Form is temporary, class is permanent: Statistical methods for predicting the career trajectories and contributions of players in the sport of cricket

Oliver George Stevenson

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Statistics, the University of Auckland, 2021.

Abstract

In the sport of cricket, player ability is generally assessed using traditional statistics, such as batting and bowling averages. However, such measures fail to account for variations in ability that may occur over the short-term, during a match, and over the long-term, between matches. As a result, batting and bowling averages are unable to distinguish between players whose abilities are declining, and those still yet to reach peak performance. This is a major shortcoming of many proposed measures of cricketing ability; coaches and selectors often cite recent performances, or form, as a reason for dropping or selecting certain players, but have no means of quantifying how such factors may impact a player's true, underlying ability.

In order to detect and quantify temporal variations in ability that may be observed over the course of a playing career, a set of Bayesian parametric models are derived to measure and predict the career trajectories of professional cricket players. Career trajectories are modelled using a Gaussian process and aim to estimate a player's past, present, and future abilities, accounting for recent form, and a number of contextual variables that are frequently ignored by alternative measures. A simulation-based method of predicting the outcome of upcoming matches is then proposed. The match-simulation algorithm takes predictions of ability obtained from the estimated batting and bowling career trajectories as inputs and attempts to quantify the likely performance and contribution of individual players in a given match.

Generally speaking, the results suggest that underlying batting and bowling ability does not fluctuate significantly in the short-term as a result of recent form. Instead, ability appears to improve and deteriorate slowly over time, likely as a result of players gaining experience in a variety of match conditions; participating in specialised coaching programmes; and due to changes in physical attributes, such as fitness and eyesight. These findings may have practical implications in the likes of player comparison, talent identification, and team selection policy, as coaches and selectors are able to better quantify player ability and understand the individual-specific risks and rewards of selecting certain players over others.

Acknowledgements

Firstly, I would like to thank the University of Auckland for funding my research through a University of Auckland Doctoral Scholarship. Without the University's support, none of this would have been possible and I am grateful to have been afforded the opportunity to pursue a topic of such personal interest during the course of my Doctoral studies.

I was also fortunate to receive financial support from the International Society for Bayesian Analysis (ISBA), who provided funding to attend both the 2018 ISBA World Meeting, hosted at the University of Edinburgh, and the 2018 Bayesian Young Statisticians Meeting (BAYSM), hosted at the University of Warwick. My trip to the UK to attend and present at each of these conferences in 2018 was undoubtedly one of the highlights of my PhD experience, except for the part where I missed my flight from Edinburgh to Birmingham.

I am grateful to New Zealand Cricket, who have not only been kind enough to share their incredibly rich data sources with me, but also contributed towards my travel to present at ISBA and BAYSM in 2018. A number of topics that were discussed when I first visited Lincoln in 2017 have provided inspiration for the content presented in this thesis. It is also my pleasure to thank Kane Williamson and Neil Wagner, for their on-field exploits with the Blackcaps over the last few years, and in particular, for their roles as main characters in Chapters 2 and 3.

I am hugely appreciative of the team at NV Play, who have done a stellar job of maintaining the database New Zealand Cricket granted me access to. Matt Smith must take a lot of credit for the fact I have had to do next to no data cleaning, and deserves significant praise for the patience shown when fielding yet another email from me on a Friday afternoon, asking why match ID: E79B3FAA-22D7-45FC-825D-C3A57BA20CF0 has been flagged as a Test match, when it is clearly a ODI match played between New Zealand and Sri Lanka, in the home summer of 2015.

Overseeing my collaboration with both New Zealand Cricket and NV Play has been Pete Mayell, the linchpin of this research project. I cannot thank Pete enough for his contributions over the last few years, particularly in the early stages, from organising initial meetings with New Zealand Cricket, to putting me in touch with NV Play, and taking me out for my first ever client lunch. Pete's background and experience in the worlds of cricket analytics and academia have been very insightful. At times he has felt like a second supervisor, going above and beyond what I ever expected of him, offering to proofread any papers I have been working on, and even helping me source additional funding to travel and present at ISBA and BAYSM in 2018. I have thoroughly enjoyed working together and hope that an opportunity to do so again may one day present itself.

I would also like to thank Scott Brooker — another talented individual who I was put into contact with by Pete — who at the time, I only ever knew as "WASP guy". Scott not only provided me with a trailblazing example of what can be achieved in the field of cricket analytics, through his own PhD research, but was also kind enough to provide feedback on a number of ideas developed throughout this thesis.

There are several people within the Department of Statistics who have made my life much easier over the last few years, to whom I would like to express my gratitude. Joei and Maria in the front office, thank you for helping me navigate the murky waters that are the university's various regulations and guidelines for PhD students. Receiving a "let me look into that for you" email from Joei, or having Maria show me which rules are slightly more bendable than others, is a sure-fire sign that any query is going to be dealt with successfully. To Ben, thank you for always taking the time — whether you had it, or not — to listen to any questions or complaints I had about various academic processes, and for always being a great sounding board for any ideas related to my research.

In particular, I would like to acknowledge the role of my supervisor, Brendon Brewer. It has been a thought-provoking experience to observe and learn first-hand how to approach statistics from a truly Bayesian perspective; I will forever remember the importance of prior sensitivity analyses whenever I review, or submit, any future papers for publication. For better or worse, I thoroughly appreciate the amount of creative freedom you gave me to take my research in whichever direction I saw fit. The level of trust you have shown in my abilities has provided me with ample self-confidence moving forward in my own career, for which I am grateful.

Lastly, a special thank you to my family, who have each contributed to this thesis in their own ways. Mum and Dad, I am grateful for your unwavering support of my academic endeavours, as niche as they may be. Steph, your love and encouragement over the past few years have been invaluable, while your (ever so slightly) increased interest in cricket comes as an added bonus. Finally, to everyone: Nana, Kate, Ben, Chris, Charlotte, Chance; thank you for the endless source of support, laughter, and good times we have shared.

Contents

Chapter 1

Introduction

1.1 Cricket: an overview

1.1.1 Laws of the game

Cricket is a bat-and-ball team sport originating in 16$^{\text{th}}$ century England and is played with two teams of 11 players per side. A standard match consists of two *innings*, with each innings consisting of a limited number of *overs* to be bowled. An over is defined by six *balls*. Teams take turns at being the *batting team* and the *bowling team* and switch roles between innings.

The team batting first aims to score as many *runs* as possible in their innings, while the team bowling first aims to restrict the number of runs scored by their opponent by getting opposing batsmen *out* in one of a number of prescribed methods (also referred to as *dismissing the batsman* or *taking a wicket*). The batting team's innings ends when either (1) they have faced all of their allotted balls, or (2) the bowling team has taken 10 wickets.

The team who bowled first then takes their turn to bat, with the aim of scoring more runs than their opponent. The match ends when the second innings is concluded in one of the manners outlined above, or, once the team batting second surpasses the target score set by the team batting in the first innings. Whichever team has scored more runs at the end of the match is declared the winner.

There have been a number of rule-changes over the course of the sport's history, leading to the three most popular variants of professional modern day cricket:

1. ***One-day* or *List A* cricket**
    - Each team bats in one innings apiece, limited to 50 overs per innings.
    - International List A cricket is referred to as ***one-day international*** or ***ODI*** cricket.

2. ***Twenty20*** **or** ***T20*** **cricket**

   - Each team bats in one innings apiece, limited to 20 overs per innings.
   - International T20 cricket is referred to as ***T20I*** cricket.

3. ***First-class*** **cricket**

   - Each team bats in two innings apiece and there is no limit to the number of overs that can be bowled per innings.
   - A match is played over four days, usually with a limit of 90 or 100 overs to be bowled per day.
   - If neither side has won after four days, the match is declared a draw.
   - In international ***Test*** cricket, a match is played over five days.

1.1.2 Statistics in cricket

Cricket is a sport with a long-standing tradition of record-keeping; the first known recorded scorecards date as far back as 1776 (Association of Cricket Statisticians and Historians, 1981). Even at the grass roots level, a considerable quantity of data are collected each match, which has lead to the sport's strong statistical culture. Individual players are often obsessive over their personal statistics; it would be a difficult task to find a cricketer who is not at least remotely aware of their batting and/or bowling average. This has lead to cricket being colloquially referred to as one of the most individual team sports.

"Cricket is a most precarious profession; it is called a team game but, in fact, no one is so lonely as a batsman facing a bowler supported by ten fieldsmen and observed by two umpires to ensure that his error does not go unpunished."

– John Arlott in An Eye for Cricket (Arlott, 1979)

Batting and bowling averages

Traditionally, a player's batting ability is recognised by their career *batting average*. The batting average represents the average number of runs a player scores for every time they are dismissed.

$$\text{Career batting average} = \frac{\sum \text{Runs scored}}{\sum \text{Dismissals}} \tag{1.1}$$

Similarly, a player's bowling ability is traditionally measured using their career *bowling average*. The bowling average represents the average number of runs a player concedes for every wicket they take.

$$\text{Career bowling average} = \frac{\sum \text{Runs conceded}}{\sum \text{Wickets taken}} \tag{1.2}$$

As shown in Equations 1.1 and 1.2, batting and bowling averages are relatively straightforward to calculate.

Economy rate and strike rates

Other commonly used statistical metrics in cricket to measure player ability are *economy rate* and *strike rate*. Economy rate is a metric associated with bowling that indicates the average number of runs conceded by a player, per over bowled.

The term strike rate has two different meanings, depending on whether it is being used in the context of batting or bowling. From a batting perspective, strike rate represents the average number of runs scored per 100 balls faced. While bowling, strike rate represents the average number of balls bowled per wicket taken.

Career records

Together with batting and bowling averages, strike rates and economy rate are generally the focal point of attention when assessing an individual player's career record. These statistics usually provide coaches, selectors, commentators and spectators with a fairly reliable assessment of a player's overall skill level. Other statistics, such as the number of times a player has passed the significant milestones of 50 and 100 runs while batting, or taken five wickets in an innings while bowling, are also of interest.

Given the author's nationality and players' interesting data, the career records for New Zealand's presently top ranked batsman (Kane Williamson) and bowler (Neil Wagner) are presented as examples in Tables 1.1 and 1.2, and are used as ongoing references throughout this thesis. Convention dictates that a player's career record be split by match format. Test cricket refers to international first-class cricket, one-day international (ODI) cricket refers to international List A cricket, while T20I cricket refers to international T20 cricket. It is worth noting, that a player's first-class, List A and T20 records include both domestic and international performances. Here, it is possible to observe that Williamson has played in all formats of cricket, while Wagner has not participated in any ODI or T20I matches.

Table 1.1. *Career batting record for Kane Williamson.*

	Matches	Innings	Not outs	Runs	High score	Average	Balls faced	Strike rate	100s	50s	4s	6s
Test	80	140	13	6476	242*	50.99	12543	51.63	21	32	706	14
ODI	151	144	14	6173	148	47.48	7551	81.75	13	39	563	49
T20I	60	58	7	1665	95	32.64	1330	125.18	0	11	170	36
First-class	148	253	20	11287	284*	48.44	21659	52.11	31	59	1319	34
List A	212	201	22	8294	148	46.33	10194	81.36	17	51	729	70
T20	181	173	21	4593	101*	30.21	3681	124.77	1	31	430	115

Table 1.2. *Career bowling record for Neil Wagner.*

	Matches	Innings	Balls	Runs	Wickets	Best bowling (innings)	Average	Economy rate	Strike rate	5 wickets (innings)
Test	48	90	10743	5480	206	7/39	26.60	3.06	52.1	9
First-class	174	319	36233	19420	729	7/39	26.63	3.21	49.7	36
List A	108	104	5276	4775	166	5/34	28.76	5.43	31.7	2
T20	76	72	1502	2198	79	4/33	27.82	8.78	19.0	0

1.2 Sports analytics

The application of statistical methods to analyse data collected and related to sporting activities is commonly referred to as *sports analytics*. When analysing data pertaining to a particular sport, the word 'sports' is often replaced with the sport in question, for example, *cricket analytics*. Sports analytics are used by a wide sector of the sporting community, ranging from professional teams competing in a diverse range of sporting codes, to individuals engaging in activities related to sports gambling. Regardless of who is using sports analytics, the ultimate goal is to gain a competitive advantage over the opposition, be that a rival team or a bookmaker. Such an advantage may manifest itself through applying a unique strategy to the game, an innovative change in approach to team selection policy, or a novel method of identifying talent in young, up-and-coming players.

Clearly, cricket is a data-rich sport, with a well-entrenched statistical culture. However, despite having all the requisite features of being an ideal sport to flourish in terms of analytics, few advanced statistical metrics or modern methods of analysing player ability have been adopted by the wider cricketing community. While batting and bowling averages might have been considered cutting-edge analytics in the early 1800s, these types of summary statistics do not hold the same weight in the 21$^{\text{st}}$ century. For a sport where a single match can continue for up to five days and a single career can span more than 20 years, it seems inadequate to use simple summary statistics to quantify the past, present, and future abilities of individual players.

It is important to note that sport is not an exact science and that there is a certain art in identifying talent and estimating player value based on *the eye test*. However, rather than using data as an additional tool in the decision making process, there are still plenty of examples in modern day cricket of coaches and selectors relying solely on gut feel. While a coach's instincts and intuitions may generally be correct — after all, this is why they are coaches — it is inevitable they will occasionally miss out some finer aspects of the game that can be explained by the data. Consequently, on the analytics front, cricket finds itself playing catch-up with a number of global sports, namely its American cousin, baseball, with whom it shares a number of similarities.

1.2.1 Origins

The mainstream popularisation of sports analytics is frequently credited to the emergence of *sabermetrics* in baseball, referring to the analysis of baseball data and statistics that measure in-game performance. The origin of sabermetrics is documented in the book *Moneyball: The art of winning an unfair game* (M. Lewis, 2003), later popularised in the 2011 Hollywood film adaptation, *Moneyball*. The book and film recount the innovative statistical approach taken by a financially limited Major League Baseball (MLB) side, the Oakland Athletics, in the scouting and analysing of players in the 2002 MLB season, which saw a revolution in how professional baseball is played, coached and watched. While the fundamental rules and objectives of baseball have remained relatively consistent since the MLB's inception in 1869, the game that is played today is vastly different to that of previous generations, partially due to the impact of sports analytics.

The introduction of sabermetrics in the early 2000s has since had a catalytic influence on the use of data-driven decision making in other sports. Initially, the primary drivers of sports analytics outside of baseball were sporting codes that had a major league with a home in the United States, such as basketball's National Basketball Association (NBA), ice hockey's National Hockey League (NHL) and American football's National Football League (NFL).

