecting consistent performers. Waldman (2005) believes a fantasy team of players performing consistently at a desired level is more valuable than higher-scoring but more erratic players. We applied our consistency measures to AFL player performance, \( X_t \) (Equation (5.6)) as an alternative measure to the player ratings from Equation (5.1). Much interest surrounded comparing the highest-rated players in Section 5.4.2 and the most
consistent players in this section. How, if at all, did they differ? A key assumption for this section was
that the consistency measure was not necessarily identifying the “best” players in a league but, rather,
those who display the least variability in their game-to-game performances.

Two consistency measures, Equations (5.14) and (5.15), were considered to measure player
\( i \)'s, consistency:

\[
\pi_1 = \frac{\sigma}{\bar{x}}
\]

(5.14)

where \( \sigma = \) the standard deviation and \( \bar{x} = \) mean of player \( i \)'s \( X_t \). Manley (1988) employed a similar
approach to Equation (5.14) when measuring performance consistency in basketball, namely points
scored, while Elderton (1909) used a coefficient of variation (\( CV = 100 \times \) standard deviation/mean) to
measure a cricket batsman’s scoring consistency, where a coefficient closer to zero implied more
consistent performances. A measure of 0.000 implies that the player has been perfectly consistent for
the measurement duration, for example, Joel Selwood scoring exactly 100 every week (\( \sigma = 0 \)). A
limitation associated with this approach is players who have played a limited number of matches
appearing in the top ranks of consistency. Based on average performance, two to three games may not
be sufficient proof that a competitor is a consistently good performer. Equation (5.15) precludes
players with limited matches in a season by expressing performance variation as a function of the
number of games played in the season, \( n \):

\[
\pi_2 = \frac{\sigma}{n\bar{x}}
\]

(5.15)

Furthermore, \( \pi_2 \) attracts higher-averaging players into the top rankings.

Table 5.4 displays the results of the consistency measures applied to the Geelong midfielders
from 2010 to 2011, ranked in ascending order of \( \pi_2 \). Although flawed (note Josh Cowan ranked third
most consistent having only played three games), \( \pi_1 \) is the most correlated with the best and fairest
voting from Table 5.2 (\( \eta = 0.55 \)), while for \( \pi_2 \), \( \eta = 0.14 \). Nevertheless, the results of \( \pi_2 \) appear more
logical due to the lower ranking of Cowan.
### Table 5.4: Consistency measures for Geelong midfielders between 2010 and 2011

<table>
<thead>
<tr>
<th>Player</th>
<th>Mean</th>
<th>Std_Dev</th>
<th>Matches</th>
<th>$\pi_1$</th>
<th>Rank_1</th>
<th>$\pi_2$</th>
<th>Rank_2</th>
<th>Votes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jimmy Bartel</td>
<td>103.9</td>
<td>19.7</td>
<td>41</td>
<td>0.190</td>
<td>2</td>
<td>0.005</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Joel Selwood</td>
<td>107.9</td>
<td>26.2</td>
<td>38</td>
<td>0.243</td>
<td>7</td>
<td>0.006</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Cameron Ling</td>
<td>86.2</td>
<td>21.6</td>
<td>36</td>
<td>0.251</td>
<td>8</td>
<td>0.007</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Joel Corey</td>
<td>92.9</td>
<td>20.1</td>
<td>30</td>
<td>0.216</td>
<td>5</td>
<td>0.007</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Gary Ablett *</td>
<td>119.0</td>
<td>19.5</td>
<td>21</td>
<td>0.164</td>
<td>1</td>
<td>0.008</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Paul Chapman</td>
<td>102.2</td>
<td>19.5</td>
<td>20</td>
<td>0.191</td>
<td>4</td>
<td>0.010</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>James Kelly</td>
<td>96.8</td>
<td>22.9</td>
<td>21</td>
<td>0.236</td>
<td>6</td>
<td>0.011</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Mitchell Duncan</td>
<td>77.2</td>
<td>21.6</td>
<td>18</td>
<td>0.280</td>
<td>9</td>
<td>0.016</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Simon Hogan</td>
<td>54.7</td>
<td>19.6</td>
<td>12</td>
<td>0.358</td>
<td>10</td>
<td>0.030</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Allen Christensen</td>
<td>68.5</td>
<td>32.7</td>
<td>16</td>
<td>0.478</td>
<td>11</td>
<td>0.030</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Josh Cowan</td>
<td>45.0</td>
<td>8.5</td>
<td>3</td>
<td>0.190</td>
<td>3</td>
<td>0.063</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

5.5 Discussion

In this chapter, an Elo-influenced adjustive player rating system was shown to adequately describe and predict player performance in the AFL, with player form and opponent strength the key determinants. This was proven using the Geelong Football Club’s midfielders’ expected performances between 2010 and 2011 as an estimator of impending match performance and then validating the final ratings using the 2011 club champion voting results (a list of the top 100 midfielder’s final ratings at the completion of the 2011 season can be found in Section 11.2). While arguments exist that rating systems in team sports are too player-centric—ignoring the important concepts of teamwork and cooperation (Gould and Gatrell, 1979/80, Duch et al, 2010)—the adjustive rating model has proven to be an interesting tool, particularly for consumption by parties interested in a particular player’s upcoming performance, for example, punters and fantasy league coaches. Simulation of player versus opponent performance was a pragmatic methodology, particularly because the player and opponent score distributions were approximately normally distributed. It is expected that the predictive properties of this model will be improved by some important additions that have been designated as future research due to the volume of work still to be covered in this dissertation. Ground advantage, travel effects and continued research into injury compensation, briefly mentioned in Section 5.4.2, will become important branches of research. Clarke (2005) used linear regression to calculate individual home advantages for each AFL club, which was necessary due to the different burdens each club faces with travel between five states in Australia and the characteristics unique to each ground. Clarke’s paper revealed important venue effects, which have implications on player
performance within our research; for example, the Docklands venue in Melbourne is covered, thereby providing protection from the elements and a scoring advantage over teams playing in matches without cover. The difference in player sample sizes continues to be a limitation in the system—Table 5.2 revealed that Simon Hogan (12 matches) and Josh Cowan (3 matches) did not have the same opportunity to reach the same high rating levels as Bartel (41 matches) and Selwood (38 matches), but perhaps this is an indication that these players did not possess the skill level to participate at the same match frequency as the better players. These players would most likely have achieved more game time if they were playing for a less successful team, one which did not contain so many talented midfielders. Rating projections may offer a solution to the problem of opportunity. Table 5.2 also revealed that players who consistently achieve a high performance score may be slightly penalised by the rating system. Gary Ablett only fell short of $X = 100$ three times in 2010, averaging 120 performance points per match. By round 4, he had scored 129, 120 and 130 and was to play Port Adelaide’s weak midfield; that round, he yielded an $Exp$ of 0.984. He scored 133 in the match, but due to the extremely high expectation of him defeating a randomly chosen Port midfielder, his rating increase was a mere 1.05. Ablett reached 570.74 rating points by the last game of 2010, but due to his outstanding year, this was deemed to be short of a realistic value. A similar issue was encountered when using exponential smoothing to create $t + 1$ performance estimates. Such penalties on players who consistently score well are to be investigated: perhaps an additional weighting is necessary.

The starting point in adjustive ratings is an important consideration; ratings need to equilibrate over time and generally require a substantial sample of matches to converge to a value which is reflective of team or player ability. For the APR system, each player in the year 2010 commenced at 500 with final ratings recorded after the ultimate Geelong match in 2011. While not a serious concern within this research, given that the majority of players commenced the ratings trial simultaneously, it does pose problems to players or teams entering competition at widely different stages. The Brooklyn Nets, for example, entered the NBA at the start of the 2013 season, winning eleven of their first fifteen matches. An Elo-style formula would assign them the starting rating of 500 under the assumption that the team will start at a skill midpoint, which clearly was not the case. While
the initial assignation of 500 is a prudent approach, there is scope for an objective and/or subjective estimation of a team or player’s skill which could then be converted into an appropriate starting rating.

We have mentioned that the performance measurement at the heart of the APR system is the result of weighting and aggregating player performance indicators and, so, is perhaps too focused on the individual exploits of each player; important aspects that reflect teamwork should conceivably be the preferred approach (Duch et al, 2010). While it would have been a pleasure to improve the APR methodology, by addressing the shortcomings mentioned in the previous paragraph, player cooperation and contributions were important subjects that required attention. In Chapter 7, we pursue the effects of player interactions and develop network-based performance measurements to determine the relationship between teamwork and success. Prior to the investigation into network methodology, the APR system benefited from further statistical analysis into the classification of players into on-field positions. In Chapter 6, a discriminant process is described, that reclassified each player into more accurate positions across the measurement period, resulting in improved predictive capabilities for the APR model.
Chapter 6

Positional Analysis

6.1 Introduction

In the previous chapter, an adjustive AFL player rating model, APR, was discussed. This model accounted for player and opponent strength in the four player positions: defenders, forwards, midfielders and rucks. A player was pitted against a randomly selected opponent from the same position—for example, a midfielder versus an opposition midfielder or a forward versus an opposition forward—with the winner’s rating rising or falling depending on the “size” of the victory or defeat respectively, against that opponent player. In calculating the ratings, it was assumed that each player’s positional allocation was uniform across the entire season; this assumption is unrealistic. Prior to the match, players are selected to play in particular defensive, attacking or midfield positions on the ground. However, the frenetic nature of the game demands that a player may be required to operate in different zones and perform a variety of duties at different match stages. Before a match, notational analysts may assume that an established forward will play entirely at the attacking end of the ground, but the forward may play a temporary defensive role at the opposition’s scoring end, demanding an additional set of rules with which to measure his performance. Figure 5.3 revealed a forward is more likely to outscore a defender by way of Equation (5.6) due to the forward’s greater exposure to scoring opportunities (a goal carries the most weight). Consider the following example: a forward is freely scoring goals, but after being repositioned in the backline in the same match, realises reduced
contributions to the value of \( X_t \). Without match vision, how is it possible to realise, let alone measure, this change in position? This chapter demonstrates a retrospective position-classification method to aid notational analysis that is dependent on data which describes a player’s position on the ground.

### 6.2 Positional Classification

Discriminant analysis for classification purposes is common in team sports. Sampaio et al. (2006) employ discriminant analysis to maximise the average dissimilarities in game statistics between guards, forwards and centres in the National Basketball Association (NBA). Also with the aid of discriminant analysis, Fratzke (1976) was able to determine basketball player ability and position using varying biographic data, while Marelic et al. (2004) observed that “block” and “spike” in volleyball were the most important predictors of team success. Pyne et al. (2006) concluded that, at the AFL draft, fitness assessments involving statistical analysis on physical qualities such as height, mass and agility were useful in determining future player position. The benefits of retrospective classification for this research were enormous, especially given the size of the dataset: 22 rounds x 8 matches x 44 players = 7744 player cases per year\(^{21}\). The chosen algorithm, which maximised the Mahalanobis distance between the four positional centroids, could effectively locate erroneous positional allocation in a matter of seconds. Another appealing aspect of the research was that only a handful of simple game-related statistics were needed to retrospectively gauge a player’s positional movements; no prior knowledge or vision of an AFL match was required.

#### 6.2.1 Mahalanobis Distance

A key concept in this dissertation has been recognising a combination of game-related skills as an important determinant in classifying player team success (Hughes and Bartlett, 2002; Ibanez et al, 2009; Nevill et al, 2002) as well as in the measurement of individual performance in team sports (Bracewell, 2003; Koop, 2002; Sampaio et al, 2006). In Equation (5.6), to derive player ratings, we used the frequencies of certain AFL player performance indicators, each with subjectively allocated weights, as a base measure of performance. An important objective within this research was to

\(^{21}\) Additional teams were added in 2011 and 2012, demanding more matches in a round, however this research was conducted prior to these additions.
improve the accuracy of our APR model by scrutinising two key areas: the calculation of the performance value, \( X \), and more efficient allocation of player positions. Chapter 7 investigates a more rigorous player performance metric while this chapter addresses the latter proposal by retrospectively assigning each player to one of \( k \) game-related positions from a particular AFL match. The classification method outlined made it possible to assign each AFL player to the group vector \( k = [D,F,M,R] \) by linearly combining thirteen recognised AFL performance indicators: \( X \): Kick (KCK); Mark (MRK); Handball (HBL); Handball Receive (HBR); Inside 50 (I50); Rebound 50 (R50); Goal (GLS); Tackle (TKL); Clearance (CLE); Loose Ball Get (LBG); Hard Ball Get (HBG); Spoil (SPL); Hit-out (HIT)\(^{22} \). These indicators were selected from an input group, numbering 50, by a stepwise model to arrive at \((k - 1)\) discriminant functions for player \( i \):

\[
d_u = b_o + \sum_{m=1}^{13} b_m X_{mt}
\]

where \( d_u \) is the \( u \)th discriminant function for player \( i \), \( X_{mt} \) is the value of player \( i \)'s performance variable \( m \) after match \( t \), \( b_o \) is a constant, and \( b_m \) are discriminant coefficients (see Section 3.2 for more detailed methodology). These coefficients were selected in the first discriminant function to maximise the Mahalanobis distance between the four positional centroids in \( k \). The second discriminant function was selected so as to be orthogonal to the first, and the third discriminant function orthogonal to the second (Johnson and Wichern, 2007). Each player was assigned to the position which his Mahalanobis distance from the positional centroids was the smallest (Sampaio et al, 2006). The Mahalanobis distance is a measure of distance between two points in the space defined by two or more correlated variables and, in some sense, is a multidimensional z-score (James, 1985) measured by:

\[
D_m(X) = \sqrt{(X - \mu)^T S^{-1} (X - \mu)}
\]

where \( X = (X_1, \ldots, X_{13}) \), \( \mu = (\mu_1, \ldots, \mu_{13}) \) and \( S \) is the common covariance matrix (see Equation (6.5)).

Mahalanobis distance measurements were used effectively by Chatterjee and Yilmaz (1999) to observe differences in the performance characteristics of MVP players in the NBA. The classification

\(^{22}\) Each indicator is defined in the Section 11.1.
judgements in this research were supported on the values of the overall structure coefficients, \( b_{im} \) from Equation (6.1); higher values were better contributors to the classification process (Sampaio et al, 2006). Table 6.1 displays the overall model classification coefficients by position for all teams in the 2009 season, with (*) indicating the strongest discriminatory predictors for each position. Table 6.1 clarifies that SPL is the primary classifier for defenders (spoiling prevents an opponent from marking the ball and kicking a goal), GLS for forwards (self-explanatory), TKL for midfielders (tackling prevents or inhibits an opponent’s ability to retain possession of the ball) and HIT for ruckmen (self-explanatory). Mark (MRK) was not a significant predictor in the classification process and was removed from the model.

![Table 6.1: Classification function coefficients by position](image)

Figure 6.1 displays how the first \( (d_1) \) and second \( (d_2) \) discriminant functions from our model have classified a random sample of 200 player matches from round 22, 2009 into the four positions. Higher values for \( d_1 \) were associated with classification of the ruck position—it was safe to conclude the ruck position was the most accurately classified based on its unique role in a match—while higher \( d_2 \) values classified the midfield positions. Defenders were best classified by negative values of \( d_2 \), while forwards fell around the origin. Figure 6.1 proved that the classification algorithm recognised the uniqueness of each player position. A manual search through the reclassified cases also supported the use of the model.
A brief case study: Tom Hawkins is an established forward for the Geelong Football Club and was assigned to this position by the AFL fantasy league for all of 2010. In round 22, Hawkins kicked two goals and was a forward target (TAR) for his teammates on five occasions; he was correctly classified as a forward by the AFL fantasy league and by Equation (6.1) for this round. In round 20, however, Hawkins was forced to play a large part of the match as a ruckman because Mark Blake, one of Geelong’s established ruckmen, was unable to play. Hawkins was a suitable stand-in due to his height advantage. For the round 20 match, Equation (6.1) identified Hawkins achieving 16 hitouts (HIT), the ruckman’s primary skill, and zero goals, the forward’s primary skill, and classified him as a ruckman accordingly. The AFL fantasy league incorrectly assumed Hawkins would play as a forward in this match. An even more apparent misclassification occurred in 2010 with Geelong player Tom Lonergan, who played mostly forward in 2009 and so was assumed to play forward again in 2010. The classification algorithm determined that he played the entire 2010 season in defence. Table 6.2 shows the pre- (“Pos”) and post- (“Pos_1”) season classifications on a match-by-match basis, as well
as the predictors from Table 6.1, GLS and SPL, that classify forwards and defenders respectively with the most significance. Equation (6.1) was clearly able to identify that Lonergan kicked zero goals in 2010 and achieved multiple spoils in most matches. Furthermore, he achieved twice as many R50—sending the ball out of the opponent’s forward line (a defender’s job)—than I50—sending the ball into his own team’s forward line (mostly executed by midfielders and forwards). Should a notational analyst be performing retrospective analyses on Geelong’s forwards for 2010 with the aid of the pre-season AFL-supplied positional data, Lonergan would have been judged a poorly performing forward given his zero match frequencies in a forward’s key indicators. An analysis of Geelong’s defenders using the reclassified data shows Lonergan to be a highly active defender, in some cases achieving up to seven and eight spoils a game, which is considered excellent.

<table>
<thead>
<tr>
<th>Year</th>
<th>Round</th>
<th>Pos</th>
<th>Pos 1</th>
<th>GLS</th>
<th>SPL</th>
<th>I50</th>
<th>R50</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>1</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>2010</td>
<td>2</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>2010</td>
<td>3</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2010</td>
<td>4</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2010</td>
<td>5</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2010</td>
<td>6</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2010</td>
<td>7</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2010</td>
<td>8</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>2010</td>
<td>9</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2010</td>
<td>10</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2010</td>
<td>11</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>2010</td>
<td>13</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2010</td>
<td>14</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2010</td>
<td>15</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2010</td>
<td>16</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>2010</td>
<td>17</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2010</td>
<td>18</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>2010</td>
<td>19</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2010</td>
<td>22</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2010</td>
<td>24</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2010</td>
<td>25</td>
<td>F</td>
<td>D</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 6.2: Tom Lonergan’s pre and post 2010 season classification

With confidence in the classification model, we decided to revisit the APR model from Chapter 5 to investigate the effect of retrospective classification of players in their observed match position(s), specifically, whether the predictive capabilities of the model improved. The sample of Geelong midfielders remained relatively unchanged after classification; however, a substantial change occurred in the total opponent midfield sample. Through 2010 and 2011, 5820 cases existed of AFL-
classified midfielders, excluding those at Geelong. Classification with Equation (6.1) yielded 7715 midfield cases, suggesting several AFL-classified defenders and forwards played midfield roles throughout the period. The same simulations were run—as with the APR in Chapter 5—on each Geelong midfielder and the opponent midfielders from the respective rounds and new Exp values generated. The relationship between Exp and \( X_t \) slightly improved \((R^2 = 0.390)\) from that of the AFL-classified positions \((R^2 = 0.341)\)—a satisfactory outcome while leaving an appetite for fine-tuning of the classification process. The classification procedure can be expanded to include a player competing in multiple positions in one match, rather than a generic classification for the whole match, through the use of probability theory.

### 6.2.2 Posterior Probabilities

Where much classification research is content to draw conclusions from the predictive path to the classified groups, a measure of classification assurance is an interesting and important extension (James, 1985). For the purposes of AFL player performance measurement, a player may be correctly assigned to a position, but how much confidence are we able to place in the result? Is it realistic to assume he played the entire match in that position? To mathematically ascertain this knowledge, a Bayesian approach was investigated where posterior probabilities were calculated for each player’s defensive, forward, midfield and ruck roles in each match. Using these probabilities it was possible to establish a player’s “time spent” in each of the four positions in each of his matches; for example, a forward who briefly played as a defender in a match may be classified as a forward with 80% probability and a defender with 20% probability, with respect to his recorded performance indicators. Once the discriminant functions derived from Equation (6.1) had been calculated, the posterior probabilities were produced by:

\[
P(\text{Pos}_k | x) = \frac{\exp[\max(d_x) - d_k]/P(\text{Pos}_k)}{\sum_{j\neq k} \exp[\max(d_x) - d_j]/P(\text{Pos}_j)}
\]  

(6.3)
where: \( \sum_{k=1}^{4} P(Pos_i | x) = 1 \) and \( P(Pos_k) \) is the prior (pre-match) probability of player \( i \) assigned to position \( k \), determined simply as the proportion of each AFL-classified position’s membership within the complete data set (see Table 6.3).

<table>
<thead>
<tr>
<th>( k )</th>
<th>( n )</th>
<th>( P(Pos_k) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>2170</td>
<td>.280</td>
</tr>
<tr>
<td>F</td>
<td>1579</td>
<td>.204</td>
</tr>
<tr>
<td>M</td>
<td>3394</td>
<td>.438</td>
</tr>
<tr>
<td>R</td>
<td>601</td>
<td>.078</td>
</tr>
<tr>
<td>Total</td>
<td>7744</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 6.3: Prior probabilities of AFL-classified positions

Returning to the Hawkins example from earlier in this chapter, he was classified as a forward in round 22 with the following probabilities: \( P(Pos_D) = 0.019; P(Pos_F) = 0.851; P(Pos_M) = 0.130; P(Pos_R) = 0.000 \). While Hawkins was a dominant forward in the match, he still exhibited midfielder skills \( (P(Pos_M) = 0.130) \). In the round 20 match, he was classified as a ruckman with the following probabilities: \( P(Pos_D) = 0.000; P(Pos_F) = 0.051; P(Pos_M) = 0.060; P(Pos_R) = 0.889 \), again exhibiting, albeit lower, levels of midfield and forward skills.

The research established interesting relationships between the probabilities of classification to positions D, F and M and the number of Disposals (KCK + HBL) the player achieved in the defensive, forward and midfield zones respectively. In Table 6.4, Jonathan Brown of the Brisbane Lions was abundantly classified as a forward for the first five rounds of the 2009 season, but in a smaller capacity for round 4 when he played a greater midfield role than in the other matches \( [P(M | x) = 38.12\%] \). The majority of his disposals were in the forward zone for higher \( P(F | x) \), but for round 4, his Forward to Midfield disposal ratio (For:Mid) decreased below 1.0 in line with his decrease in \( P(F | x) \) and increase in \( P(M | x) \). Conversely, Brown’s most prominent forward performance \( [P(F | x) = 99.31\%] \) in round 3 returned his highest For:Mid ratio for the five rounds. This data augments our discriminant model, adding richer meaning to the classifications.
A logical step in the development of the APR system is to adjust each player’s match score, \( X_t \), for the “time” each player spent in the position of interest, as determined by the player’s posterior probabilities for match \( t \). This adjusted metric may provide a more accurate insight into the different roles performed by each player and improve a player’s performance prediction \( (\text{Exp}) \) for match \( t + 1 \).

This project is outside the current research parameters, but will be pursued at a later stage.

### 6.2.3 Intra-position Analysis

In rugby, James et al (2005) used medians with confidence limits in developing performance profiles which were dependent on player position. He extended his research to include intra-positional profiles; a “Prop”, “Hooker” and “Lock”, all forward positions, each displayed differing frequencies of selected indicators, for example, successful and unsuccessful tackles. By examining the squared Euclidian distance between 2 sets of covariates \([\text{MRK, GLS}], [\text{HBR, GLS}]\) in a two-dimensional space (Gordon, 1981), we were able to investigate intra-position differences between AFL skills, further demonstrating the differing roles assumed in the four positional groups. The complete analysis was carried out on every player’s \( t \times X_m \) performance covariance matrices (PCM) from the 2009 season, where \( X_m \) is performance variable \( m = 1, \ldots, 13 \) for match \( t \). A case study is provided using the covariance matrices of players allocated to the forward group. The information drawn from this research becomes important for coaching staff and pundits alike because immediate post-hoc deductions can be made, not only about a player’s influence on the match but also about a player’s influence within each position. Because performance variable weights can be adjusted to reflect the relative influence of covariates in the position in which a player is classified, player rating systems become more accurate.
With the discriminant model accurately classifying each player into position by his accumulation of performance variables, the research considered a pair of performance variables \( \{X_i, X_j\} \) for player \( i \) to match \( m \), with the covariance \( \text{Cov}(X_i, X_j) \) as a measure of the linear coupling between \( X_i \) and \( X_j \) (James, 1985). If entries in the column vector:

\[
X = \begin{bmatrix} X_1 \\ \vdots \\ X_m \end{bmatrix}
\]  

(6.4)

are random variables, each with finite variance, then the covariance matrix \( \Sigma \) is the matrix whose \((i,j)\) entry is the covariance:

\[
\Sigma_{ij} = \text{cov}(X_i, X_j) = E[(X_i - \mu_{ikm})(X_j - \mu_{jkm})]
\]  

(6.5)

where \( \mu_{ikm} = E(X_i) \) in position \( k \) in match \( t \), and \( \mu_{jkm} = E(X_j) \) in position \( k \) in match \( t \) are the expected values of the \( i^{th} \) and \( j^{th} \) entry in Equation (6.4). Incorporating the match mean vector in Equation (6.5) for variable \( X_o \) rather than the league mean at round \( n \), standardises the distances from the performance variable mean vectors for matches that may exhibit unusually high or low variable means, for example, wet weather having a negative impact on total disposals. In performance measurement, Chatterjee and Yilmaz (1999) favour the use of covariance matrices, rather than correlation matrices, because they express variability in the performance variables’ commonly used scales. With Equation (6.5), covariance matrices were established at a league, position (from Equation (6.1)) and player level, allowing analysis of matrix elements, for example, the covariance between Kicks and Goals \([\text{KCK, GLS}]\). Table 6.5 displays Brisbane Lions key forward, Jonathan Brown’s performance covariance matrix from 2009. Notice the high covariance, \([\text{MRK, KCK}] = 19.22, [\text{KCK, GLS}] = 9.38\) and \([\text{MRK, GLS}] = 7.13\), which is logical given his role as primary goal kicker at the club. Figure 6.2 illustrates how it is possible to compare sets of covariate couplets such as \([\text{KCKGLS, HBRGLS}]\) and, hence, allowing analysis of four covariates in a two-dimensional space (Gordon, 1981). Furthermore, by assessing the Squared Euclidean distances between these covariate couplets, the positions classified by Equation (6.1) could be further segmented to enhance the knowledge of intra-position performance characteristics.
The Squared Euclidean distance formula is defined as:

\[ d_{ij} = \sum_{p=1}^{n} (x_{ip} - x_{jp})^2 \]  

where \( x_{ip} \) and \( x_{jp} \) denote the values taken by the \( p^{th} \) player on covariate couplet \( i \) and \( j \) respectively.

Table 6.5: Performance covariance matrix (PCM) for Jonathan Brown

<table>
<thead>
<tr>
<th></th>
<th>KCK</th>
<th>HBL</th>
<th>MRK</th>
<th>HBR</th>
<th>GLS</th>
<th>TKL</th>
<th>HIT</th>
<th>I50</th>
<th>R50</th>
<th>CLE</th>
<th>HBG</th>
<th>LBG</th>
<th>SPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>KCK</td>
<td>28.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBL</td>
<td>2.97</td>
<td>5.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRK</td>
<td>19.22</td>
<td>0.92</td>
<td>16.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBR</td>
<td>-2.54</td>
<td>1.54</td>
<td>-3.4</td>
<td>3.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLS</td>
<td>9.38</td>
<td>-0.07</td>
<td>7.13</td>
<td>-1.94</td>
<td>5.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TKL</td>
<td>-2.67</td>
<td>-0.07</td>
<td>-2.33</td>
<td>1.28</td>
<td>-1.69</td>
<td>1.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIT</td>
<td>1.15</td>
<td>0.58</td>
<td>0.15</td>
<td>0.39</td>
<td>-0.06</td>
<td>0.11</td>
<td>0.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I50</td>
<td>4.24</td>
<td>1.72</td>
<td>2.46</td>
<td>0.31</td>
<td>0.48</td>
<td>-0.14</td>
<td>0.58</td>
<td>2.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R50</td>
<td>-1.05</td>
<td>-0.21</td>
<td>-0.91</td>
<td>0.34</td>
<td>-0.59</td>
<td>0.24</td>
<td>-0.01</td>
<td>-0.13</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLE</td>
<td>1.41</td>
<td>0.79</td>
<td>0.24</td>
<td>0.33</td>
<td>0.43</td>
<td>0.08</td>
<td>0.00</td>
<td>0.38</td>
<td>0.04</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBG</td>
<td>1.76</td>
<td>0.28</td>
<td>0.79</td>
<td>-0.52</td>
<td>0.85</td>
<td>-0.19</td>
<td>0.23</td>
<td>0.36</td>
<td>-0.09</td>
<td>0.22</td>
<td>1.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LBG</td>
<td>5.04</td>
<td>1.93</td>
<td>2.27</td>
<td>0.00</td>
<td>0.72</td>
<td>0.17</td>
<td>0.38</td>
<td>0.69</td>
<td>-0.22</td>
<td>0.55</td>
<td>0.48</td>
<td>3.03</td>
<td></td>
</tr>
<tr>
<td>SPL</td>
<td>0.47</td>
<td>0.46</td>
<td>0.21</td>
<td>0.16</td>
<td>0.19</td>
<td>0.03</td>
<td>-0.06</td>
<td>0.03</td>
<td>0.03</td>
<td>-0.10</td>
<td>0.23</td>
<td>0.29</td>
<td></td>
</tr>
</tbody>
</table>

Figure 6.2: Reclassification of forwards using Squared Euclidean distance

Having used Equation (6.1) to classify all players after round 22, a dissimilarity matrix was defined using Equation (6.6) to determine robust classifiers within the forward position (Table 6.6).
The largest distance (*) was between MRKGLS and HBRGLS, implying forwards could be classified into two further groups: discrete play forwards who predominantly kick goals after taking a mark (MRKGLS) and continuous play forwards who set up or kick goals through handball receives (HBRGLS).