For example, the way in which basketball teams operate their offences in the modern NBA has changed dramatically ever since data analysts put the concept of shot efficiency (expected points per shot) under the microscope. Between the 1980s and early 2000s, teams shot very few three-point shots, instead favouring mid-range, two-point shots from the baseline. While the underlying three-point shooting abilities of players do not appear to have improved dramatically in the modern era of basketball — the league average three-point field goal percentage has consistently hovered between 34% and 36% in the last 20 years — only in the last decade have teams started to exploit the mathematics behind shooting and shot selection. With the application of sports analytics to basketball, teams have recognised that in many situations, mid-range two-point shots, tend to have a similar probability of success as longer-ranged three-point shot attempts, while being worth one fewer point. Therefore, to maximise scoring efficiency, modern NBA teams generally aim to take a two-point shot, under or very close to the hoop, or to take a three-point shot, which has lead to the current *three-point revolution* in basketball (Young,

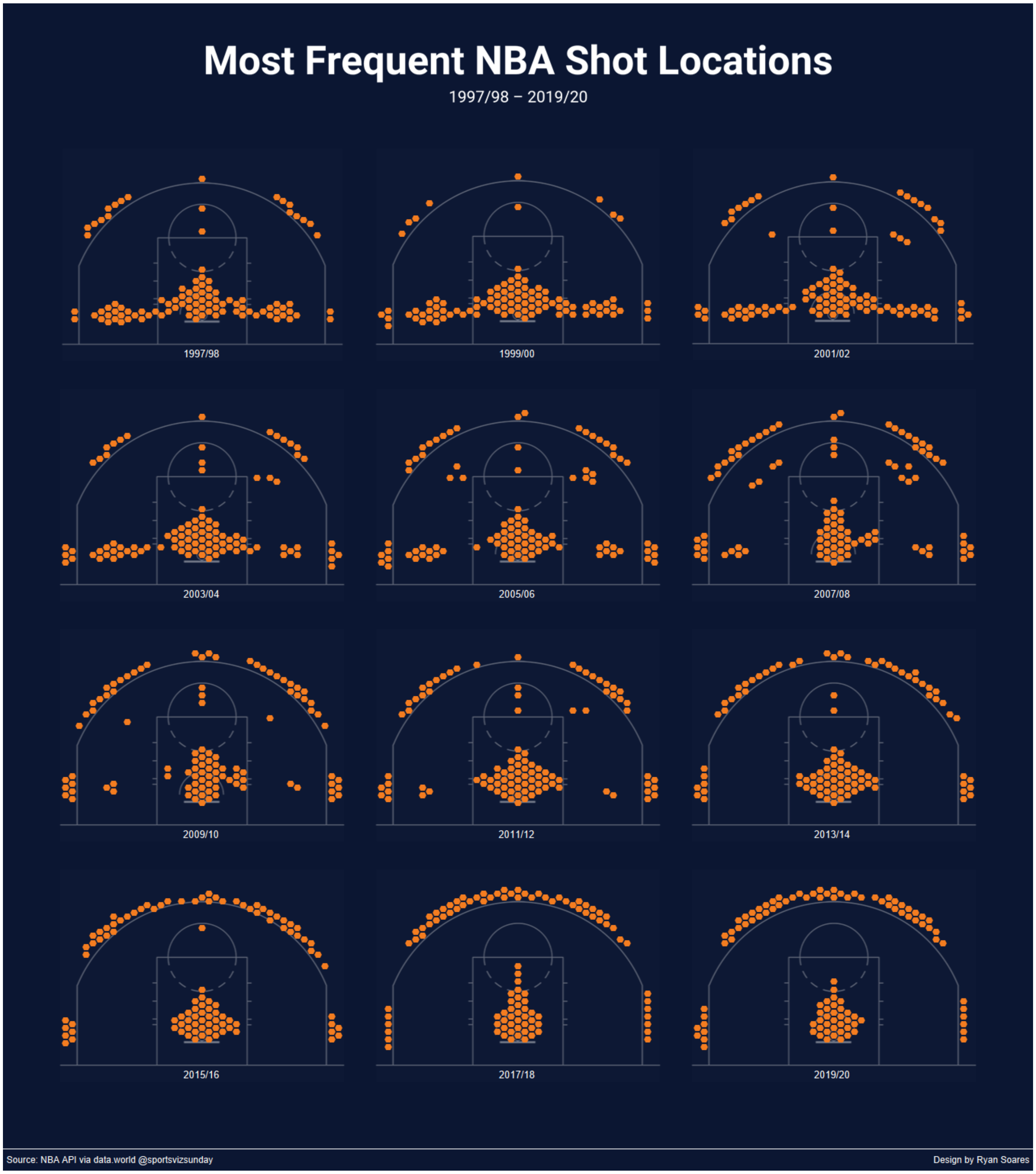


Figure 1.1. *Most common shot locations in the NBA between the 1997/1998 and 2019/2020 seasons (Soares, 2020).*

2019). In the 2010/2011 season, there was an average of 18.0 three-point attempts per game. This number has increased in every subsequent season, to the point there was an average of 34.1 three-point shots taken per game in the most recent 2019/2020 season. No team personifies this more than the Houston Rockets, who in the 2019/2020 season, shot a league-leading average of 45.3 three-point shots per game. This gradual shift in scoring mentality over the years is illustrated in Figure 1.1.

1.2.2 Popularity

Sports analytics in the US have become so popular that the use of advanced metrics has emerged from behind the closed doors of secret boardroom meetings and coaching sessions, to public broadcasts and live coverage of matches. There are numerous examples of advanced statistical measures initially purposed for use by coaches, general managers and players, becoming the gold standard for comparing players, even by more casual fans of the sport. This is not to say that teams are not continually developing their own metrics for private use. However, with the increasing number of data analysts finding employment in the likes of the MLB and NBA, it is probable that many franchises are at least somewhat aware of the various analytical tools being developed and utilised by their opponents. Eventually, metrics that were once considered advanced statistics become commonplace among all teams in a league, eliminating the need for secrecy in how they were derived. Such metrics then trickle down to broadcasters and the general public, where they are adopted and become another means of comparing player performance.

Part of the success and growth in the field of sports analytics can be attributed to the respective sporting administrative bodies, who have provided public access to large quantities of data, allowing for collaboration between data analysts and fostering an environment that promotes the growth of statistical analysis in sport. In terms of public engagement with statistics, credit must also be given to the creators of many of these popular advanced statistical metrics for maintaining an easy to understand, intuitive interpretation, which can be understood by fans both with and without formal statistical training.

The factors behind what defines a reliable, advanced sports metric were considered in Franks et al. (2016), who identified three key criteria: (1) *stability*, whether the metric measures the same thing in different contexts; (2) *discrimination*, whether the metric is able to differentiate between individual performers; and (3) *independence*, whether the metric provides new information. These criteria are closely related to the concept of construct validity in psychology, introduced by Cronbach & Meehl (1955) to define how well a certain test measures what it claims to be measuring. Sporting metrics that have been successful in their transition from advanced statistic to publicly available measure, all tend to satisfy the three criteria identified by Franks et al. (2016). Some metrics have become so popular they are even recognised as an official statistic by the sport's administrative body, which in turn, has resulted in fans consuming more analytical content than ever.

For example, in baseball, *win shares* aims to assess a player's value in terms of how many wins the player contributed to over the course of a season (James & Henzler, 2002). In a similar vein, *wins above replacement player*, and *value over replacement player*, aim to estimate how much a player contributes to their team in comparison to a replacement level player, whom a team could acquire at minimal cost.

In basketball, metrics such as *effective field goal percentage* and *true shooting percentage* allow for the fair comparison of shooting ability by adjusting for the types of shot taken by individual players. Individual *plus-minus (+/-)* aims to reflect a player's individual contribution to a team performance, by measuring the points difference between teams, while the player is on the court. Similarly, *offensive* and *defensive ratings* measure the expected points scored and conceded by a player's team per 100 possessions respectively, while that player is participating in a game. The difference between a player's offensive and defensive ratings defines their *net rating*, which provides an overall assessment of how much better or worse a team is when a specific player is on the court (Oliver, 2004). A positive net rating indicates that a player's team is likely to outscore their opponent while the player is on the court; a negative net rating indicates a player's team is likely to be outscored. All of these are examples of metrics that were once considered trade secret advanced statistics, but are now publicly available on numerous websites.

Developments in sport-related analytics are not strictly limited to on-field performance either. Dubbed *fanalytics*, teams are now looking ways to increase revenue through improving and optimising fan experiences when attending live sport. For example, the San Francisco 49ers NFL franchise have developed an in-stadium application to analyse real-time data, helping ground staff identify which bathrooms need cleaning and ensure busy concession stalls are well-supplied with hot dogs and beer (Leuty, 2018).

1.2.3 Barriers to growth

While the sports analytics industry in the US grew rapidly across sporting codes following the introduction of sabermetrics, team sports more popular in Europe, Africa and Australia, such as association football (soccer), rugby and field hockey have taken longer to adapt and take advantage of using analytics. One could speculate this to be a result of the open nature of the skills required in these sports, compared with their American counterparts.

Skills in physical education and sport have been proposed to exist on a continuum, from *closed* to *open* skills (Poulton, 1957; Knapp, 1963; Whiting, 1969). Closed skills take place in a stable, predictable environment, have a clear beginning and end and are often self-paced. Examples of closed skills include a tennis serve, a 100 metre sprint, and a penalty kick in football.

Open skills occur in more variable environments, where the participant has to consider a number of external factors in choosing how and when to execute the skill. Examples of open skills include passing in netball, shooting at goal in football, and tackling in rugby.

In baseball, the actions of pitching and hitting are examples of relatively closed skills; each execution of the skill takes place in a fairly constant environment. Of course some factors, such as the crowd, game situation, and level of pressure felt by players will vary, meaning these are not purely closed skills, but overall the skill is performed under similar conditions during each repetition. In basketball, the action of shooting can vary in how open or closed the skill is classified, depending on the type of shot. A free throw is almost a purely closed skill, while taking a shot that is contested by an opposing defender is slightly more toward the open end of the continuum. Taking an uncontested shot in open play lies somewhere between these two shot types — there are some variable factors at play, such as the position the shooter receives the ball in — but ultimately a player must simply catch and shoot the ball.

In contrast, consider the skills required in playing the likes of football and rugby. The environment in which a player executes the skill of shooting at goal in football will rarely be the same between any two shots. Instead, when taking a shot, a player must consider the positions of the goalkeeper and any defenders in front of and behind them, as well as their particular orientation and the ideal foot to shoot with. Similarly, when making a tackle, a rugby player must consider whether the ball carrier has any passing options, the defensive position of their own team mates, and the type of tackle most appropriate for the situation.

There are numerous advantages to analysing data that are collected in sports consisting of skills that are mostly closed in nature, a significant one being the ease of comparison between players. It is far easier to determine who is the best free throw shooter in basketball, than it is to tell who is the best goal kicker in rugby. In the first example, career *free throw percentage* is a direct measure of ability, as each free throw is shot in almost identical circumstances. However, in the latter example, one cannot simply use career goal kicking percentage as a measure of ability, as the data do not take into account the degree of difficulty of the kicks a player has taken in their career. For this reason, analysts in team sports outside the US have taken some time to grapple with how best to analyse the data available.

1.3 Cricket analytics

Considering the number of similarities between cricket and baseball, it may come as a surprise that cricket is only now beginning to experience the same statistical revolution baseball underwent in the early 2000s. In each sport, batting and bowling (pitching in baseball) are both relatively closed skills in nature. Additionally, the sequence of play in cricket is neatly divided into individual balls, much like baseball's pitches, suggesting cricket is a sport that could stand to gain a lot from a well-developed analytics scene.

One might attribute the lack of attention to developing more advanced statistical metrics in cricket to the absence of a major international cricket league — and therefore financial value — until the inception of the Indian Premier League in 2007. A cynical reader will not be surprised to hear that after the estimated value of the IPL surpassed the US $5 billion mark in 2017 (Duff & Phelps, 2017), franchises began to invest more into cricket analytics in order to gain a competitive advantage over their opponents and subsequently, a larger slice of the financial pie. For comparison, MLB's most popular franchise, the New York Yankees, were valued at US $1.3 billion in 2007, and in 2020 are estimated to be worth US $5 billion alone (Forbes, 2008, 2020).

Due to the sensitive nature of the information, it is difficult to know to what extent professional domestic and international cricket teams are using data analysis in their day-to-day routines. However, while the modern scene of cricket analytics is still in its relative infancy, due to the abundance of data available, cricket has been the subject of a number of academic studies, which can generally be grouped into one of several categories.

1.3.1 Achieving a fair result in interrupted matches

Many professional sports must contend with the natural elements and at times will require cancellation and rescheduling. As an outdoor sport where a single day's play can span over eight hours of the day, it is inevitable that many games of cricket will be interrupted with seasonal rains and poor weather. However, due to cricket's extensive time commitments, cancellation is undesirable and rescheduling is often impossible. This has lead to an area of research particularly unique to cricket, in developing methods of achieving a fair result in an interrupted match (Carter & Guthrie, 2004; Duckworth & Lewis, 1998; Duckworth et al., 2019; Ian & Thomas, 2002; Jayadevan, 2002; Perera & Swartz, 2013; Stern, 2016).

Many followers of the game will be familiar with the Duckworth-Lewis-Stern (DLS) method, which has become the established method of determining the winner of interrupted one-day and T20 matches. The method was first derived in Duckworth & Lewis (1998), where the underlying principal is to consider the number of resources available to the team batting at the time of the interruption (namely wickets and overs remaining), and to adjust their final score accordingly. For example, if the second innings of a match is interrupted to the point a match cannot be completed, the DLS method will consider the number of wickets and overs the batting team had remaining at the point of interruption. A calculation is then done to estimate the number of runs the batting team would have scored, which determines the final result of the match. In order to maintain its relevance, the method is frequently updated to adapt to modern scoring rates and new formats of the game, such as T20 cricket (Stern, 2016).

1.3.2 Optimising playing strategies

Prior to the boom of T20 cricket in the late-2000s, a number of studies looked at means of optimising playing strategies in both one-day and Test cricket. These studies generally aim to challenge the conventional strategies used by teams in their approach to the game.

Several studies have suggested that teams might expect to score more runs by making certain tweaks to their batting orders (Clarke & Norman, 2003; Norman & Clarke, 2010; Preston & Thomas, 2000; Scarf et al., 2011; Swartz et al., 2006), while others suggest teams could stand to gain from established batsmen making minor changes in their approach to scoring runs at the end of an innings, when batting with weaker batting partners (Clarke & Norman, 1998, 1999). However, the proposed findings tend to require a very specific set of match conditions to be put into practice, which limits the ability to apply these methods in a broader scope.

A more robust method of optimising team lineups, as well as batting and bowling orders in T20 cricket was proposed in Perera et al. (2016), who derived a method of finding the optimal matchday team via a simulated annealing algorithm. The model attempts to estimate the lineup that maximises the difference between expected runs scored while batting and expected runs conceded while bowling. The approach and potential output of this model is closer to something that could have an application in real-world cricket. One drawback is the expensive computational nature of the method. The model fitting process requires roughly 24 hours for sufficient convergence, which would make adapting batting and bowling orders on the fly during a match very difficult.

1.3.3 Match outcome prediction

As with many sports that have a heavy statistical focus, numerous past studies have attempted to identify key elements that can impact a team's chances of winning a match. More recently, this has become an area of cricket analytics that has also garnered attention from a commercial perspective.

At present, several commercial products are employed by modern broadcasters to facilitate and enhance viewer experience, by providing in-game analysis relating to the current state of the match and updated estimates predicting the likely victor. These products include the winning and scoring predictor (WASP) tool and the WinViz model. The former is the subsequent output of an academic study, which investigated the factors that impact batting conditions in one-day cricket (Brooker & Hogan, 2011), while the latter as been developed in a private capacity by UK-based cricket analytics company CricViz. Unfortunately, due the commercial sensitivity surrounding these products, little information is publicly available regarding the specific factors and variables that are considered in these models and the modelling process used to obtain such predictions is unknown.