<table>
<thead>
<tr>
<th></th>
<th>HBRGLS</th>
<th>MRKGLS</th>
<th>LBGGLS</th>
<th>HBGGLS</th>
<th>TKLGLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBRGLS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRKGLS</td>
<td>335.74</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LBGGLS</td>
<td>108.53</td>
<td>257.77</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBGGLS</td>
<td>72.67</td>
<td>229.84</td>
<td>48.06</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TKLGLS</td>
<td>84.07</td>
<td>283.08</td>
<td>65.02</td>
<td>41.13</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 6.6: Dissimilarity matrix of Squared Euclidean distance between forwards’ cov($X_i, X_j$)

Figure 6.2 displays how all forwards are positioned after the final round in 2009, based on MRKGLS and HBRGLS. By maximising the Mahalanobis distance between the centroids for these covariate couplets using Equation (6.2), it was possible to segment the forwards into sub-positions to the left and right of the dotted line, where the left contains the discrete play forwards and right contains the continuous play forwards. A small cluster of players around [HBRGLS=0] and [MRKGLS=0] proved difficult to classify, implying little variability through the season. Recognised forwards, however, show the highest variability and largest distance from the couplet centroids. Jonathan Brown could be considered the best discrete play forward, based on his largest distance from the MRKGLS centroid. Akermanis could be considered the best continuous play forward resulting from his distance from the HBRGLS centroids. However, Brown and Akermanis are contrasting forwards given the large distance in Table 6.7, measured by Equation (5.6). This hypothesis is supported by the AFL community’s consensus that Brown is a highly rated key forward and that Akermanis is a highly rated small or roving forward. The right mixture of these two types of forwards in a team is important from a coaching perspective. Successful sides generally have two key forwards and at least two small forwards. This modelling can assist in the selection process.

As with the posterior probabilities, the APR system could be augmented to include intrapositional components. For example, instead of rating all forwards as one entity, key forwards and small forwards could be retrospectively indentified with the covariance and distance measures.
mentioned above and rated separately given their specialised roles in the forward line. Moreover, it is
common to hear midfielders described as “inside” and “outside” players. “Inside” suggests the player
is at the nucleus of a group of players trying to gain possession of the ball, usually the result of a tap
down from a ruckman to a teammate after an umpire restarts play.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Player</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12 St Riewoldt</td>
<td>2.933</td>
</tr>
<tr>
<td>2</td>
<td>16 Br Brown</td>
<td>0.000</td>
</tr>
<tr>
<td>3</td>
<td>12 Ri Richardson</td>
<td>4.897</td>
</tr>
<tr>
<td>4</td>
<td>25 Ca Fevola</td>
<td>7.708</td>
</tr>
<tr>
<td>5</td>
<td>23 St Koschitzke</td>
<td>10.814</td>
</tr>
<tr>
<td>6</td>
<td>16 Po Tredrea</td>
<td>14.671</td>
</tr>
<tr>
<td>44</td>
<td>21 Wb Akermanis</td>
<td>48.317</td>
</tr>
</tbody>
</table>

Table 6.7: Dissimilarity matrix of Squared Euclidean distance: Brown and Akermanis [MRKGLS, HBRGLS]

The midfielder is physically trying to overcome other midfielders to handball or knock the ball out of
the contested situation to the advantage of a teammate in the clear. Should the inside midfielder
achieve this, he is credited with a “clearance” of the ball from the ruck. It is common for “outside”
midfielders to receive the ball from the inside player and then run and kick the ball to a forward or
towards the goal posts. The outside players are less likely to have muddy knees at the end of the
match. We anticipate that midfielders will be neatly divided into inside and outside players by
substituting contested possessions (CP)—higher frequency for inside players—and uncontested
possessions (UCP)—higher frequency for outside players—into Equation (6.6).

6.3 Discussion

The development of a rigorous team sport player-rating system has been a modulated process.
An early obstacle was that additive performance equations tended to bias midfielders. The top 10
players, by average score, in the 2012 AFL fantasy competition were midfielders, suggesting the
calculated score is correlated with ball possession. Rather than introduce weights to rescale scoring
based on a player’s position, which dictates the player’s exposure to the ball, four separate rating
systems were developed—one to be applied to each positional group. Initially, the ratings were
generated for Geelong midfielders, where the opponent strength was an approximately normal
distribution of like-position opponent scores in a match. For the sake of convenience, positional
classifications from the AFL fantasy site were incorporated. After noticing these positional allocations were uniform across the season—and, hence, not always reflective of where players actually played—relevant performance indicators were sought to reclassify each player’s position on a match-by-match basis. Once a more accurate set of positional data was available, there was a higher degree of confidence that the player being rated was facing off against the correct pool of opponent players with respect to position. While some misclassification occurred in the data, for example, a defender without R50 and SPL statistics being classified as a midfielder, resultant misleading effects were negligible. Misclassification could be justified logically as a player not effectively performing his regular duties within that position so should not be compared with other defenders. Misclassification is also dependent on the notational analyst having knowledge that the player was indeed a defender for that match, which may be impossible to detect from the data alone.

Sample size pending, further research will focus on intra-positional classification or on employing distance techniques to reclassify, for example, midfielders into contested (inside) and uncontested (outside) possession players and noting any improvement in the predictive properties of the ratings. Any number of retrospective positional analyses can be undertaken in the knowledge that positional classification has been improved: for example, which defenders prefer contested situations (high frequency of SPL) to running the ball out of the defensive zone (high frequency of R50). Chapter 7 introduces network analysis which detects players who are more likely to give the ball (defenders and inside midfielders) than receive it (forwards and outside midfielders) and vice versa; positions are semi-established by the network algorithms focusing on the usual location and function of a player in the dynamic network flow. Correlations are evident between the positional classification from this chapter and the network position allocations, and will be discussed with the aid of network diagrams. As mentioned in earlier stages of this dissertation, the investigation into network analysis was necessary in response to an important hypothesis that greater degrees of teamwork improved the likelihood of team success in the AFL. To now, only individual performance metrics have been analysed with respect to player ratings, and while we have proven that these variables contribute
effectively to the development of adjustable player ratings, the potential for model improvement in response to teamwork metrics was too attractive to ignore.
Chapter 7

Link Plays and Player Interaction Simulation

7.1 Introduction

Chapter 5 investigated the questions that inspired this research, namely, the adaption of an Elo-influenced player ratings system to a continuous team sport. A parametric methodology adequately rated and predicted AFL player performance with respect to player position and opponent strength, however, the discussion flagged a potential shortcoming in the performance metric; the oversight of the key foundation of success in team sport, teamwork. This chapter is concerned with the measurement of on-field cooperation, or interaction, using transactional data, and how it is possible to simulate levels of player interaction within a match and observe the consequences on likelihood of victory. It will be proven statistically that cooperative metrics have a stronger relationship with team success than individual ones, evidence that they might be more informative inputs in the APR model than individual ones, such as from Equation (5.6). Understanding this cooperative methodology, however, requires a thorough review of how the players interact on the field and how these interactions, when aggregated, form link plays.

7.2 Link Plays

Ball movement in the AFL, like other invasion games, is the result of a series of discrete critical events—in previous chapters of this dissertation, we defined these as performance
indicators—executed by the individuals involved in the contest (Nevill et al, 2002). In modern sports media, player performance indicators are intensively collected and published online across an ever-increasing number of sports—prior to, during and after a match—in an attempt to describe the level of a player’s involvement in a match. In the AFL, there has been rapid growth in the consumption of game-related player data by coaches, pundits and the general public. The latter’s interest in this data is driven by fantasy football: fantasy coaches can receive relevant player statistics that update in real time, on portable devices, giving them an idea on how their team is performing against opposition teams. During a match, an AFL coach is supplied real-time feeds of performance indicator frequency data to remain informed about the performance of his and the opposition players. The information guides certain coaching decisions: for example, if the live data reveals that a certain opposition player is receiving too much of the ball (high KCK/MRK and/or HBL/HBR frequency), then the coach can instruct one of his players to “tag” the opposition player. Furthermore, the data will reveal players who aren’t impacting the game, who can subsequently be substituted off. Figure 7.1 is a screen shot of the type of software employed by coaches to gain a competitive edge over the opposition; each Geelong player’s performance indicator frequencies from the round 8 match against Collingwood in 2011 are displayed. The data communicates to the Geelong or Collingwood coach that Corey and Selwood are the main distributors of the ball for Geelong, with 33 disposals (14 KCK + 19 HBL) and 30 disposals (20 KCK + 10 HBL) respectively. A coach, therefore, can consider tactics to suppress the impact on the game of one or both of these players. Geelong won this match by 3 points, but the match was low scoring, with Bartel the only multiple goal kicker (2 GLS), a rarity for Geelong given their success in that year. The software also allows spatial and temporal analysis of the performance indicators, that is, where and when each on-field event was recorded. Teams are going to greater lengths than ever before to secure the latest technology to remain as best informed as possible about their progress.

23 Tagging is a popular strategy whereby a player legally impedes his direct opponent, making it difficult for him to receive and distribute the ball.
24 Software provided by Prowess Sports (http://www.pro-stats.com.au)
As detailed in earlier sections of this dissertation, it is common across different sports for player $i$’s indicators from past matches to be weighted and linearly combined, resulting in numerical performance appraisals from which player ratings can be derived. The indicator data from Figure 7.1, for example, was substituted into Equation 5.6 to generate the score distributions from which the APR model was ultimately derived in Chapter 5. A statistician would be quick to identify a number of more technically advanced approaches just by looking at Figure 7.1: for example, a season’s data could be entered into a stepwise logistic regression model to determine the main predictors of team success, at the player level, with associated weights. For the early stages of the research, however, we maintained a simple approach to the player performance equation, taking advantage of the fantasy league score data that had already been made publicly available by the AFL. This baseline approach ensured model improvement could be logically pursued at any developmental stage.
7.2.1 Transactional Data

A criticism of individual performance indicator methodologies is that they can be too player-centric, ignoring an important underlying concept that a team is supposed to be more than the sum of the individual players (Gould and Gatrell, 1979/80). We decided the research should shift beyond individual on-field performance exploits, such as those displayed in Figure 7.1 and used in the APR system, towards a measurement of each player’s performance within a dynamic system of team play. The major concern was that individual performance indicators lacked complete information about the effect of a particular player’s contribution to a match; for example, a player may have achieved an impressive 20 kicks in a match, but what if most were turnovers or kicks that put the ball out of bounds, both resulting in immediate opposition possession? His contribution would therefore be viewed as negative. Hughes and Bartlett (2002) noted that notational analysis of team and match-play sports focuses on the movements and behaviours of the individual players as well as the interactions between the players. Hence, our interest shifted from the analysis of performance indicators executed by the individuals to those executed between the individuals. We termed these cooperative match moments link plays, where each link play was any sequence or set of relations involving two or more cooperating players from team \( a \), where the ball’s movement effectively increased that team’s scoring likelihood. Such relations express the on-field actions of a player as linkages which run between the other players on the team (Scott, 2000) and are defined in this research as effective transactions, or two events \( (t_1, t_2) \) that occur between two players from team \( a \):

\[
T_q = \{t_1, t_2\} \quad (7.1)
\]

where \( T_q \) is transaction \( q \) so events \( t_1, t_2 \in T_q \) involving two players on team \( a \). Each \( T_q \) is termed a link node, with \( t_1 \) a send indicator and \( t_2 \) a receive indicator (see Section 4.2). Effective transactions are built from send-receive events such as movement of the ball by foot (kick by player \( i \) proceeded by catch or mark by player \( j \)) or a closed hand (handball by player \( i \) proceeded by handball received by player \( j \)). Transactions (Equation (7.1)) that occur consecutively for team \( a \) form link plays, expressed as:
where $L_r$ is link play $r$ for team $a$, such that $T_q \in L_r$ and $n \geq 2$. Each link play continues until one of these terminal points is reached:

i) a score is realised (play requires a restart)

ii) play is dead (for example, the ball going “out of bounds” or the end of a quarter)

iii) team $a$ relinquishes possession of the ball to team $b$ (turnover) so that $L_{rb}$ may commence.

The latter is the most frequent in the AFL. Reep and Benjamin (1968) suggested the position of the players involved in passing a soccer ball, including the opponent defenders, and the relative skills of the players involved in the passing sequence are two important factors that determine the likelihood of a successful pass and, hence, passing sequence. These factors also influence points i) to iii) above, the length and player membership of the link plays. The ability of players to not only play in their position but, also, to manoeuvre themselves into space, free of any opposition players, is crucial in maintaining possession and establishing scoring opportunities. This research hypothesised that the more a player is involved in link plays during a match, the greater his contribution to the team, except if a player’s turnover frequency reaches a certain level. How a player contributes to his team through transactions with other players will be conceptualised as this chapter progresses. Transactional data analysis is especially popular in invasion sports—Hughes and Franks (2005) analysed passages of play leading to goals in the 1990 FIFA World Cup, while Moura et al (2007) performed similar team play analysis on Brazilian soccer teams. Brillinger (2007) analysed one particular sequence of 25 successful passes that resulted in a goal by Argentina in the 2006 World Cup, and Nevill et al (2002) applied general linear methodology to measure the distribution of passes by France in the 2000 European Championships through the length and width of the pitch. Published research regarding the application of transactional data for statistical analysis in the AFL is virtually non-existent. The probable cause of this is the difficulty in obtaining AFL event data. O’Shaugnessy (2006) studied transactions in AFL matches in the 2004 and 2005 seasons to estimate scoreboard “equity” when possessing the ball at any location on the field. Link plays in our research were produced from event
data from each match, that is, time-stamped data representing every \( t_p \) in a match. Figure 7.2 gives the events that comprise a single passage of play in the round 8, 2011 match between Geelong and Collingwood. The events were recorded in a CSV file that was exported from the software in Figure 7.1. The Geelong passage starts 2 minutes and 10 seconds into the third quarter (event 1508)\(^{25}\), with Cameron Ling gathering the disputed ball (BG 7.1. Collingwood) and handballing (HBL) to Harry Taylor (HBR). The passage continues effectively through five more Geelong players before ending in a goal (event 1522, 02min 34sec). Each send and receive indicator is visible in the “Stat” column.

<table>
<thead>
<tr>
<th>Trans</th>
<th>Geelong</th>
<th>Qtr</th>
<th>Elapsed</th>
<th>Collingwood</th>
<th>Details</th>
<th>Stat</th>
<th>Player</th>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1506</td>
<td>3</td>
<td>2</td>
<td>02:07</td>
<td>Geelong</td>
<td>HBL</td>
<td>31</td>
<td>C Davies</td>
<td>HBL To Cont</td>
<td></td>
</tr>
<tr>
<td>1507</td>
<td>3</td>
<td>2</td>
<td>02:07</td>
<td>Geelong</td>
<td>KCK</td>
<td>36</td>
<td>D Swan</td>
<td>Kick Ground To Space</td>
<td></td>
</tr>
<tr>
<td>1508</td>
<td>3</td>
<td>2</td>
<td>02:10</td>
<td>Geelong</td>
<td>BG</td>
<td>45</td>
<td>C Ling</td>
<td>Ball Get UCont</td>
<td></td>
</tr>
<tr>
<td>1509</td>
<td>3</td>
<td>2</td>
<td>02:11</td>
<td>Geelong</td>
<td>HBL</td>
<td>45</td>
<td>C Ling</td>
<td>Ball Get Adv</td>
<td></td>
</tr>
<tr>
<td>1510</td>
<td>3</td>
<td>2</td>
<td>02:11</td>
<td>Geelong</td>
<td>HBR</td>
<td>7 H Taylor</td>
<td>HBL Recov</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1511</td>
<td>3</td>
<td>2</td>
<td>02:11</td>
<td>Geelong</td>
<td>HBL</td>
<td>7 H Taylor</td>
<td>HBL Recov</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1512</td>
<td>3</td>
<td>2</td>
<td>02:15</td>
<td>Geelong</td>
<td>HBR</td>
<td>4 A Mackie</td>
<td>HBL Recov</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1513</td>
<td>3</td>
<td>2</td>
<td>02:15</td>
<td>Geelong</td>
<td>KCK</td>
<td>4 A Mackie</td>
<td>Kick Short To Adv Rebound 50 Corridor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1514</td>
<td>3</td>
<td>2</td>
<td>02:16</td>
<td>Geelong</td>
<td>MRK</td>
<td>9 J Kelly</td>
<td>Mark Adv Ply On</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1515</td>
<td>3</td>
<td>2</td>
<td>02:20</td>
<td>Geelong</td>
<td>KCK</td>
<td>9 J Kelly</td>
<td>Kick Long To Adv Cross Cent Right</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1516</td>
<td>3</td>
<td>2</td>
<td>02:23</td>
<td>Geelong</td>
<td>MRK</td>
<td>20 S Johnson</td>
<td>Mark Adv Ply On</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1517</td>
<td>3</td>
<td>2</td>
<td>02:25</td>
<td>Geelong</td>
<td>KCK</td>
<td>20 S Johnson</td>
<td>Kick Long To Adv Inside 50 Right</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1519</td>
<td>3</td>
<td>2</td>
<td>02:29</td>
<td>Geelong</td>
<td>DM</td>
<td>10 D Mensel</td>
<td>Dropped Mark</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1520</td>
<td>3</td>
<td>2</td>
<td>02:30</td>
<td>Geelong</td>
<td>BG</td>
<td>5 T Varcoe</td>
<td>Ball Get UCont</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1521</td>
<td>3</td>
<td>2</td>
<td>02:35</td>
<td>Geelong</td>
<td>KCK</td>
<td>5 T Varcoe</td>
<td>Kick Short</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1522</td>
<td>3</td>
<td>2</td>
<td>02:34</td>
<td>Geelong</td>
<td>GLS</td>
<td>5 T Varcoe</td>
<td>Goal General 10-30m In Front Corridor Regulation</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

Figure 7.2: Transactional data from round 8, 2011

[Image of Production Menu]

Figure 7.3: LINK Production Menu

Link plays were generated by our Visual Basic program, LINK (see Figure 7.3 and Section 11.3). The primary production module loops through the thousands of events in any single match to

---

\(^{25}\) Most matches exceed 2500 single events.
identify and separate each effective link play for each side from moments of disputed ball possession, dead play and any scores resulting from the actions of a single player only; typically, these are goals or behinds from a free kick awarded within scoring range. Events are imported into \( LINK \) in a CSV file (see Figure 7.2), then each team’s link plays are generated from three routines that recognise events existing conjointly within the event space. The first routine \((l_1)\) recognises the primary link play events \((t_p)\) occurring between players on team \(a\):

\[
l_1 = \{\text{KCK}; \text{MRK}\};\{\text{HBL}; \text{HBR}\};\{\text{KCK}; \text{GLS}\};\{\text{KCK}; \text{BHS}\};\{\text{HIT}; \text{BG}\}
\]  

(7.3)

where event \(t_p\) is a kick (KCK) and \(t_{p+1}\) is a catch or mark (MRK), or \(t_p\) is a handball (HBL) and \(t_{p+1}\) is a handball receive (HBR), or \(t_p\) is a kick and \(t_{p+1}\) is a goal (GLS) or behind (BHS), or \(t_p\) is a hit-out (HIT) and \(t_{p+1}\) is a ball get (BG). The events comprising \(l_1\) are regarded as primary covariates because MRK, GLS and BHS cannot occur without KCK, HBR cannot occur without HBL and HIT is mostly followed by BG and, when carried out between two players of the same team, form the foundations of all link plays. Covariance between any of these indicators can be calculated using Equation (6.5). The source code establishes the start of link plays by identifying these indicator sets.

The secondary routine, \(l_2\) flags link play events that may occur closely before or after \(l_1\):

\[
l_2 = \{\text{BG}\};\{\text{KNK}\};\{\text{FF}\}
\]  

(7.4)

where \(t_p = l_1\) and \(t_{p-1}\) or \(t_{p+1}\) = ball get (BG); knock-on (KNK); free kick for (FF) (see Section 11.1). The code will work recursively to locate these indicators in proximity to \(l_1\) as these act as joiners between each \(l_1\). The third routine, \(l_3\), adds miscellaneous events to the link plays that occur within the link plays created by combinations of \(l_1\) and \(l_2\) but do not dispossess team \(a\) of the ball. Such variables include defensive measures like tackles (TKL), spoils (SPL) and smother (SMT) and are often executed by the team without possession (team \(b\)). Where these variables do affect a turnover and commence a team \(b\) link play, they are included at the beginning of that link play for team \(b\).
A typical AFL match for team $a$, $m$ can, therefore, be modelled as:

$$m_a = \sum_{r=1}^{n} \sum_{k=1}^{n} L_r I_k + \epsilon$$

(7.5)

where $L_r$ = link play $r$, $I_k$ = score from individual ($i$) performance $k$ and $\epsilon = (L_r \cap I_k)$ = a noise term representing disputed ball (no team has clean possession). The individual performance, $I_k$ is usually represented by a player receiving a free kick within scoring range, then kicking a goal or behind. This is not associated with a link play as only one player is involved in the transaction, but must be included in Equation (7.5) as a score is the result. The error term poses interesting questions for coaching staff: is the length and frequency of disputed ball moments indicative of skill levels or opposition pressure? What is the probability that our team will regain control of the ball? We were concerned primarily with the link plays, so have set the distribution of the error term (disputed ball analysis) aside for future research.

### 7.2.2 LINK Output

Figure 7.4 is a snapshot of the output file from LINK’s principal production module, and demonstrates the data that drives some of the practical features that will be described shortly. Events 1508 through 1522 (denoted as $t_{1508,...,1522}$) were recognised as one Geelong link play (see cells with “Geelong” in the “Link” column). The source code detected [HBL\textsubscript{Ling}, HBR\textsubscript{Taylor}] ($t_{1509,...,1510}$) as the origin, searched recursively for any preliminary Geelong indicators (BG\textsubscript{Ling}) ($t_{1508}$), then forward until a terminal point was detected, in this case, a Geelong goal (GLS\textsubscript{Varcoe}) ($t_{1522}$). The events leading up to the link were identified as noise ($\epsilon$) because no two teammates could execute an effective transaction, $T_q$. Additional variables are created by LINK: the link number sequence in relation to all other $L_r$ (“Link#”), the link membership or number of events comprising each link (“Order”), the field position of each event in the link (“Pos”), the period in which the link occurred (“Period”) and the time elapsed for the link (“Time”). From Figure 7.4, the Geelong link, $L_{90}$ was the 90th link in the match, comprising 14 events moving through four field positions in the first half of the third quarter and lasting 24 seconds.
A suite of tools was built in to LINK for the user’s specific analysis. The link map in Figure 7.5 is a visual program output, allowing the user to visualise a selected link in the match. Each $T_q$ or node on the map is labelled with the player responsible for the event, with each transaction in \{Receive, Send\} format as is the order of events for each player. A unique feature of link plays is that although a player may not be effective in the link, his inclusion is warranted provided there is an effective transaction prior to and proceeding his ineffective event. For example, in Figure 7.5, although Menzel dropped the mark kicked by Johnson, he is included in the link because Geelong
retained possession of the ball through Varcoe. If a Collingwood player had taken possession of the ball after Menzel’s drop, he would have been negatively represented as contributing to the termination of the link.

The map output is displayed graphically as a typical AFL field comprising a defensive \((D)\) and forward \((F)\) zone—each bound by an arc extending 50 metres from the goal centre—and a midfield \((M)\) zone, or the area between the forward and defensive zones. The midfield zone is further divided into forward \((M_f)\) and defensive \((M_d)\) sectors depending on whether the ball is forward of or behind the centre circle respectively. By executing a series of conditional string searches on the event data description at each stage, \(\text{LINK}\) subdivides these zones, by broad \(x\)-\(y\) coordinates, into left \((L)\), centre or corridor \((C)\) and right \((R)\) sections, generating the zone set:

\[
m = \{DL; DC; DR; M_dL; M_dC; M_dR; M_fL; M_fC; M_fR; FL; FC; FR\} \tag{7.6}
\]

where \(DL\) is defensive-left, \(M_dC\) is midfield-defensive side-centre, and so on.

An additional benefit provided by the data is that each event is time-stamped, allowing the user to observe the performance of the team and/or any player at different match stages. It is common for a team to perform at a similar level in certain quarters from match to match, so the ability for a coach to detect these trends was deemed important. We added code so that \(\text{LINK}\) could determine each quarter’s midpoint and extend the temporal analysis from quarters to eighths for greater analytical accuracy. In Figure 7.6, “Period = 3 1” implies the Geelong link occurred in the first half of the third quarter, or 5/8 of the way through the match. Moreover, what if teams, or players, were displaying performance trends at a certain match period(s) in a certain field position(s)? Another output from \(\text{LINK}\) is a simple descriptive tool that attempted to capture such features, a spatial-temporal matrix (see Figure 7.6). This was generated from the zone and period data where element \(a_{mn}\) describes the number of events achieved by team \(a\) in zone \(m\) at period \(n\), with the proportion of \(a_{mn}\) that resulted in a score (score efficiency) in shaded font. Figure 7.6 is a match by the Brisbane Lions in 2009 showing the team achieved 21 link events during period 3_2 in zone F_C with 76% score efficiency; the following period, 4_1, yielded 12 events in zone F_C with 100% score efficiency. These were productive periods in the match for Brisbane. An interesting development
might involve transferring the data from the matrix on to a “heat map”, where periods and locations of higher activity would be a deeper red than other elements.

<table>
<thead>
<tr>
<th>Zone m</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>2</th>
<th>2</th>
<th>2</th>
<th>2</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>4</th>
<th>4</th>
<th>4</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>D L</td>
<td>8</td>
<td>11</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>20</td>
<td>2</td>
<td>23</td>
<td>91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D C</td>
<td>20</td>
<td>8</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D R</td>
<td>10</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Md L</td>
<td>17</td>
<td>20</td>
<td>2</td>
<td>5</td>
<td>11</td>
<td>14</td>
<td>7</td>
<td>13</td>
<td>89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Md C</td>
<td>10</td>
<td>5</td>
<td>4</td>
<td>11</td>
<td>13</td>
<td>9</td>
<td>5</td>
<td>11</td>
<td>68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Md R</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>13</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mf L</td>
<td>18</td>
<td>20</td>
<td>20</td>
<td>4</td>
<td>42</td>
<td>6</td>
<td>3</td>
<td>6</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mf C</td>
<td>5</td>
<td>9</td>
<td>13</td>
<td>16</td>
<td>14</td>
<td>18</td>
<td>17</td>
<td>3</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mf R</td>
<td>0.06</td>
<td>0.25</td>
<td>0.00</td>
<td>1.00</td>
<td>0.33</td>
<td>0.00</td>
<td>0.33</td>
<td>0.50</td>
<td>0.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F L</td>
<td>0.67</td>
<td>0.50</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.50</td>
<td>0.40</td>
<td>0.60</td>
<td>0.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F C</td>
<td>0.00</td>
<td>0.20</td>
<td>1.00</td>
<td>0.00</td>
<td>0.75</td>
<td>0.76</td>
<td>1.00</td>
<td>1.00</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F R</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>12</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>116</td>
<td>111</td>
<td>76</td>
<td>88</td>
<td>113</td>
<td>108</td>
<td>80</td>
<td>94</td>
<td>786</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 7.6: Spatial-temporal matrix of Brisbane’s events (score proportions shaded)**

### 7.3 Player Interactions

An ever-present complication in the calculation of team sport player ratings is how and when to equitably assign credit for a player’s performance. Is it realistic to assume that “teamwork” can be truly quantified at an individual player level? Oliver (2004) developed a “difficulty theory” for distributing credit to players in basketball—the more difficult the contribution, the more credit it deserves—illustrating how expected points (scoring shots) change when players of varying skill level and position become involved in the chain of play. An AFL “difficulty theory” would certainly prove to be problematic in that there are 26 more players to account for on the AFL field than on the basketball court, and that scoring in AFL occurs far less frequently. This chapter is concerned with AFL player performance as contributions to a dynamic and complex team network, rather than as additions to the individual’s performance variable vector.
where each \( t \) is an on-field event, or indicator, with finite variance. The APR model was dependent on player equations (Equation 5.6). Duch et al (2010) argue that the real measure of player performance is “hidden” in the team plays and not derived from strictly individual events associated with player \( i \). Furthermore, in their research on football-passing patterns from EURO 2004, Lee et al (2005) measured passing between players at a group level rather than at an individual level, demonstrating how a player’s passing patterns determined his location in the team’s on-field “network”. Discussions about network analysis commonly refer to the use of relational data or the interactions that relate one agent (player) to another and, so, preclude the properties of the individual agents themselves (Scott, 2000). An initial hypothesis in this research was that the players’ acts of cooperation would determine the team’s network strength during a match, yielding a more appropriate rating measure than an “individual” rating such as from the APR model. This hypothesis will be proven later in this chapter by recursively comparing the individual and cooperative ratings’ match prediction ability for a particular team.

### 7.3.1 Player Interaction Matrix

Prior to arriving at a team’s network structure, it was necessary to reduce matches of interest to their transactional properties at the player level (Equation (7.1)). Gould and Gatrell (1979/80) analysed the structure of the 1977 Liverpool–Manchester soccer cup final by looking at the interaction between players on each team. By analysing passes between any two players, they developed an interaction matrix with ball ‘senders’ (rows) and ball ‘receivers’ (columns) from team \( a \), and a second matrix representing balls sent from team \( a \) and received by players from team \( b \). We termed such linking play between player \( i \) and player \( j \) on team \( a \) as effective; link plays involving a player from team \( a \) relinquishing possession to a player on team \( b \) were seen as ineffective. Moreover, three forms of interaction were recognised and programmed in \( \text{LINK} \) to isolate each within the interaction data:
i) primary interaction: the most efficient ball movement achieved through \( \{KCK; MRK\} \), \( \{HBL; HBR\} \) or \( \{HIT; BG\} \)

ii) secondary interaction: less efficient ball movement, namely, player \( j \) gaining possession of the ball by means other than a mark or handball receive (“Ball Get”) due to an inaccurate player \( i \) event; team \( a \) retains possession of the ball

iii) negative interaction: inefficient ball movement, where player \( i \) relinquishes possession of the ball to player \( k \) from team \( b \) (“Turnover”).

| Figure 7.7: Geelong’s interaction matrix from 2011’s grand final |

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>11 J Corey</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12 T West</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13 T Longman</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14 J Redwood</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>20 S Johnson</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>23 M Duncan</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>26 T Hawkins</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>27 M Stokes</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>28 A Christensen</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>39 M Scarlett</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>31 J Podsiadly</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>35 P Chapman</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>4 A Mackie</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>44 D Wijesuriya</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>44 C Enright</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>45 C Ling</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5 T Vanice</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>6 B Ottens</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>7 H Taylor</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8 J Hunt</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9 J Kelly</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Behind</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>4</td>
<td>2</td>
<td>20</td>
<td>16</td>
<td>7</td>
<td>14</td>
<td>8</td>
<td>14</td>
<td>20</td>
<td>5</td>
<td>2</td>
<td>18</td>
<td>12</td>
<td>8</td>
<td>7</td>
<td>15</td>
<td>14</td>
<td>4</td>
<td>17</td>
<td>8</td>
<td>11</td>
</tr>
</tbody>
</table>

The link plays for any given match were established by locating the primary, secondary and negative interactions, disregarding periods of noise (disputed ball). An effective interaction frequency between any pair of players \( [i, j] \) from team \( a \) in a match was represented by the discrete random variable \( r_{ij} \) forming each element in our interaction matrix (Geelong’s grand final match in Figure 7.7). Given the directional nature of the data within the link plays, the initial interaction matrices were asymmetric as there was interest in the send/receive ratios, or in and out degrees (Borgatti, 2005).
From Figure 7.7, a notable interaction from the match was [Scarlett, Chapman] = 4 and [Chapman, Scarlett] = 1. This is logical because Scarlett is a defender and Chapman a midfielder/forward; one would expect a good team’s defenders to be sending the ball from the back line more than receiving the ball in the defensive zone. This latter scenario implies the opponent is moving the ball into the defenders’ zone and the defenders aren’t repelling the attack; the opponent is setting up scoring opportunities more than the defenders are repelling them.

Having classified the Geelong grand final players into their position using Equation (6.1), Figure 7.8 validated the relationship between those positions and the send/receive averages in that match. Defenders were more likely to send, forwards were more likely to receive and there was a marginal receive advantage for midfielders, which is probably correlated with the high send frequency from ruckmen \{\text{HIT}_{i}; \text{BG}_{j}\}. This ruckman-midfielder relationship is evident in Figure 7.7, with [Ottens, Selwood] = 5 and [Selwood, Ottens] = 0. Ottens is a ruckman whose job it is, at the hit-outs, to tap the ball, bounced in the air by the umpire, down to a Geelong midfielder whose main objective is to clear the ball to construct a scoring opportunity. An opponent coach might be forced to react to the high interaction frequency of \{\text{HIT}_{\text{Ottens}}, \text{BG}_{\text{Selwood}}\} by instructing a player to (legally) impede the efforts of Selwood to gather the tap from Ottens.

![Figure 7.8: Send/Receive event averages by position for Geelong’s grand final match](image)

The interaction methodology outlined in this section so far is similar to “r-pass movement” in world football (Reep and Benjamin, 1968) in which a player from team \( a \) is responsible for a series of
r successful passes resulting in either a shot at goal by player r, an infringement or an intercepted pass (r + 1). Reep and Benjamin (1968) defined the probability of an r-pass movement, p, such that \( p_1 > p_2 \) \( > p_3 \ldots p_r \) \( > p_{r+1} \), that is, a decay in the likelihood that the sequence will extend beyond pass r. While the same logic is applicable to AFL link plays, this chapter is primarily concerned with the frequency of “passes” between AFL players, rather than with the membership and length of each link.

7.3.2 Network Diagrams

The link maps produced by the LINK program (Figure 7.5) were useful for tracking ball movement and player membership of particular links, but they did not provide adequate information about the interacting players’ contribution to total team relations. A valuable network analysis tool is the network diagram which allows a graphical comprehension of the interaction between agents and the prominence of each player in a network. Duch et al (2010) produced network diagrams to determine the most important players in matches from the 2008 European Cup—these players were centrally located in the network diagram and possessed high passing accuracy in the match. Figure 7.9 illustrates Geelong’s 2011 grand final match and is an abundant source of information for decisions in the coach’s box. The diagram was generated by Ucinet network software (www.analytictech.com/ucinet) having imported Geelong’s grand final directed interaction matrix (Figure 7.7). Each circular node in the network diagram represents a player, with each one colour-coded depending on the position of each player in the match (Defender, Forward, Midfielder and Ruckman). Positions were retrospectively determined by Equation (6.1) and coloured accordingly. However, the algorithm driving the network diagram has automatically grouped the players in positional clusters, dependent on interactions with “neighbouring” players, that neatly reflect the positions in which they were classified earlier; a visual correlation is evident between the classification results outlined in Chapter 6 and the network positioning in this chapter. The network algorithm grouped Defenders at the left of the diagram, midfielders in the middle and forwards to the right. This is logical given that, traditionally, midfielders are the bridge between the defenders and the forwards. From this diagram, it could be concluded that Joel Corey, classified as a midfielder by Equation (6.1), interacted more with the defenders than, say, James Kelly, given his proximity to the
classified defenders at the left. The goal and behind nodes will always be at the far right as they only
“receive” (link plays are terminated after either of these events). The lines joining the nodes represent
the interactions between the players; arrows point to the receiver in the interaction. The density of the
lines is proportional to the number of interactions between any two players. The diameter of each
node is representative of a player’s prominence in the network (see eigenvector centrality in Section
7.4), where wider diameters are reserved for more central, influential players in the network.

Figure 7.9: Network diagram of Geelong’s 2011 grand final match

The first observation on the network is that it is visually compact, which is common with the
victorious teams. We deduced, in simple terms, that the greater number of central players a team has
in its network, the greater its chances of victory. Apart from James Podsiadly, who was injured early
and substituted out of the match, all players were relatively active in the network. It is possible to
conceptualise a central square in the network, which holds the most central players in the match, with
Corey, Christensen, Johnson and Mackie at the corners. The three players in the centre of the square
are Bartel, Ling and Selwood, with Bartel, seemingly, the nucleus in the network. This is appropriate
given that he was awarded the best player on the ground. Note the heavy Bartel interactions: [Ottens,
Bartel], [Bartel, Duncan], [Bartel, Selwood] and [Bartel, Goal]. [Selwood, Varcoe], [Selwood,
Johnson], [Ottens, Ling], [Scarlett, Wojcinski] and [Scarlett, Mackie] are other frequent interactions. Interaction analysis can be extended to tri-nodal, with frequent interactions occurring between [Scarlett, Wojcinski, Chapman], [Ottens, Bartel, Selwood], [Ottens, Bartel, Duncan], [Selwood, Johnson, Goal] and [Selwood, Varcoe, Goal]. Although a coach would be well aware of the level of impact Bartel, Selwood and Ling can have on a match, the advantage of these diagrams is identifying from and to whom each of these players mostly receives and sends the ball; the detection of these “usual suspects” means coaches can employ tactics to minimise their presence in the link plays that constitute the network.

If we now compare Geelong’s grand final winning network diagram to a loss to Sydney in round 23 (Figure 7.10), there are a number of interesting distinctions. The centrality “square”, visible in the victorious match, has now widened and become, seemingly, non-existent. Of the three players who formed the nucleus of the network in the victory, only Ling is seen to play a central role in this match; Bartel and Selwood are peripheral players in this match, an indication of their value to Geelong’s chances of victory. A central sub-group does not exist for defenders as in the victory; Lonergan, Taylor, Hunt and Mackie appear to be more involved in “sending” to Scarlett and Enright, suggesting the ball is coming into Geelong’s defensive zone too often. Enright’s role is interesting from the diagram; he appears to be the principal link between defenders and the midfielders (notably, Selwood and Chapman). Furthermore, the lines connecting the forwards to the goals are not dense, meaning Geelong did not have multiple forward goal kickers. Finally, the [Ottens, Selwood] combination—so effective in the grand final—is not as prominent here, instead eclipsed by [Ottens, Chapman] and [Ottens, Kelly].

The network diagram algorithm is not publicly available, but efforts are being made, as a branch of this research, to reproduce a similar product for inclusion in the LINK program. The final product will also add a function enabling the user to remove players such as Bartel and observe the effect on the revised network. The next section establishes such an approach.
7.3.3 Interaction Simulation

The AFL network diagram was driven by a directed network (see Figure 7.7) as we were interested in a graphical account of any player’s send/receive ratio, mainly for use in the network diagram; for example, because he is mostly attempting to score, a forward would receive the ball from teammates more than he would send the ball (see Figure 7.8). An important development in the research was the ability to remove a player from the network, replace him with another, and note the change in the network structure. Simulating these player effects was the new direction; this necessitated an undirected network—that is, any and all relations between players regardless of the directional flow (Scott, 2000) (see Figure 7.11). The undirected network required each matrix to be symmetrised, using:

\[ r_{ij} = r_{ji} = A_{ij} + A_{ji}, \quad i, j = 1, \ldots, 22 \]  

(7.8)
Frequency distributions could then be calculated for each \([i, j]\) in each of Geelong’s 25 matches (22 regular season games and three finals matches). Geelong fielded 34 players throughout the season, so a total of \((34 \times (34-1))/2 = 561\) distributions were computed. In this calculation, the subtraction of 1 removed player \(i\)’s interaction with himself and the divisor of 2 halved the
distributions to be calculated because \( r_{ij} = r_{ji} \). Figure 7.12 displays the observed interaction, \( f(r) \), between Geelong’s Jimmy Bartel and Andrew Mackie for all 2011 season matches. This player pair was more likely to interact between one and six times in a match than not at all. The maximum number of interactions measured in the season between any pairing from the team was eight.

If the frequency of discrete events that occurs between two players within an AFL match remained constant over its course, the events could be described with a Poisson distribution (Nevil et al, 2002). However, interaction rates between any \([i, j]\) are stochastic, depending on factors such as the position of the two players, their skill levels and the defensive quality of the opposition. For this reason, the negative binomial distribution (\( nbd \)) was deemed more appropriate than Poisson. Pollard et al (1977) determined that performances of individual players do not give close fits to the \( nbd \), adding that the fit improves as more players become involved; they provided an example of an improved fit for batting partnerships in cricket, rather than from individual batsman scores. This research also seemed to suit the \( nbd \) approach due to multiple, cooperating players. From the negative binomial distribution, the probability of \( r \) interactions for each \([i, j]\) is:

\[
P(r) = \binom{k + r - 1}{k - 1} p^k q^r, \quad r = 0, \ldots, 8
\]

where \( k > 0, 0 < p < 1 \) and \( q = 1 - p \). The parameters \( k \) (the threshold number of successes) and \( p \) (the probability of a success) were estimated so as to minimise the Pearson’s chi-squared statistic, \( \chi^2 \), for each \([i, j]\), by using the observed \((O)\) and expected \((E)\) probabilities derived from Equation (7.9), or:

\[
\min \chi^2 = \sum_{r=0}^{8} \frac{(O - E)^2}{E}
\]

where \( r \) is the number of failures (interactions). Fitting \( nbd \) to various sports, Pollard et al (1977) estimated \( k \) and \( p \) by a method of moments, so:

\[
k = m^2/(s^2 - m), \quad p = m/s^2
\]

where \( m \) is the sample mean and \( s^2 \) is the sample variance. We concluded that Equations (7.9) and (7.11) more adequately fitted the interaction data, providing lower \( \chi^2 \) values for the majority of
Geelong’s \([i,j]\). The [Bartel, Mackey] example is displayed in Table 7.1, where \(k\) and \(p\) in each \(P(r)_1\) were estimated using Equations (7.9) and (7.10) and each \(P(r)_2\) using Equations (7.11). The \(\chi^2\) is noticeably lower for the approach in this dissertation.

\[
\begin{array}{|c|c|c|c|}
\hline
r & f(r) & P(r)_1 & P(r)_2 \\
\hline
0 & 0.3333 & 0.3333 & 0.2897 \\
1 & 0.1905 & 0.2222 & 0.2675 \\
2 & 0.1429 & 0.1481 & 0.1853 \\
3 & 0.1429 & 0.0988 & 0.1141 \\
4 & 0.0952 & 0.0658 & 0.0659 \\
5 & 0.0476 & 0.0439 & 0.0365 \\
6 & 0.0476 & 0.0293 & 0.0197 \\
7 & 0 & 0.0195 & 0.0104 \\
8 & 0 & 0.013 & 0.0054 \\
\hline
\end{array}
\]

Table 7.1: Probabilities and \(\chi^2\) values for [Bartel, Mackey]

A Visual Basic module was written to fit the optimised \(nbd\) to all combinations of players in the Geelong club and to simulate the players’ interactions for any chosen team list in the 22 x 22 team matrix. The initial routine produced a random probability, \(u \sim U(0,1)\), for each \([i,j]\) in the match, with \(r_{ij}\) determined by the cumulative distribution function:

\[
F(r) = P(R \leq r)
\]

where \(R\) represents the cumulative probability. For example, a randomly generated probability of \(u = 0.3000\) would produce \(r_{[\text{Bartel, Mackie}]} = 0\) as \(u < P(R \leq 1) = 0.0000 + 0.3333\) (see Table 7.1). For each simulation, all \((22 \times (22 - 1))/2 = 231\) elements of the interaction matrix assumed a value for \(r\) as determined by \(u\) and Equation (7.12), enabling the eventual calculation of cooperative player ratings from the simulated matrix. These player ratings needed to be derived from a player performance metric which forms the next section of this chapter.

### 7.4 Eigenvector Centrality

After reading the rich resources on network analysis in various sports (Duch et al, 2010; Gould and Gatrell, 1979/80; Lee et al, 2005; Lusher et al, 2010), a network algorithm was introduced for player performance purposes to better understand the causality of player \(i\)’s performance with respect to that of his teammates. Centrality is one of the most widely studied concepts in network
analysis and allows implicit assumptions about the prominence of an individual in a network (Lusher et al, 2010; Borgatti, 2005). A specific type, eigenvector centrality, was trialled as a valid player performance model, under the assumption that the higher a player’s centrality in the Geelong network, the greater his interaction with other central players; that is to say, the more interactions a player has with a highly central player, the higher his centrality will be (Borgatti, 2005). The eigenvector centrality rating, \( e \) for player \( i \), was measured using:

\[
e_i = \frac{1}{\lambda} \sum_j r_{ij} x_j
\]  

(7.13)

expressed in matrix form as: \( Ax = \lambda x \), where \( x \) is the corresponding eigenvector from our interaction matrix, \( A \), and the eigenvalue, \( \lambda \), was solved using an automated power method: following \( n \) multiplications of \( A \) and \( x \), the point at which \( \lambda_{n-1} \) and \( \lambda_n \) converged prompted calculation of the ratings (Equation (7.13)) for all players within the actual or simulated interaction matrix (see Section 11.3 for code sample). Figure 7.13 offers the eigenvector centrality metrics for each Geelong player in the 2011 grand final match. It is clear, with respect to this metric, that Selwood, Bartel, Christensen, Varcoe, Chapman and Ottens were the most prominent players in the network. This was a very encouraging result, given that Bartel (13 votes) and Selwood (9 votes) were voted the two best players on the ground by the voting panel (see Section 4.2). For this Geelong example, at least, we can conclude that eigenvector centrality is well correlated with player performance, so is a compelling alternative to Equation (5.6). Recall in the network diagram that Ling, visually, appeared to be a central player, but he is ranked eighth by eigenvector centrality in Figure 7.13. This can be explained by Ling’s lower frequency interaction with the other key players—only one interaction each with both Bartel and Selwood—keeping in mind that the definition of eigenvector centrality states a player’s interaction with other central players has a more positive outcome on his own centrality. Each Geelong player’s prominence in the grand final network can be observed visually in Figure 7.9; the diameter of each node is proportional to that player’s eigenvector centrality.
Figure 7.13: Eigenvector centrality for Geelong’s 2011 grand final team

The process for each simulation can hence be summarised by the following routines in $LINK$:

1. $r_{ij}$ is generated for each player pairing in the match from the parent probability matrix (using Equation (7.9) and Equation (7.10))

2. $\lambda$ is calculated by the power iteration method

3. $\lambda$ and $r_{ij}$ are substituted into Equation (7.13) to arrive at $e$ for each player

4. $r_i$ and $e_i$, ($i = 1,..,22$) are recorded after each simulation with statistics calculated at $n = 100$.

7.4.1 Team Strength

The simulated network and corresponding ratings detailed in this research provided a pragmatic framework for estimating player $i$’s utility within a selected side. An important step in this procedure was calculating team $\alpha$’s network “strength”, $\pi$, after each match, by:

$$\pi_\alpha = \frac{1}{n} \sum_{i=1}^{n} e_i, \ r = 1,\ldots,8$$  (7.14)

for $n = 22$ players where each $e$ is derived from Equation (7.13). We compared Geelong’s 25 network indices—one for each match measured—from Equation (7.14) with each match’s final score margin,
and discovered a linear regression line effectively approximates the margin \( (R^2 = 0.5302) \) (see Figure 7.14). In practical terms, a team increases its likelihood of winning if more players force themselves to be central in the match network. This is analogous to the finding that soccer teams, skilful enough to retain possession for longer periods than their opposition, have a greater chance of scoring (Hughes and Franks, 2005). It is important to note that retrospective attempts at gauging team strength should also account for opponent strength. In the case that a coach rests important players against weaker sides, the relationship between team strength and margin of victory is jeopardised. This research did not consider Geelong’s opponents throughout the season because a stronger team commonly maintains a high skill level even with a rotation of particular players (see Figure 7.14). Continued network research will require a parameter for opponent skill, particularly for sides whose player depth is unable to cover the absence of stronger players.

To further validate the centrality ratings, an “individual” rating equation, \( Y_i \), was developed, ignoring network methodology and focusing solely on player \( i \)’s post-match performance indicator totals—the same six indicators (\( m \)) as in the primary interaction data: \([KCK, MRK], [HBL, HBR]\) and \([HIT, BG]\). The equation was of the form:

\[
Y_i = b_o + \sum_{m=1}^{6} b_m X_m
\]

(7.15)

where \( X_m \) is the frequency of performance indicator \( m \) for player \( i \), \( b_m \) are weights and \( b_o \) is the intercept. The weights were optimised to maximise the linear relationship between the mean ratings and final score margin in each Geelong match, a similar approach to Equation (5.7) but without the bias of the scoring variables (GLS and BHS). Substituting \( Y \) for \( e \) in Equation (7.14) produced a comparable measure of team strength for the individual ratings. Figure 7.15 confirms team strength was not as accomplished at predicting score margin when each player was assessed individually by Equation (7.15) \( (R^2=0.3953) \) rather than as an agent within a team’s network by Equation (7.13) \( (R^2 = 0.5302) \). This was a significant finding in the research, proving an original hypothesis that cooperative rather than individual performance might be a better indicator of a team’s likelihood of victory. This finding offered huge encouragement for future development and implementation of the LINK
application. It also provided confidence for the final stages of the network research—observing individual player effects by simulation.

![Figure 7.14: Relationship between Geelong’s mean network rating and final score margin](image)

![Figure 7.15: Relationship between Geelong’s mean individual rating and final score margin](image)

### 7.5 Model Validation and Results

Before investigating player effects within the network, we performed a preliminary examination on our simulator, testing the hypothesis of similar means between the observed and simulated interaction totals, $\Sigma r_i$, $(i = 1,\ldots,22)$ from Geelong’s 22 regular season matches. One hundred
simulations were run on each round’s totals and the mean and standard deviation of each distribution was compared with the total observed interactions in each match. Figure 7.16 displays the simulation module in LINK prior to running the 100 test simulations on the 2011 grand final data.

Figure 7.16: Interaction simulation module in LINK

Figure 7.17 reveals a satisfactory fit for the model, with no significant difference between the simulated and observed series means ($p = 0.764$, $\alpha = 0.05$). Moreover, the majority of observed totals fell within 95% confidence intervals associated with each simulated match mean. Match 13 was considered an anomaly in the series—Geelong fielded their weakest side for the season, as acknowledged by the simulator, but managed to achieve almost 600 interactions and to win by 52 points, most likely due to their home-ground dominance. The outlier at Match 18 was Geelong beating Melbourne by 186 points—the second-highest margin in AFL history—yet the simulator acknowledged the strength of this side, offering the largest simulated interaction mean of all matches ($\Sigma r_i = 653.980$). This was an encouraging result given opponent effects have not been introduced
Melbourne were recognised as a poorly performing team in 2011). The overall fit gave us confidence to proceed to analysis of individual player effects.

7.5.1 Player Effects

A case study was undertaken on Geelong’s 2011 grand final team list, beginning with one thousand network simulations. Using the regression line in Figure 7.14 ($y = 7.04x - 146.41$), final score margins were predicted and logged after each simulation. The black curve in Figure 7.18 represents the normal distribution ($\mu_1 = 47.031, \sigma_1 = 14.315$) of predicted margins given Geelong’s actual grand final network. Geelong won the game by 38 points, which is a good reflection of the model’s predictive properties. Another one thousand simulations were run on the same side, but we replaced Bartel with a player of lesser skill, Shannon Byrnes. The light grey curve in Figure 7.18 represents the normal distribution ($\mu_2 = 31.960, \sigma_2 = 13.151$) of margins after Byrnes replaced Bartel in the side. Interpretation of this result is important; we concluded that, given his replacement (Byrnes), Bartel’s estimated net contribution to the selected team was $\mu_1 - \mu_2 = 15.071$ points. Emphasising the selected side was necessary as it could be hypothesised that Byrnes replacing Bartel in a stronger side may have less impact on margin due to the contribution of the other high-calibre players. To conceptualise the importance of selecting the best replacement player, we ran a third
iteration in which we replaced Bartel with Darren Milburn—a highly regarded player but not as skilful as Bartel—and again ran one thousand simulations. The normal distribution ($\mu_3 = 42.850, S_3 = 14.875$) is represented by the dark grey curve in Figure 7.18, from which we concluded that, given his replacement (Milburn), Bartel’s estimated net contribution to the selected team was $\mu_1 - \mu_3 = 4.181$ points. The difference between the mean of the Byrnes and Milburn distributions ($\mu_2 - \mu_3 = -10.890$) implied a coach would be more inclined to replace Bartel with Milburn in that side because the negative effect on margin is reduced. It is logical that a player may be selected on grounds other than his net effect on margin—for example, Byrnes’s style of play may be more suited than Milburn’s to the game-day conditions—but this is outside the concerns of this dissertation. This is a single demonstration of how a player affects a team network; many other permutations could have been explored, for example, multiple player exclusions, but the results from this example were considered conclusive enough.

![Figure 7.18: Margin distributions with and without Bartel](image)

7.6 Discussion

Player-based statistical analysis is as important in today’s sporting environments as ever before, with coaches continually searching for the right mix of players to include in a team. In the
AFL, the decision to include in a team one player over another can have serious repercussions on the outcome of the game. This chapter outlined an important model to assist in such selection decisions by simulating different players’ interactions with one another and by measuring the effect of such networks on final score margin. Negative binomial distributions were fitted to all pairs of players within a side so that interactions between players could be simulated prior to a match. It was discovered that the strength of the Geelong team’s networks was predictive of its final score margin; therefore, it was possible to measure the contribution any player could make to the final margin. Hence, when a team’s line-up is revealed, so too is the likelihood of the team winning. From a pre-match betting perspective, it is possible to calculate the odds of the selected team “covering the line”.

The agreed approach for this section of the research was to advance beyond individual performance measurement (see Chapter 5) and to introduce teamwork as a platform for ratings generation and win likelihood. Given the complexity of this subject, a baseline model was the research objective, ensuring model improvement could be logically observed at subsequent developmental stages. Ongoing research will also focus on improving the predictive power of the networks by weighting the three forms of player interactions in Section 7.3.1 with respect to the levels of efficiency, scoring capacity and ground and opponent effects. Some other limitations in this section of the research require discussion. If a prominent player is removed from the network, remaining $r_{ij}$ distributions are not recalculated—that is, we assume teammates do not improve their performance to cover the absence of the excluded player. This phenomenon of players exceeding expectation will be explored further in ongoing research. Furthermore, this research has not considered the presence of covariance between any $r_{ij}$. The initial stages of this research governed that each $r_{ij}$ is independent even though degrees of interaction covariance between sets of $[i, j]$ are almost certain. The thousands of $[i, j]$ permutations and covariance between each would command an additional research paper.

It is anticipated that an in-play simulation model will add further value because coaches and punters can make informed decisions with knowledge of live match scenarios. For example, if particular pairs of players have high interaction frequencies at half-time, simulations could be run to estimate final interactions and score margin. Finally, the application of this simulation model to
Australian Rules football was like diving straight into the deep end; any trepidation was overcome by our love of the game. There is massive potential in adapting a similar model to basketball and football, particularly the former, with only ten people on the court (not 36 like in AFL) and continuous incremental scoring. Duch et al (2010) have modelled teamwork in football with network analysis with positive results; this work would certainly augment their research.

Finally, an agreed research direction was to substitute the centrality data into the APR to confirm the hypothesis that cooperative measures, such as centrality, were more reflective and predictive of team performance than individual measures as were trialled in Chapter 5—that a team of champions may not be worth anything unless all of them are cooperating and contributing together. The centrality data from Geelong’s midfield from the 2011 season was approximately normal (see Figure 7.19), adhering to a primary assumption for use of the APR. The major setback was organising the opponent data. We did not possess the resources to be able to run the LINK program and associated nbd simulations and centrality generation on every team’s matches for the 2011 season. This is a massive task but certainly one worth undertaking because it will provide the official sign-off on the AFL player ratings research. It is anticipated that the project may require a programming platform beyond the capabilities of Visual Basic, one with a much faster runtime.

Figure 7.19: Eigenvector centrality approximately normal distribution for Geelong’s midfield (2011)
Chapter 8

Nonparametric Performance Forecasting

8.1 Introduction

The simulations from Chapter 5 produced a probability that a Geelong midfielder would outscore randomly selected opponents playing in the same position with normally distributed performances, therefore offering a player performance prediction \( (Exp) \) and a form guide for the paired contest. What modifications might be necessary to such an approach if one was unable to fit a traditional statistical distribution to the data, such as the normal distribution to the APR model data? In this chapter, we offer a nonparametric smoothing approach to performance prediction, where player performance for \( t + 1 \) assumes the form of a forecasted performance score, \( X \), where performance is measured by separate equations for comparative purposes: Equation (5.6) and Equation (7.13). Using a random sample of twenty-two players from the 2008 AFL season, this chapter details the arrival at an optimal performance forecast score for each player’s final match of the AFL season. Three models for predicting the player performance scores are compared: naive averaging \( (Mean) \), exponential smoothing \( (EXP) \) and Tukey-influenced, nonlinear smoothing \( (T-EXP) \). The \( Mean \) was calculated as each player’s average performance score to round 22, and is still the only online descriptive statistic available for aiding AFL fantasy team player selections. This section of the research improves on the unsophisticated assumption that a player will score approximately his fantasy league average the next time he plays. The nonparametric approach investigated in this chapter was inspired by the need for a
tool with which to select players when trading in fantasy football; while the common reference point for picking a player in most fantasy competitions is their season average, we developed an original forecasting technique which was more indicative of a player’s recent form and was less susceptible to the effects of outliers. The Tukey-smoothed forecasts provided a lower season and final match $RMSE$ per player than the other methods outlined.