Thankfully, papers published via the peer-review process lend themselves to a more collaborative environment. Unsurprisingly, home ground advantage is consistently identified as being one of the most significant predictors of match outcome, across T20, one-day, and Test cricket (Akhtar & Scarf, 2012; Bailey & Clarke, 2006; Bandulasiri, 2008). Other potentially significant variables are often related to the strength of individual teams and historical team performances. However, when predicting the outcome of a hypothetical match, the majority of previously proposed models tend to quantify the strength of teams based on the team itself, either via past performances or using metrics such as the official International Cricket Council team ratings, rather than inferring team strength from the individual players that make up a side. As a result, few methods are able to easily make adjustments to predictions of match outcome if key players are injured, or are not selected to play in a certain match.

Therefore, when making predictive statements regarding the most likely outcome of a match, it would be advantageous to develop a means of accounting for the individual batting and bowling strengths of the participating players. Such methods are proposed in this thesis, along with a novel simulation-based model of predicting match outcome in the context of Test cricket. Of course, this approach requires the ability to obtain accurate estimates of player ability in the first instance, which itself is no easy feat, as discussed in Section 1.3.4.

1.3.4 Evaluating player performance and ability

This category of sports analytics is relevant in almost every sport and tends to have the largest impact on the careers of professional athletes. In the long run, advancements in this category of analytics enable teams to improve their overall performance by selecting players with skill sets that tend to result in success, while also saving money by identifying value in previously overlooked players, and to avoid overpaying players whose reputation is not worth their price tag. Sports such as baseball and basketball have successfully managed to innovate and develop numerous advanced statistics in this field, allowing organisations to continually evolve their underlying strategies and policies in terms of team selection and talent identification.

On the other hand, cricket analytics in this area is still relatively uncharted territory, in both the public and academic spheres. By not providing a publicly available, standardised data source, the sport's governing body, the International Cricket Council (ICC), are incentivising analysts to keep their work and data sources to themselves and lend their expertise in a private capacity to organisations that may benefit from (and pay for) their advice. Consequently, due to the lack of a collaborative environment, there are few advanced metrics that have been accepted as alternatives to batting and bowling averages among the wider cricketing community.

Alternative proposals to the batting and bowling average

At the time of their introduction, batting and bowling averages fulfilled the three criteria of a successful sporting metric identified by Franks et al. (2016); they consistently measure batting and bowling ability in all contexts; allow for differentiation in ability between players; and at the time, provided new information. As a result, they were willingly adopted by the cricketing community, as they provide a general overview of a player's abilities, whilst being easy to understand by all viewers of the game. This last point should not be overlooked; many proposed alternatives to batting and bowling averages have ultimately failed to gain any traction in terms of uptake and acceptance, due to their lack of an intuitive cricketing interpretation. With this in mind, when considering the value of newly proposed metrics in cricket, one could justify adding a fourth criteria of *interpretability* to the three previously identified by Franks et al. (2016).

A series of papers published by Tony Lewis, of Duckworth-Lewis-Stern fame, proposed a method of evaluating a player's net contribution to a one-day match that utilised the established Duckworth-Lewis methodology at the time (A. Lewis, 2005, 2008). As batting and bowling averages provide no means of objective comparison, the method considers the number of resources consumed (balls faced and wickets lost) and resources contributed (runs scored) to quantify the overall contribution made in a match by individual players. A strong batting contribution is defined by a high ratio of resources contributed to resources consumed. The inverse is considered while evaluating bowling; the fewer resources the opposition batsman contributes to their team per ball bowled, the better a player's bowling contribution. Additionally, the rate at which runs are scored is considered in the context of the match. As scoring tends to be slower in early stages of a team's innings, batsmen are rewarded for scoring quickly at the beginning of an innings. However, later in a team's innings, scoring is expected to be higher and consequently, batsmen who face a large number of balls for few runs in the latter stages of an innings are penalised. Similar to plus-minus in basketball, a player's batting and bowling contributions can be combined to give an indication as to whether the player had a positive or negative overall impact on the match. While this method allows for a comparison of performance between players, the units of *average run contribution per unit of resource consumed* lack a clear and concise interpretation and will mean little to even the most seasoned cricket viewer. This lack of interpretation is likely a major contributing factor as to why player contribution is not a metric that has been adopted by the wider cricketing audience.

More recently, ESPNcricinfo (an organisation widely considered the primary news source for all things cricket) introduced a number of new, advanced statistics, dubbed *Smart Stats*, to better measure how individual players were performing in the 2018 and 2019 editions of the IPL (ESPNcricinfo, 2018, 2019). The underlying reasons for introducing Smart Stats were sound; it is generally accepted that traditional metrics of player ability, such as batting and bowling averages, are lacklustre when it comes to T20 cricket, as these measures fail to account for the match context of an individual performance. Scoring 30 runs off 15 balls is rarely going to have a game-changing impact in a Test match, however in many cases this could be a match-defining innings in T20 cricket. Therefore, rather than using the usual metrics of runs scored, batting average and strike rate to measure player batting performance, *smart runs* and *smart strike rate* measures were introduced. Similarly, *smart wickets* and *smart economy rate* were introduced to better quantify individual bowling performances.

In principle, Smart Stats are great on paper; they aim to account for a number of contextual variables not considered by traditional metrics, such as the strength of the batsman and bowler, the general scoring rate in a match and comparing individual batting and bowling performances by players at similar stages of a match. However, similar to the proposal in A. Lewis (2005, 2008), a major drawback of Smart Stats is their lack of interpretability. There is no clear cut answer as to what a smart run represents, what a smart strike rate of 200 actually means, or what the value of one smart wicket is. Furthermore, there is no publicly available methodology of how to compute Smart Stats, which limits any form of collaboration or engagement from other experts in the field of cricket analytics. As a result, the usage of Smart Stats are limited to ranking and comparing player performances, but have no real meaning when it comes to quantifying differences in ability between players in more real terms.

Outside of batting and bowling averages, the method that has garnered the most public attention when it comes to ranking the best players in the world, is the official ICC ratings. This system utilises a broader range of information to rate and rank player performances, including opposition strength and places emphasis on more recent performances, often referred to as *form* in cricket and other sports. As such, the ICC rankings are widely considered a better means of ordering the current abilities of players, compared with simply ranking players by their career batting and bowling averages. However, a look at the ICC ratings for the current[1] top 10 Test batsmen and bowlers in Tables 1.3 and 1.4, highlights the shortcomings of this approach and raises several valid questions.

What does a batting rating of 911 for Steve Smith tell us about his underlying batting ability? What does it mean that Pat Cummins has 61 more bowling rating points than Neil Wagner? Like Smart Stats, the ICC ratings are useful for ranking players, but lacks an intuitive cricketing interpretation and fails to inform about the differences between players' underlying abilities. Similarly, the closed source nature of the ICC rating formula means it is unknown as to how each factor impacts a player's rating, making it difficult to compare with other methods. Finally, it is unclear whether the ICC ratings attempt to provide inferential or predictive accuracy, or instead try to formalise expert judgement about who is in and out of form. These two goals may

[1] as of 1$^{\text{st}}$ December 2020

Table 1.3. *ICC Test batting ratings as of 1st December 2020.*

Rank	Player		ICC rating
1.	Steve Smith	(AUS)	911
2.	Virat Kohli	(IND)	886
3.	Marnus Labuschagne	(AUS)	827
4.	Kane Williamson	(NZ)	812
5.	Babar Azam	(PAK)	797
6.	David Warner	(AUS)	793
7.	Cheteshwar Pujara	(IND)	766
8.	Ben Stokes	(ENG)	760
9.	Joe Root	(ENG)	738
10.	Anjinkya Rahane	(IND)	726

Table 1.4. *ICC Test bowling ratings as of 1st December 2020.*

Rank	Player		ICC rating
1.	Pat Cummins	(AUS)	904
2.	Stuart Broad	(ENG)	845
3.	Neil Wagner	(NZ)	843
4.	Tim Southee	(NZ)	812
5.	Jason Holder	(WI)	810
6.	Kagiso Rabada	(SA)	802
7.	Mitchell Starc	(AUS)	797
8.	James Anderson	(AUS)	781
9.	Jasprit Bumrah	(IND)	779
10.	Trent Boult	(NZ)	770

not be entirely compatible.

Limitations of the batting and bowling average

To date, batting and bowling averages have yet to be supplanted as the primary method of gauging the overall abilities of an individual player. However, as single point estimates, these simple averages fail to inform about variations in ability on two scales.

Firstly, averages fail to measure short-term changes in ability that occur during or within a single match. Compared to many sports, cricket is distinct in the sense that the physical differences between matches, such as the local weather and pitch conditions, have a significant impact on how a match will be played out. Therefore, short-term variation in ability can be attributed to the need to adjust to these external factors, which are rarely the same between any two innings. Additionally, certain innings will require a unique mental approach in terms of how aggressive or defensive to play, depending on the context of the match. Adapting to the specific match conditions can impact player performance — particularly while batting — where the process of adjustment is commonly referred to as *getting your eye in*. Subsequently, batsmen are frequently dismissed early in their innings, on a low score, before they have gotten their eye in.

Secondly, averages provide no information in regards to how ability varies over the long-term, between matches. A playing career can last as long as 20 years; it would be naïve to think that an individual's underlying abilities remain constant over such a long period of time. Instead, it is likely that many players' career trajectories tend to follow an anecdotal description of a typical sporting career. Young players are generally assumed to begin their careers with some raw talent and ability, which improves over time as a result of coaching, gaining experience in a variety of match conditions, as well as general improvement or deterioration in the likes of technique, eyesight and agility. Eventually players reach the pinnacle of their career, after which ability tends to decline.

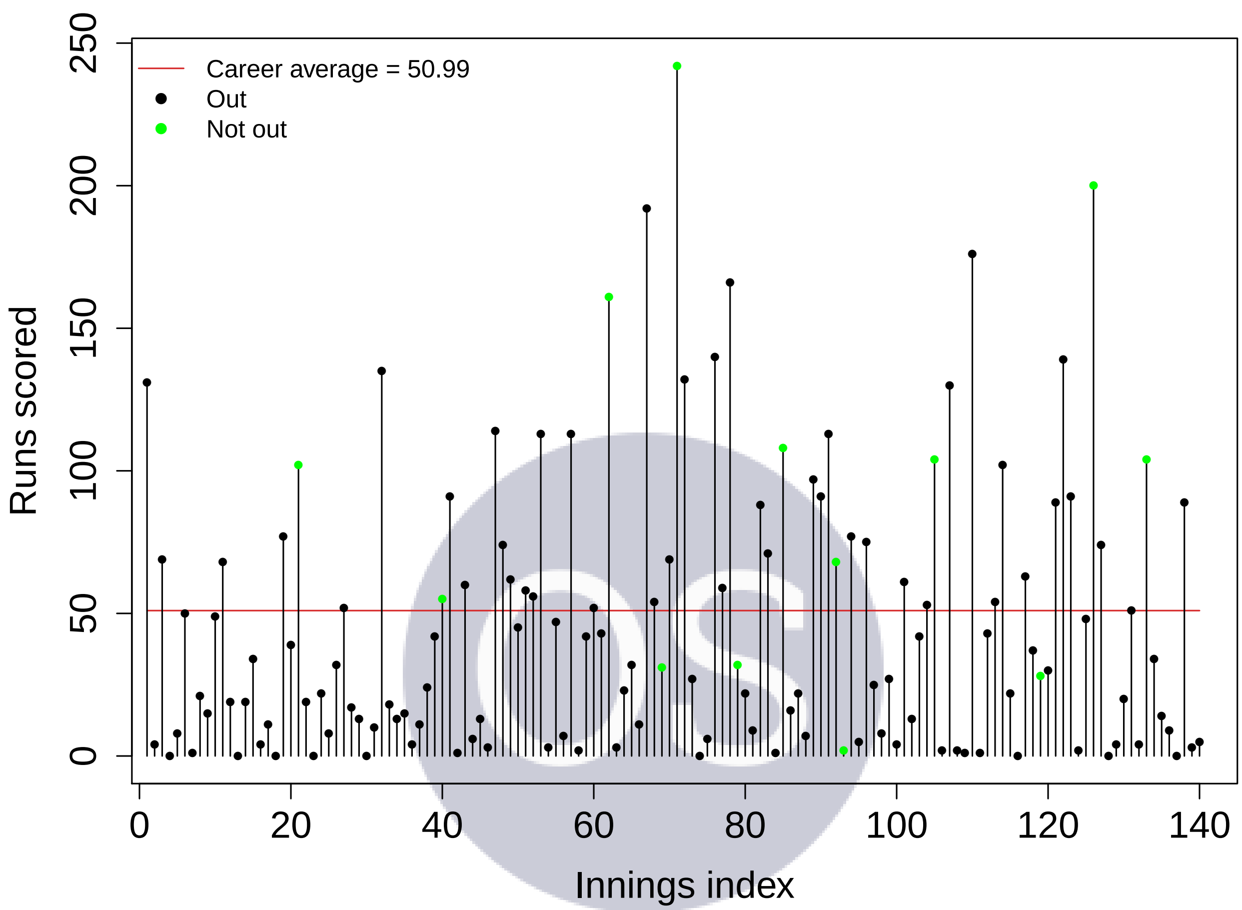


Figure 1.2. *Career batting data for Kane Williamson in Test matches.*

To illustrate how a typical career batting record can be visualised, a graphical representation of Kane Williamson's Test match batting career is presented in Figure 1.2. A feature of cricket that makes it difficult to ascertain the true underlying ability of a player, is the significant amount of variation observed between individual batting and bowling performances. Or, speaking in a more statistical manner, it can be difficult to distinguish the signal from the noise. While batting, due to the notion of getting your eye in, even the best batsmen in the world are more likely to fail than succeed in any given innings. As a result, many players only score higher than their career batting average in roughly 30–40% of all completed innings. Most professional cricketers are aware of this, as suggested by Rahul Dravid, India's second highest run scorer of all time.

"In cricket you fail a lot more than you succeed. In batting, in general, you fail a lot more. If you consider a fifty as a success point, you don't cross fifty in the majority of your innings, so you do learn to fail a lot in cricket, and a guy who has an average of 50 in international cricket has failed a lot more times than he has succeeded."

– Rahul Dravid (ESPNcricinfo, 2020)

In the case of Williamson, fewer than 50% of his completed Test innings to date have resulted in a score of 30 or higher, a score Williamson himself would no doubt consider underwhelming in many circumstances. While his overall career batting average of 50.99 provides a glowing reference of his historical batting ability, it does not necessarily indicate how good he is at present, or, how good he is likely to be in the future. Nor does it inform whether his more recent scores are any indication of his current ability.

The methods discussed in this thesis aim to provide insight regarding the presence of temporal variation in individual player batting and bowling abilities, on both short and long-term scales. As there has been a shift in terms of funding and viewership to shorter form Twenty20 cricket in the last decade, few modern metrics have been, or can be, applied to longer form domestic first-class or international Test cricket. Given the recent attention given to shorter form cricket and the number of situational and contextual complications it brings (Davis et al., 2015), the primary application of these methods is in first-class and Test cricket, where batting and bowling performances depend less on the context of a match and more on the underlying abilities of the players competing.