### 8.2 Nonlinear Smoothing

Apart from being applicable when Gaussian assumptions are violated, nonlinear smoothing reduces noise in performance data sets and removes the misleading effects of performance outliers (Tukey, 1971), for example a poor performance due to an injury sustained in the early stages of a match. The usual mathematical operator in nonlinear smoothing is the $median$, an appropriate alternative if there is doubt surrounding the use of the mean to describe non-normal data sets. Tukey (1971) offers a diverse mix of nonlinear smoothers, using running median combinations through contiguous values (see Section 3.3), with which to remove unwanted noise and outliers from data sets while Velleman (1980) discusses nonlinear smoothing as a method of reducing the misleading effect of abrupt features in data sets, prior to data exploration. James et al (2005) overcame non-normal performance indicator distributions by using the median, rather than the mean, as an approximation of rugby player performance. Gebski and McNeil (1984) examine the use of nonlinear smoothers, as opposed to linear ones, identifying three important properties: resistance to outliers; retention of peaks and troughs (still resisting outliers); and repeatability until no further change occurs (still preserving peaks and troughs). They suggest linear smoothers are appropriate when Gaussian assumptions are met, but repeated use may “over-smooth” the data and erode explanatory peaks and troughs, as in cyclical economic data or seasonal data. The distribution of the majority of AFL performance variables examined in previous chapters of this dissertation, for example kicks (KCK) and handballs (HBL), are non-Gaussian (see Figure 8.1), possessing a right skew, given that competitive sports require a select few players in the “elite” category (right tail). It is also rational to assume that individual player samples and any unweighted aggregate of these performance variables from the
league distribution will also possess a right skew. The same distribution feature is evident in cricket batsmen’s scores (see Chapter 9).

![Figure 8.1: Frequency distributions of kicks (KCK) and handballs (HBL) for all players in 2011](image)

The non-normal distributions, evident in Figure 8.1, as well as the informative existence of outliers (such as the early removal of a player from a match) and peaks and troughs (above/below expected player performance) motioned this research into trialling nonlinear smoothers to arrive at player performance expectations. Literature on the application of nonlinear smoothing in sporting performance is not as accessible. Sargent and Bedford (2007) employed a Tukey 4253H smoother (defined in Section 3.3) to remove noise from player performance data when calculating simple AFL player ratings, while Shepherd and Bedford (2010) used the same smoother to eliminate noise in the probabilities of winning a medal in a pistol-shooting competition. The 4253H smoother was found to be effective as a smoothing tool, however, great interest surrounded the possibility of calculating different median combinations for different datasets to account for characteristics unique to that set (see Section 8.3.1). The remainder of this chapter seeks to explain how these optimal median combinations were arrived at.
8.3 Performance Score Forecasts

This chapter moves beyond the typical justifications for nonlinear smoothing to determine the value of forecasts from a nonlinear-smoothed series. Each player’s season performance data was “Tukey-smoothed” using an optimised combination of running medians spanning one to five, with a \( t + 1 \) forecast produced by exponentially smoothing the optimal Tukey-smoothed series. Each player’s final Tukey smooth was decided upon by running optimisations on a combination of median orders around each match score and the associated exponential smoothing parameter \( \theta \) so as to minimise the root mean-squared error (RMSE) of that player’s match forecasts. The median combinations were optimised on data arrays in Excel using a publicly available macro; the functions of the macro will be detailed later in this chapter with additional theory in Section 3.3. Figure 8.2 is a flowchart of how forecasts are arrived at for each player.

Figure 8.2: Events leading to the optimal Tukey-smoothed AFL player performance forecast

As discussed earlier in this chapter, three models were compared in generating forecasts of fantasy league player performance scores and associated root mean-squared error (RMSE): a season average (Mean), simple exponential smoothing (EXP) and an exponential smooth of a Tukey-smoothed series (T-EXP). The exponential smooth was a necessary step because Tukey smoothing is a recursive procedure, lacking a predictive element; the smooth does not extend beyond the final observed data point, \( t + 1 \) unlike with regression, moving averages or exponential smoothing. When an AFL fantasy “coach” is selecting a new player for the team at any stage in the season where historical data is available, a season average, \( \bar{x}_i \) of performance scores is the only publicly available measure of form for player \( i \).
Recalling Equation (5.6), the mean for player $i$ is simply:

$$
\bar{x}_i = \frac{\sum_{t=1}^{n} x_{it}}{n} 
$$  \hspace{1cm} (8.1)

where $n$ is the number of matches from player $i$ prior to the forecast period. The forecast for match $t + 1$ is $\bar{x}_i$, with residuals, $r_i = x_i - \bar{x}_i$ calculated after each match for player $i$'s $t = 1, \ldots, n$. While the average is a useful indicator of a performance of a player, it fails to accurately portray recent performance, which is important given the length of a season: a player may start strongly then fade as the season progresses. For example, if $\bar{x}_i = 150$ for $t = 1$ to 11 (the first half of the season) and $\bar{x}_i = 50$ for $t = 12$ to 22, his $\bar{x}_i$ for $t = 23$ is $(150 + 50)/2 = 100$, which is considered excellent in the AFL fantasy competitions; however, his form since the middle of the year has been poor, only averaging 50. Exponential smoothing gives weight to recent performance and is capable of recognising the form slump.

### 8.3.1 Smoothing the Smoother

The work outlined in this section demonstrates how exponentially smoothing players’ performance scores to produce $t + 1$ forecasts produced smaller errors than using simple averages (Equation (8.1)). Low average forecast error is considered valuable when selecting fantasy league players because a player with low error is more likely to consistently contribute to the overall fantasy team performance than a similarly performing player with high associated error (Waldman, M, 2005).

Exponential smoothing (EXP) is denoted as:

$$
\hat{x}_{i,t+1} = x_t + \theta_t (\hat{x}_t - x_t) 
$$  \hspace{1cm} (8.2)

where $\hat{x}_{i,t+1}$ is the performance score forecast for player $i$ after match $t$, and $\theta_t$ is the smoothing parameter ($0 \leq \theta_t \leq 2$), optimised to minimise the average of $r_t = |x_t - \hat{x}_t|$ for each of player $i$’s $t = 1, \ldots, n$. Exponential smoothing is common in sporting performance prediction. Clarke (1993) employed exponential smoothing to aid in the prediction of match results for the Australian Football
League, while Bedford and Clarke (2000) investigated a ratings model for tennis players based on exponentially smoothing margins of victory. Although not related to sport, Goodman (1974) was able to reduce mean-squared error in exponentially smoothed data sets by twicing (see Section 3.3) the smoothed residuals. This approach inspired a thought: instead of exponentially smoothing residuals, was it possible to exponentially smooth a brand-new nonlinear-smoothed series? Where Goodman applied Tukey to exponential smoothing to reduce error, we propose exponential smoothing to Tukey, defined as:

\[ \hat{y}_{t+1} = f(x_t) + \theta_2(\hat{y}_t - f(x_t)) \]  

(8.3)

where \( f(x_t) \) is the Tukey-smoothed performance score for the most recent match, \( \hat{y}_{t+1} \) is the exponentially smoothed one-step forecast and \( \theta_2 \) is the exponential smooth parameter which is optimised in conjunction with the median smooth sequence (defined by \( f(x) \)) for each player, to minimise their \( RMSE \) at \( t = 1, \ldots, n - 1 \) (Equation (8.4)). The optimisation objective function is defined as:

\[ \min[RMSE_{(x_t - \hat{y}_t)}] \]  

(8.4)

s.t. \( f(x) \in 0,\ldots,7 \) and \( 0 \leq \theta_2 \leq 2 \)

where \( f(x) \) can assume median spans of \([0,\ldots,5]\), \( 6 = S \) and \( 7 = R \) (see Section 3.3), with equal probability at each iteration. Although there is evidence for effective use of any number of nonlinear smooth components (Velleman (1980) goes as far as a \( 43RSRS2H,T \) smooth!), we have settled on a maximum of four components, finished with \( H \), then twiced for each player’s optimal sequence allocation, adhering to the logic of \( 4253H,T \) (Velleman (1980)). The inclusion of zeros and ones in the sequence (see conditions for Equation (8.4)) allows a smaller number of components, such as \( 3H,T \) (the smooth macro treats 0 and 1 medians as “blank”).

For the optimisation procedure, we used a publicly available Excel macro that allows near-boundless median span combinations and features \((R,H,S)\) combinations on arrays of data. A data array is smoothed by specifying the median combination and any additional arguments in the formula bar, for example, =SMOOTH(B2:B30, "3RSSH") where the data array is in column B, rows 2 to 30.
Figure 8.4 shows a variety of nonlinear smooths as a single player case study. The macro was written by Quantitative Decisions (http://www.quantdec.com/Excel/smoothing.htm). A snapshot of the code is in Section 11.3. The macro formula references five adjacent smoothing order cells for each player (see Figure 8.3), the first four cells undergoing 50,000 Monte Carlo iterations (see Section 3.4.4.1) of median span, S and R combinations, with H constant in the 5th cell.

![Image of Figure 8.3: The T-EXP optimisation model in Excel](image)

Using as an example the data of an AFL player (Chris Mayne), Table 8.1 details the arrival at match score forecasts and minimised RMSE. The \( x_t \) column provides his score for match \( t \), with each \( y_t \), a smoothed array of scores to match \( t \). Smoothing starts from match seven. The rationale for this is provided by Velleman and Hoaglin (1981, p.183) who, when applying a \( 4253H,T \) smoother, recommend a minimum sample size of 7 to allow for smoothing of the highest median span of five, plus the first and last points. Janosky et al. (1997) argue a larger sample results in a more robust smoother. For this research, one season yielded player data arrays with at least seven matches and no greater than twenty-two matches. A simple exponential smooth would suffice for a player’s forecast had they not achieved seven performance scores.

Each smooth sequence iteration in Equation (8.4) is applied to every smooth array from the seventh match to the player’s penultimate match, keeping in mind the minimum of seven data points.
(match scores) and that we are attempting to forecast the final match. Equation (8.3) is simultaneously run on each nonlinear smooth iteration to arrive at the one-step forecast (bold numbers in Table 8.1) for each match and associated residuals:

$$\hat{r} = x_{t+1} - \hat{y}_{t+1}$$  

(8.5)

<table>
<thead>
<tr>
<th>$t$</th>
<th>$x_1$</th>
<th>$y_7$</th>
<th>$y_8$</th>
<th>$y_9$</th>
<th>$y_{10}$</th>
<th>$y_{11}$</th>
<th>$y_{12}$</th>
<th>$y_{13}$</th>
<th>$y_{14}$</th>
<th>$y_{15}$</th>
<th>$y_{16}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>39</td>
<td>72.7</td>
<td>67.6</td>
<td>67.6</td>
<td>67.6</td>
<td>67.6</td>
<td>67.6</td>
<td>67.6</td>
<td>67.6</td>
<td>67.6</td>
<td>67.6</td>
</tr>
<tr>
<td>2</td>
<td>68</td>
<td>62.8</td>
<td>58.8</td>
<td>58.8</td>
<td>58.8</td>
<td>58.8</td>
<td>58.8</td>
<td>58.8</td>
<td>58.8</td>
<td>58.8</td>
<td>58.8</td>
</tr>
<tr>
<td>3</td>
<td>62</td>
<td>52.8</td>
<td>51.5</td>
<td>51.5</td>
<td>51.5</td>
<td>51.6</td>
<td>51.6</td>
<td>51.6</td>
<td>51.6</td>
<td>51.6</td>
<td>51.6</td>
</tr>
<tr>
<td>4</td>
<td>45</td>
<td>48.7</td>
<td>48.7</td>
<td>48.7</td>
<td>49.1</td>
<td>49.1</td>
<td>49.1</td>
<td>49.1</td>
<td>49.1</td>
<td>49.1</td>
<td>49.1</td>
</tr>
<tr>
<td>5</td>
<td>27</td>
<td>48.7</td>
<td>52.8</td>
<td>52.8</td>
<td>52.8</td>
<td>54.8</td>
<td>54.8</td>
<td>54.8</td>
<td>54.8</td>
<td>54.8</td>
<td>54.8</td>
</tr>
<tr>
<td>6</td>
<td>85</td>
<td>48.9</td>
<td>65.1</td>
<td>65.1</td>
<td>65.1</td>
<td>67.2</td>
<td>67.2</td>
<td>67.8</td>
<td>68.3</td>
<td>68.3</td>
<td>68.3</td>
</tr>
<tr>
<td>7</td>
<td>49</td>
<td>49.0</td>
<td>77.1</td>
<td>77.1</td>
<td>77.1</td>
<td>74.9</td>
<td>74.9</td>
<td>78.3</td>
<td>78.3</td>
<td>78.3</td>
<td>78.3</td>
</tr>
<tr>
<td>8</td>
<td>81</td>
<td>71.9</td>
<td>81.0</td>
<td>81.0</td>
<td>81.0</td>
<td>75.9</td>
<td>75.9</td>
<td>75.9</td>
<td>80.4</td>
<td>80.4</td>
<td>80.4</td>
</tr>
<tr>
<td>9</td>
<td>98</td>
<td>67.3</td>
<td>81.0</td>
<td>81.0</td>
<td>75.0</td>
<td>75.0</td>
<td>75.0</td>
<td>79.8</td>
<td>79.8</td>
<td>79.8</td>
<td>79.8</td>
</tr>
<tr>
<td>10</td>
<td>64</td>
<td>67.4</td>
<td>81.0</td>
<td>73.0</td>
<td>73.0</td>
<td>73.0</td>
<td>76.9</td>
<td>76.9</td>
<td>76.9</td>
<td>76.9</td>
<td>76.9</td>
</tr>
<tr>
<td>11</td>
<td>72</td>
<td>67.5</td>
<td>72.0</td>
<td>72.0</td>
<td>72.0</td>
<td>69.8</td>
<td>69.8</td>
<td>69.8</td>
<td>69.8</td>
<td>69.8</td>
<td>69.8</td>
</tr>
<tr>
<td>12</td>
<td>72</td>
<td>67.4</td>
<td>72.0</td>
<td>72.0</td>
<td>72.0</td>
<td>60.0</td>
<td>60.0</td>
<td>60.0</td>
<td>60.0</td>
<td>60.0</td>
<td>60.0</td>
</tr>
<tr>
<td>13</td>
<td>53</td>
<td>67.5</td>
<td>72.0</td>
<td>72.0</td>
<td>53.4</td>
<td>53.4</td>
<td>53.4</td>
<td>54.1</td>
<td>54.1</td>
<td>54.1</td>
<td>54.1</td>
</tr>
<tr>
<td>14</td>
<td>42</td>
<td>67.5</td>
<td>51.8</td>
<td>51.8</td>
<td>52.3</td>
<td>51.8</td>
<td>52.3</td>
<td>52.3</td>
<td>52.3</td>
<td>52.3</td>
<td>52.3</td>
</tr>
<tr>
<td>15</td>
<td>66</td>
<td>67.3</td>
<td>51.8</td>
<td>51.8</td>
<td>52.1</td>
<td>52.1</td>
<td>52.1</td>
<td>52.1</td>
<td>52.1</td>
<td>52.1</td>
<td>52.1</td>
</tr>
<tr>
<td>16</td>
<td>52</td>
<td>67.2</td>
<td>52.0</td>
<td>52.0</td>
<td>52.0</td>
<td>52.0</td>
<td>52.0</td>
<td>52.0</td>
<td>52.0</td>
<td>52.0</td>
<td>52.0</td>
</tr>
<tr>
<td>17</td>
<td>55</td>
<td>67.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8.1: T-EXP player performance matrix (Chris Mayne example)

The diagonal forecasts derived for each of Mayne’s smooth iterations from matches eight to sixteen ($x_{8,...,x_{16}}$) are entered into a traditional error formula:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}|x_i - y_i|^2}{n}}$$  

(8.6)

From 50,000 iterations, the minimum $RMSE$ (excluding $\hat{r}_{17}$) was produced by a $3225H,T$ smooth, with $\theta_2 = 0.0$ (see Figure 8.4 (i)). The same process was run on a match-by-match basis for Mayne from matches seven to fifteen to show the evolution of the optimised smooth (Figure 8.4). The main features of the progressive smooth are important. First, we see the smoother is more responsive to matches two and three after a switch to odd leading median spans (Figure 8.4 (c)). Secondly, there is resistance to one-off poor and excellent performances in matches five and nine respectively. Thirdly, the smoother never exceeds the match eight performance (Figure 8.4 (c) onwards) and, finally, the
smoother is pulled down by consecutive average games in matches thirteen and fourteen (Figure 8.4 (h)), where the combination switches from two medians to four.

As discussed earlier in this chapter, an attractive property of median smoothing is the resistance to outliers (Gebbski and McNeil, 1984). This is especially the case when an outlier is the last data point in an array, and a one-step forecast is to be generated. In his 17th match in the 2011 season, Joel Corey of the Geelong Football Club scored 164 points, considered an outlier at over 5 standard deviations from his performance mean ($\bar{x}_{t-1} = 89.4, \sigma_{t-1} = 13.8$). Figure 8.5 illustrates the benefit of exponentially smoothing the Tukey-smoothed series ($T-EXP$) rather than the raw performance data—the $EXP$ curve where $\theta$ is optimised to minimise RMSE. The $t + 1$ EXP forecast (for match 18) of 140 points is erroneously influenced by Corey’s excellent performance in match 17; it implies that Corey will continue his (well) above-average scoring but does not acknowledge that Corey’s record performance was against Melbourne, a far weaker side. Interestingly, match 18 was played against the Gold Coast—the worst side in the competition—with Corey only scoring 79 points—his worst score.
in six matches. The $T$-$EXP$ forecast of 109 is far more prudent, smoothing through the final outlying performance but managing to recognise Corey’s improving form since match 11. The error between the observed and forecasted performance at $t_{18}$ is halved by $T$-$EXP$ relative to $EXP$, which is the preferred outcome.

![Figure 8.5: Joel Corey’s match 18 score prediction after a peak in round 17](image)

### 8.4 Results

Table 8.2 provides details of the key results of the three forecasting models applied to the AFL player data. Each player’s twenty-second match score ($x_{22}$) and forecast ($y_{22}$) are given to calculate an overall (fantasy team) mean absolute deviation ($MAD$) for the final game. Each player’s season $RMSE$ ($e$ in Table 8.2), minimised in Equation (8.4), is also provided. The $Mean$ and $EXP$ approach revealed very similar $RMSE$ and final match $MAD$ per player. $T$-$EXP$ reduced $RMSE$ per player by 16.2% (26.5 to 22.2) and $MAD$ by 7% (20.3 to 18.9) when compared to the $Mean$ model. This supports $T$-$EXP$ as a more reliable fantasy player selection tool than coaches expecting a player to achieve his season average for the next match. Prior to optimising each player’s Tukey smooth, we trialled a $4253H,T$ smooth (recommended by Velleman (1980)) with optimised $\theta_z$. It was found to possess a much higher $RMSE$ per player (27.47) and final match $MAD$ (28.6) than that of the $mean$,
EXP and T-EXP, supporting our decision to progress to optimising each player’s Tukey and exponential smoothers.

The optimised smooth sequences are listed in the Smooth column in Table 8.2. The optimised exponential smoothing parameters for EXP (θ₁) and T-EXP (θ₂) are also listed. On average, θ₂ is higher than θ₁, suggesting that the exponential smooth is more responsive when the data array has reduced noise due to the Tukey smooth. Individual player analysis reveals no significant reductions in the value of θ₂ from θ₁. Conversely, Goodwin, Mitchell, Gray and Cotchin each progress from near-zero θ₁ to the maximum θ₂ \approx 2.0 under the T-EXP model.

<table>
<thead>
<tr>
<th>Player i</th>
<th>Team</th>
<th>X₀₂₂</th>
<th>y₀₂₂</th>
<th>E</th>
<th>θ₁</th>
<th>y₀₂₂</th>
<th>E</th>
<th>Smooth</th>
<th>θ₂</th>
<th>y₀₂₂</th>
<th>E</th>
<th>CI₀</th>
<th>CIₚ</th>
<th>P(θ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Knights</td>
<td>Ad</td>
<td>149</td>
<td>92.4</td>
<td>44.2</td>
<td>0.2</td>
<td>100.2</td>
<td>46.4</td>
<td>2S5H</td>
<td>0</td>
<td>100.5</td>
<td>39.6</td>
<td>67.1</td>
<td>101.5</td>
<td>21.2%</td>
</tr>
<tr>
<td>2 Goodwin</td>
<td>Ad</td>
<td>62</td>
<td>92.6</td>
<td>16.8</td>
<td>0.2</td>
<td>80.6</td>
<td>13.9</td>
<td>23SH</td>
<td>2</td>
<td>76.5</td>
<td>9.9</td>
<td>71.1</td>
<td>89.2</td>
<td>11.7%</td>
</tr>
<tr>
<td>3 Hooper</td>
<td>Br</td>
<td>43</td>
<td>61.6</td>
<td>16.4</td>
<td>0.2</td>
<td>62.5</td>
<td>17.2</td>
<td>423SH</td>
<td>0</td>
<td>62.6</td>
<td>15.6</td>
<td>51</td>
<td>66</td>
<td>12.2%</td>
</tr>
<tr>
<td>4 Bradshaw</td>
<td>Br</td>
<td>79</td>
<td>79.2</td>
<td>18.9</td>
<td>0.2</td>
<td>78.5</td>
<td>19.8</td>
<td>4S4RH</td>
<td>0</td>
<td>78</td>
<td>13.9</td>
<td>72.1</td>
<td>82.3</td>
<td>98.7%</td>
</tr>
<tr>
<td>5 Scotland</td>
<td>Ca</td>
<td>71</td>
<td>92.8</td>
<td>20.1</td>
<td>0.2</td>
<td>94.5</td>
<td>20.3</td>
<td>5RH</td>
<td>0.3</td>
<td>93.2</td>
<td>9.3</td>
<td>72.2</td>
<td>105.9</td>
<td>21.9%</td>
</tr>
<tr>
<td>6 Shaw</td>
<td>Co</td>
<td>55</td>
<td>91.1</td>
<td>22.6</td>
<td>0.1</td>
<td>87</td>
<td>22.5</td>
<td>5R3H</td>
<td>0.4</td>
<td>82.8</td>
<td>20.1</td>
<td>41.1</td>
<td>88.3</td>
<td>30.3%</td>
</tr>
<tr>
<td>7 Welsh</td>
<td>Es</td>
<td>63</td>
<td>82.3</td>
<td>42.7</td>
<td>0</td>
<td>89</td>
<td>40.7</td>
<td>3RS2H</td>
<td>0.1</td>
<td>86.6</td>
<td>40.5</td>
<td>66.7</td>
<td>101.5</td>
<td>21.0%</td>
</tr>
<tr>
<td>8 Wayne</td>
<td>Fr</td>
<td>55</td>
<td>60.9</td>
<td>20.3</td>
<td>0.5</td>
<td>55.4</td>
<td>19.2</td>
<td>322SH</td>
<td>0</td>
<td>67.5</td>
<td>16.4</td>
<td>47.1</td>
<td>62.6</td>
<td>94.3%</td>
</tr>
<tr>
<td>9 Stokes</td>
<td>Ge</td>
<td>77</td>
<td>95.3</td>
<td>22.5</td>
<td>0.3</td>
<td>102.1</td>
<td>22.2</td>
<td>54S5H</td>
<td>0</td>
<td>95.7</td>
<td>21.1</td>
<td>94.8</td>
<td>105</td>
<td>0.0%</td>
</tr>
<tr>
<td>10 Sewall</td>
<td>Ha</td>
<td>95</td>
<td>87.9</td>
<td>29.5</td>
<td>1.3</td>
<td>128.6</td>
<td>23.3</td>
<td>2224H</td>
<td>1.9</td>
<td>125.4</td>
<td>23.6</td>
<td>84.7</td>
<td>126.7</td>
<td>64.7%</td>
</tr>
<tr>
<td>11 Dew</td>
<td>Ha</td>
<td>41</td>
<td>69.1</td>
<td>35.4</td>
<td>0.2</td>
<td>75.5</td>
<td>37</td>
<td>235SH</td>
<td>1.3</td>
<td>79</td>
<td>31.5</td>
<td>64.3</td>
<td>98</td>
<td>28.6%</td>
</tr>
<tr>
<td>12 Mitchell</td>
<td>Ha</td>
<td>98</td>
<td>93.4</td>
<td>30</td>
<td>0</td>
<td>98.6</td>
<td>29.7</td>
<td>55S5SH</td>
<td>2</td>
<td>106</td>
<td>27</td>
<td>93.7</td>
<td>110</td>
<td>98.3%</td>
</tr>
<tr>
<td>13 Unghart</td>
<td>Ka</td>
<td>30</td>
<td>58.7</td>
<td>27.2</td>
<td>0.4</td>
<td>44.4</td>
<td>26.4</td>
<td>4SH</td>
<td>1.4</td>
<td>38.3</td>
<td>23.7</td>
<td>22.6</td>
<td>56.4</td>
<td>89.7%</td>
</tr>
<tr>
<td>14 Simpson</td>
<td>Ka</td>
<td>73</td>
<td>92.1</td>
<td>17.1</td>
<td>0.2</td>
<td>90.4</td>
<td>17.4</td>
<td>3H</td>
<td>0.4</td>
<td>92.1</td>
<td>16.3</td>
<td>75.1</td>
<td>92.2</td>
<td>34.5%</td>
</tr>
<tr>
<td>15 Johnson</td>
<td>Me</td>
<td>45</td>
<td>81.1</td>
<td>57.4</td>
<td>0.1</td>
<td>63.5</td>
<td>51</td>
<td>5S24H</td>
<td>0</td>
<td>65.6</td>
<td>45</td>
<td>13.4</td>
<td>101</td>
<td>89.6%</td>
</tr>
<tr>
<td>16 Gray</td>
<td>Po</td>
<td>80</td>
<td>62.2</td>
<td>17</td>
<td>0</td>
<td>64</td>
<td>13</td>
<td>224H</td>
<td>2</td>
<td>75.2</td>
<td>0.8</td>
<td>57.5</td>
<td>70</td>
<td>90.2%</td>
</tr>
<tr>
<td>17 Cotchin</td>
<td>Rl</td>
<td>69</td>
<td>72.5</td>
<td>23.6</td>
<td>0.2</td>
<td>73.5</td>
<td>23.6</td>
<td>5R2H</td>
<td>2</td>
<td>68</td>
<td>21.5</td>
<td>67.3</td>
<td>79.9</td>
<td>99.9%</td>
</tr>
<tr>
<td>18 Fisher</td>
<td>St</td>
<td>95</td>
<td>94.4</td>
<td>23.2</td>
<td>0</td>
<td>95</td>
<td>20.7</td>
<td>5S4RH</td>
<td>0.3</td>
<td>100.4</td>
<td>19.9</td>
<td>81.6</td>
<td>108.5</td>
<td>78.9%</td>
</tr>
<tr>
<td>19 Kirk</td>
<td>Sy</td>
<td>40</td>
<td>101.4</td>
<td>24.6</td>
<td>0.3</td>
<td>101.3</td>
<td>25.6</td>
<td>5S4SSH</td>
<td>0</td>
<td>100.9</td>
<td>22.4</td>
<td>93</td>
<td>110.3</td>
<td>0.0%</td>
</tr>
<tr>
<td>20 Hargrave</td>
<td>Wb</td>
<td>95</td>
<td>74.7</td>
<td>28.6</td>
<td>0.7</td>
<td>62.6</td>
<td>25</td>
<td>4S4RH</td>
<td>1.4</td>
<td>77</td>
<td>22.3</td>
<td>40</td>
<td>78.2</td>
<td>7.2%</td>
</tr>
<tr>
<td>21 Cross</td>
<td>Wb</td>
<td>95</td>
<td>102.3</td>
<td>15.7</td>
<td>0.1</td>
<td>98.1</td>
<td>14.7</td>
<td>54RH</td>
<td>0</td>
<td>96</td>
<td>13.9</td>
<td>93</td>
<td>105.2</td>
<td>89.2%</td>
</tr>
<tr>
<td>22 Jones</td>
<td>Wc</td>
<td>63</td>
<td>67.0</td>
<td>29.8</td>
<td>0.2</td>
<td>70.5</td>
<td>31.6</td>
<td>445SH</td>
<td>0.9</td>
<td>74.5</td>
<td>26</td>
<td>67</td>
<td>85.4</td>
<td>99.8%</td>
</tr>
</tbody>
</table>

| Average | 26.5 | 0.26 | 25.5 | 0.75 | 22.2 | 65.3 | 92   | 53.70% |
| MAD     | 20.3 | 19.9 | 18.9 |      |      |      |      |      |

Table 8.2: Comparison of forecast error from a Mean, EXP and T-EXP model

An important step in assessing the strength of our findings was to establish 95% confidence intervals for the forecast statistic. Monte Carlo simulations (50,000 iterations) of feasible smoothing sequences and associated θ₂ were run for each player with lower and upper intervals recorded for the match twenty-two forecasts (see CI₀, CIₚ in Table 8.2). High-averaging players with relatively narrow confidence intervals represent players who have shown little variation in their match-to-match scores.