1.4 Data sources

The data analysed in this thesis consist of two distinct formats.

1. Summary data
   - Career records of batting and bowling performances, summarised on a per-innings basis, for individual players.
2. Ball-by-ball data
   - Data pertaining to the outcome of each individual ball bowled in a match.

The data come from three distinct sources, each of which provides the data in a slightly different format. Data from each source are cleaned, formatted and stored in a standardised structure using the R computer language (Ihaka & Gentleman, 1996; R Core Team, 2020). This maintains a consistent means of accessing, analysing, and visualising the data from each source, using the same R code and functions.

1.4.1 ESPNcricinfo

A primary data source for international cricket data is Statsguru, the statistics database hosted on the ESPNcricinfo website[2], which provides summary data for every international match played in the modern era. Test career batting records for every player to have played in a Test match since the year 2000 were scraped from Statsguru using R and stored in individual CSV files. These player records are continually updated as matches are played out in real-time.

Table 1.5. *Excerpt of Test match summary batting data scraped from Statsguru for Kane Williamson.*

Innings index	Runs	Innings	Venue	BF	Pos	Dismissal	Opposition	Ground	Start date
132	4	1	home	20	3	caught	England	Hamilton	29 Nov 2019
133	104*	3	home	234	3	not out	England	Hamilton	29 Nov 2019
134	34	2	away	70	3	caught	Australia	Perth	12 Dec 2019
135	14	4	away	8	3	caught	Australia	Perth	12 Dec 2019
136	9	2	away	14	3	caught	Australia	Melbourne	26 Dec 2019
137	0	4	away	9	3	lbw	Australia	Melbourne	26 Dec 2019
138	89	2	home	153	3	caught	India	Wellington	21 Feb 2020
139	3	2	home	8	3	caught	India	Christchurch	29 Feb 2020
140	5	4	home	8	3	caught	India	Christchurch	29 Feb 2020

A range of variables are available for each batting performance, including the usual information recorded on match scorecards, such as runs scored, balls faced, and whether the batsman remained on a not out score. Innings-specific variables are also recorded, such as which innings of the match the performance took place in; whether the player was batting at a home, neutral or away venue; and the player's position in their team's batting order. Table 1.5 provides an example of how the summary batting data are stored.

1.4.2 New Zealand Cricket

A second data source has been provided by New Zealand Cricket (NZC), the country's national cricket board. The data are made available through NV Play, a private company based in New Zealand, who offer a global cricket technology platform for professional and recreational cricketers. NV Play manage and store the data in a SQL database on behalf of NZC, which can then be accessed by authenticated users.

[2]https://www.espncricinfo.com/

Given NZC's close connections with the wider cricketing community, the data provided are incredibly rich and include ball-by-ball information for all professional domestic and international matches played in New Zealand, and a number of international matches played overseas. Table 1.6 provides a non-exhaustive overview of the range of variables present in the data for many balls bowled.

Table 1.6. *Examples of variables available for each ball in the ball-by-ball data provided by NZC.*

General match information		
• Home team	• Away team	
Innings-specific information		
• Innings #	• Batting team	• Bowling team
Ball-specific information		
• Date of delivery	• Time of delivery	
• Over #	• Ball #	
• Bowler	• Batsman	
• Wicket type	• Batsman dismissed	
• Extras scored	• Types of extra scored	
• Runs scored	• Shot type	
• Batsman footwork	• Batsman connection	
• Batsman score	• Team score	

The depth of the data provides almost limitless avenues of exploration for any user with an interest in cricket. The content of this thesis primarily focuses on the analysis of player performances and match outcome prediction, and does not even scratch the surface of what could be done with certain elements of the data. For example, ball tracking data are provided for a large number of international matches and provides significant insights as to where certain bowlers bowl and where certain batsmen tend to hit the ball. Based on over 300,000 deliveries for both left and right-handed batsmen, Figures 1.3 and 1.4 clearly illustrate bowlers' tendencies to bowl in the *corridor of uncertainty*, on, or just outside the top of a batsman's off-stump, with the occasional short-pitched bouncer thrown in.

The results presented in this thesis focus tend to focus on international Test cricket. While the NZC data source is incredibly rich in terms of the amount of information provided for every ball bowled, in the context of Test cricket the data are incomplete, as data for a number of Test matches played outside of New Zealand in the last decade or so are missing. However, while not specifically discussed, the methods and statistical applications detailed in this thesis were first applied to data pertaining to New Zealand's various domestic cricket leagues, such as the

Plunket Shield, the country's premier first-class cricket competition. As such, it is important to acknowledge the critical role this data source played in the development and implementation of the models discussed in later chapters.

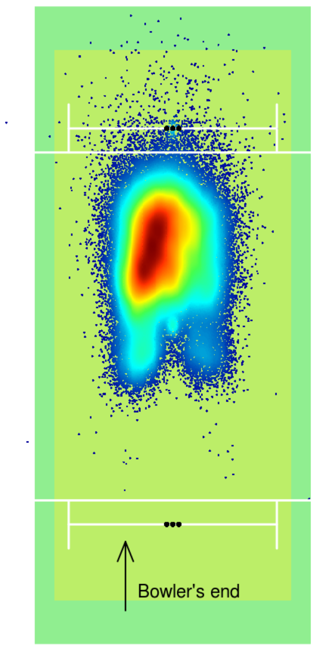


Figure 1.3. *A bird's eye view heat map showing the locations of where balls tend to pitch on the wicket. Blue indicates areas of low density and red indicates areas of high density.*

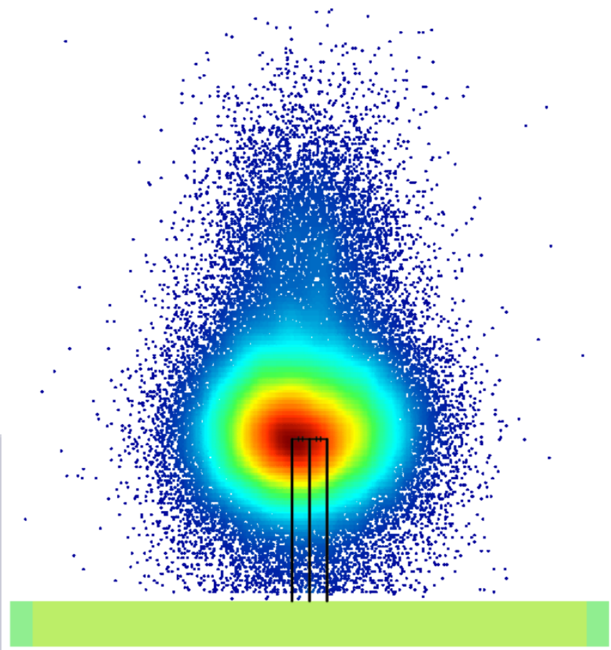


Figure 1.4. *Heat map showing the locations of where balls tend to pass the stumps. Blue indicates areas of low density and red indicates areas of high density.*

1.4.3 Cricsheet

This data source consists of ball-by-ball data for Test matches, downloaded from the Cricsheet website[3]. The data set consists of ball-by-ball data for 532 Tests played between 1st January 2008 and 1st December 2020. This covers all Test matches played over the period, with the exception of four identified missing matches that occurred in 2008. Given very few international players from 2008 still feature in international fixtures in 2020 and beyond, the Cricsheet data includes ball-by-ball information for every delivery faced, or bowled by, the majority of current Test cricketers.

Each ball-by-ball match file is stored in YAML format, which is then cleaned and formatted using R. The level of information recorded for each ball is not as in-depth as the NZC data provided in Table 1.6; namely, the data omits personal player characteristic information, such as

[3] https://cricsheet.org/

batting and bowling handedness and individual bowling type. However, all essential information is available, such as the individual batsman/bowler matchup for each ball, the outcome of the ball and all innings and venue-specific data.

1.5 Bayesian inference

The methods and models proposed in this thesis have been developed and implemented under a Bayesian framework. Working within a Bayesian context allows for the expression of any uncertainty in terms of probability (O'Hagan & Forster, 2004), while providing the tools to continually update estimates and opinions that are formed about individual players, as more matches are played and more data are collected.

1.5.1 Bayes' theorem

When applying the principles of Bayesian inference, the goal is to derive meaningful inference in regards to a set of model parameters, $\boldsymbol{\theta}$, through the observation of data, $\boldsymbol{D}$. Prior to observing data, certain choices must be made in regards to the questions one is trying to answer and the assumptions one is willing to make. An initial part of this process includes specifying *prior distributions* for the parameters of interest, $\boldsymbol{\theta}$, representing one's initial or current state of beliefs regarding a certain parameter. Given a proposed model and corresponding set of assumptions, M, a prior distribution can be expressed as $P(\boldsymbol{\theta}|M)$. An advantage of working within the Bayesian paradigm when dealing with statistical problems that have real-world applications is the ability to formally impart expert judgement on the topic, via the assignment of subjective prior distributions. In the case of cricket, an experienced viewer will have an intuitive idea about how the game works and realistic expectations in regards to the feasible range of underlying abilities of professional cricketers.

Upon observing data, $\boldsymbol{D}$, it is possible to update one's beliefs regarding $\boldsymbol{\theta}$, by expressing them as a *posterior distribution*, $P(\boldsymbol{\theta}|\boldsymbol{D}, M)$. In order to obtain a posterior distribution, one must first consider the likelihood of observing the specific data, given the specified prior beliefs, $P(\boldsymbol{D}|\boldsymbol{\theta}, M)$. If the observed data are different to what was expected under the prior assumptions of the model, $P(\boldsymbol{\theta}|M)$, then the posterior distribution will be notably different to the prior distribution, reflecting the plausibility of observing similar data in the future. It is possible to obtain a posterior distribution, $P(\boldsymbol{\theta}|\boldsymbol{D}, M)$, from a prior distribution, $P(\boldsymbol{\theta}|M)$, and observed data, $\boldsymbol{D}$, using Bayes' theorem, which is derived in Equation 1.4 below.

Let A and B define two events. Using the product rule, the probability of A and B can be expressed as follows

$$\begin{aligned} P(A \cap B) &= P(A)\ P(B|A), \\ P(B \cap A) &= P(B)\ P(A|B). \end{aligned} \tag{1.3}$$

As $P(B \cap A) = P(A \cap B)$, it follows that

$$\begin{aligned} P(B)\ P(A|B) &= P(A)\ P(B|A), \\ \therefore P(A|B) &= \frac{P(A)\ P(B|A)}{P(B)}. \end{aligned} \tag{1.4}$$

Therefore, using the result of Bayes' theorem from Equation 1.4, the posterior distribution for $\boldsymbol{\theta}$ can be expressed as

$$P(\boldsymbol{\theta}|\boldsymbol{D}, M) = \frac{P(\boldsymbol{\theta}|M)\ P(\boldsymbol{D}|\boldsymbol{\theta}, M)}{P(\boldsymbol{D}|M)}. \tag{1.5}$$

The denominator, $P(\boldsymbol{D}|M)$, in Equation 1.5 is often referred to as the *evidence* or *marginal likelihood*. This quantity represents the likelihood of observing the data, given the prior model assumptions, and is computed by integrating over the parameter space, $\boldsymbol{\theta}$, weighted by the prior plausibility about each possible value for the parameter, $P(\boldsymbol{\theta}|M)$. More simply, one can express Bayes' theorem as follows

$$\text{Posterior} = \frac{\text{Prior} \times \text{Likelihood}}{\text{Marginal likelihood}}. \tag{1.6}$$

As the marginal likelihood, $P(\boldsymbol{D}|M)$, is often difficult to obtain, the posterior distribution is often simplified and expressed as

$$\text{Posterior} \propto \text{Prior} \times \text{Likelihood}, \tag{1.7}$$

or, more formally

$$P(\boldsymbol{\theta}|\boldsymbol{D}, M) \propto P(\boldsymbol{\theta}|M)\ P(\boldsymbol{D}|\boldsymbol{\theta}, M). \tag{1.8}$$

1.5.2 Monte Carlo sampling methods

When dealing with multiple parameters in the parameter space, $\boldsymbol{\theta}$, the joint posterior distribution, $P(\boldsymbol{\theta}|\boldsymbol{D}, M)$, can be of high dimensionality, which can be difficult to express numerically. A common solution is to summarise the model parameters by sampling from the joint posterior distribution, using numerical sampling techniques such as Markov chain Monte Carlo (MCMC).

Popular MCMC methods include the Metropolis-Hastings algorithm (Hastings, 1970; Metropolis et al., 1953), Gibbs sampling (Geman & Geman, 1984), slice sampling (Neal, 2003) and Hamiltonian MCMC (Neal, 2011).

Nested sampling

Nested sampling is a sampling scheme developed by physicist John Skilling, with the primary purpose of computing the marginal likelihood or evidence, $P(\boldsymbol{D}|M)$, hereafter denoted as Z (Skilling, 2006). Given a set of prior assumptions about some model parameters, $P(\boldsymbol{\theta})$, the model likelihood, $P(\boldsymbol{D}|\boldsymbol{\theta}, M)$, and the posterior distribution, $P(\boldsymbol{\theta}|\boldsymbol{D})$, an expression for Z can be obtained by rearranging Bayes' theorem in Equation 1.4.

$$\begin{aligned} P(\boldsymbol{\theta}|\boldsymbol{D}) &= \frac{P(\boldsymbol{\theta})\ P(\boldsymbol{D}|\boldsymbol{\theta}, M)}{Z} \\ P(\boldsymbol{\theta}|\boldsymbol{D}) \times Z &= P(\boldsymbol{\theta}) \times L(\boldsymbol{\theta}) \\ \text{Marginal likelihood} = Z &= \int P(\boldsymbol{\theta}) \times P(\boldsymbol{D}|\boldsymbol{\theta}, M)\ d\theta \end{aligned} \tag{1.9}$$

The ability to compute the marginal likelihood is a significant advantage of nested sampling, which, depending on the problem at hand, can be achieved at minimal extra cost. The computation of the marginal likelihood does not hinder the method's ability of estimating the posterior distributions for the model parameters, $\boldsymbol{\theta}$, which can be obtained by taking weighted samples of $\boldsymbol{\theta}$ from the output of a nested sampling run. Rather, the marginal likelihood is the primary output of nested sampling and *"samples from the posterior distribution are an optional by-product"* (Skilling, 2006).

A major benefit of possessing the marginal likelihood values of a Bayesian model is the ease of which model comparison becomes. Within a Bayesian framework, model comparison is often performed by computing a Bayes factor, which is the ratio of marginal likelihoods of two proposed models (Kass & Raftery, 1995). Consider two models, M_1 and M_2, and two corresponding sets of model parameters, $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$. When fitting each model to the same data, $\boldsymbol{D}$, the posterior distributions for each set of model parameters can be obtained using Bayes' theorem.