With this knowledge, we can assume the “best” selection for a fantasy team for match twenty-two is Cross (54RSH,T), averaging 102.3 with intervals [93.0,105.2]. The next surest is Kirk (54SSH,T):
101.4, [93.0,110.3]. A risky selection would be Johnson: 81.1, [13.4,101.0]. The large width of this interval is evidence of volatile performance scores through the season. Mayne’s simulated forecast score distribution is illustrated in Figure 8.7.

Modelling every player’s residuals (using Equation (8.5)) from each match was a necessary step to determine underlying distributions in the performance score data. Figure 8.6 shows the error terms as normally distributed for our league, providing evidence that player sample residuals should be of similar form. With a $t$-distribution, 95% confidence intervals for the average residual were established for $\hat{y}_{t+1}$ for every smooth sequence recorded in each player’s simulation. The probability of the final match score, $x_{t+1}$, falling within these limits over the 50,000 iterations was calculated (see $P(x)$ in Table 8.2) to evaluate our findings after match twenty-two. Our best pick, Cross, scored 95 in his final match, which 89.2% of our iterations predicted, hence justifying his selection in the fantasy team. Cotchin, averaging 72.3 prior to the final game, reached 99.9% with a score of 69. Conversely, Kirk, averaging 101.4 was injured early in the final game and only achieved 40 points (0.0%). Predictions outside certain levels of confidence reflect the uncertainties surrounding a player’s week-to-week performance. The performances of consistent players that stay uninjured are more likely to be predicted with the least error.

![Figure 8.6-7: Histogram of sample players’ residuals / Distribution of Mayne’s simulated score forecasts](image)

163
8.4.1 Eigenvector Centrality Forecasts

In Chapter 7, centrality measures were generated for each Geelong player in the 2011 grand final using Equation (7.13). In the discussion section, it was noted that future research would see these performance measures substituted into the APR model to produce network-influenced adaptive player ratings. The nonparametric alternative will also be pursued. Subsequently, we decided to test the $T$-EXP forecast model on the Geelong midfield eigenvector centrality data from 2011 to ensure our research findings were not a fluke or relevant only to the performance data generated by Equation (5.6). Mean absolute error terms were again calculated and compared across the Mean, EXP and $T$-EXP models, with resulting MAE of 7.8, 7.1 and 6.0 respectively (see Table 8.3). That Tukey smoothing reduced error rates when run on different teams and in different years was a pleasing result, moreover its reduction effect when transferred to the centrality data. Perhaps the most interesting finding from this supplementary analysis was the median combinations for the optimised smooths: all smoothing sequences possessed a maximum of two median elements (excluding Hanning), whereas quite a few of the previous performance data smooths in Table 8.2 used four medians, plus additional features such as splicing. This we put down to the lower variance in the centrality data set due to the reduced range of the data: from Table 8.2, $Max(x_{it}) - Min(x_{it}) = 149 - 30 = 119$, for all $i$ and $t$. For the centrality data in Table 8.3, the range is $68.19 - 0.03 = 68.15$.

<table>
<thead>
<tr>
<th>Player</th>
<th>$e_i$</th>
<th>$y_{i1}$</th>
<th>$E$</th>
<th>$\theta_i$</th>
<th>$y_{i22}$</th>
<th>$E$</th>
<th>Smooth</th>
<th>$\theta_2$</th>
<th>$y_{i22}$</th>
<th>$E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allen Christensen</td>
<td>30.5</td>
<td>23.7</td>
<td>6.8</td>
<td>0.24</td>
<td>26.4</td>
<td>4.2</td>
<td>53H</td>
<td>0.27</td>
<td>26.8</td>
<td>3.7</td>
</tr>
<tr>
<td>Cameron Ling</td>
<td>24.3</td>
<td>32.5</td>
<td>8.1</td>
<td>0.15</td>
<td>31.0</td>
<td>6.7</td>
<td>34H</td>
<td>1.47</td>
<td>21.6</td>
<td>2.8</td>
</tr>
<tr>
<td>James Kelly</td>
<td>24.0</td>
<td>34.3</td>
<td>10.3</td>
<td>0.09</td>
<td>35.4</td>
<td>11.4</td>
<td>4SH</td>
<td>0.00</td>
<td>35.1</td>
<td>11.1</td>
</tr>
<tr>
<td>Jimmy Bartel</td>
<td>23.5</td>
<td>31.6</td>
<td>8.1</td>
<td>0.17</td>
<td>34.3</td>
<td>10.8</td>
<td>SSH</td>
<td>0.00</td>
<td>35.1</td>
<td>11.6</td>
</tr>
<tr>
<td>Joel Corey</td>
<td>23.7</td>
<td>37.7</td>
<td>14.0</td>
<td>0.02</td>
<td>36.3</td>
<td>12.6</td>
<td>23H</td>
<td>0.03</td>
<td>38.2</td>
<td>14.5</td>
</tr>
<tr>
<td>Joel Selwood</td>
<td>35.3</td>
<td>38.5</td>
<td>3.3</td>
<td>0.21</td>
<td>37.4</td>
<td>2.1</td>
<td>2SH</td>
<td>0.00</td>
<td>36.8</td>
<td>1.5</td>
</tr>
<tr>
<td>Mitch Duncan</td>
<td>16.6</td>
<td>26.3</td>
<td>9.7</td>
<td>0.04</td>
<td>24.7</td>
<td>8.1</td>
<td>H</td>
<td>0.56</td>
<td>18.5</td>
<td>1.9</td>
</tr>
<tr>
<td>Paul Chapman</td>
<td>36.8</td>
<td>36.1</td>
<td>0.7</td>
<td>0.10</td>
<td>36.5</td>
<td>0.3</td>
<td>H</td>
<td>0.14</td>
<td>36.1</td>
<td>0.7</td>
</tr>
<tr>
<td>Taylor Hunt</td>
<td>14.0</td>
<td>23.2</td>
<td>9.2</td>
<td>0.28</td>
<td>22.2</td>
<td>8.2</td>
<td>3H</td>
<td>1.98</td>
<td>20.2</td>
<td>6.2</td>
</tr>
</tbody>
</table>

Table 8.3: Comparison of forecast error of 2011 Geelong midfield

As mentioned in Chapter 3, the fewer running medians involved, the less resistant the smooth will be to abrupt changes in the data; the low centrality range implies abrupt changes were uncommon in Geelong’s centrality data.
8.5 Discussion

This section of the research has not been undertaken solely to achieve the “best” forecast of Australian Football League player performance but, also, to highlight the benefits of nonlinear smoothing prior to fitting a $t + 1$ forecasting model, such as exponential smoothing. With a sample of twenty-two players, exponentially smoothing a Tukey-smoothed series has delivered a significantly smaller average forecast error than on an unsmoothed series. This was a satisfactory outcome, especially because important match factors were not included in the final model. Opponent strength proved to be an important input for the APR ratings in Chapter 5, improving the relationship between the probability of a player outscoring a randomly selected opposition player and the player’s observed match score. Bailey (2000) proved the importance of opponent effects on a team’s margin of victory by analysing the predictive properties of the competing teams’ differences in age, weight, experience and number of kicks and handballs with linear regression. An important development for the $T-EXP$ forecasting model will be the inclusion of a parameter that accounts for any opponent effects that increase or decrease the likelihood of an individual player achieving approximately his average score. For example, it is common for the best midfielder in a side to be “tagged” by an opponent, whereby the opponent attempts to reduce the midfielder’s influence on the game through legal physical (and verbal) means. Similarly, there are cases of forwards who are not as effective against particular teams when a particular defender is playing. For these reasons, it will be necessary to access the team lists and detect incidences of significant player match-ups that may reduce the player’s impact on the match. Unfortunately, the computational requirements to simulate the team in Table 8.2 were enormous meaning the incorporation of their opponents would demand an improved platform.

The use of optimisation to build the smoothing model is only recently possible with the advent of computer simulation, specifically, Visual Basic programming for Excel. The remarkable work of Tukey can now be improved upon through the use of near-boundless median smoothers responsive to the data, rather than using the opinion of experience of types of data. The use of Tukey smoothing, modified at each stage, is an important tool to the sport forecaster in the prediction of
player performance. Such an approach would be an interesting adaption for online fantasy sites, rather than publishing the players’ average performance for the season.
Chapter 9

Player Performance in Limited Overs Cricket

9.1 Introduction

The previous four chapters of this dissertation have been concerned with the statistical analysis of player performance in the AFL. As stated in Chapter 4, the AFL is a continuous game, able (unrealistically) to progress fluidly for four quarters without any breaks in play. This chapter introduces player performance analysis in a discrete sport, limited overs cricket, where each match consists of 600 legitimate trials (300 for each team) of bowlers and fielders attempting to dismiss as many batsmen as possible while limiting run scoring. The nature of the contest attracts a less ambiguous methodology for player performance measurement than continuous sports as the efforts of the batsman and bowler can be assessed after each trial by two outcomes: whether a run(s) was scored or whether a batsman dismissal occurred. This chapter details how each batsman’s outcomes can be forward simulated during the match to arrive at a total score for himself and his team based on conditional probabilities describing the ball-by-ball likelihoods of runs and a dismissal. These probabilities were dependent on the match state (number of dismissals and deliveries bowled) and the skill of the non-dismissed (currently batting or yet to come) batsmen. Using a multiple regression model, the aggregate simulated runs were adjusted to reflect any difference in opponent strength as well as the presence of innings effect (the advantage of batting first or second) and venue effects. This chapter will ultimately demonstrate the benefits of simulation by fitting log-normal distributions to
500 score iterations by Australia’s Ricky Ponting in the 2011 ODI World Cup quarterfinal, at the 0-, 10-, 20- and 30-over mark of the match, providing a statistical context for his observed score.

9.2 Discrete Sports

The “one-on-one” nature of cricket and baseball—that is, a specialist bowler/pitcher competing directly with a specialist batsman—offers countless directions for the analysis of the players’ performance and the pragmatic application of predictive models. Lewis (2005) went so far as to describe the game of cricket as a “sporting statistician’s dream”. Moreover, the ongoing publishing of official International Cricket Council (ICC) team and player ratings (ICC Rankings 2011) is testament to cricket’s quantitative elegance. The player ratings are of particular interest, calculated with respect to each player’s role (batsman/bowler/all-rounder) using limited performance measures: runs scored / wickets taken; quality of bowling/batting line-up; level of run-scoring within the match (for batsman / against bowler); and match result. Weighted moving averages are then employed to arrive at the final ratings for each player (see Section 11.4 for the top 50 batsmen as at November 2012). The statistical goldmine, born from the discrete composition of a baseball game, allowed the development of a brand-new field of statistical analysis by Bill James, titled Sabermetrics. Sabermetrics is concerned with a multitude of quantifiable categories within a baseball game (in excess of twenty), for example, Base Runs (BsR), Defence independent pitching statistics (DIPS), Pythagorean Projection and player ratings, all of which attempt to explain and/or predict team and player outcomes in a baseball match. While these statistical categories are mostly derived from efforts concerning run scoring and prevention, their sheer number and depth is analogous to the argument posed in Section 4.2.3. that games where scoring is relatively infrequent demand further statistical attention to understand how scoring frequency, or likelihood of victory, can be increased.

In Chapter 4, One-Day International (ODI) cricket and baseball were defined as “discrete” sports, or sports comprising a series of contests, occurring at close intervals, between a player from the offensive (batting) and defensive (bowling/pitching and fielding) teams. This definition contrasted

---

26 The methodology is published in words only; no algorithm is provided.
27 The name is derived from the Society of American Baseball Research, or SABR.
sports with a continuous, dynamic flow such as football or hockey. Cricket and baseball are two of the most recognised bat and ball team sports, with each providing an interesting study regarding a batter’s approach to the match objectives. Koop (2002) mentioned the difficulty in comparing baseball batters due to baseball being, fundamentally, a multiple-output sport. He used the example of the difference in approaches between power hitters like Mark McGwire and single hitters like Tony Gwynn. Similarly, in the 2012 baseball World Series, Marco Scutaro, batting second for the San Francisco Giants, excelled at singles and doubles, averaging 0.500 in the seven-game 2012 NLCS (National League Championship Series) against St Louis but without any home runs (or triples). Pablo Sandoval, however, batting third, was a slugger, hitting three home runs in Game 1 of the World Series. These batting styles were in direct contrast with each other but, ultimately, beneficial for the team—Scutaro would likely achieve a base hit, and Sandoval would bat him in with consistent power hitting. The difference in approaches between a second and third batsman in ODI cricket, however, is not as pronounced: each is slightly more defensively inclined than the fourth or fifth batsman in an attempt to retain resources in the early stages of the innings so that a platform for runs can be laid. The main difference between the second and third batsmen, for example, would be the rate at which they score runs. Bailey and Clarke (2004) capitalised on these batting styles by designing strategies to maximise profits derived from wagering on one batsman outscoring another (“head-to-head”) during the 2003 ODI World Cup. They provided the probability of Ganguly, the third-order batsman for India, outscoring Kaif, the fourth-order batsman (Pr(Ganguly > Kaif) = 0.605). Section 9.3.2 classifies batsmen using the different batting styles adopted by each player, using similar approaches to those in Chapter 6.

Batting styles, through each position, are dependent on the type of cricket being played. Chapter 2 distinguished between three different types of cricket currently played at the international level: test match cricket, one day international (ODI) cricket and T20 cricket. While each format possesses the “one-on-one” property discussed earlier, their durations differ, demanding a unique set of performance measurement and prediction parameters. Test match cricket is played over a

---

28 Sandoval became the fourth player in history to achieve three home runs in a World Series game.
maximum of five days, with deliveries following a continuous distribution: prior to the match, it is unknown how many overs will be bowled in one of a possible four innings (two innings per team during the match); wickets are the only discrete resource. ODI matches run for 50 overs per side, where an over is made up of a minimum of six deliveries\textsuperscript{29}. T20 cricket, as the name suggests, allows each team 20 overs—a maximum of 120 legitimate deliveries—to accumulate runs (innings 1) or chase the opposition’s total (innings 2). Test match batsmen are able to patiently bat for up to and exceeding 90 overs in a full day’s play, however, the explosive hitting of T20 cricket occurs in one innings. While it may be a simpler task to predict total team and player runs, with lower error rates, in T20 cricket because of the shorter duration, this chapter is concerned with ODI cricket where runs are accumulated within 50 overs. In developing optimal scoring rates for ODI matches, Clarke (1988) worked under the commonly shared principle that batsmen are more cautious in early and middle stages of their team’s innings in an attempt to preserve team resources, while the latter stages of the innings are prone to aggressive batting to maximise team runs and to increase the probability of victory. A batsman, however, adopts the latter tactic at his own peril because sustained aggression increases the risk of dismissal (Swartz et al 2006). In T20 cricket a batsman can realise a trade-off: adopting an aggressive approach earlier in the match than ODI, in the knowledge that there are fewer delivery resources available, but knowing the probability of dismissal is higher. The common thread running between each format is that run-scoring avenues are almost identical (see Section 2.2.3). A notable rule change occurred in 2005, however, in response to conservative middle-innings scoring in the ODI format, when increased fielding restrictions were introduced. ‘Powerplays’ see extra fielders inside the 30-metre circle for a brief period, increasing the opportunity of batsmen achieving scores of four or six by hitting over the fielders. Hence, scoring rates typically increase during these Powerplays, although risk taking (big hitting to the boundary) by the batsmen can also lead to higher dismissal rates. Test match cricket, due to its multi-day strategies, is not subjected to the Powerplay. Section 9.3 discusses in-play statistical modelling for ODI matches. A predictive model would become more robust being capable of recognising when a Powerplay is introduced—fluctuations in

\textsuperscript{29} Matches usually extend beyond 300 deliveries (50 overs x six deliveries) because illegitimate deliveries, such as ones deemed by the umpire to be too wide of the batsman, are required to be bowled again.
the probability of run scoring and dismissals would need to be addressed—however, for the purposes of this research, a post hoc adjustment to the in-play predictions was observed as the Powerplay data was fed through the model.

9.2.1 Statistical Modelling in Cricket

Numerous efforts have been made over the last century to retrospectively fit statistical models to batsmen’s scores. Early work by Elderton (1945) proposed test match cricket batting scores followed a geometric distribution, but subsequent research revealed this distribution provided an inadequate fit for zero and extreme scores (Wood 1945). This analysis was performed on test match batsman data—limited overs cricket was not even a thought at this time—but Figure 9.1 reveals that the issue extends to the ODI format. This graph depicts 7,295 batsmen’s innings from ODI matches played between 2008 and 2011; well over 5% of scores are zeros. This phenomenon is due to a number of factors, but most probably the unfamiliarity of the batsman with the pitch conditions and bowler performance as he comes to the crease. Nerves are another likely explanation. The difficulty involved with fitting distributions to this data is also clear; the research trialed an exponential distribution fit, but the likelihood of a zero score was underestimated at slightly over 4%, approximately the same as scoring one run, which clearly is not the case. Bailey and Clarke (2004) log-transformed batting scores to alleviate this problem in the tail of the distribution, while Bracewell and Ruggiero (2009) employed a beta distribution to model zero scores, separate to non-zero scores, that were modeled with a geometric fit. There is enough evidence to suggest a binomial distribution may explain zero scores and then a geometric (or similar) fit for non-zero scores, but this is to be included in the ever-growing list of future research. Rather than letting a pre-chosen model dictate the analysis, the observed data drove the appropriate choice of model structure. This was achieved by expressing each batsman i’s runs in a match, $r$, as a fraction of the team’s match total, $R$, simply denoted as:

$$C_i = r/R$$

(9.1)

where $C$ is described as the batsman’s contribution to the team total (see Figure 9.2). Equation (9.1) offers a fractional player contribution estimate—a far less ambiguous approach than player
contributions in a continuous sport, such as AFL, which required a whole chapter (Chapter 7) of examination. Reapplication of the exponential distribution to the transformed contribution data reveals a better fit for zero scores; interestingly, the distribution over-fit the zero contribution data rather than under-fitting, as in the raw zero score data. Such methodology could also be applied to individual batsmen to determine how many runs he may score in an upcoming match. An attractive property of in-play prediction models in cricket is that distributions can be fit after the batsman scores his first run, removing some of the modelling flaws presented by the zero score.

Figure 9.1: ODI batsman scores (2008–2011) fitted with an exponential distribution

Figure 9.2: ODI batsman contributions (2008–2011) fitted with an exponential distribution
9.3 In-play Simulation

While there is extensive published statistical research on cricket in all forms, there is a scarcity of models that can describe a batsman’s scoring progress and expectation while a match is in play. Swartz et al (2006) applied a log-linear approach to simulate runs scored during any stage in an ODI match for a proposed batting order. This research was motivated by the potential to assess, and then reassess, the probability of batting outcomes while an ODI event was in play. For this, “ball-by-ball” run and dismissal estimations were investigated for “in” and remaining batsmen in team \(j\), at any stage in the first innings\(^{30}\), by forward simulating the outcome of a discrete random variable, \(X\). The distribution of \(X\) was described by probabilities that were conditional on the following match resources: batsman type; a classification reflecting batting ability and style; and the delivery number, \(b\), an indication of time remaining in the first innings. The joint distribution of runs and dismissals, given batsman type and delivery number, effectively described the match “state” by offering a dynamic set of scoring and dismissal likelihoods at any delivery; these probabilities were critical for observing and predicting a batsman’s scoring strategy. A batsman’s runs, simulated from each trial, were added from \(b_1, \ldots, b_n\), where \(n\) is either the batsman’s simulated dismissal or his team’s simulated innings completion point. Swartz et al (2009) employed a similar conditional process to simulate ODI cricket match outcomes, with outcome probabilities dependent on batsman/bowler combinations, the number of dismissals and deliveries, and the current match score. Figure 9.3 illustrates the simulation process while the remainder of the chapter discusses it in detail.

\[\text{Figure 9.3: ODI run simulation flowchart}\]

\(^{30}\) The second innings is a more complicated task because the model must incorporate runs required for the second batting team to win, in conjunction with the decay of delivery and batting resources.
9.3.1 Data

The dearth of practical match data quantifying ODI match delivery outcomes is a hurdle for in-play modelling. This research benefited from extensive ball-by-ball ODI match commentary found at [www.espncricinfo.com](http://www.espncricinfo.com). Using a tailored Visual Basic script, relevant data was extracted from the commentary of all completed ODI matches played between sixteen nations from January 2006 until October 2011. The final dataset comprised approximately 215,000 independent trials from the first innings of these matches. The script prepared key simulation variables for import into statistical software: bowler, batsman, batsman order, delivery ($b$), runs scored (from $b$) and a binary dismissal (from $b$) variable ($d$) (see Table 9.1). The batsman’s team, opponent, innings number and match venue were also included for post-simulation adjustments.

Table 9.1: First two overs of data from South Africa v India ODI match (12.01.2011)

Once collated, the ball-by-ball data revealed extensive match information—indiscernible in match summary data—to complement decision-making for wagering, coaching or player-rating calculations. After collapsing the dismissal data from deliveries into overs, a simple dot chart was generated (Figure 9.4(a)) and presented important trends. Predictably, dismissals possessed a positive trend in the final 10 overs because batsmen are inherently more aggressive knowing delivery resources are decaying. However, there was a need to further examine the noticeable increase in dismissals between the 18th and 20th overs. By introducing a second factor, batting order, it was possible to extrapolate additional dismissal information. Figure 9.4(b) shows a modal point for fourth-order dismissals at the 20-over mark. The data confirmed that fourth-order batsmen were most exposed in this over, accounting for approximately 3.4% of balls faced in their innings, increasing the
likelihood of dismissal relative to other overs. Figure 9.4(b) also confirms a first-order batsman’s susceptibility in the first over. He faces the first delivery of the innings largely uncertain of any influential match factors—for example, pitch condition—and, as a result, has a higher likelihood of dismissal. As the game progresses, other batsmen are advantaged because they can witness the conditions as well as receive first-hand information from dismissed batsmen. Section 9.3.3 will explain how in-play conditional probabilities of runs (0,…,6) and dismissals were derived from the approach that generated the data in Figure 9.4(b).

![Figure 9.4: Dot charts displaying dismissal frequencies (a) per over, (b) per over by batting order](image)

### 9.3.2 Batsman Classification

Classification of batsmen according to their “type” was an extension of the batting-order analysis, discussed in the previous section. Using simulated annealing, Swartz et al (2006) addressed the importance of a well-selected batting order by analysing the many permutations within a team’s batting line-up. Prior to this study, Bailey and Clarke (2004) concluded that batting order was a highly significant predictor of runs scored ($p < .001$) when analysing batsman head-to-head wagering. While it is well established that batting order is correlated with batting ability (see Figure 9.5), it is flawed to assume that batsmen of the same order possess the same run-scoring approach. For example, third-order batsmen from different teams will be among the strongest batsmen in their team but almost
certainly will possess differing approaches to their batting, dependent on their skill level. This research categorised batsmen with respect to their order (from 1 to 11), the rate at which he scores (“strike rate”) in that order at early (first third of his innings), middle (second third) and latter (last third) stages, and his derived contribution to his team’s run total in each match (see Equation (9.1)).

Figure 9.5: Mean runs scored per batting order

The strike rate for any batsman \( p \) can be simply calculated at any match stage by:

\[
SR_p = \frac{\sum_{d=1}^{n} (x \mid d = 0)}{n}
\]

where \( x \) is each run outcome from the first \((b = 1)\) to final delivery \((n)\) faced in his innings. A discriminant classification process, similar to that used to classify AFL players in Chapter 6, was engaged to assign each player in the dataset to the type vector \( t = [F, M, S] \), where \( F \) = “high risk, fast scorer”, \( M \) = “low risk, medium scorer” and \( S \) = “low risk, slow scorer”. Batsmen of order 9, 10 and 11 in a match were automatically assigned to a fourth type, \( B \) (“bowler”) because it was considered safe to assume they were specialist bowlers of consistently limited batting ability\(^{31}\). Type \( F \) batsmen played aggressively from the early stage of their innings; these batsmen were likely to achieve higher strike rates but also a higher rate of dismissals than other types of batsmen (Swartz et al 2006). Type

\(^{31}\) While bowlers may be further classified on batting ability—some achieve a batting average above 10, others below 5—a single classification was deemed appropriate for this research.

176
M batsmen were more cautious in early stages, with strike rates improving as their innings progressed; they generally played longer innings, with higher average scores than other batting types. Type S were less able batsmen (lower strike rates and average runs) but more able bowlers than F and M. For reference in this chapter, a batsman of order 3 and type F, for example, is represented by the notation $\alpha = F_3$. With an intrinsic knowledge of these batting styles, a training sample was formulated, with 50 batsmen manually assigned to the $t$ vector. The equation for classifying the remaining batsmen cases ($n = 1,036$) was:

$$d_{up} = b_\alpha + \sum_{i=1}^{3} b_i X_{ik} + b_j X_{jk}$$

(9.3)

where $d_{up}$ is the $u^{th}$ discriminant function for batsman $p$, $X_{ik}$ is the mean value of his strike rate at innings period, $i$ after match $k$, $X_{jk}$ is the value of his contribution to team runs, $b_\alpha$ is a constant, and $b_{ij}$ are coefficients selected to maximise the Mahalanobis distance between the three type centroids in $t$ (see Section 3.2 and Section 6.2.1).

Figure 9.6 provides a snapshot of the classification results with respect to the mean strike rates by batsmen of each type (excluding bowlers). The most interesting outcome was the similarity in strike rates between $F_3$ and $M_3$, and $F_4$ and $M_4$. An ANOVA post-hoc test confirmed $F_3$ and $M_3$ ($p = .845$) and $F_4$ and $M_4$ ($p = .551$) strike rate means were not statistically different ($\alpha = 0.05$). This phenomenon is probably due to the requirement of batsmen in these positions to avoid batting in a high-risk fashion because they typically bat through the middle stages of an innings, protecting batting resources for aggressive batting in the latter overs. Subsequent orders displayed increasing $F$ strike rates suggesting it is usual for these batsmen to bat more aggressively to increase the run rate (runs/overs bowled). From Figure 9.6 it appears that $M$ batsmen possess a more consistent strike rate through the orders whereas $F$ and $S$ both realise increasing mean strike rates from order four through eight. Table 9.2 offers further insight, with a breakdown of each batsman type’s mean strike rates in early-, middle- and latter-innings stages. $F_7, 8$ batsmen possessed higher strike rates at the three innings stages than any other order because it is common for these batsmen to be “in” at latter team innings stages, where it is customary to bat aggressively in an attempt to score as many runs as possible (Clarke 1988). The ANOVA test confirmed $F_8$ ($p = .000$) strike rate was significantly
different to $M_8$ and $S_8$ strike rates, while $M_8$ and $S_8$ strike rates were statistically similar ($p = .343$). Moreover, the $F_{5, 6}$ strike rates systematically increase at a higher rate than any other type, suggesting these batsmen were capable of rapid scoring when the innings enters an aggressive phase. The proceeding section reveals how batsman type data gives the conditional probability estimations further dimension.

Figure 9.6: Boxplots of mean strike rate by batsman type

<table>
<thead>
<tr>
<th>Orders</th>
<th>Type</th>
<th>Innings stage</th>
<th>Early</th>
<th>Middle</th>
<th>Latter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 2</td>
<td>$F$</td>
<td>0.74</td>
<td>0.80</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$M$</td>
<td>0.51</td>
<td>0.62</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$S$</td>
<td>0.29</td>
<td>0.38</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>3, 4</td>
<td>$F$</td>
<td>0.53</td>
<td>0.65</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$M$</td>
<td>0.60</td>
<td>0.66</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$S$</td>
<td>0.31</td>
<td>0.43</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>5, 6</td>
<td>$F$</td>
<td>0.70</td>
<td>0.98</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$M$</td>
<td>0.52</td>
<td>0.63</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$S$</td>
<td>0.31</td>
<td>0.41</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>7, 8</td>
<td>$F$</td>
<td>0.82</td>
<td>1.08</td>
<td>1.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$M$</td>
<td>0.50</td>
<td>0.61</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$S$</td>
<td>0.41</td>
<td>0.59</td>
<td>0.65</td>
<td></td>
</tr>
</tbody>
</table>

Table 9.2: Mean strike rates for batsmen types at early-, middle- and latter-innings stages
9.3.3 Conditional Probability Estimation

As discussed in Section 9.1, this research investigated in-play run (x) and dismissal (d) estimations by forward simulating the outcome of a discrete random variable, X. Probabilities for the joint distributions given by $X = \{x, d\}$ were derived from a frequency matrix, $A_{mn}$, using data from the first innings of the collected matches, where each $mn$ is a unique match state describing a batsman of type $m$’s scoring/dismissal outcomes at delivery $n$. Swartz et al (2009) employed a similar frequency approach to estimate conditional probabilities associated with runs scored from illegitimate deliveries in ODI cricket. The frequency distribution in this research provided 8,100 unique match states (300 deliveries x 27 groups$^{32}$) and, in turn, 56,700 independent raw probabilities (8,100 match states x 7 outcomes of $X$). Each $\{x, d\}$ in each $mn$ was then assigned to a conditional probability distribution, such that:

$$\sum_{x=0}^{a} P(X = x, d | \alpha_m, b_n) = 1$$

(9.4)

where $a \in \{(0,...,4) \cup (6)\}$ and $\alpha$ is the type of batsman facing delivery $b$ ($n = 1,...,300$).