The quantity computed in Equation 1.10 is the posterior odds ratio, which is simply the prior odds multiplied by the Bayes factor. Values greater than 1 suggest that M_1 is preferred over M_2, while values less than 1 suggest the opposite. Assuming each model is equally likely under the prior model assumptions, the posterior odds ratio is simply equal to the Bayes factor. Put simply, whichever model has the higher marginal likelihood is the preferred model, which makes for trivial model comparison.

$$\frac{P(M_1|\boldsymbol{D})}{P(M_2|\boldsymbol{D})} = \frac{P(M_1)}{P(M_2)} \times \frac{P(\boldsymbol{D}|M_1)}{P(\boldsymbol{D}|M_2)} \tag{1.10}$$

Unless stated otherwise, the majority of the model fitting process in this thesis is performed using classical nested sampling, as proposed by Skilling (2006). The nested sampling algorithm is implemented in C++ (Stroustrup, 2013) with the relevant model output files saved in a TXT or CSV file format. These output files are then post-processed in R (Ihaka & Gentleman, 1996; R Core Team, 2020) for the purpose of visualising the results and obtain meaningful inference.

1.6 Thesis overview

In this thesis, a series of models that employ the use of machine learning algorithms are proposed for estimating and predicting the past, present, and future batting and bowling abilities of professional cricket players. The resulting model outputs are used to investigate the plausibility of various cricket-related concepts, such as whether there is any evidence to suggest that recent form is really an accurate predictor of future performance. A novel method of combining these estimates in order to accurately predict the outcome of any given match is then discussed, along with potential real-world applications of these predictive outputs. A list of references is provided immediately after Chapter 5, followed by an Appendix containing relevant R code referred to throughout Chapter 4.

This section provides a brief overview of each chapter, with the exception of Chapter 5.

1.6.1 Estimating batting career trajectories

As a point estimate, the batting average is a metric that is unable to effectively describe underlying changes in a player's batting ability. Such fluctuations in ability may exist on both short and long-term scales, both during and between individual innings. A method for quantifying potential variation in batting ability observed within a single innings was discussed in Stevenson & Brewer (2017). However, like the batting average, this method makes the assumption that a player's underlying ability remains constant across all innings of a career. Recent form is often cited as an indicator of current batting ability and is frequently used as a justification for dropping or selecting specific players from domestic and international teams. Therefore, if form is truly a predictor of future performance, it should be possible to develop a model which can identify temporal deviations in ability across the careers of individual players.

In Chapter 2, the model developed in Stevenson & Brewer (2017) is extended to have the functionality to detect and measure changes in underlying batting ability, both within a single innings, and between multiple innings, across entire playing careers. The proposed model estimates a player's batting career trajectory, with the intention of (1) estimating how a player's underlying batting ability has varied over the course of their career to date; (2) predicting a player's current and future batting ability; and (3) quantifying any short or long-term temporal effects that may exist at the individual level, due to a player's past performances and recent form. The impact of several external, innings-specific factors is also discussed. The model output is provided in units of a batting average, which allows for an easy to understand cricketing interpretation and can be directly compared with the batting average. It is shown that the proposed model and corresponding predictions of ability generally provide more accurate and intuitive predictions of future player performance than other metrics, such as the batting average.

1.6.2 Estimating bowling career trajectories

Like the batting average, the bowling average fails to inform about variations in ability that occur over time. As such, it is of practical use to develop a class of model that can estimate the bowling career trajectory of individual players, to measure and predict how a player's underlying bowling ability might vary over the course of their career. However, quantifying the value of specific bowling performances comes with several additional challenges, not present in a batting context.

Firstly, while batting performances can be effectively summarised by the number of runs scored, bowling performances are typically reported using two variables: runs conceded and wickets taken. This can make it difficult to objectively compare two distinct career performances, or to compare performances between two players in a particular innings. Secondly, the quality of batsmen bowled to during an innings can vary significantly between separate bowling efforts, further compounding the difficulty of comparing performances.

In Chapter 3, a novel method of summarising bowling performances is proposed, removing the need to quantify bowling efforts using multiple variables. Furthermore, a new metric, *standardised runs*, is introduced to adjust for the quality of batsmen that bowlers have conceded runs against during their career. Under this specification, individual bowling performances can be compared in a more objective manner, as the quality of opposition has been accounted for. The adjusted data are then used to develop a model, similar to that detailed in Chapter 2, to estimate bowling career trajectories, measuring and predicting the past, present, and future bowling abilities of individual players.

1.6.3 A simulation-based method of match outcome prediction

As previously identified, few proposed methods of predicting the outcome of a cricket match consider the individual strengths of the participating players. Instead, estimates of relative team strength are typically based on historic team performance, or a team rating system, such as the official ICC team ratings. Therefore, such methods have difficulty adapting in situations when key players are ruled out with injury, or where team lineups are significantly different to what has been observed in the past.

In Chapter 4, a simulation-based method of predicting the outcome of Test matches, given two proposed playing XIs, is discussed. The proposed methodology utilises the estimates of player batting and bowling ability derived in Chapters 2 and 3, which provides an additional level of detail to the estimated probabilities, not considered by other methods. An overview of the simulation process is provided, detailing how the various intricacies and complexities of Test cricket are accounted for.

Several practical applications of the simulation tool are discussed. These include a potential public interest, in terms of possible uses in the broadcasting of Test cricket, and a private interest, whereby the results can be applied by coaches and selectors of professional teams to gain a competitive advantage over their opposition.

Chapter 2

Estimating batting career trajectories

2.1 Introduction

As discussed in Section 1.3.4, when it comes to evaluating the batting abilities of individual players there are a number of limitations with the batting average and many of its proposed alternatives, such as the ICC rating system. Such methods struggle to maintain an intuitive cricketing interpretation while also lacking the ability to inform about variations in ability that occur on two scales (1) short-term changes that occur *during* or *within* a single innings, due to factors relating to the concept of getting your eye in; and (2) long-term changes that occur *between* innings, over the course of entire playing careers, which are a result of players participating in specialised coaching programmes, gaining experience in various match conditions, as well as general improvements or deteriorations in the likes of technique, eyesight and agility. Furthermore, as identified in Section 1.3.4, many modern methods have tended to focus on Twenty20 cricket, given the format's global rise in popularity over the last decade. The methods presented in this chapter focus on longer form first-class and Test cricket, where batting decisions are less reliant on match context and are therefore more indicative of a player's true underlying ability.

2.1.1 Estimating short-term variation in ability

An implicit assumption of batting and bowling averages is that a player performs with some constant ability during a given innings. While this assumption may hold some validity when bowling, there is plenty of anecdotal evidence to suggest that, within each innings, batting ability develops over time as a player scores runs and gets their eye in. One of the earliest documented statistical studies of cricket provided empirical evidence to suggest that a batsman's set of career scores could be modelled by a geometric progression, supporting the idea of an approximately constant batting ability during a single innings (Elderton & Wood, 1945). However, multiple studies have since shown that the geometric assumptions do not hold for numerous players, due to the inflated number of low scores, particularly scores of zero *(ducks)*, that are present in many players' career records (A. C. Kimber & Hansford, 1993; Brewer, 2008; Bracewell & Ruggiero, 2009; Stevenson & Brewer, 2017).

The findings of these later studies provide a theoretical justification to the concept of getting your eye in and suggest that many players exhibit some form of short-term variation in batting ability, which is not expressed by the batting average. Rather than directly model batting scores, as was the approach in Elderton & Wood (1945), A. C. Kimber & Hansford (1993) introduced the concept of a hazard function, $H(x)$, while batting, representing the probability that a batsman currently on score x gets out, rather than progressing to a higher score. The results found that hazard functions generally have larger values for low values of x, indicating that batsmen are more likely to get out on low scores, early in their innings, supporting the idea of getting your eye in. However, as the underlying method used to estimate the hazard function was nonparametric, the estimates for $H(x)$ become unreliable for larger scores, due to the sparsity of the data at these scores for many players. Cai et al. (2002) applied a parametric smoother to the empirical dismissal probabilities implied by the hazard function, which helped partially address the issues of data sparsity. However, as the underlying hazard function was still modelled using a nonparametric estimator, the issues of data sparsity remained and continued to distort the hazard function at higher scores.

To account for the inflated number of scores of zero present in many career records, Bracewell & Ruggiero (2009) proposed a *Ducks 'n' runs* distribution for modelling batting scores. The distribution aims to estimate a batsman's likely contribution towards their team's total and consists of a mixture of a beta and geometric distribution. The beta component estimates the probability of a player failing to contribute any runs (getting a duck), while the geometric component represents the distribution of non-zero scores. While this method addresses the inflated number of ducks present in a player's career record, the model assumes that once avoiding a score of zero, a batsman plays out the remainder of their innings with a constant ability. This implicitly assumes that the getting your eye in process is complete upon scoring just a single run, which is probably not a reasonable assumption in the context of most innings.

To better quantify the effect of getting your eye in for individual batsmen, Brewer (2008) and Stevenson & Brewer (2017) introduced the concept of the *effective batting average*, $\mu(x)$, representing a batsman's underlying batting ability on score x, in units of a batting average. Given the prevalence of the batting average in cricket, it is far more intuitive for coaches, players and viewers of the game to think of batting ability in terms of batting averages, rather than dismissal probabilities implied by the hazard function, $H(x)$. The hazard function and effective average function can be expressed in terms of one another, as shown in Equations 2.1 and 2.2:

$$H(x) = \frac{1}{\mu(x) + 1}, \tag{2.1}$$

$$\mu(x) = \frac{1}{H(x)} - 1. \tag{2.2}$$

The effective average has a relatively simple interpretation. For example, a value of $\mu(0) = 20$, indicates that when on a score of zero, a player bats with the average ability of somebody who has a career batting average of 20. Similarly, a value of $\mu(50) = 60$, suggests that when on a score of 50, the player bats with the ability of somebody with a career batting average of 60.

In this series of papers, the authors fitted parametric models in a Bayesian context in order to estimate how the underlying effective average, $\mu(x)$, varies over the course of an innings for individual batsmen (Brewer, 2008; Stevenson & Brewer, 2017). The model consists of three parameters, $\{\mu_1, \mu_2, L\}$, which aim to quantify:

1. A player's initial batting ability when starting a new innings (μ_1).

2. A player's eye-in, equilibrium or peak batting ability, once familiar with the specific match conditions (μ_2).

3. How long it takes a player to transition from their initial batting state to their peak state (L).

Both μ_1 and μ_2 are expressed in units of a batting average. The timescale parameter, L, measures the speed of transition between μ_1 and μ_2 and is formally the *e-folding time*, a measurement usually seen in theoretical physics. This quantity represents the number of runs scored for approximately 63% (formally $1 - \frac{1}{e}$) of the transition between μ_1 and μ_2 to take place and can be understood by analogy with a half-life.

Constructing the effective average function, $\mu(x)$, from the posterior distributions of the model parameters, allows for the batting ability of any player to be quantified, while batting on any score x. This helps answer questions about specific players, such as (1) how well a player performs when beginning a new innings, (2) how well a player performs once they have their eye in, and (3) how long it takes a player to get their eye in. For the vast majority of past and present Test match players analysed, Stevenson & Brewer (2017) found overwhelming evidence to suggest that players are far more likely to get out early in their innings, while on a low score, than later on in their innings, further supporting the notion of getting your eye in.

To provide a practical example, the posterior predictive effective average function, $\mu(x)$, for Kane Williamson is presented in Figure 2.1. As per the authors' recommendation, the posterior

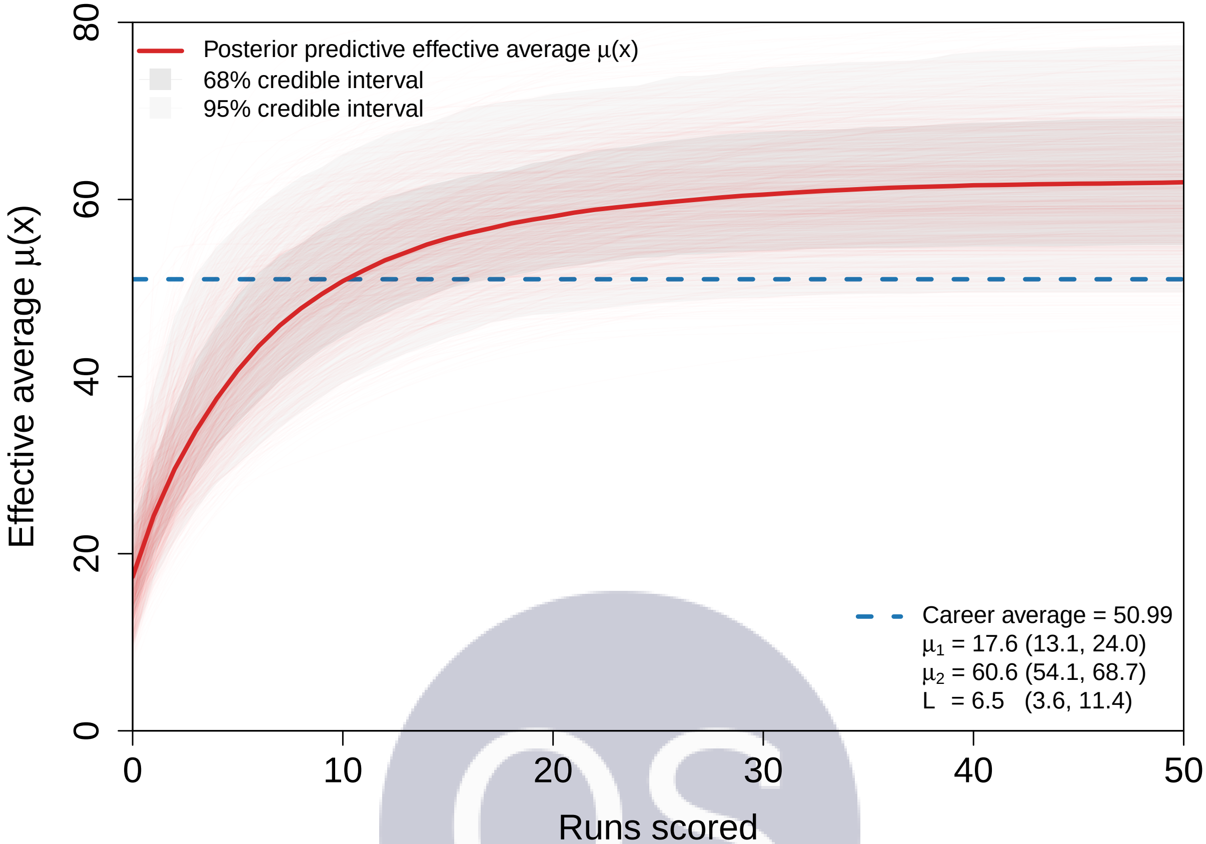


Figure 2.1. *Estimated posterior predictive effective average function, $\mu(x)$, for Kane Williamson, including the 68% and 95% credible intervals. Each faint red line represents a posterior sample that is used in the calculation of the posterior predictive estimate. Posterior parameter summaries and 68% credible intervals are also provided, as is Williamson's career average, which infers a constant ability during an innings.*

median is used to summarise the parameter distributions, as the posterior distributions are not necessarily symmetric and can have heavy tails. Clearly the estimated values for $\mu(x)$ are smaller for low scores, illustrating the effect of getting your eye in for Kane Williamson. The estimated posterior value for $L = 6.5$ suggests that after scoring 7 or so runs, Williamson has transitioned almost two thirds of the way between his initial and peak batting states. A quick glance at Figure 2.1, suggests Williamson is batting near his peak ability after scoring roughly 30 to 40 runs. The difference between the expected distribution of scores using the effective average, $\mu(x)$, versus a geometric distribution is then presented in Figure 2.2.

Teams may benefit from this type of analysis, as they can identify opposition players who are particularly vulnerable at the beginning of an innings and may be able to capitalise on such a weakness by setting more attacking fields than usual, early in the specific player's innings. It

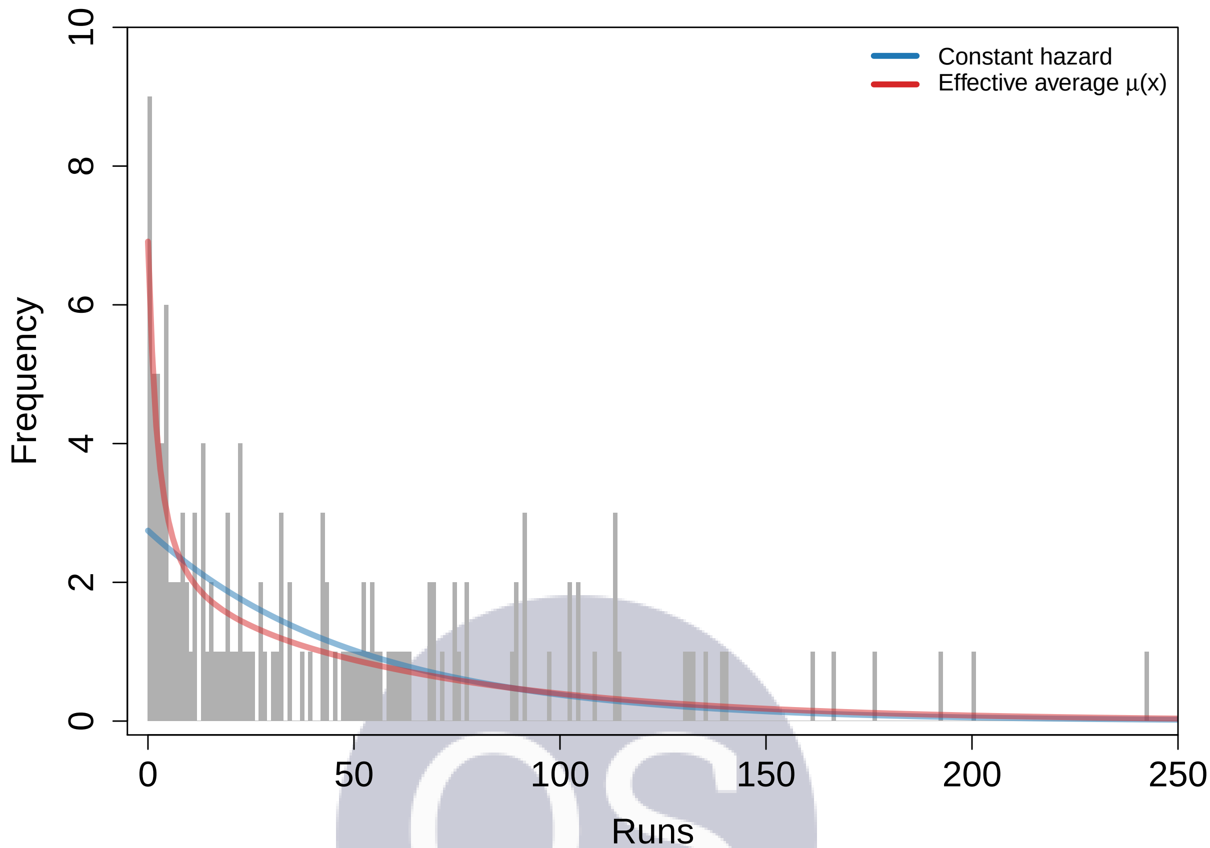


Figure 2.2. *Histogram of Kane Williamson's Test career scores (grey). The implied distribution of scores based on a constant hazard model that assumes a geometric distribution (blue) and effective average model (red) are provided for comparison. The constant hazard model under predicts the probability of being dismissed on a low score as it ignores the effect of getting your eye in, while the effective average model does a better job of quantifying this effect.*

is worth noting that the model does not consider potentially important factors such as balls faced, boundaries scored, or the quality of bowlers faced. However, it has been shown to provide far more accurate estimates of how player batting ability can vary during the course of a single innings than the batting average and allows for some interesting player comparisons.

For example, in Figure 2.3 and Table 2.1, the posterior predictive effective average functions, $\mu(x)$, and relevant parameter summaries are presented for a group of batsman — colloquially referred to as *the big four* (J. Kimber, 2017) — who have dominated world cricket in recent years. The posterior parameter point estimates obtained from the method detailed in Stevenson & Brewer (2017) suggest that despite having the lowest career average of the big four, English captain Joe Root appears to be a stronger batsmen at the start of an innings and gets his eye in faster than Kane Williamson. However, once the players have got their eye in and reached their peak ability, Williamson is likely the superior batsman. Similarly, Indian captain Virat Kohli appears to begin an innings with a lesser ability than Root, but gets his eye in very quickly and has a peak ability somewhere between that of Root and Williamson. On the other hand, former Australian captain Steve Smith appears to begin an innings already batting at a high ability, but takes slightly longer to reach his peak ability, compared with the other players in the big four.

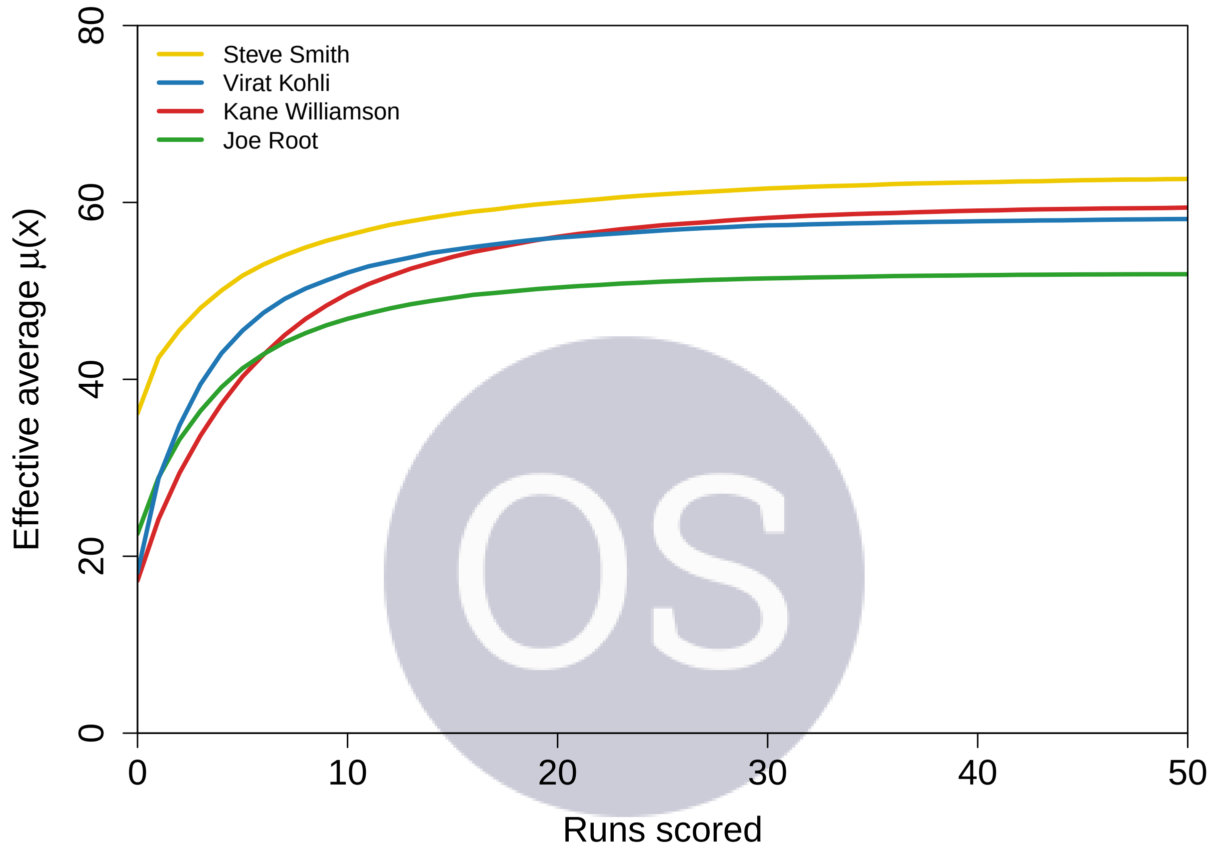


Figure 2.3. *Estimated posterior predictive effective average function, $\mu(x)$, for the big four.*

These findings are consistent with Root's oft-touted ability of making a solid start to an innings, but failure to regularly convert scores of 50+ into scores of 100+. Root has made 17 scores of 100+ but has been dismissed 46 times between scores of 50 and 100. Ignoring not out scores between 50 and 100, Root has converted just 27.0% of his 50+ scores into centuries, which is considerably lower than the conversion rates of Smith (52.0%), Kohli (57.4%), and Williamson (41.2%). The results may also suggest that opposition teams could benefit by setting more attacking fields early on to Williamson and Kohli, before they have their eye in. Additionally, it is possible that teams could set aggressive fields for a prolonged period to a player like Steve Smith, who appears to take longer to get his eye in.

As with any statistical model there is a degree of uncertainty associated with these estimates.

Table 2.1. *Posterior parameter summaries for the within innings effective average function, $\mu(x)$, for the big four, including point estimates and 68% credible intervals.*

Player	Career average	μ_1	μ_2	L
Steve Smith (AUS)	62.8	36.2 (26.2, 46.4)	63.1 (56.5, 71.6)	6.9 (1.5, 17.2)
Virat Kohli (IND)	53.6	18.1 (12.9, 26.0)	58.3 (52.2, 65.9)	4.3 (1.4, 10.9)
Kane Williamson (NZ)	51.0	17.2 (12.6, 23.5)	59.8 (52.7, 68.0)	6.6 (3.6, 11.5)
Joe Root (ENG)	48.0	22.5 (16.6, 29.9)	52.0 (47.2, 58.0)	5.2 (2.4, 10.1)

However, an advantage of performing the analysis in a Bayesian context is the ease of which one can make probabilistic statements in regards to the parameters of interest by drawing a number of posterior samples. For example, comparing the posterior distributions for μ_2 suggests there is an 78.4% chance that Kane Williamson has a superior peak batting ability compared with Joe Root, a 55.3% chance compared with Virat Kohli, and a 36.6% chance compared with Steve Smith.

In a related study, Stevenson (2017) applied a set of more flexible models that allow for score-based deviations in the effective average function at any score, not just at the beginning of an innings when a batsman is getting their eye in. The aim of this analysis was to determine whether there is any evidence to suggest that batsmen are more likely to get out on certain scores — as suggested by the cricketing superstition *the nervous 90s* — whereby players are thought to bat with inferior ability due to nerves that arise when nearing the significant milestone score of 100. Although there is plenty of anecdotal evidence to suggest that players do get nervous in the 90s, Stevenson (2017) found no conclusive evidence to suggest batting ability is affected by these nerves. Instead, some evidence was found to suggest that players are more likely to get out immediately after passing significant milestones such as 50 and 100, suggesting that perhaps the *fallible 50s* and *hazardous 100s* would be more justified clichés.

2.1.2 Estimating long-term variation in ability

Section 2.1.1 has established that there is plenty of statistical evidence to suggest that players do not bat with some constant ability during an innings or match, due to the concept of getting your eye in. However, it is equally unlikely, if not more unlikely, that players perform with some constant ability throughout an entire playing career, which can span up to 20 years. Instead, variations and fluctuations in a player's underlying abilities are likely to occur between innings and matches. Viewers of the game, ranging from armchair spectators, to newspaper columnists, to national selectors, will often cite a player's recent performances, or form, as a reason for selecting a new player or dropping an incumbent. As previously noted, due to the nature of the sport and the process of getting your eye in, the majority of batsmen are more likely to fail than succeed in any given innings, resulting in a probability distribution over scores that is relatively heavy-tailed, and data sets that are noisy. Consequently, it is common to observe players string together a number of low scores in a row, even if their underlying ability has not changed.

If recent batting form were to have a considerable impact on player performance then it should be possible to view a player's career batting record and identify sustained sequences of both high scores (indicating a player is *in form*) and low scores (indicating a player is *out of form*) that are statistically significant. On the contrary, an analysis of 16 Test match batsmen by Durbach & Thiart (2007) found little empirical evidence to suggest that past performances were predictive of future scores, with the majority of players' scores being well described by an independent, identically distributed sequence. Instead, the authors concluded that public perceptions of batting form tend to be overstated, possibly due to a case of recency bias, which is known to impact sports betting markets (Bailey & Clarke, 2006) and results in more importance given to recent performances and outcomes. Similarly, A. C. Kimber & Hansford (1993) justified the use of a batting average to quantify ability as they found no significant evidence to suggest the presence of autocorrelation in a player's career record. However, it is worth noting that the sample size of players analysed in each of these studies is relatively low.

Related concepts have been investigated in other sports, such as the idea of the *hot hand* in basketball, which suggests that players are more likely to make subsequent shot attempts after making several shots in a row. For a long time, the existence of such effects have been treated with scepticism (Gilovich et al., 1985; Tversky & Gilovich, 1989). However, more recent studies have suggested that players on a hot streak tend to start taking more difficult shots, which may mask any hot hand effect due to recent form, even if players are making these tough shots at a higher rate than usual (Csapo et al., 2015). A more modern and pragmatic viewpoint is that for many players there is no evidence of a hot hand effect, but this varies from player to player and such an effect may exist in certain circumstances (Shea, 2014; Wetzels et al., 2016). An investigation into the hot hand in professional darts yielded similar findings; there is some weak support for such an effect, but the evidence is generally inconclusive (Ötting et al., 2020).

Returning to a cricketing context, there is further evidence from psychological studies to indicate that a player's recent form can impact mood, anxiety and stress levels, which in turn can affect player performance. A study which canvassed the opinions of professional cricketers found that the majority of players surveyed believed a small amount of stress and tension is beneficial for performance (Sahni & Bhogal, 2017). On the other hand, excessive external pressure has been proposed to impact performance negatively (Totterdell, 1999). Consider the following quotes from Jeet Raval, who was an incumbent opening batsman for the New Zealand Test side between 2017 and 2019, before being dropped in early 2020.

"I would have loved to get not just a hundred, but a big hundred because that would have helped us get into a winning position. That's what this Black Cap team is about."

– A confident Jeet Raval after scoring two fifties in his first two Test matches (Moonda, 2017)

"I hope I don't get out this ball."

– Raval opens up about his batting mindset while playing England in late-2019 (Kishore, 2020)

Raval appeared full of confidence after a successful start to his Test career and continued to score consistent runs for New Zealand. However, by his own admission, Raval was simply in survival mode while batting after a turbulent 2019 where he struggled to score runs. In hindsight, it is likely that Raval's negative thought patterns contributed towards his ongoing difficulties with the bat before being dropped in 2020. This sentiment is echoed below by Indian great Rahul Dravid.

"If you're switched on or too intense all the time, it drains you of a lot of mental energy and when you need that to play, you won't have any of it because you'd already be so tired mentally."

– Rahul Dravid on how too much intensity and pressure can be detrimental (ESPNcricinfo, 2020)

Therefore, if recent performances can have an influence on a player's mental state, it is plausible that recent form may impact current batting ability. Consequently, the underlying assumption made by numerous studies that each innings in a player's career record can be treated as independent and identically distributed, may be violated to some extent.