In certain $A_{mn}$ elements, particularly the left tails of lower-order batsmen’s score distributions, missing values and an extremely small sample size for certain X outcomes produced unrealistic zero and one raw probabilities. To correct this, a fourth-order polynomial smooth, $f’(X)$, was applied through each $f(X) = P(X = x,d | \alpha, b)$ at each $b$ to interpolate missing values and to reduce extreme probabilities in the relevant distributions. Non-parametric median smoothing, as detailed in Chapter 8, was also tried but failed the requirement of a low-response smooth. The preferred smooth for $f(X)$ took the form:

$$f’(X) = \frac{vb^4 + wb^3 + xb^2 + yb + z}{\sum_{x=0}^{d} f(X)}$$

(9.5)

where $v, w, x$ and $y$ are smoothing order coefficients, $z$ is the intercept, and the denominator ensured:

$$\sum_{x=0}^{d} f’(X) = 1$$

$^{32}$ Three batsman types x eight orders + three bowlers
Figure 9.7 displays the polynomial smooth of probabilities for a batsman $\alpha = F_1$ scoring $x = 4$ at each $b$ in an innings. Interestingly, the smooth achieved two maximal points around $b = 36$ (Over 6) and $b = 276$ (Over 46), suggesting opening batsmen of type $F$ were more likely to engage in offensive strategies at the early and latter stages of their innings, rather than in the middle stages. A discussion on the contribution of delivery resources and batting-type combinations to several unique match scenarios is worthwhile. The influence of delivery number on scoring in a match was evident from the data; Figure 9.7 exemplifies how commonplace aggressive batting is in the last 10 overs ($b > 240$), particularly by type $F$ batsmen. When two higher-order batsmen are active towards the end of the innings, a state of sturdy team progress is observed, whereas lower-order batsmen active at a relatively early match stage reflects a poor innings due to the loss of top-order resources. The generated probability distributions reflected the various match states: for example, the difference in a $M_4$ batsman’s probability of dismissal in the first three overs, $P(d = 1 \mid M_4, b \leq 18) = 0.008$ ($\sigma = 0.003$), compared with the final three overs, $P(d = 1 \mid M_4, b \geq 282) = 0.055$ ($\sigma = 0.002$), was statistically significant ($p < .001$) because the former scenario required the batsman to play a defensive role in response to the loss of top-order resources.

Table 9.3 is a very small subset of $A_{mm}$ distributions of type $M$ batsmen, but reveals the dynamic nature of the probabilities from an early stage in the innings for each $(b = 1 \mid \alpha = M_1), (b =$
20 \mid \alpha = M_5), (b = 100 \mid \alpha = M_{11}), and from the final delivery in the innings (b = 300 \mid \alpha = M_{1,5,11}). The opening batsman, for example, was noticeably more defensive (zero runs scored from a delivery) at the first delivery of the innings ($P(x = 0 \mid M_{1,1}) = 0.760$) than at the final delivery of the innings ($P(x = 0 \mid M_{1,300}) = 0.200$). This was because, initially, he had the luxury of 100% batting and delivery resources available, allowing him the most time of any batsman to adapt to the physical conditions of the match (such as bowling quality, climate and pitch condition). A bottom-order batsman was almost certainly one of the team’s specialist bowlers and, therefore, a poor batsman, which was reflected in the marginal difference in the probability of dismissal at a relatively early match stage ($P(d = 1 \mid M_{11,100}) = 0.120$) and at the final delivery ($P(d = 1 \mid M_{11,300}) = 0.165$).

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$b$</th>
<th>$P(x=0)$</th>
<th>$P(x=1)$</th>
<th>$P(x=2)$</th>
<th>$P(x=3)$</th>
<th>$P(x=4)$</th>
<th>$P(x=6)$</th>
<th>$P(d=1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$</td>
<td>1</td>
<td>0.760</td>
<td>0.156</td>
<td>0.032</td>
<td>0.006</td>
<td>0.047</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>$M_1$</td>
<td>300</td>
<td>0.200</td>
<td>0.444</td>
<td>0.189</td>
<td>0.001</td>
<td>0.069</td>
<td>0.011</td>
<td>0.064</td>
</tr>
<tr>
<td>$M_5$</td>
<td>20</td>
<td>0.814</td>
<td>0.109</td>
<td>0.009</td>
<td>0.002</td>
<td>0.046</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>$M_{11}$</td>
<td>100</td>
<td>0.111</td>
<td>0.339</td>
<td>0.174</td>
<td>0.015</td>
<td>0.139</td>
<td>0.082</td>
<td>0.139</td>
</tr>
<tr>
<td>$M_{11}$</td>
<td>300</td>
<td>0.245</td>
<td>0.444</td>
<td>0.058</td>
<td>0.006</td>
<td>0.063</td>
<td>0.009</td>
<td>0.165</td>
</tr>
</tbody>
</table>

Table 9.3: Probabilities for $\alpha=M_{1,5,11}$ at early- and late-innings stages

### 9.3.4 Simulation and Run Projection

A Visual Basic simulator, SimScore (see Figure 9.8), was written specifically for this research (see Section 11.3 for code sample) to estimate unknown run quantities for each “in” and remaining batsman by calling the conditional probability matrix from Section 9.3.3. An attractive property of the simulated approach is that batsmen’s run estimates can be revised following any ball in a live match scenario, reflecting the match state at $A_{mn}$. Furthermore, live match scores feed into SimScore, allowing the user to refresh the simulated batting scores after critical events, such as a dismissal, occur in the observed match. The initial routine in SimScore produces a random probability, $u_1 \sim U(0,1)$, at each $b$, with the outcome, $X = \{x, d\}$, determined by the cumulative distribution function:

$$F'(x, d) = P(X \leq x, d \mid \alpha, b) \quad (9.6)$$

To demonstrate, batsman $p$ is facing the first delivery of an innings: if $u_1 = 0.70$ and $F'(0) = 0.75$, then $x_p = 0$. Each trial is repeated at a specified delivery in a match, through the last specified delivery.
(usually \( b = 300 \)) for each batsman. Batsman \( p \)'s revised total runs, \( R \), are easily recalculated at any delivery, \( b \), by:

\[
R_{pb} = \sum_{b+1}^{n} (x \mid d = 0) + r
\]  

(9.7)

where \( r \) is total observed runs at \( b \), and \( n \) is the final delivery of batsman \( p \)'s simulated innings.

\[
S_{b_{t+1}} = \begin{cases} 
p & \text{if } x_p = 0,2,4,6 \\
q & \text{if } x_p = 1,3 \\
r & \text{if } d_p = 1
\end{cases}
\]  

(9.8)

where \( S_{b_{t+1}} \) is the rule for the last delivery of each over, after which a new bowler bowls from the opposite end to the previous over. It should be noted that byes and run outs also impact on strike but were not included in \( X = \{x, d\} \). Model improvements will account for these outcomes.

Figure 9.8: SimScore screen shot

For SimScore to properly reflect a match state, its logic had to extend to assigning strike (facing the next delivery) to each “in” batsman, given certain outcomes of \( X \), for example, batsman \( q \) assuming strike after batsman \( p \) scored one run (the batsmen crossed once). Two sets of rules were written for assigning strike, \( S \), if batsman \( p \) and \( q \) were in and batsman \( p \) faced the previous delivery, \( b_t \):

\[
S'_{b_{t+1}} = \begin{cases} 
p & \text{if } x_p = 1,3 \\
q & \text{if } x_p = 0,2,4,6 \\
q & \text{if } d_p = 1
\end{cases}
\]
9.4 Match-day Adjustments

As in most contests, an ODI match consists of two teams of varying strengths, usually competing at one of the team’s home ground. The concept of home-ground advantage in ODI is not as distinct as, say, domestic limited overs cricket, where each of six Australian states are aligned with a ground on which they train and play regularly, hence breeding familiarity. As the Australian ODI team comprises players from each of these states, it is more reasonable to describe any advantage the national team realises from playing in Australia as simply “home advantage” (Stefani and Clarke 1992), or the advantage from not having to travel internationally for the match. Table 9.4 displays winning percentage for the Australian ODI team between 2011 and 2012; of venues where more than three matches have been played, Australia has won more matches in Brisbane (80%), Hobart (80%) and Perth (78%) and the least in Melbourne (59%) and Sydney (59%). If these figures were roughly similar to the Victorian domestic limited overs side, for example, it be would be more appropriate to say Victoria had little home-ground advantage (playing in Melbourne) than the national side, given the low win percentage, especially if no Victorian players were representing Australia in this period.

<table>
<thead>
<tr>
<th>Venue</th>
<th>Matches</th>
<th>Wins</th>
<th>Mean Runs</th>
<th>Win_%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adelaide</td>
<td>13</td>
<td>8</td>
<td>247</td>
<td>0.62</td>
</tr>
<tr>
<td>Brisbane</td>
<td>10</td>
<td>8</td>
<td>282</td>
<td>0.80</td>
</tr>
<tr>
<td>Darwin</td>
<td>3</td>
<td>3</td>
<td>226</td>
<td>1.00</td>
</tr>
<tr>
<td>Hobart</td>
<td>5</td>
<td>4</td>
<td>266</td>
<td>0.80</td>
</tr>
<tr>
<td>Melbourne</td>
<td>17</td>
<td>10</td>
<td>247</td>
<td>0.59</td>
</tr>
<tr>
<td>Perth</td>
<td>9</td>
<td>7</td>
<td>257</td>
<td>0.78</td>
</tr>
<tr>
<td>Sydney</td>
<td>17</td>
<td>10</td>
<td>277</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Table 9.4: Australia ODI team winning percentage and mean first innings runs by venue

Ground dimensions, climatic conditions and crowd effects differ from state to state, so each venue needs to be considered separately when calculating run-scoring potential for the national team. The Mean Runs column in Table 9.4 is the average runs scored by Australia in the first innings of matches between 2011 and 2012 for venues where more than three matches had been played. The team had been the most offensive in Brisbane (averaging 282 runs) and the least in Melbourne and Adelaide.

33 Many more limited overs matches are played at the domestic level than at the international level in a calendar year.
34 It might be of interest to test for a relationship between match outcome by ground and the proportion of players in the national team playing at their home-state ground in each match.
(247 runs each), which corresponds with the winning percentages at these grounds. Sydney was an exception for such a correlation; although the venue was ranked second for runs scored, the team was ranked last for win percentage. This suggests that there may have been an advantage attached to batting in the second innings in Sydney, where high first-innings totals were eclipsed by the team batting second with greater ease than at other venues.

The difference in team totals relative to venue, team strength and various other factors dictates that batsmen’s run-scoring ability must also differ, given the team total is the sum of the individual batsmen’s scores (plus runs from illegitimate deliveries and byes / leg byes). For example, an opening batsman from Australia—a recognised cricket nation—is expected to possess a greater scoring ability, and be supported by more able batting partners, when playing Canada—not ranked at international level—than an opening batsman from Canada facing Australia. For the in-play model, the conditional probabilities associated with each delivery outcome were generated from a full sample, so as to include all team strengths and match factors; no distinction was made between a game, say, between Australia and Canada, where Australia’s likelihoods of dismissal would be lower and run scoring would be higher. For the probabilities and resulting batsmen’s simulated scores to be indicative of the match characteristics, each non-dismissed batsman’s projected score was adjusted after each simulated delivery, so as to sensibly reflect a projected team runs total. This approach was more appropriate than expanding the \( A_{nx} \) matrix so that relevant match conditions could be accounted for probabilistically. One can imagine the difficulties in computing the probability of dismissal for an opening Australian batsman on the sixth delivery against South Africa in Durban on a hot day with a hostile crowd! Sample size is the obvious limitation. Estimated total team runs (\( T \)) required an efficient calculation able to be revised at any stage during a match. Total team runs were calculable at any stage prior to or during a match with the aid of the Duckworth-Lewis method (Duckworth and Lewis 1998), a resource-based approach—resources being 50 overs and 10 wickets per team—to setting a revised run target for the team batting second when a match has been shortened (fewer than 50 overs are able to be bowled in the second innings), typically by rain. Clarke and Allsopp (2001) and de Silva et al (2001) made use of the Duckworth-Lewis (D/L) rain interruption rules to project a
second-innings winning score, after a match’s completion, to calculate a true margin of victory for the team batting second (team \(b\)): if team \(b\) reached the run target set by team \(a\), and still had available resources, the D/L rules were used to project the team’s runs, should they (hypothetically) have kept batting. Team \(a\)’s runs were then subtracted from the estimated team \(b\) total to arrive at the margin of victory. If team \(a\) won (team \(b\)’s resources were exhausted in the run chase), its winning margin was simply team \(a\) total runs minus team \(b\) total runs. The equation de Silva et al (2001) used for the projected number of runs \((T)\) set by team \(b\) to estimate a true margin of victory was:

\[
T_b = \frac{100Y}{(100 - \lambda)}
\]

(9.9)

where \(Y\) is the actual runs scored by \(b\), and \(\lambda\) is the remaining resources from the Duckworth-Lewis table (Duckworth and Lewis 1998). Say team \(b\) beat team \(a\), having achieved the target of 219 + 1 runs with 4 wickets in hand and 10 overs left, the run projection for \(b\) would be set at \(100/(100 - 22.7) = 285\), suggesting \(b\) had enough resources left to potentially score a further 285 - 220 = 65 runs, should play have continued after the target was reached. Should they have had one wicket and 10 overs left, \(T\) would have been computed at 230, a more modest margin of victory of 10 runs, given the limited resources (\(\lambda = 4.5\)). Adhering more closely to the revised run target equation in Duckworth and Lewis (1998) for when \(\lambda_b > \lambda_a\), it is possible for a team’s total runs to be estimated in either innings at any stage in the match, using:

\[
T = S + \frac{\lambda G_{50}}{100}
\]

(9.10)

where \(S\) is the score achieved by the batting team at any innings stage, and \(G_{50}\) is the average score in an uninterrupted ODI match, dependent on innings number. Duckworth and Lewis (1998) originally set this value at 225 for the first innings, but it is common today to see a value of 235, with run rates slowly increasing in the game, possibly due to the cross-pollination from the more aggressive T20 cricket. An argument arising from the research was that when using Equation (9.10) for \(T\) projections, a standard \(G_{50}\) value across all games provided a flawed projection, especially in pre-match (\(\lambda = 100\%\)), and instead should be weighted to account for differing team strengths and match conditions. For example, in a match where Australia (ranked 4th in ODI cricket as at November 2012) played Kenya (ranked 13th), should either team have batted first and be one wicket down for ten runs after 3
overs, their projection, from Equation (9.10), would be $10 + \frac{(90.9 \times 235)}{100} = 224$. It is unlikely that Australia would score this few runs against Kenya, and Kenya this many against Australia. A unique, pre-match total runs projection ($TR$) was calculated for each team, for varying circumstances, by running a multiple regression analysis on games between 2011 and 2012:

$$TR = b_0 + b_1G + b_2T + b_3O + b_4H + \varepsilon$$

(9.11)

where $G$ is ground effect, $T$ is team strength in first and second innings, $O$ is opponent strength, $H$ is home advantage in first and second innings, $b_1$ to $b_4$ are optimised weights, $b_0$ is the intercept and $\varepsilon$ is an error term. The new $TR$ figure was substituted into Equation (9.10) for team $a$, so run projections better reflected the competitive environment. To prove the worth of weighting $G_{50}$, mean absolute error rates were calculated between the observed final runs and projected final runs in all innings before matches (Pre) and at five different match stages: 0–10 overs, 11–20 overs, 21–30 overs, 31–40 overs and 41–50 overs. Compared with the $G_{50}$ approach, the $TR$ approach realised lower average errors at all periods, with a sharper decay in error as the innings progressed, which validated the procedure.

Figure 9.9: $TR$ v $G_{50}$ mean absolute error rates in pre- and intra-match states

Figure 9.10 portrays the $T$ run projection at each delivery for an ODI match between Australia and South Africa at Port Elizabeth (SA) on 23 October 2011. Australia batted in the second innings,
chasing 304 runs for victory. Substituting the relevant pre-innings data into Equation (9.10), Australia was expected to score 228 runs, represented by the broken line in Figure 9.10. Note that the $T$ estimate equals $TR$ at the commencement of the innings—this can be proven by substituting $\lambda = 0$ (full resources available) into Equation (9.10)—and observed runs at the end of the innings, but only when team $b$ loses or draws the match—this can be proven by substituting $\lambda = 100$ into Equation (9.10). Australia started the innings well, scoring 34 runs from 22 deliveries (see $T$ line increasing to roughly a 250-run estimate) before Ricky Ponting was dismissed at the 23rd delivery; the loss of wicket resources meant the $T$ estimate fell to 240, which remained steady until Michael Clarke was dismissed at the 37th delivery (see further drop in $T$ line to 221). Australia eventually lost all wicket and delivery resources—Xavier Doherty, a specialist bowler, was dismissed on the last delivery of the innings—scoring 223 runs, providing South Africa with a margin of victory of $303 - 223 = 80$ runs. From an Australian supporter’s perspective, the result of the match was not pleasing, but from a statistical researcher’s perspective, predicting the South African victory along with a $TR$ prediction error of only 5 runs was a more than satisfactory outcome.

![Figure 9.10: TR projection for Australia v South Africa (23/10/2011)](image)

Having established the match day adjustment model in Equation (9.10), a demonstration at the player level is now required. To revise batsman $p$’s runs, $R$, at any delivery, $b$, $q + r$ was used,

---

35 It is rare that the 10th wicket is taken on the 300th delivery.
where \( q = \Sigma(x | d = 0) \) for \( b + 1, \ldots, n \), and \( r \) is his observed runs to \( b \). For batsmen who have been dismissed, \( q = 0 \), and for batsmen to come, \( r = 0 \). A team’s total revised runs, \( S \), at any \( b \) is simply \( \Sigma(q + r) \) for all \( p = 1, \ldots, 11 \). The adjustment of each non-dismissed (two current and those yet to come) batsman’s score to reflect team strengths and match conditions required a redistribution of the difference between estimated and revised runs among these batsmen, or \( R_{adj} = \alpha(T - S) \) where \( \alpha \) is the redistribution coefficient, defined as \( 1/(w_1 - w_2 + 2) \), where \( w_1 \) is the simulated number of wickets and \( w_2 \) is the actual number of wickets at delivery \( b \). This ensures that the non-dismissed batsmen receive the run difference, so at any stage in the match, \( S = T \). Table 9.5 offers a simulated match scenario for Australia in a first innings: say the score is 3 wickets down for 150 runs; Watson (74 runs), Haddin (37) and Ponting (6) have been dismissed; Clarke (20) and M Hussey (13) are currently batting (*) with simulated totals of 58 and 16 respectively; a \( T \) total of 235 and simulated total of 6 wickets for 224 runs means a difference of +11 runs would be redistributed through Clarke, M Hussey, White, D Hussey and Johnson. After redistribution, the simulated match total equals the projected run total (\( T \)).

<table>
<thead>
<tr>
<th>Batsmen</th>
<th>Type</th>
<th>Runs</th>
<th>Adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 SR Watson</td>
<td>1</td>
<td>74</td>
<td>74</td>
</tr>
<tr>
<td>2 BJ Haddin</td>
<td>1</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>3 RT Ponting</td>
<td>3</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>4 MJ Clarke*</td>
<td>3</td>
<td>58</td>
<td>60</td>
</tr>
<tr>
<td>5 MEK Hussey*</td>
<td>1</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>6 CL White</td>
<td>2</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>7 DJ Hussey</td>
<td>3</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>8 MG Johnson</td>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>9 B Lee</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 JJ Krejza</td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 SW Tait</td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>224</td>
<td>235</td>
</tr>
<tr>
<td><strong>Wickets</strong></td>
<td></td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td><strong>Ball</strong></td>
<td></td>
<td>300</td>
<td></td>
</tr>
<tr>
<td><strong>Projected Runs (T)</strong></td>
<td></td>
<td>235</td>
<td></td>
</tr>
</tbody>
</table>

Table 9.5: Simulated match scenario showing run redistribution

9.5 Model Validation and Results

Validation of the simulation model was performed in two stages, with post-simulation adjustments for teams and match conditions. A generalised validation involved a measurement of average simulated runs at each delivery stage, irrespective of batsman type. One thousand match simulations were performed and each batsman’s score was logged at each delivery. The overall fit of
the simulated runs was compared with a holdout sample of observed data from 10 recent matches. Preliminary analysis revealed no significant difference between the mean simulated and holdout samples of batsman runs through each delivery, assuming equal variances ($p = .186$). The means of pre-smooth ($Sim\_mean$) and post-smooth ($Sim\_S\_mean$) run means was compared with the holdout sample, irrespective of batsman type, at three match phases: early, middle and latter (see Figure 9.11). Intraclass correlation coefficients were calculated to measure the resemblance of means within each group to the holdout mean at each phase. A negligible difference in means existed between the smoothed and unsmoothed run means in the early phase. However, the decision to smooth the probabilities was justified in the latter stages of the game, with $Sim\_mean$ ($\rho = 0.632$) shifting further from the holdout mean than $Sim\_S\_mean$ ($\rho = 0.798$). These results confirmed a good overall fit for the smoothed $SimScore$ model. Increasing variance in the final 10 overs could be explained by the combination of top-order (higher run scoring) and lower-order (lower run scoring) batsmen alternating strike on the latter deliveries.

A second-stage validation involved a parametric approach, designed to address any anomalies resulting from a particular batsman type’s run simulations. An investigation into runs scored by Australia’s Ricky Ponting in the 2011 ODI World Cup quarterfinal against India provided interesting results. Ponting was classified as $\alpha = M_3$ and scored 104 runs, considered a milestone for any batsman. $SimScore$ ran 500 iterations of Ponting’s innings from four match stages: pre-match ($b = 0$); after 10 overs were bowled ($b = 60$); after 20 overs ($b = 120$); and after 30 overs ($b = 180$). His score was calculated after each iteration from each stage, using Equation (9.7) where $r$ was Ponting’s observed runs at the particular simulation start stage: $(r = 0 \mid b = 0)$; $(r = 0 \mid b = 61)$; $(r = 22 \mid b = 121)$; and $(r = 48 \mid b = 181)$. At the end of the 10th over ($b = 60$), the opening batsman, Watson, was dismissed, so Ponting faced delivery 61 without any runs to his name, or $r = 0$. The primary interest lay in the distribution of how many more runs Ponting would score, after $b = 60$, 120 and 180, or:

$$\hat{R} = \sum_{b+1}^{n} (x \mid d = 0) - r \quad (9.12)$$
A log-normal distribution, LN(\(\mu, \sigma^2\)), was fitted at each stage to approximate the likelihood of Ponting achieving his observed runs, \(r = 104\), in the match. Figure 9.12 displays the fit at each stage, with Figure 9.12(A) a pre-match approximation—that is, starting the simulation procedure from \(b = 1\).

Figure 9.11: Intraclass correlation of Sim_mean and Sim_S_mean with holdout mean

Figure 9.12(a) confirmed a good fit at the early and latter simulated run stages and, also, appears to be a solution to run modelling problems in the left and right tails, encountered as early as Elderton (1945). The probability of Ponting reaching \(R = 104\) prior to match commencement was \(1 - P(R < 104) = 0.075\). After 10 overs (Figure 9.12(b)), the score expectation increased to \(1 - P(R < 104) = 0.091\), as more certainty in the outcome had been established by the in-play conditional estimations.

An improved fit around the 30- and 60-run curve areas was also observed. After 20 overs (Figure 9.12(c)), a further improvement in fit was observed as Ponting gained momentum during his innings. Ponting’s new simulated run target, \(R\), was \(104 \times (r = 22) = 82\), with probability \(1 - P(R < 82) = 0.098\). After 30 overs (Figure 9.12(d)), an increased likelihood \((1 - P(R < 56) = 0.125)\) met the research expectations because Ponting’s observed run target of 104 was incrementally closer than at previous deliveries. Moreover, the log-normal fit was considered satisfactory after 30 overs, given the negligible difference between \(P(R < 56) = 0.125\) and \(P(X < 56) = 0.135\). The application of this
methodology to the other batsman scoring with Ponting and with subsequent batsmen in the order would help ascertain a strategy with which to determine the highest-scoring batsman in the innings.

![Figure 9.12: Ricky Ponting’s simulated runs fitted with log-normal distributions after (a) 0 overs, (b) 10 overs, (c) 20 overs, d) 30 overs](image)

9.6 Consistency

Recall in Chapter 5 the discussion on player performance consistency, namely, how consistent performance can be attributed to skill levels more so than luck (Marquardt, 2008). Cricket batsmen can similarly be rated on the consistency in their run scoring from match to match. This is certainly not a new concept in cricket: Elderton (1909) used a coefficient of variation to measure a
batsman’s scoring consistency, where a coefficient closer to zero implied more consistent performances:

\[ CV_1 = \frac{\sigma}{\bar{x}} \]  

(9.13)

where sigma and \( \bar{x} \) are a batsman’s score standard deviation and average, respectively. It was simple to apply such a formula to ODI batsmen scores between 2008 and 2011, expressing the consistency coefficient (\( CV_1 \)) in descending order and in decimal format (see Table 9.6). The 30 batsmen with the lowest \( CV \) who had played at least 10 games in the time period are displayed. On close inspection, this list is not as meaningful as was hoped.