To determine whether a batsman is in or out of form in the context of one-day cricket, Koulis et al. (2014) employed a Bayesian hidden Markov model. Here, the authors define a number of underlying batting states for individual players while estimating the number of runs to be scored when in each state. Parameters that measure *availability*, the probability a player is in form for the next match; *reliability*, the probability that a player is in form for the next n matches; and *mean time to failure*, the expected number of innings a player will bat in before they are deemed out of form, are also estimated for each player. While the findings suggest that players experience periods of being in and out of form, the model is limited by its requirement for an explicit and discrete judgement of what constitutes an in form and out of form state.

In the paper, the authors specify a batting state that has a posterior expected median number of runs scored of less than 25, as being out of form, and all other states as being in form. From the perspective of T20 cricket this is not necessarily an unreasonable specification, however there are plenty of cases in one-day cricket where a score of 25 or greater, scored at a low strike rate, may be considered an unsuccessful innings. Additionally, this approach does not consider the underlying ability of a batsman. For example, a player who typically averages above 40 might consider a state with a posterior expected median of 25 runs as an out of form state, whereas a similar state for a player who averages below 20 would certainly indicate strong recent form.

A more recent study in the context of Test cricket utilised a Poisson random-effects model for the purpose of estimating the abilities of batsmen over the course of a Test career (Boys & Philipson, 2019). The primary aims of this study were to find a method of fairly comparing batsmen across eras and identifying the age at which individual batsmen are believed to have peaked. The model considers a number of important external factors for each career innings, including the strength of the opposition, as well as innings and venue-specific effects. Similar to the concept of the effective average adopted by Stevenson & Brewer (2017), the authors interpret batting ability in units of an expected batting average for a given innings. The model output provides a means of approximating the differences in ability between batsmen in a far more meaningful manner than the ICC's rating points system, allowing for more insightful player comparisons and quantifications of batting ability.

The methods presented in this chapter build on the Bayesian parametric model developed in Stevenson & Brewer (2017), allowing for the estimation of past, present, and future batting abilities from Test career batting data (Section 2.2). The model considers both short-term (Section 2.2.2) and long-term (Section 2.2.3) variations in ability that may occur as a result of the factors outlined earlier in Sections 2.1.1 and 2.1.2. Similar to Boys & Philipson (2019), the model estimates batting ability in terms of an expected number of runs to be scored in a given innings, which helps address the limitations of the ICC rating method. The model output allows for the construction of individual player batting career trajectories, illustrating how player ability has evolved over time, which are presented alongside the general findings and practical applications in Section 2.3. In Section 2.4, the model predictions are compared with other various stationary and non-stationary methods of estimating batting ability, including the traditional batting average. Of all methods, the present model is shown to provide the most accurate estimates of future ability. The methodology detailed in this chapter formed the basis of publications Stevenson & Brewer (2018) and Stevenson & Brewer (2021).

2.2 Model specification

2.2.1 Model likelihood

The derivation of the model likelihood for a single innings, or set of career innings, follows the method detailed in Stevenson & Brewer (2017, 2018, 2021). If $X \in \{0, 1, 2, 3, ...\}$ represents the number of runs scored in a given innings, then the *hazard function*, $H(x) \in [0,1]$, defines the probability of a batsman getting out while on score x (Equation 2.3).

$$H(x) = P(X = x | X \geq x) \tag{2.3}$$

Assuming a functional form for $H(x)$, conditional on a set of parameters, $\boldsymbol{\theta}$, the probability distribution of scores, X, can be expressed in terms of the hazard function, as shown in Equation 2.4.

$$H(x) = \begin{cases} H(0), & \text{if } x = 0 \\ H(x) \prod_{r=0}^{x-1} [1 - H(r)], & \text{otherwise} \end{cases} \tag{2.4}$$

For any given value of x, Equation 2.4 is the conditional probability of a batsman surviving until score x, then being dismissed. When inferring model parameters, $\boldsymbol{\theta}$, from data, Equation 2.4 provides the likelihood function for a single innings. However, in cricket there are certain instances where a batsman's innings may end without being dismissed, referred to as a *not out* score. In the case of not out scores the likelihood is computed as $P(X \geq x)$, rather than $P(X = x)$. Comparable to right-censored observations in survival analysis, the computation of $P(X \geq x)$ for not out scores assumes a batsman would have scored some unobserved score, conditional on their current score and assumed hazard function. Treating not out scores in this manner implies that the sequence of out and not out flags in the data, without the associated score, provides no information about the model parameters, $\boldsymbol{\theta}$.

Therefore, if T is the total number of innings a player has batted in and N is the total number of not out innings, the probability distribution for a set of conditionally independent out scores, $\boldsymbol{x} = \{x_1, x_2, ..., x_{T-N}\}$, and not out scores, $\boldsymbol{y} = \{y_1, y_2, ..., y_N\}$, can be expressed as

$$P(\{\boldsymbol{x}, \boldsymbol{y}\}) = \prod_{t=1}^{T-N} \Big(H(x_t) \prod_{r=0}^{x_t - 1} \big[1 - H(r)\big] \Big) \times \prod_{t=1}^{N} \Big(\prod_{r=0}^{y_t - 1} \big[1 - H(r)\big] \Big). \tag{2.5}$$

When data $\{\boldsymbol{x}, \boldsymbol{y}\}$ are fixed and known, Equation 2.5 provides the likelihood for any proposed form of the hazard function. Therefore, conditional on the set of parameters, $\boldsymbol{\theta}$, governing the form of the hazard function, $H(x; \boldsymbol{\theta})$, one can derive the log-likelihood, $\ell(\boldsymbol{\theta})$, from Equation 2.5 as follows

$$\ell(\boldsymbol{\theta}) = \sum_{t=1}^{T-N} \log \big[H(x_t)\big] + \sum_{t=1}^{T-N} \sum_{r=0}^{x_t - 1} \log \big[1 - H(r)\big] + \sum_{t=1}^{N} \sum_{r=0}^{y_t - 1} \log \big[1 - H(r)\big]. \tag{2.6}$$

The equations presented in Section 2.2.1 define the likelihood function for any proposed theory of how a player's batting ability varies as a function of their current score. The specific parameterisation of $H(x)$ is defined in Section 2.2.2 and is similar to the model used in Stevenson & Brewer (2017). However, the model is extended to accommodate both short-term and long-term variation in batting ability by allowing several model parameters to vary over time, as well as accounting for a number of innings-specific factors that are described in Section 2.2.3.

2.2.2 Within-innings and short-term effects

Estimating how a batsman's ability changes during an innings, as they score runs and get their eye in, requires an explicit parameterisation of the hazard function, $H(x)$. As discussed earlier, this is achieved by introducing the effective average function from Stevenson & Brewer (2017), which is denoted by $\mu(x)$ and represents a player's underlying batting ability while on a score of x, in units of a batting average. The hazard function can then be expressed in terms of the effective average function as per Equation 2.7.

$$H(x) = \frac{1}{\mu(x) + 1} \tag{2.7}$$

Therefore, the hazard function is dependent on the parameterisation of the effective average function, $\mu(x)$. As detailed in Section 2.1.1, the function contains three parameters, $\boldsymbol{\theta} = \{\mu_1, \mu_2, L\}$ and takes an exponential functional form (Equation 2.8).

$$\mu(x) = \mu_2 + (\mu_1 - \mu_2) \exp\left(-\frac{x}{L}\right) \tag{2.8}$$

The definitions of these parameters have been outlined previously in Section 2.1.1 but are reiterated in Table 2.5.

2.2.3 Between-innings and long-term effects

To this point, the model is equivalent to that defined in Stevenson & Brewer (2017) and can estimate how individual batting ability changes during an innings, as function of current score. However, the model in Stevenson & Brewer (2017) assumes that a player's underlying batting ability is constant throughout their career, since the same parameter values are are used in each innings. Realistically, a player's expected scores should vary from innings to innings, due to various individual-specific factors and several innings-specific effects discussed in this section.

Individual-specific effects

In order for the model to allow for temporal variation in ability between individual innings, the effective average function in Equation 2.8 is extended by introducing a time dependence, t, giving

$$\mu(x,t;\boldsymbol{\theta}) := \text{batting ability on score } x\text{, in player's } t^{\text{th}} \text{ career innings.}$$

Again, $\mu(x,t)$ is expressed in units of a batting average. Taking the expectation over all possible x values for a given innings, t, provides the quantity $\nu(t)$, which describes the expected number of runs to be scored by a batsmen in their t^{th} career innings, given the parameters, $\boldsymbol{\theta}$.

$$\nu(t;\boldsymbol{\theta}) := \text{expected number of runs to be scored in } t^{\text{th}} \text{ career innings.}$$

Conditional on the model parameters, $\boldsymbol{\theta}$, the quantity $\nu(t;\boldsymbol{\theta})$ is equivalent to

$$\nu(t;\boldsymbol{\theta}) := \text{expected batting average in } t^{\text{th}} \text{ career innings.}$$

If all parameters defining $\mu(x,t)$ at a particular innings, t, are known, $\nu(t)$ is given by a deterministic function of those parameters; it is the expectation of the implied distribution over scores x_t. As the aim of the model is to consider temporal variation in batting ability between innings, when estimating estimating $\nu(t)$ for an individual player one must account for individual-specific factors such as a player's past performances and recent form. This may be achieved by fitting a unique eye-in or peak batting ability parameter, μ_2, for each of the t innings in a player's career, which is denoted by μ_{2_t}.

Restricting the set of parameters that are free to vary with time to only include μ_2, implies that the model assumes the process of getting your eye in occurs at a similar rate in each of a specific player's career innings. It seems a reasonable assumption that the getting your eye in process is more closely related to an individual's playing style, which for many players remains relatively constant throughout a playing career, rather than their current underlying batting ability. Conceptually, μ_{2_t} can be thought of as an innings-specific skill ceiling, which is assumed to be more likely to vary over time than a player's initial batting ability, μ_1, or the rate at which they get their eye in L. While it would be possible to allow both μ_1 and L to also vary with time, this would require the introduction of a number of additional parameters, which given the already high dimensional parameter space, would result in severely decreased computational efficiency for minimal improvement in terms of predictive accuracy.

Therefore, under this model specification the innings-specific effective average, $\mu(x,t)$, has parameters $\boldsymbol{\theta} = \{\mu_1, \{\mu_{2_t}\}, L\}$, and can be expressed as per Equation 2.9, where the dependence of μ_2 on time is explicitly noted.

$$\mu(x,t) = \mu_{2_t} + (\mu_1 - \mu_{2_t}) \exp\left(-\frac{x}{L}\right) \tag{2.9}$$

Gaussian processes

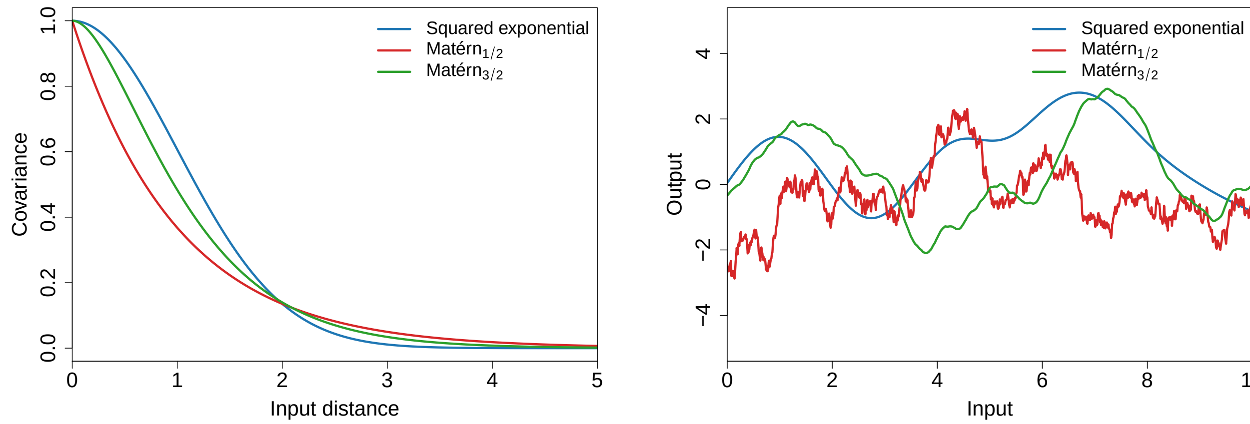
When estimating the set of $\{\mu_{2_t}\}$ terms that explain the variation in a player's underlying batting ability between innings, one must consider classes of models that afford a reasonable amount of flexibility between terms, to account for the noisy data often observed in cricket. A range of smoothing techniques and functions were contemplated, including splines and both autoregressive and moving average models, commonly used in time-series modelling. Ultimately, the prior for the set of $\{\mu_{2_t}\}$ terms was chosen to be specified using a Gaussian process, which is a class of function that can identify trends in data, while also providing a decent amount of flexibility to account for the high variance often present in a sequence of batting scores. A Gaussian process is fully specified by an underlying mean value, λ, and a covariance function $K(t_j, t_k)$, where t represents the index of a player's j^{th} and k^{th} career innings (MacKay, 2003; Rasmussen & Williams, 2006). Therefore, the selected form of the covariance function, $K(t_j, t_k)$, defines how much a player's ability is free to vary from innings to innings and is the key factor when it comes to identifying whether recent form is significantly associated with current and future ability.

Several families of covariance functions are available to choose from, depending on desired temporal relationship and nature of between-innings variation one wishes to afford a player's set of underlying eye-in batting abilities, $\{\mu_{2_t}\}$, and consequently their underlying effective batting average, $\nu(t)$. A common choice is the squared exponential covariance function (Equation 2.10), which was the choice of covariance function in Stevenson & Brewer (2018). Most covariance functions have at least two parameters: (1) a *scale* parameter, σ, which controls the amount of variation allowed from the mean value, λ; and (2) a *length-scale* parameter, ℓ, which can be conceptualised as the approximate distance one must move in the input space before the function value can change significantly (Rasmussen & Williams, 2006).

$$K(t_j, t_k) = \sigma^2 \exp\left(-\frac{(j-k)^2}{\ell^2}\right) \tag{2.10}$$

Other popular choices include the Matérn class of covariance functions, which has a number of special cases, such as the Matérn$_{\frac{1}{2}}$ and Matérn$_{\frac{3}{2}}$ functions that can be selected, depending on how closely related one wishes observations within the input and output space to be. Figure 2.4 illustrates how the choice of covariance function determines the relationship between two points and how this can affect the flexibility of the corresponding Gaussian process.

As noted in Rasmussen & Williams (2006), it is important to restrict the covariance function to only allow for Gaussian process functions that agree with the corresponding data points. This requires careful consideration of the modelling problem at hand and an understanding of possible Gaussian processes that could feasibly explain any signal in the data. With this in



(a) **(b)**

Figure 2.4. *(a) Covariance functions and corresponding relationships between points in the input space. (b) Gaussian processes with a stationary mean centred on 0, drawn using different covariance functions. Each covariance function has fixed parameters:* $\sigma = 1$, $\ell = 1$.

mind, since the the publication of Stevenson & Brewer (2018) it has been determined that the squared exponential covariance function does not allow for enough short-term variation in ability, as it weights smooth functions with higher prior probability. As such, the covariance function is overly restrictive of the plausible functions that can effectively model an individual's career batting data — short-term variability is ruled out from the beginning. On the other hand, a Matérn$_{\frac{1}{2}}$ covariance function provides too much flexibility and tends to assume that a player's underlying ability fluctuates wildly from innings to innings, as illustrated in Figure 2.4.