<table>
<thead>
<tr>
<th>Player</th>
<th>Matches</th>
<th>Mean</th>
<th>Std Dev</th>
<th>( CV_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>TN de Grooth</td>
<td>13</td>
<td>12.69</td>
<td>7.78</td>
<td>0.61</td>
</tr>
<tr>
<td>M Vijay</td>
<td>11</td>
<td>17.82</td>
<td>10.97</td>
<td>0.62</td>
</tr>
<tr>
<td>TLW Cooper</td>
<td>11</td>
<td>45.73</td>
<td>28.16</td>
<td>0.62</td>
</tr>
<tr>
<td>AN Cook</td>
<td>10</td>
<td>37.50</td>
<td>23.31</td>
<td>0.62</td>
</tr>
<tr>
<td>JF Mooney</td>
<td>19</td>
<td>26.47</td>
<td>16.57</td>
<td>0.63</td>
</tr>
<tr>
<td>GA Lamb</td>
<td>10</td>
<td>16.30</td>
<td>10.37</td>
<td>0.64</td>
</tr>
<tr>
<td>IK Pathan</td>
<td>14</td>
<td>16.50</td>
<td>10.68</td>
<td>0.65</td>
</tr>
<tr>
<td>AC Botha</td>
<td>22</td>
<td>18.32</td>
<td>12.30</td>
<td>0.67</td>
</tr>
<tr>
<td>MV Boucher</td>
<td>28</td>
<td>18.75</td>
<td>12.74</td>
<td>0.68</td>
</tr>
<tr>
<td>MEK Hussey</td>
<td>71</td>
<td>37.10</td>
<td>25.46</td>
<td>0.69</td>
</tr>
<tr>
<td>N Deonarine</td>
<td>14</td>
<td>29.36</td>
<td>20.35</td>
<td>0.69</td>
</tr>
<tr>
<td>GEF Barnett</td>
<td>11</td>
<td>20.27</td>
<td>13.93</td>
<td>0.69</td>
</tr>
<tr>
<td>RS Bopara</td>
<td>35</td>
<td>24.66</td>
<td>18.58</td>
<td>0.70</td>
</tr>
<tr>
<td>SE Bond</td>
<td>12</td>
<td>7.67</td>
<td>5.37</td>
<td>0.70</td>
</tr>
<tr>
<td>MS Dhoni</td>
<td>73</td>
<td>40.08</td>
<td>28.42</td>
<td>0.71</td>
</tr>
<tr>
<td>Misbah-ul-Haq</td>
<td>38</td>
<td>36.05</td>
<td>25.89</td>
<td>0.72</td>
</tr>
<tr>
<td>MF Maharoof</td>
<td>14</td>
<td>16.64</td>
<td>12.28</td>
<td>0.74</td>
</tr>
<tr>
<td>S Chanderpaul</td>
<td>38</td>
<td>41.00</td>
<td>30.65</td>
<td>0.75</td>
</tr>
<tr>
<td>HS Baidwan</td>
<td>14</td>
<td>12.57</td>
<td>9.39</td>
<td>0.75</td>
</tr>
<tr>
<td>SPD Smith</td>
<td>13</td>
<td>19.31</td>
<td>14.52</td>
<td>0.75</td>
</tr>
<tr>
<td>RR Sarwan</td>
<td>35</td>
<td>34.20</td>
<td>26.37</td>
<td>0.77</td>
</tr>
<tr>
<td>CJ Ferguson</td>
<td>24</td>
<td>27.50</td>
<td>21.09</td>
<td>0.77</td>
</tr>
<tr>
<td>CO Obuya</td>
<td>28</td>
<td>27.36</td>
<td>20.99</td>
<td>0.77</td>
</tr>
<tr>
<td>Asad Shafiq</td>
<td>12</td>
<td>22.17</td>
<td>16.99</td>
<td>0.77</td>
</tr>
<tr>
<td>Naved-ul-Hasan</td>
<td>10</td>
<td>16.50</td>
<td>12.68</td>
<td>0.77</td>
</tr>
<tr>
<td>AB de Villiers</td>
<td>55</td>
<td>46.18</td>
<td>35.87</td>
<td>0.78</td>
</tr>
<tr>
<td>CL White</td>
<td>58</td>
<td>31.62</td>
<td>24.81</td>
<td>0.78</td>
</tr>
<tr>
<td>JR Hopes</td>
<td>51</td>
<td>21.29</td>
<td>16.57</td>
<td>0.78</td>
</tr>
<tr>
<td>HM Amla</td>
<td>44</td>
<td>52.84</td>
<td>41.27</td>
<td>0.78</td>
</tr>
<tr>
<td>KD Karthik</td>
<td>24</td>
<td>28.25</td>
<td>21.96</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Table 9.6: Consistency coefficient (\( CV_1 \)) for 30 ODI batsmen (2008–2011)

Firstly, the \( CV \) method favours players who played a small number of matches—the top seven players all played fewer than 20 games. Secondly, it provides a mix of skilled and less skilled batsmen—11 of
the 30 batsmen average fewer than 20 runs per match, which is not considered to be at the level of a
top-class batsman. Moreover, quite a few bowlers are present in the list, for example, Shane Bond
from New Zealand, who averages 7.67 runs a match. A more useful list of consistent batters may be
generated by multiplying the denominator in $CV$ by the number of matches played, $n$, or:

$$CV_2 = \frac{\sigma}{\bar{x}n}$$  \hspace{1cm} (9.14)

<table>
<thead>
<tr>
<th>Player</th>
<th>Matches</th>
<th>Mean</th>
<th>Std_Dev</th>
<th>$CV_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEK Hussey</td>
<td>71</td>
<td>37.1</td>
<td>25.46</td>
<td>0.0097</td>
</tr>
<tr>
<td>MS Dhoni</td>
<td>73</td>
<td>40.08</td>
<td>28.42</td>
<td>0.0097</td>
</tr>
<tr>
<td>KC Sangakkara</td>
<td>75</td>
<td>38.97</td>
<td>31.54</td>
<td>0.0108</td>
</tr>
<tr>
<td>MJ Clarke</td>
<td>68</td>
<td>35.35</td>
<td>29.13</td>
<td>0.0121</td>
</tr>
<tr>
<td>CL White</td>
<td>58</td>
<td>31.62</td>
<td>24.81</td>
<td>0.0135</td>
</tr>
<tr>
<td>SK Raina</td>
<td>64</td>
<td>30.09</td>
<td>26.78</td>
<td>0.0139</td>
</tr>
<tr>
<td>RT Ponting</td>
<td>66</td>
<td>34.65</td>
<td>31.83</td>
<td>0.0139</td>
</tr>
<tr>
<td>AB de Villiers</td>
<td>55</td>
<td>46.18</td>
<td>35.87</td>
<td>0.0141</td>
</tr>
<tr>
<td>SR Watson</td>
<td>60</td>
<td>41.55</td>
<td>35.31</td>
<td>0.0142</td>
</tr>
<tr>
<td>G Gambhir</td>
<td>66</td>
<td>40.7</td>
<td>38.86</td>
<td>0.0145</td>
</tr>
<tr>
<td>Tamim Iqbal</td>
<td>70</td>
<td>32.27</td>
<td>33.06</td>
<td>0.0146</td>
</tr>
<tr>
<td>Yuvraj Singh</td>
<td>69</td>
<td>32.23</td>
<td>32.69</td>
<td>0.0147</td>
</tr>
<tr>
<td>DPMD Jayawardene</td>
<td>70</td>
<td>27.87</td>
<td>28.87</td>
<td>0.0148</td>
</tr>
<tr>
<td>JR Hopes</td>
<td>51</td>
<td>21.29</td>
<td>16.57</td>
<td>0.0153</td>
</tr>
<tr>
<td>PD Collingwood</td>
<td>56</td>
<td>30.41</td>
<td>26.04</td>
<td>0.0153</td>
</tr>
<tr>
<td>Mahmudullah</td>
<td>50</td>
<td>20.64</td>
<td>17.22</td>
<td>0.0167</td>
</tr>
<tr>
<td>JP Duminy</td>
<td>49</td>
<td>34.9</td>
<td>29.64</td>
<td>0.0173</td>
</tr>
<tr>
<td>HM Amla</td>
<td>44</td>
<td>52.84</td>
<td>41.27</td>
<td>0.0178</td>
</tr>
<tr>
<td>MJ Guptill</td>
<td>46</td>
<td>32.65</td>
<td>28.23</td>
<td>0.0188</td>
</tr>
<tr>
<td>Misbah-ul-Haq</td>
<td>38</td>
<td>36.05</td>
<td>25.89</td>
<td>0.0189</td>
</tr>
<tr>
<td>GC Smith</td>
<td>43</td>
<td>36.4</td>
<td>30.25</td>
<td>0.0193</td>
</tr>
<tr>
<td>S Chanderpaul</td>
<td>38</td>
<td>41</td>
<td>30.65</td>
<td>0.0197</td>
</tr>
<tr>
<td>JH Kallis</td>
<td>41</td>
<td>41.68</td>
<td>33.88</td>
<td>0.0198</td>
</tr>
<tr>
<td>RS Bopara</td>
<td>35</td>
<td>26.46</td>
<td>18.58</td>
<td>0.0201</td>
</tr>
<tr>
<td>DL Vettori</td>
<td>40</td>
<td>19.63</td>
<td>16.39</td>
<td>0.0209</td>
</tr>
<tr>
<td>RR Sarwan</td>
<td>35</td>
<td>34.2</td>
<td>26.37</td>
<td>0.0220</td>
</tr>
<tr>
<td>OA Shah</td>
<td>36</td>
<td>31.44</td>
<td>25.71</td>
<td>0.0227</td>
</tr>
<tr>
<td>Umar Akmal</td>
<td>33</td>
<td>31.67</td>
<td>24.99</td>
<td>0.0239</td>
</tr>
<tr>
<td>MV Boucher</td>
<td>28</td>
<td>18.75</td>
<td>12.74</td>
<td>0.0243</td>
</tr>
<tr>
<td>IR Bell</td>
<td>34</td>
<td>29.53</td>
<td>24.6</td>
<td>0.0245</td>
</tr>
</tbody>
</table>

Table 9.7: Consistency coefficient ($CV_2$) for 30 ODI batsmen (2008–2011)

Table 9.7 reveals a more impressive list of batsmen, with most considered to be among the world’s
best ODI batsmen. The $CV_2$ method isolates consistent players who have played more matches and
have higher batting averages than the $CV_1$ method—batsmen in the latter average 27.1 games and 26.8
runs, while in the former, 52.1 matches and 33.6 runs. While the batsmen in Table 9.6 are consistent
in their scoring, they are not necessarily good batsmen. There are still advantages to this; should a
coach be unsure about where to play two average batsmen in the order, he may insert the batsman
with the lowest $CV_1$ ahead of the other, in the knowledge that he is more likely to achieve his average.

9.7 Discussion

In this chapter, a batsman’s runs in an ODI cricket match were effectively estimated in a real-
time scenario using forward simulation methodology. A Visual Basic simulator, calling a conditional
probability matrix, described how likely a batsman was to score runs, or to be dismissed, at certain
stages in the match. The shape of these distributions was conditional on the type of the batsman and
on the delivery number at the point of execution. The most attractive feature of the simulator was its
ability to reassess a batsman’s scoring likelihoods from any point in a live match; this enabled log-
normal distribution fits on multiple iterations of Ricky Ponting’s innings in the 2011 ODI World Cup
quarterfinal. The distributions were appropriate for estimating likelihoods of Ponting reaching a
particular score. The in-play approach also improves the fit in the tails of log-normal score
distributions because there is added certainty surrounding the momentum of the player’s innings. The
in-play model goes some way to alleviating problems encountered when applying traditional
statistical models to batsmen’s scores, namely, how zero scores defy such models (Wood 1945). In
the simulations, once a batsman has scored a run, this problem is removed. Although only really
accurate for the first batsman in this research, the probability of a batsman of a particular order
achieving zero was assessed; future research will provide more certainty for batsmen of subsequent
orders. Further research should also investigate individual bowler effects on each batsmen’s simulated
scores (Swartz et al, 2009). Should a team’s best bowler replace a lesser bowler at a particular match
stage, the expectation would be for the dismissal likelihood to increase proportionally with the skill of
the batsman on strike. The additional work associated with the inclusion of bowler strength was
outside the scope of this dissertation, however remains an enticing concept.

The ever-increasing online gambling space provides ample opportunity for the application of
such developments, particularly wagering on markets such as lowest- and highest-scoring batsmen.
Bailey and Clarke (2004) designed strategies to maximise profits derived from wagering on one
batsman outscoring another (“head-to-head”) during the 2003 ODI World Cup. Moreover, with various betting markets established for wagering on total team and batsman runs while a match is in play, there is increasing focus on extending these methodologies to live scenarios. Easton and Uylangco (2007) even provided some evidence of the ability of market odds to predict the outcomes of impending deliveries in ODI matches. The in-play simulation methodology outlined in this chapter adds further appeal to wagering on markets such as “highest-scoring batsman” because there is a real-time knowledge of the match conditions with which to generate likelihoods of a batsman outscoring all others in his team; market inefficiencies are able to be detected in real-time match scenarios, not just through pre-match approximation. It is expected that this research will augment work on retrospective modelling of ODI batting scores for wagering purposes, such as that undertaken by Bailey and Clarke (2004). The authors anticipate quantifying the gain from the application of SimScore in the betting space to be a complex project, one which will be documented in future research papers. SimScore estimates are also expected to provide a valuable footing for real-time player-rating developments as it becomes unnecessary to wait until match completion to update a player’s rating.

Pollard et al (1977) inspired the central concept of the AFL research in Chapter 7, namely, that statistical modelling—specifically, the negative binomial distribution (nbd)—is sometimes more effective when applied to performance at a group level rather than at an individual one. While able to estimate individual batsmen’s scores at different stages of an ODI match with the conditional probabilities, an interesting development for this research would be to simulate partnership scores using the nbd. Sometimes in ODI matches, situations arise where the top batsmen have been dismissed and a middle-order batsman is batting with a tailender (bowler). While the conditional distribution would naturally assume that the bowler will achieve a low score and predict a quick end to the innings, it is almost always the case that these two batsmen will strategise to keep the bowler off strike and the more recognised batsman facing the opponent bowling as much as possible. This strategy may even require the two batsmen to not run after a delivery, where an easy run is available, to protect the bowler’s wicket. At the end of the over, the batsmen will do everything in his power to
run a single, or three runs, so the recognised batsman retains the strike at the change of ends. Should this tactic be successful, a fruitful partnership can result, one which may be overlooked by simulating the individual batsmen’s scores. There is support for the hypothesis that the nbd would have a better chance of predicting the lower-order partnership.
Chapter 10

Conclusions and Further Research

This dissertation has detailed a series of mathematical models and Visual Basic programming approaches that, in conjunction, have provided a pragmatic framework with which to rate the performance of players involved in continuous and discrete team sports, specifically, the Australian Football League (AFL) and limited overs cricket (ODI). The early chapters defined the different types of sports under observation as well as the player performance indicators relevant to each sport. A simple adjustive AFL player rating system (APR) was then investigated in Chapter 5 and was found to be an adequate predictor of player performance in match $t + 1$ but unable to answer the primary research question of whether a player’s contribution to a team’s success could be reasonably estimated. A network approach was employed in Chapter 7 and interactions between players simulated, with resulting centrality measures quantifying a player’s utility in a side. An ODI simulator, SimScore, produced Chapter 9, allowing a batsman’s contribution to be assessed then reassessed once in-play match information became available. The remainder of this chapter summarises each of the previous chapters and the opportunities for future research.

10.1 Player Performance

Chapter 4 introduced notational analysis—or the collection and/or analysis of ‘performance indicators’, or open skills—by analysts and coaches to assess the performance of an individual, team or elements of a team (Hughes and Bartlett, 2002). This was conceptualised by defining different
types of sports with an emphasis on “invasion” and “bat and ball” sports, the former describing AFL and the latter, ODI cricket. The AFL was classified as a “continuous” sport, or having the ability to (unrealistically) continue without any breaks in play for the entire match. Cricket was described as a “discrete” sport, where matches are comprised of a series of one-on-one events at close intervals between a bowler and batsman. For the AFL, any player performance indicator that resulted in a i) score, ii) successful pass or iii) turnover was termed a send, while any performance indicator that resulted from a send was termed a receive. The primary send indicators in an AFL game were kicks (KCK) and handballs (HBL), while primary receives were marks (MRK) and handball receives (HBR); ball gets (BG) were secondary receives. Using a t-test, these five player indicators were found to be significantly higher ($\alpha = 0.05$) for victorious teams than for losing teams. The indicators were not only quantified as individual actions but, also, as determinants of player interaction, where a $[KCK, MRK]$ or $[HBL, HBR]$ between players $i$ and $j$ were recognised as primary interactions. This early research was the foundation for the player rating systems explored in Chapter 5 and Chapter 7.

In cricket, the performance indicators of interest for each batsman were runs scored and batting average, while for the bowler, wickets taken, runs conceded and bowling average were investigated. There were far fewer notational descriptors in cricket than in AFL, given cricket’s higher frequency of scoring. Observations from this research suggested that sports with infrequent scoring have a greater focus on the actions that form the plays resulting in scoring opportunities. The number of runs scored / wickets taken is of far more interest than the technical fashion in which they are done, unlike in baseball where the rarity of run scoring encourages close technical scrutiny of the batters’ and pitchers’ mechanics in an attempt to predict how, when and why the next run will come. Also discussed were batter and bowler contributions, or the fraction of the total team runs/wickets a batsman/bowler achieved in a match.

10.2 Ratings Systems

In Chapter 5, we used notational analysis to develop an AFL player rating system. Player ratings were explained as being derived from evaluations of the performance of a team or individual
player, most often with prior performances in mind. Stefani (2012) mentioned the adjustive rating system produces ratings that increase, decrease or remain constant depending on above, below and met performance expectations respectively. An important attribute of these ratings was that expected team or player performance, $Exp$, was predictive in nature, providing a pre-match approximation of that team’s or player’s performance with respect to such match variables as field position and opponent strength. The Elo rating is arguably the most recognised adjustive system, originally developed for rating chess players but more recently adapted for rating football teams. This genesis of this doctorate was a discussion on adapting Elo ratings for AFL players. Certain assumptions were not satisfied for a direct application, but it was still possible to establish an adjustive AFL player rating system (APR) which shared some important features with the Elo system. Each player’s performance was measured by a linear equation that combined weighted performance indicator frequencies—achieved by each player during a single match—to arrive at a match “score”, $X_i$. Players were then pitched in simulated “head-to-head” contests where player $i$’s score and opponent $j$’s score were randomly generated from two independent normal score distributions from that season, prior to the impending match. Opponent $j$ represented a player in the opposition team, playing in the same game position (for example, midfielder) as $i$. After 1000 simulations, a pre-match expectation ($Exp$) of $i$ outscoring $j$ was generated, where a low $Exp$ value implied a stronger opponent and vice versa. A ratings increase or decrease was dependent on the difference between the observed result ($Obs$) of the player contest ($1 = \text{win}, \ 0.5 = \text{draw}, \ 0 = \text{loss}$) and $Exp$, subsequently weighted by the “size” of the win by that player against his opponent. A satisfactory linear relationship between the $Exp$ and $X_i$ values ($R^2 = 0.341$) suggested performance could be predicted with modest confidence. A case study was offered in which the Geelong Football Club’s midfielders were rated from the 2010 and 2011 seasons. These player ratings are of benefit to AFL and fantasy league coaches in player performance prediction and, also, guide betting on AFL player awards. AFL player performance consistency was also investigated as a rating system, inspired by a coefficient of variation that was employed by Elderton (1909) to measure the consistency of cricket batsmen’s scores.
It is anticipated that the predictive power of the APR model will be improved by some important additions that have been designated as future research due to the prioritised network methodology in Chapter 7. Ground advantage, travel effects and continued research into injury compensation will become a focus in an attempt to improve the $[\text{Exp}, X_t]$ relationship as well as the correlation between the final ratings and club champion voting. Further validation for the APR model might be the AFL Coaches Association voting system where votes are assigned on a “5,4,3,2,1” basis—five is the best performance—by the coaches of the two competing teams at the completion of their match. A common consensus exists among AFL players and pundits that this system is the most reliable subjective measure of AFL player performance, particularly because the votes are assigned by the coaches, experts who are directly involved in each match.

10.3 Positional Analysis

Chapter 6 investigated a problem with the positions of AFL players within our datasets, as allocated by the AFL fantasy site and the Prowess sports database. When calculating the ratings in Chapter 5, players’ positions, for example, Jimmy Bartel as a midfielder, were uniform for each of their matches across the entire season—we deemed this uniform allocation to be unrealistic. Prior to an AFL match, players are selected to play in particular defensive, attacking or midfield positions on the ground, but may play in different positions from week to week for a variety of reasons. This chapter demonstrated a retrospective position classification method to aid notational analysis that may be reliant on data that describes where a player was positioned on the ground. Specifically, the algorithm maximised the Mahalanobis distance between thirteen historical player performance indicators and their positional centroids to optimally allocate each player to one of four positions [Defence (D); Forward (F); Midfield (M); Ruck (R)] for any completed match. Moreover, Bayesian probabilities established a player’s “time spent” in each of the four positions in each of his matches; for example, a forward who briefly played as a defender in a match may be classified as a forward with 80% probability and a defender with 20% probability, given his recorded performance indicators. The APR model was rerun for the Geelong midfielders and found to improve the relationship between $\text{Exp}$ and $X_t$. Moreover, covariance between two performance indicators was investigated as a measure.
of linear coupling to describe performance indicator relationships within the allocated positions. By assessing the Squared Euclidean distances between these covariate couplets, the positions could be further segmented to enhance the knowledge of intra-positional performance characteristics.

Further research will focus on intra-positional classification, or employing distance techniques to reclassify midfielders into contested and uncontested possession players and noting any improvement in the predictive component of the APR ratings. An opportunity also exists to substitute the performance indicator frequencies with interaction data. We proved in this dissertation that the interaction data is a better predictor of margin than the individual performance variable frequencies, suggesting that an improvement in classification accuracy may also result.

10.4 Link Plays and Interaction Simulation

In Chapter 7, we attempted to answer the key research question behind this dissertation: is it possible to adequately estimate players’ contributions to their team’s performance in a match? Having decided that the APR ratings were too player-centric, lacking sufficient evidence of a player’s cooperation with his teammates, we decided to expand our notational analysis to include the transactions between the players, not just each player’s performance indicator frequencies. With access to every transaction in any AFL match, from the Prowess Sports database, we set about writing a looping Visual Basic program that could isolate periods of effective play between two or more players from team $a$, where the ball’s movement effectively increased that team’s scoring likelihood. Effective transactions were built from send-receive events, discussed in Chapter 4. Links continued until a score, dead play—ball out of bounds, for example—or a turnover was realised, with player membership, transaction description, ground position, match period and link length—number of transactions and duration of link—all recorded by the program, even generating a graphical output. The send and receive data fashioned an interaction matrix, revealing frequencies of player interaction. Network diagrams offered a graphic portrayal of the interactions between players in a match, with node diameter indicating the prominence of a player in the network and connecting line density proportional to the frequency of interactions between $i$ and $j$. The interaction matrix was then
symmetrised so that interactions between players could be simulated to estimate a player’s influence on a team’s network. Interactions between pairs of players followed a negative binomial distribution (Pollard et al, 1977), with parameters estimated using a Pearson chi-squared approach. Player performance in a match was quantified using eigenvector centrality, an important network metric, indicating a player’s level of interaction with other central players, while team strength was calculated by averaging the player’s centrality in the match network. The team index for any match was adequately related to the score margin for that match, making it possible to observe different players’ contribution to team performance (margin) when included and excluded from a simulated network.

Furthermore, team strength, rather than individual assessment, was more accomplished at predicting score margin using the network approach ($R^2 = 0.5302$), than the position-adjusted (Chapter 6) APR ratings from Chapter 5 ($R^2 = 0.3953$). This was a significant finding in the research, supporting an original hypothesis that grouping cooperative rather than individual performances would be a better indicator of a team’s chances of success. The importance of Jimmy Bartel’s presence in the 2011 grand final was treated as a case study; after running 1000 network simulations, the average winning margin when Bartel was in the chosen side ($\bar{x}_1 = 47.031$) was significantly greater than when replaced with a lesser quality player, in this case, Shannon Byrnes ($\bar{x}_2 = 31.960$).

The most immediate progression for the network research is to substitute the centrality metrics into the APR equation in an attempt to improve the correlation between the player ratings and real-life performances. Unfortunately, we did not possess the resources to be able to run the LINK program, associated nbd simulations and centrality generation on every team’s matches for the 2011 season. The data extraction is the main obstacle, so the consensus is that the project may require a programming platform beyond the capabilities of Visual Basic, one with a much faster run time.

Furthermore, this research did not consider the presence of covariance between elements in the interaction matrix. The initial stages of this research governed that each element was independent, even though degrees of interaction covariance between sets of $[i, j]$ are almost certain. The thousands of $[i, j]$ permutations and covariance between each $[i, j]$ has been designated as important future research. Ongoing research will also focus on improving the predictive power of the networks by
weighting the forms of player interactions detailed in Chapter 7 with respect to the levels of efficiency, scoring capacity, ground effects and, most importantly, an opponent strength indicator.

10.5 Nonparametric Performance Forecasting

In Chapter 8, an alternative player performance expectation methodology was assessed. We investigated a nonparametric smoothing approach to performance prediction, where player performance for match $t + 1$ assumed the form of the (forecasted) player performance score $X$, from Chapter 5. This approach proved especially useful when normality assumptions were violated under the APR approach and, also, in reducing noise in performance data sets and quelling the misleading effects of performance outliers. Each player’s season performance data was smoothed using median smoothing (Tukey, 1971), an optimised combination of running medians spanning one to five, with a $t + 1$ forecast produced by exponentially smoothing the optimal Tukey-smoothed series. The optimised Tukey smooth, unique to each player in a match, was decided upon by running optimisations on a combination of median orders—made possible in Excel by a smoothing macro allowing experimentation with median combinations—around each match score and an associated exponential smoothing parameter, $\theta$, to minimise the root mean-squared error (RMSE) of that player’s match forecasts. Three models for predicting performance scores were compared: naive averaging (Mean), exponential smoothing (EXP) and exponentially smoothed Tukey smoothing (T-EXP). The Mean and EXP approach revealed very similar RMSE values per player. T-EXP reduced RMSE per player by 16.2% when compared to the Mean model. The Geelong midfielders’ centrality was also forecasted for the 2011 grand final, with mean absolute error rates recorded after the observed centrality figures were available, again proving the worth of the T-EXP method: Mean = 7.8; EXP = 7.1; T-EXP = 6.0.

This chapter was an introduction on how Tukey’s smoothing can benefit sport forecasting models at the base level. An important development for the T-EXP forecasting model will be the inclusion of a parameter that accounts for any opponent effects that increase or decrease the likelihood of an individual player achieving approximately his average score. Travel effects must also be investigated; travel takes its toll on the players, particularly an eastern-based team that must fly to
Western Australia and back to eastern Australia in the same week—a return flight from Melbourne to Perth is roughly 8.5 hours.

10.6 Cricket

Chapter 9 was dedicated to an ODI cricket simulator, developed for in-play prediction. It was possible to simulate any batsman’s runs at any point during a match using a Visual Basic program, SimScore, which called a conditional probability matrix that described the ball-by-ball likelihood of runs or a dismissal. These probabilities were dependent on the match state (number of dismissals and deliveries bowled) and the skill of the non-dismissed batsmen (currently batting or yet to come). Our frequency distribution provided 8,100 unique match states and, in turn, 56,700 independent raw probabilities for a regular match. A fourth-order polynomial smooth was applied to the probabilities when missing values and extremely small sample sizes for certain $X$ outcomes produced unrealistic zero and one raw probabilities. SimScore could be executed at any match stage and produced a run estimate for non-dismissed batsmen, which was adjusted after simulation for team strength, innings and venue effects using a generalised linear procedure. This chapter ultimately demonstrated the benefits of simulation by fitting log-normal distributions to 500 score iterations by Australia’s Ricky Ponting in the 2011 ODI World Cup quarterfinal, at the 0-, 10-, 20- and 30-over mark of the match, providing a statistical context for his observed score. The probability of Ponting reaching his observed score of 104 increased at each interval as more certainty in the outcome had been established by the in-play conditional estimations. The in-play methodology provided a solution to a long-running issue in modelling batsmen’s scores: that traditional statistical distributions, such as the exponential distribution, provide inadequate fits for zero scores (zero is the most frequent score in cricket). By waiting until a batsman scores a run and then simulating, the problem is avoided.

The ever-increasing online gambling space provides ample opportunity for further research into SimScore’s ability to guide wagering on markets, for example the lowest- and highest-scoring batsmen. As information comes to hand during a match, the program can be run and re-run to produce odds which can help to identify betting market inefficiencies. The negative binomial distribution
(nbd) will also be investigated as an alternative to the conditional probability matrix detailed in this chapter. The influence Pollard, Benjamin and Reep (1977) and Reep and Benjamin (1968) have had on this dissertation is undeniable. Further research into in-play nbd modeling is a legacy to their outstanding work and is anticipated to be highly fruitful for SimScore, particularly if we decide to model partnerships. Bowler effects may also prove to be significant. The probability of a particular bowler dismissing a particular batter may differ considerably to the generic probabilities in our matrix. It is well known that some batters are more susceptible to spin bowling than pace bowling, for example, which can be modelled using our database with relative ease.

10.7 Summary

The research outlined in this dissertation was the result of a series of (mostly) productive ideas and their application, born from a passion for sport and statistics and driven by an opportunity to provide the sport statistics field and the AFL and cricket communities with information that can educate and guide decision-making in a professional and recreational capacity. The APR system may not have satisfied the ultimate aim of this research, but the utility for fantasy league coaches is conclusive. The network measurements—although yet to be trialled in an adjutant system—are exciting, with interaction weighting and opponent effects expected to further improve the model. There is massive potential for these player ratings, not just for use in the AFL coaching boxes but in any continuous sport where network methodologies are applicable, for example, basketball and soccer. The publically available ICC cricket team and player ratings, and Sabermetrics, the empirical study of baseball statistics, suggests discrete sports are more conclusively analysed than continuous sports. The ODI cricket SimScore program from this research offers an in-play batsman performance expectation methodology from which live player ratings can be derived, moving beyond conventional rating systems which are updated with completed match data. Much debate surrounds the future of ODI cricket, with suggestions it may be phased out to make way for T20 cricket, the 20-over format. The SimScore program is flexible enough to translate to the shortened format.
Bibliography


Chapter 11

Appendix

11.1 AFL Performance Indicator Glossary

The following is a list of AFL performance indicators included in the relevant analyses in this dissertation. The abbreviations are in brackets:

Kick (KCK): striking the ball with the foot in an attempt to pass the ball to a teammate.

Mark (MRK): catching a kicked ball without it hitting the ground.

Handball (HBL): striking the ball with a closed hand in an attempt to pass the ball to a teammate.

Handball Receive (HBR): catching a handball without the ball hitting the ground.

Ball Get (BG): gathering the in-dispute ball into the hands. A loose ball get (LBG) is gathering the ball in an uncontested situation while a hard ball get (HBG) is gathering the ball in a contested situation.

Goal (GLS): the result of the ball being kicked between the two taller posts; six points is awarded.

Behind (BHS): the result of the ball being kicked between the shorter and taller posts; one point is awarded.
Hit Out (HIT): tapping the ball, that has been bounced or thrown in to restart play, with a hand, down to the advantage of a teammate.

Tackle (TKL): wrapping the hands and arms around an opposition player to prevent him from advancing with the ball; the arms must be above the opponent’s knees and below his shoulders.

Inside 50 (I50): kicking, handballing or carrying the ball into one’s attacking zone (forward 50 metre arc).

Rebound 50 (R50): kicking, handballing or carrying the ball out of one’s defensive zone (opponent’s 50 metre arc).

Kick In (KIN): a kick by a player from team a from team b’s goal square to restart play after a behind (one point) has been registered by team b.

Knock On (KNK): hitting a ball that is “in dispute” with the hand to the advantage of yourself or a teammate.

Spoil (SPL): punching the ball in a marking contest to prevent an opposition player from the marking the ball.

Smother (SMT): using the arms and/or hands to block the ball as it is being kicked by an opposition player.

Forward Target (TAR): attributed to a forward when the ball is kicked to him to execute a shot on goal.

Free Kick For (FF): a free kick awarded to player i.

Free Kick Against (FA): a free kick awarded against player i.

36 No team has possession of the ball
## 11.2 APR Midfielder Ratings

The following table displays the top 100 rated midfielders across all teams in the AFL at the end of the 2011 season, as determined by the APR model outlined in Chapter 5. The rating period commenced at the start of the 2010 season.

<table>
<thead>
<tr>
<th>Player</th>
<th>Team</th>
<th>Opp</th>
<th>Round</th>
<th>Score</th>
<th>Score</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marc Murphy</td>
<td>CAR</td>
<td>STK</td>
<td>24</td>
<td>105</td>
<td>0.74</td>
<td>39.667</td>
</tr>
<tr>
<td>Scott Pendlebury</td>
<td>COL</td>
<td>GEE</td>
<td>24</td>
<td>106</td>
<td>0.65</td>
<td>50.909</td>
</tr>
<tr>
<td>Nick Dal Santo</td>
<td>STK</td>
<td>CAR</td>
<td>24</td>
<td>75</td>
<td>0.62</td>
<td>22.833</td>
</tr>
<tr>
<td>Dayne Beams</td>
<td>COL</td>
<td>GEE</td>
<td>24</td>
<td>129</td>
<td>0.47</td>
<td>33.909</td>
</tr>
<tr>
<td>Daniel Wells</td>
<td>NOR</td>
<td>STK</td>
<td>23</td>
<td>76</td>
<td>0.62</td>
<td>39.300</td>
</tr>
<tr>
<td>Dane Swan</td>
<td>COL</td>
<td>GEE</td>
<td>24</td>
<td>84</td>
<td>0.79</td>
<td>30.909</td>
</tr>
<tr>
<td>Jack Redden</td>
<td>BRI</td>
<td>SYD</td>
<td>24</td>
<td>86</td>
<td>0.67</td>
<td>36.200</td>
</tr>
<tr>
<td>Joel Selwood</td>
<td>GEE</td>
<td>COL</td>
<td>24</td>
<td>118</td>
<td>0.67</td>
<td>17.636</td>
</tr>
<tr>
<td>David Mundy</td>
<td>FRE</td>
<td>NOR</td>
<td>22</td>
<td>93</td>
<td>0.57</td>
<td>87.273</td>
</tr>
<tr>
<td>Sam Mitchell</td>
<td>HAW</td>
<td>GST</td>
<td>24</td>
<td>90</td>
<td>0.77</td>
<td>17.667</td>
</tr>
<tr>
<td>Luke Ball</td>
<td>COL</td>
<td>GEE</td>
<td>24</td>
<td>82</td>
<td>0.49</td>
<td>2.091</td>
</tr>
<tr>
<td>Jobe Watson</td>
<td>ESS</td>
<td>POR</td>
<td>23</td>
<td>116</td>
<td>0.74</td>
<td>14.111</td>
</tr>
<tr>
<td>Kieren Jack</td>
<td>SYD</td>
<td>BRI</td>
<td>24</td>
<td>79</td>
<td>0.51</td>
<td>4.455</td>
</tr>
<tr>
<td>Andrew Swallow</td>
<td>NOR</td>
<td>RIC</td>
<td>24</td>
<td>86</td>
<td>0.69</td>
<td>71.333</td>
</tr>
<tr>
<td>Brad Sewell</td>
<td>HAW</td>
<td>WBG</td>
<td>23</td>
<td>100</td>
<td>0.46</td>
<td>4.818</td>
</tr>
<tr>
<td>Mitch Robinson</td>
<td>CAR</td>
<td>STK</td>
<td>24</td>
<td>95</td>
<td>0.48</td>
<td>8.333</td>
</tr>
<tr>
<td>Chris Judd</td>
<td>CAR</td>
<td>STK</td>
<td>24</td>
<td>79</td>
<td>0.64</td>
<td>9.333</td>
</tr>
<tr>
<td>Ryan Griffen</td>
<td>WBG</td>
<td>FRE</td>
<td>24</td>
<td>110</td>
<td>0.76</td>
<td>43.667</td>
</tr>
<tr>
<td>James Kelly</td>
<td>GEE</td>
<td>COL</td>
<td>24</td>
<td>101</td>
<td>0.57</td>
<td>30.636</td>
</tr>
<tr>
<td>Luke Hodge</td>
<td>HAW</td>
<td>WBG</td>
<td>23</td>
<td>101</td>
<td>0.58</td>
<td>18.818</td>
</tr>
<tr>
<td>Liam Shiels</td>
<td>HAW</td>
<td>GST</td>
<td>24</td>
<td>90</td>
<td>0.61</td>
<td>1.667</td>
</tr>
<tr>
<td>Matt Priddis</td>
<td>WCE</td>
<td>ADE</td>
<td>24</td>
<td>121</td>
<td>0.73</td>
<td>37.500</td>
</tr>
<tr>
<td>Jordan Lewis</td>
<td>HAW</td>
<td>WBG</td>
<td>23</td>
<td>114</td>
<td>0.54</td>
<td>50.818</td>
</tr>
<tr>
<td>Ben Howlett</td>
<td>ESS</td>
<td>POR</td>
<td>23</td>
<td>91</td>
<td>0.58</td>
<td>46.111</td>
</tr>
<tr>
<td>Lenny Hayes</td>
<td>STK</td>
<td>RIC</td>
<td>24</td>
<td>65</td>
<td>0.93</td>
<td>31.818</td>
</tr>
<tr>
<td>Josh Kennedy</td>
<td>SYD</td>
<td>BRI</td>
<td>24</td>
<td>111</td>
<td>0.42</td>
<td>13.455</td>
</tr>
<tr>
<td>David Zaharakis</td>
<td>ESS</td>
<td>POR</td>
<td>23</td>
<td>133</td>
<td>0.58</td>
<td>35.111</td>
</tr>
<tr>
<td>Kade Simpson</td>
<td>CAR</td>
<td>STK</td>
<td>24</td>
<td>163</td>
<td>0.63</td>
<td>52.667</td>
</tr>
<tr>
<td>Nathan Jones</td>
<td>MEL</td>
<td>POR</td>
<td>24</td>
<td>76</td>
<td>0.60</td>
<td>7.000</td>
</tr>
<tr>
<td>Heath Hocking</td>
<td>ESS</td>
<td>WCE</td>
<td>22</td>
<td>112</td>
<td>0.51</td>
<td>36.700</td>
</tr>
<tr>
<td>Jake Melksham</td>
<td>ESS</td>
<td>POR</td>
<td>23</td>
<td>86</td>
<td>0.43</td>
<td>16.111</td>
</tr>
<tr>
<td>Daniel Hannebery</td>
<td>SYD</td>
<td>BRI</td>
<td>24</td>
<td>92</td>
<td>0.40</td>
<td>2.545</td>
</tr>
<tr>
<td>Simon Black</td>
<td>BRI</td>
<td>SYD</td>
<td>24</td>
<td>63</td>
<td>0.66</td>
<td>5.200</td>
</tr>
<tr>
<td>Jack Ziebell</td>
<td>NOR</td>
<td>STK</td>
<td>23</td>
<td>109</td>
<td>0.46</td>
<td>71.700</td>
</tr>
<tr>
<td>Joel Corey</td>
<td>GEE</td>
<td>ADE</td>
<td>21</td>
<td>81</td>
<td>0.64</td>
<td>15.714</td>
</tr>
<tr>
<td>Jarrad McVeigh</td>
<td>SYD</td>
<td>BRI</td>
<td>24</td>
<td>98</td>
<td>0.45</td>
<td>18.455</td>
</tr>
<tr>
<td>Steele Sidebottom</td>
<td>COL</td>
<td>GEE</td>
<td>24</td>
<td>112</td>
<td>0.44</td>
<td>7.091</td>
</tr>
<tr>
<td>Ben Johnson</td>
<td>COL</td>
<td>GEE</td>
<td>24</td>
<td>73</td>
<td>0.38</td>
<td>28.091</td>
</tr>
<tr>
<td>Jack Trengove</td>
<td>MEL</td>
<td>POR</td>
<td>24</td>
<td>104</td>
<td>0.65</td>
<td>16.000</td>
</tr>
<tr>
<td>Domenic Cassisi</td>
<td>POR</td>
<td>MEL</td>
<td>24</td>
<td>84</td>
<td>0.60</td>
<td>15.333</td>
</tr>
<tr>
<td>Jude Bolton</td>
<td>SYD</td>
<td>BRI</td>
<td>24</td>
<td>78</td>
<td>0.56</td>
<td>41.455</td>
</tr>
<tr>
<td>Shane Savage</td>
<td>HAW</td>
<td>GST</td>
<td>24</td>
<td>68</td>
<td>0.48</td>
<td>17.333</td>
</tr>
<tr>
<td>Scott Selwood</td>
<td>WCE</td>
<td>ADE</td>
<td>24</td>
<td>114</td>
<td>0.61</td>
<td>15.500</td>
</tr>
<tr>
<td>Matthew Pavlich</td>
<td>FRE</td>
<td>NOR</td>
<td>22</td>
<td>39</td>
<td>0.71</td>
<td>14.727</td>
</tr>
<tr>
<td>Dustin Martin</td>
<td>RIC</td>
<td>NOR</td>
<td>24</td>
<td>91</td>
<td>0.57</td>
<td>21.333</td>
</tr>
<tr>
<td>Leigh Adams</td>
<td>NOR</td>
<td>RIC</td>
<td>24</td>
<td>119</td>
<td>0.54</td>
<td>15.333</td>
</tr>
</tbody>
</table>
11.3 Visual Basic Code

Visual Basic for Excel was a critical component of this research, allowing complex looping procedures that could circumvent the use of cumbersome simulation and optimisation software. The following examples were recognised as the most important to the direction of this research. The code that generated the link plays was the genesis of the AFL network analysis in Chapter 7. Figure 11.1 is a screenshot of the first routine: establishing the primary link play variables. Figure 11.2 is a sample of the code that performs eigenvector centrality calculations as part of the AFL network simulations. Figure 11.3 is a sample of the nonlinear smoothing code provided for public use by Quantitative Decisions (www.quantdec.com). This was executed in Chapter 8 as an alternative to the Exp methodology in Chapter 5. Finally, Figure 11.4 is a sample of the SimScore code that performed in-play batsman run simulations in Chapter 9. The additional feature of this code was the ability to import live match scenarios with which to reassess batsman and team runs by simulation.
Figure 11.1: Link play isolation code
Private Sub SolveButton_Click()
' This code is attached to the "Run Simulations" button on the main menu.
' It runs a specified number of network simulations and calculates eigenvector
' centrality for each player in the network with a power iteration method.

For i = 1 To SimNo
' Sets up random number generator
Worksheets("Interactions").Select
Range("BY2").Select
ActiveCell.FormulaR1C1 = “=MATCH(RAND(),RC[-21]:RC[-13],1)-1”
Range("BY2").Select
Selection.AutoFill Destination:=Range("BY2:BY1000")
Range("BY2:BY1005").Select

' Applies random numbers to the simulation
Range("BY2:BY1005").Copy
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
1= False, Transpose:=False
Sheets("Symmetric").Select

' Resets the eigenvector
Range("A1A1:A25").Value = 1
Range("A25").Value = Range("A25").Value
Range("A25").Value = ""

' Power iterations to find largest eigenvalue
Do
Range("A25").Value = Range("A25:A25").Value
Range("A25").Value = Range("A25:A25").Value
Loop Until Range("A25").Value = Range("A25").Value

' Logs margin for each simulation
If LogSim = True Then
Range("I25").Copy
Worksheets("Simulations").Select
Range("L1").End(xlDown).Offset(1, 0).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
1= False, Transpose:=False
End If

' Logs total interactions for each simulation
If LogTotal = True Then
Worksheets("Simulations").Select
Range("Q1").End(xlDown).Offset(1, 0).Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
1= False, Transpose:=False
End If
End Sub

Figure 11.2: Eigenvector centrality code
Figure 11.3: “Tukey” smoothing code
Figure 11.4: SimScore code
Other Player Rating Systems

The following table is a list of the top 50 ranked batsmen in ODI cricket as determined by the ICC player ratings as at November, 2012. Note the similar rating values as the APR system. It could be assumed that each player starts with a rating of 500.

<table>
<thead>
<tr>
<th>ID</th>
<th>Rating</th>
<th>Name</th>
<th>Nation</th>
<th>Career Best Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>901</td>
<td>H.M. Amla</td>
<td>SA</td>
<td>901 v England, 05/09/2012</td>
</tr>
<tr>
<td>2</td>
<td>858</td>
<td>V. Kohli</td>
<td>IND</td>
<td>866 v Sri Lanka, 31/07/2012</td>
</tr>
<tr>
<td>3</td>
<td>852</td>
<td>A.B. de Villiers</td>
<td>SA</td>
<td>852 v England, 05/09/2012</td>
</tr>
<tr>
<td>4</td>
<td>749</td>
<td>I.J.L. Trott</td>
<td>ENG</td>
<td>781 v India, 23/10/2011</td>
</tr>
<tr>
<td>5</td>
<td>746</td>
<td>R.C. Sangakkara</td>
<td>SL</td>
<td>781 v India, 28/07/2012</td>
</tr>
<tr>
<td>6</td>
<td>745</td>
<td>M.S. Dhoni</td>
<td>IND</td>
<td>836 v Australia, 31/10/2009</td>
</tr>
<tr>
<td>7</td>
<td>718</td>
<td>M.J. Clarke</td>
<td>AUS</td>
<td>750 v Sri Lanka, 22/02/2008</td>
</tr>
<tr>
<td>8</td>
<td>707</td>
<td>A.N. Cook</td>
<td>ENG</td>
<td>752 v West Indies, 19/06/2012</td>
</tr>
<tr>
<td>9</td>
<td>689</td>
<td>G. Gambhir</td>
<td>IND</td>
<td>722 v New Zealand, 04/12/2010</td>
</tr>
<tr>
<td>10</td>
<td>688</td>
<td>M.E.K. Hussey</td>
<td>AUS</td>
<td>857 v New Zealand, 28/01/2007</td>
</tr>
<tr>
<td>11</td>
<td>683</td>
<td>Shakib Al Hasan</td>
<td>BAN</td>
<td>683 v Pakistan, 22/03/2012</td>
</tr>
<tr>
<td>12</td>
<td>682</td>
<td>T.M. Dilshan</td>
<td>SL</td>
<td>765 v New Zealand, 29/03/2011</td>
</tr>
<tr>
<td>13</td>
<td>680</td>
<td>S.R. Watson</td>
<td>AUS</td>
<td>773 v Sri Lanka, 14/08/2011</td>
</tr>
<tr>
<td>14</td>
<td>674</td>
<td>E.J.G. Morgan</td>
<td>ENG</td>
<td>690 v Australia, 30/06/2010</td>
</tr>
<tr>
<td>15</td>
<td>674</td>
<td>Umar Akmal</td>
<td>PAK</td>
<td>702 v Afghanistan, 10/02/2012</td>
</tr>
<tr>
<td>16</td>
<td>667</td>
<td>B.R.M. Taylor</td>
<td>ZIM</td>
<td>667 v New Zealand, 09/02/2012</td>
</tr>
<tr>
<td>17</td>
<td>664</td>
<td>J.H. Kallis</td>
<td>SA</td>
<td>817 v West Indies, 04/02/2004</td>
</tr>
<tr>
<td>18</td>
<td>661</td>
<td>R.L. Taylor</td>
<td>NZ</td>
<td>693 v Australia, 10/02/2009</td>
</tr>
<tr>
<td>19</td>
<td>641</td>
<td>P.R. Stirling</td>
<td>IRE</td>
<td>650 v Scotland, 12/07/2011</td>
</tr>
<tr>
<td>20</td>
<td>636</td>
<td>J.P. Duminy</td>
<td>SA</td>
<td>668 v Netherlands, 03/03/2011</td>
</tr>
<tr>
<td>21</td>
<td>634</td>
<td>V. Sehwag</td>
<td>IND</td>
<td>774 v New Zealand, 11/01/2003</td>
</tr>
<tr>
<td>22</td>
<td>633</td>
<td>C.H. Gayle</td>
<td>WI</td>
<td>804 v Zimbabwe, 30/11/2003</td>
</tr>
<tr>
<td>23</td>
<td>633</td>
<td>D.P.M.D. Jayawardene</td>
<td>SL</td>
<td>738 v West Indies, 19/12/2001</td>
</tr>
<tr>
<td>24</td>
<td>628</td>
<td>G.C. Smith</td>
<td>SA</td>
<td>784 v India, 25/11/2005</td>
</tr>
<tr>
<td>25</td>
<td>626</td>
<td>B.B. McCullum</td>
<td>NZ</td>
<td>664 v Pakistan, 09/11/2009</td>
</tr>
<tr>
<td>26</td>
<td>622</td>
<td>M.J. Guptill</td>
<td>NZ</td>
<td>684 v Zimbabwe, 09/02/2012</td>
</tr>
<tr>
<td>27</td>
<td>618</td>
<td>Tamim Iqbal</td>
<td>BAN</td>
<td>629 v England, 28/02/2010</td>
</tr>
<tr>
<td>28</td>
<td>617</td>
<td>Misbah-ul-Haq</td>
<td>PAK</td>
<td>623 v West Indies, 28/04/2011</td>
</tr>
<tr>
<td>29</td>
<td>612</td>
<td>S. Raina</td>
<td>IND</td>
<td>658 v Sri Lanka, 30/05/2010</td>
</tr>
<tr>
<td>31</td>
<td>591</td>
<td>I.R. Bell</td>
<td>ENG</td>
<td>702 v India, 27/08/2007</td>
</tr>
<tr>
<td>32</td>
<td>588</td>
<td>L.D. Chandimal</td>
<td>SL</td>
<td>667 v Australia, 06/03/2012</td>
</tr>
<tr>
<td>33</td>
<td>588</td>
<td>K.P. Pietersen</td>
<td>ENG</td>
<td>833 v Australia, 08/04/2007</td>
</tr>
<tr>
<td>34</td>
<td>581</td>
<td>K.A. Pollard</td>
<td>WI</td>
<td>596 v Australia, 25/03/2012</td>
</tr>
<tr>
<td>35</td>
<td>575</td>
<td>D.J. Hussey</td>
<td>AUS</td>
<td>630 v Sri Lanka, 02/03/2012</td>
</tr>
<tr>
<td>36</td>
<td>571</td>
<td>A.D. Russell</td>
<td>WI</td>
<td>571 v New Zealand, 16/07/2012</td>
</tr>
<tr>
<td>37</td>
<td>569</td>
<td>A.D. Mathews</td>
<td>SL</td>
<td>598 v Australia, 14/08/2011</td>
</tr>
<tr>
<td>38</td>
<td>568</td>
<td>D.A. Warner</td>
<td>AUS</td>
<td>589 v England, 29/06/2012</td>
</tr>
<tr>
<td>40</td>
<td>559</td>
<td>W.U. Tharanga</td>
<td>SL</td>
<td>663 v New Zealand, 28/12/2006</td>
</tr>
<tr>
<td>41</td>
<td>558</td>
<td>L.M.P. Simmons</td>
<td>WI</td>
<td>629 v India, 02/12/2011</td>
</tr>
<tr>
<td>42</td>
<td>557</td>
<td>K.J. O’Brien</td>
<td>IRE</td>
<td>582 v England, 02/03/2011</td>
</tr>
<tr>
<td>43</td>
<td>553</td>
<td>Mohammad Hafeez</td>
<td>PAK</td>
<td>583 v Sri Lanka, 18/11/2011</td>
</tr>
<tr>
<td>44</td>
<td>547</td>
<td>C. Kieswetter</td>
<td>ENG</td>
<td>547 v South Africa, 05/09/2012</td>
</tr>
<tr>
<td>45</td>
<td>539</td>
<td>Shahid Afridi</td>
<td>PAK</td>
<td>663 v India, 02/10/1997</td>
</tr>
<tr>
<td>46</td>
<td>538</td>
<td>T.L.W. Cooper</td>
<td>NED</td>
<td>576 v West Indies, 28/02/2011</td>
</tr>
</tbody>
</table>
Table 11.2: ICC ranked batsmen as at November, 2012.

This table is a list of the top 50 ranked soccer players in the Major Soccer League as determined by the Castrol player ratings as at September, 2013. Note the similar rating values as the ICC and APR system.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Player</th>
<th>Team</th>
<th>Minutes</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Marco Di Vaio</td>
<td>Montreal Impact</td>
<td>2653</td>
<td>819</td>
</tr>
<tr>
<td>2</td>
<td>Mike Magee</td>
<td>Chicago Fire</td>
<td>2700</td>
<td>803</td>
</tr>
<tr>
<td>3</td>
<td>Donovan Ricketts</td>
<td>Portland Timbers</td>
<td>2790</td>
<td>781</td>
</tr>
<tr>
<td>4</td>
<td>Robbie Keane</td>
<td>LA Galaxy</td>
<td>1892</td>
<td>770</td>
</tr>
<tr>
<td>5</td>
<td>Chris Wondolowski</td>
<td>San Jose Earthquakes</td>
<td>2520</td>
<td>766</td>
</tr>
<tr>
<td>6</td>
<td>Jack McInerney</td>
<td>Philadelphia Union</td>
<td>2178</td>
<td>743</td>
</tr>
<tr>
<td>7</td>
<td>Eddie Johnson</td>
<td>Seattle Sounders FC</td>
<td>1648</td>
<td>742</td>
</tr>
<tr>
<td>8</td>
<td>Blas Pérez</td>
<td>FC Dallas</td>
<td>1930</td>
<td>742</td>
</tr>
<tr>
<td>9</td>
<td>Nick Rimando</td>
<td>Real Salt Lake</td>
<td>2340</td>
<td>739</td>
</tr>
<tr>
<td>10</td>
<td>Diego Fagundez</td>
<td>New England Revolution</td>
<td>2335</td>
<td>728</td>
</tr>
<tr>
<td>11</td>
<td>Tim Cahill</td>
<td>New York Red Bulls</td>
<td>2152</td>
<td>727</td>
</tr>
<tr>
<td>12</td>
<td>Graham Zusi</td>
<td>Sporting Kansas City</td>
<td>2266</td>
<td>699</td>
</tr>
<tr>
<td>13</td>
<td>Lamar Neagle</td>
<td>Seattle Sounders FC</td>
<td>2324</td>
<td>690</td>
</tr>
<tr>
<td>14</td>
<td>Luis Robles</td>
<td>New York Red Bulls</td>
<td>2970</td>
<td>683</td>
</tr>
<tr>
<td>15</td>
<td>Thierry Henry</td>
<td>New York Red Bulls</td>
<td>2434</td>
<td>670</td>
</tr>
<tr>
<td>16</td>
<td>Will Bruin</td>
<td>Houston Dynamo</td>
<td>2505</td>
<td>668</td>
</tr>
<tr>
<td>17</td>
<td>Drew Moor</td>
<td>Colorado Rapids</td>
<td>2703</td>
<td>660</td>
</tr>
<tr>
<td>18</td>
<td>Ryan Johnson</td>
<td>Portland Timbers</td>
<td>1740</td>
<td>659</td>
</tr>
<tr>
<td>19</td>
<td>Camilo Da Silva Sanvezzo</td>
<td>Vancouver Whitecaps</td>
<td>2325</td>
<td>657</td>
</tr>
<tr>
<td>20</td>
<td>José Gonçalves</td>
<td>New England Revolution</td>
<td>2970</td>
<td>623</td>
</tr>
<tr>
<td>21</td>
<td>Dominic Oduro</td>
<td>Columbus Crew</td>
<td>2762</td>
<td>622</td>
</tr>
<tr>
<td>22</td>
<td>Giles Barnes</td>
<td>Houston Dynamo</td>
<td>2475</td>
<td>619</td>
</tr>
<tr>
<td>23</td>
<td>Conor Casey</td>
<td>Philadelphia Union</td>
<td>2108</td>
<td>618</td>
</tr>
<tr>
<td>24</td>
<td>Shea Salinas</td>
<td>San Jose Earthquakes</td>
<td>1889</td>
<td>618</td>
</tr>
<tr>
<td>25</td>
<td>Jon Busch</td>
<td>San Jose Earthquakes</td>
<td>2970</td>
<td>616</td>
</tr>
<tr>
<td>26</td>
<td>Zac MacMath</td>
<td>Philadelphia Union</td>
<td>2970</td>
<td>608</td>
</tr>
<tr>
<td>27</td>
<td>Aurélien Collin</td>
<td>Sporting Kansas City</td>
<td>2520</td>
<td>607</td>
</tr>
<tr>
<td>28</td>
<td>Marcelo Sarvas</td>
<td>LA Galaxy</td>
<td>2803</td>
<td>604</td>
</tr>
<tr>
<td>29</td>
<td>Benny Feilhaber</td>
<td>Sporting Kansas City</td>
<td>1770</td>
<td>600</td>
</tr>
<tr>
<td>30</td>
<td>Federico Higuain</td>
<td>Columbus Crew</td>
<td>2593</td>
<td>599</td>
</tr>
<tr>
<td>31</td>
<td>Diego Valeri</td>
<td>Portland Timbers</td>
<td>2228</td>
<td>599</td>
</tr>
<tr>
<td>32</td>
<td>Clinton Irwin</td>
<td>Colorado Rapids</td>
<td>2784</td>
<td>597</td>
</tr>
<tr>
<td>33</td>
<td>Nathan Sturgis</td>
<td>Colorado Rapids</td>
<td>1946</td>
<td>597</td>
</tr>
<tr>
<td>34</td>
<td>Deshorn Brown</td>
<td>Colorado Rapids</td>
<td>2089</td>
<td>593</td>
</tr>
<tr>
<td>35</td>
<td>Kyle Beckerman</td>
<td>Real Salt Lake</td>
<td>2191</td>
<td>592</td>
</tr>
<tr>
<td>36</td>
<td>Matt Besler</td>
<td>Sporting Kansas City</td>
<td>1980</td>
<td>588</td>
</tr>
<tr>
<td>37</td>
<td>Brad Davis</td>
<td>Houston Dynamo</td>
<td>2112</td>
<td>586</td>
</tr>
<tr>
<td>38</td>
<td>Dilly Duka</td>
<td>Chicago Fire</td>
<td>1714</td>
<td>586</td>
</tr>
<tr>
<td></td>
<td>Player</td>
<td>Club</td>
<td>Rating</td>
<td>Overall</td>
</tr>
<tr>
<td>---</td>
<td>------------------------</td>
<td>---------------------</td>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>39</td>
<td>Javier Morales</td>
<td>Real Salt Lake</td>
<td>2236</td>
<td>578</td>
</tr>
<tr>
<td>40</td>
<td>George John</td>
<td>FC Dallas</td>
<td>2070</td>
<td>577</td>
</tr>
<tr>
<td>41</td>
<td>Omar González</td>
<td>LA Galaxy</td>
<td>2277</td>
<td>577</td>
</tr>
<tr>
<td>42</td>
<td>Chad Marshall</td>
<td>Columbus Crew</td>
<td>2547</td>
<td>575</td>
</tr>
<tr>
<td>43</td>
<td>Sebastien Le Toux</td>
<td>Philadelphia Union</td>
<td>2204</td>
<td>574</td>
</tr>
<tr>
<td>44</td>
<td>Amobi Okugo</td>
<td>Philadelphia Union</td>
<td>2790</td>
<td>573</td>
</tr>
<tr>
<td>45</td>
<td>Patrick Nyarko</td>
<td>Chicago Fire</td>
<td>2092</td>
<td>568</td>
</tr>
<tr>
<td>46</td>
<td>Gyasi Zardes</td>
<td>LA Galaxy</td>
<td>2082</td>
<td>564</td>
</tr>
<tr>
<td>47</td>
<td>Michael Gspurning</td>
<td>Seattle Sounders FC</td>
<td>2385</td>
<td>564</td>
</tr>
<tr>
<td>48</td>
<td>Kenny Miller</td>
<td>Vancouver Whitecaps</td>
<td>1643</td>
<td>564</td>
</tr>
<tr>
<td>49</td>
<td>Fabian Espindola</td>
<td>New York Red Bulls</td>
<td>2051</td>
<td>562</td>
</tr>
<tr>
<td>50</td>
<td>Landon Donovan</td>
<td>LA Galaxy</td>
<td>1785</td>
<td>562</td>
</tr>
</tbody>
</table>

Table 11.3: Top 50 rated players in the MLS