When attempting to model data that exhibits more than one type of feature, such as both short and long-term variation, it is possible to consider multiple covariance functions at once. This is achieved by multiplying or adding covariance matrices that have been computed from two separate covariance functions (Duvenaud et al., 2011). Addition is equivalent to an *or* operation and is useful when *at least one* of the covariance functions is capturing a certain feature, while multiplication can be thought of as an *and* operation and works well when *both* covariance functions are capturing a feature present in the data.

Combining multiple covariance functions was an approach that was considered and trialled during the model building process. However, increasing the number of covariance functions used to produce Gaussian processes results in increased model complexity and uncertainty, due to the increased number of parameters in the parameter space. Instead, a practical compromise between the smoother squared exponential covariance function and more volatile Matérn$_{\frac{1}{2}}$ covariance function, is achieved by considering the γ-exponential family of covariance functions (Rasmussen & Williams, 2006) presented in Equation 2.11, which was the choice of covariance function used in Stevenson & Brewer (2021).

$$K(t_j, t_k) = \sigma^2 \exp\left(-\frac{|j-k|^\gamma}{\ell^\gamma}\right) \tag{2.11}$$

This covariance function provides a means of identifying both short and long-term variation in ability, at the cost of just one additional smoothing parameter, γ. As $\gamma \to 1$, the γ-exponential covariance function is equivalent to an Ornstein-Uhlenbeck process (Uhlenbeck & Ornstein, 1930), or AR(1) process, which is equivalent to the Matérn$_{\frac{3}{2}}$ covariance function in Figure 2.4. As $\gamma \to 2$, this function converges to the smoother squared exponential function. Thus, the γ-exponential covariance function makes it possible to determine whether or not an individual exhibits significant short-term, innings-to-innings variation in ability, or more long-term variation, by observing the posterior distributions for ℓ and γ respectively.

Innings and venue-specific effects

In addition to the individual-specific effects that might affect underlying batting ability, such as recent form, there are also a number of important factors that are worth considering in the context of each individual innings. As noted briefly in Chapter 1, a unique characteristic of cricket is the significant role that pitch and weather conditions can have on how a match plays out. Vastly difference approaches to batting and bowling are observed between the dusty, spin-friendly pitches of the sub-continent; the hard, flat pitches commonly prepared in Australia; and the green pitches that offer additional assistance to pace bowlers, frequently found in countries such as England and New Zealand. No two pitches are the same and as such, it can be difficult to adjust to a new pitch from match to match, particularly when playing a series away from home, in a foreign environment.

Therefore, due to the variable nature of cricket pitches, it is reasonable to assume that many players will perform better in their home country, where they have grown up playing cricket and are more familiar with the local conditions. The concept of home ground advantage is widely observed across many sports (Pollard, 1986; Nevill & Holder, 1999). However, anybody who has been to a cricket match — especially a Test match — can attest to the generally subdued nature of the crowd, compared with the attendees of many other sporting events. It is therefore not unreasonable to hypothesise that any home ground effect is more likely to be due to familiarity with the local pitch and weather conditions, rather than an effect due to the crowd itself (Morley & Thomas, 2005).

A second worthwhile consideration is the manner in which pitches tend to deteriorate over the course of a match. Domestic first-class and international Test matches can span up to four and five days respectively, with each team batting twice, for a total of four innings in a match.

In many countries, batting is widely considered to become far more difficult as a match goes on, as the condition of the pitch will often deteriorate due to exposure to the elements and general wear and tear that occurs as a result of hundreds of deliveries being bowled. This notion is summarised succinctly by W. G. Grace, a former English Test captain during the 19$^{\text{th}}$-century.

"When you win the toss — bat. If you are in doubt, think about it, then bat. If you have very big doubts, consult a colleague, then bat."

– W. G. Grace

While not every pitch will deteriorate at the same rate, the data do suggest that batting tends to become more difficult as a match continues. The empirical data presented in Tables 2.2, 2.3 and 2.4, show the differences in batting averages across all Test matches since January 1$^{\text{st}}$ 2000, split by venue, innings in match and team innings in match.

Table 2.2. *Test match batting averages since January 1st 2000, split by venue.*

Venue	Innings	Runs	Dismissals	Average
Home	18,945	560,003	16,404	34.14
Away	21,021	537,251	18,451	29.12
Neutral	288	6,557	253	25.92

Table 2.3. *Test match batting averages since January 1st 2000, split by innings #.*

Innings #	Innings	Runs	Dismissals	Average
1$^{\text{st}}$	11,839	363,001	10,544	34.43
2$^{\text{nd}}$	11,840	349,449	10,617	32.91
3$^{\text{rd}}$	10,427	258,978	9,029	28.68
4$^{\text{th}}$	6,148	132,383	4,918	26.92

Table 2.4. *Test match batting averages since January 1st 2000, split by team innings # in match.*

Team innings #	Innings	Runs	Dismissals	Average
1$^{\text{st}}$	23,679	712,450	21,161	33.67
2$^{\text{nd}}$	16,575	391,361	13,947	28.06

These data suggest the hypothesis that players will generally bat better in home conditions, early on in a Test match, is reasonable. Player batting averages are 17% higher at home venues, compared with away venues; and 20% higher in a team's first innings of a match, compared with their second innings. The batting average for neutral venues should be interpreted with caution; aside from the greater implicit sampling error, many of these 'neutral' innings take place in UAE where Pakistan hosted the majority of their international home fixtures during the 2010s. As such, many of these innings are likely to be closer to away conditions for all but the Pakistan players.

To account for these venue and innings-specifc effects, two indicator variables are introduced, i_t and v_t, representing whether it is a team's first or second innings of a Test match and the match venue, for a player's t^{th} career innings.

$$i_t = \begin{cases} 1, & \text{if team's first innings of a match} \\ -1, & \text{if team's second innings of a match} \end{cases} \tag{2.12}$$

$$v_t = \begin{cases} 1, & \text{if batting at a home venue} \\ 0, & \text{if batting at a neutral venue} \\ -1, & \text{if batting at an away venue} \end{cases} \tag{2.13}$$

To estimate the innings and venue-specific effects, two new parameters, ϕ and ψ, are introduced to the effective average function, $\mu(x, t)$, such that

$$\mu(x,t) = \left[\mu_{2_t} + (\mu_1 - \mu_{2_t}) \exp\left(-\frac{x}{L}\right)\right] \times \phi^{i_t} \times \psi^{v_t}. \tag{2.14}$$

While the empirical data in Table 2.3 suggest that batting gets progressively more difficult as a match goes on, given the relatively small sample size of career data for some players, the innings-effect variable is considered on a binary scale. As shown in Table 2.4, the largest difference between innings-by-innings average is between a team's first and second innings of a match. Therefore, considering team innings as a binary variable generally provides larger sample sizes within each innings group, while still accounting for an effect related to deteriorating pitch conditions. It is also worth noting an interaction effect between the innings and venue effects, ϕ and ψ was considered. However, for the same reason the innings variable was considered on a binary scale, the inclusion of such an effect was deemed to add little inferential value to the model.

The venue variable, v_t, has three levels, with the baseline, $v_t = 0$, indicating a neutral venue. The structure of this variable allows for the venue-specific parameter to be interpreted as a set of multiplicative values, $\{\frac{1}{\psi}, 1, \psi\}$, affecting run scoring at away, neutral, and home venues respectively. An estimate $\psi > 1$ indicates that a player performs better in home conditions, while $\psi < 1$ indicates a player performs better at venues away from home. Similarly, for the innings effect, $\phi > 1$ indicates that a player performs better in their team's first innings of a match, while an estimate $\phi < 1$ indicates a player performs better in their team's second innings. For both ψ and ϕ, a value of 1 indicates that the player performs equally between away, neutral and home venues, as well as between their team's first and second innings of a match.

Under the specification of the effective batting average, $\mu(x, t)$, defined in Equation 2.14, it is possible to obtain an estimate for a player's batting ability at any given point of an innings, for any innings of the player's career. An estimate for $\nu(t)$ can then be obtained for every innings in a player's career, given both the venue and specific innings in the match, by taking the expectation over all scores, x, which can be computed analytically via Equation 2.14. Plotting $\nu(t)$ against time, t, provides a player's batting career trajectory, illustrating how the model estimates underlying batting ability to have varied over the course of a playing career.

2.2.4 Prior distributions

As per Sections 2.2.2 and 2.2.3, the model contains the following set of parameters:

$$\boldsymbol{\theta} = \{\mu_1, \{\mu_{2_t}\}, \phi, \psi, L, \lambda, \sigma, \ell, \gamma\}. \tag{2.15}$$

To facilitate the underlying assumption $\{\mu_{1_t}, L_t\} < \mu_{2_t}$ imposed by Stevenson & Brewer (2017), parameters $\{C, D\} \in [0, 1]$ are introduced with the following specifications,

$$\begin{aligned} \mu_{1_t} &\leftarrow C\mu_{2_t}, \\ L_t &\leftarrow D\mu_{2_t}. \end{aligned} \tag{2.16}$$

This assumption is made to avoid considering effective average functions, $\mu(x, t)$, with excessively large transition timescales, L, while also implying that we do not expect a batsman to get any worse during the getting your eye in process.

As detailed in Section 2.2.3, the prior for the set of $\{\mu_{2_t}\}$ terms is specified by a Gaussian process, with an underlying mean value, λ, and covariance function $K(t_j, t_k; \sigma, \ell, \gamma)$, given by Equation 2.11. As the effective average function measures batting ability in units of a batting average — which by definition must be positive — it is actually the set of $\log\{\mu_{2_t}\}$ terms that are modelled using a Gaussian process, which are then back-transformed appropriately. This ensures positivity for the entire parameter space of $\{\mu_{2_t}\}$.

The model parameters, prior distributions and relevant definitions are summarised in Table 2.5. These priors are generally informative, but conservative, loosely reflecting expert judgement in regards to the batting abilities of professional cricket players. The beta priors for C and D,

Table 2.5. *The batting career trajectory model hyperparameters, parameters, data, and effective average functions, including the prior distribution for each quantity where relevant.*

Quantity	Interpretation	Prior
Data		
t	Career innings index (time)	
o_t	Out/not out flag in t^{th} career innings	
i_t	Team innings # in t^{th} career innings	
v_t	Venue in t^{th} career innings	
x_t	Runs scored in t^{th} career innings	Likelihood function given in Equation 2.6
Within-innings effects		
$\mu_{1,t}$	Initial batting ability in t^{th} career innings	$C \sim \text{Beta}(1,2); \mu_{1,t} \leftarrow C\mu_{2,t}$
L_t	Transition parameter in t^{th} career innings	$D \sim \text{Beta}(1,5); L_t \leftarrow D\mu_{2,t}$
Innings and venue-specific effects		
ϕ	Team innings # effect	$\log(\phi) \sim \text{Normal}(\log(1), 0.25^2)$
ψ	Venue effect	$\log(\psi) \sim \text{Normal}(\log(1), 0.25^2)$
Between-innings effects		
$\{\mu_{2,t}\}$	Eye-in batting ability in t^{th} career innings	$\log\{\mu_{2,t}\} \sim \text{GP}(\lambda, K(t_j, t_k; \sigma, \ell, \gamma))$
λ	Mean value of Gaussian process	$\log(\lambda) \sim \text{Normal}(\log(25), 0.75^2)$
σ	Scale parameter of covariance function, $K(t_j, t_k)$	$\log(\sigma) \sim \text{Normal}(\log(0.2), 1^2)$
ℓ	Length parameter of covariance function, $K(t_j, t_k)$	$\log(\ell) \sim \text{Normal}(\log(20), 1^2)$
γ	Smoothing parameter of covariance function, $K(t_j, t_k)$	$\gamma \sim \text{Uniform}(1,2)$
Covariance and effective average functions		
$K(t_j, t_k)$	Covariance function for Gaussian process	Functional form given in Equation 2.11
$\mu(x,t)$	Batting ability on score x, in t^{th} career innings, in units of a batting average	Functional form given in Equation 2.9
$\nu(t)$	Expected number of runs scored in t^{th} career innings	Computable from $\mu(x,t)$

which affect how a player is estimated to get their eye in, are chosen as per the recommendation in Stevenson & Brewer (2017). The log-normal priors over the innings and venue-specific effects are centred on a value of 1, indicating the model's prior assumption that a player bats equally as well in home and away conditions in any innings of a match, unless the data suggest otherwise.

The log-normal prior for λ suggests that the median player will have an eye-in or peak batting ability, μ_{2_t}, equivalent to an average of roughly 25 runs, which in the context of Test cricket is a reasonable assumption. The conservative log-normal prior for the length-scale parameter, ℓ, and uniform prior for γ, allow for a flexible range of functions to be fitted to a player's career data, in order to measure any effects that may exist due to short and long-term form.

The log-normal prior for the scale parameter, σ, is the most restrictive and has been selected for the specific modelling problem at hand. The proposed prior implies that the median player's eye-in batting ability in a given innings, μ_{2_t}, can deviate by approximately 20% from some underlying mean value, λ, over the course of their playing career. If the prior for σ is left unchecked, or is too wide, the model can have a tendency to fit values for σ that are too large, resulting in Gaussian processes that suggest underlying batting ability fluctuates wildly from innings to innings. While some of these proposed Gaussian processes may theoretically fit the data better, in many cases it is simply not believable that a player's underlying ability can vary by such a large amount between two innings.

2.2.5 Model fitting

Data

The batting career trajectory model has been fitted to the career data of all players who have participated in at least one Test match innings since 1$^{\text{st}}$ Jaunary 2000, as per the ESPNcricinfo data set discussed in Section 1.4.1. Due to a combination of law changes and technological advancements, the pace and format of Test cricket has changed considerably since the first Test match was played in 1877. The analysis of all players who have batted in the 21$^{\text{st}}$ century allows for the results to be interpreted while maintaining a modern outlook on the game. The full data set corresponds to a total of 1,018 players from 12 different countries, who have batted in a combined 40,273 innings.

An excerpt of Kane Williamson's Test career batting data are provided as an example in Table 2.6, illustrating how each of the relevant variables in the likelihood and effective average functions are stored. The not out dummy variable, o_t, indicates whether an innings was concluded on a not out score, with a value of 1 indicating a not out innings and a value of 0 indicating an out innings. Auxiliary information such as the opposition and ground the match was played at is also included to provide context to each innings.

Table 2.6. *Test match batting data for Kane Williamson's 10 most recent Test innings, as of* 1st *December 2020.*

Innings index (t)	Runs	Not out (o_t)	Innings	Team innings (i_t)	Venue (v_t)	Opposition	Ground
131	51	0	2	1	1	England	Mount Maunganui
132	4	0	1	1	1	England	Hamilton
133	104	1	3	-1	1	England	Hamilton
134	34	0	2	1	-1	Australia	Perth
135	14	0	4	-1	-1	Australia	Perth
136	9	0	2	1	-1	Australia	Melbourne
137	0	0	4	-1	-1	Australia	Melbourne
138	89	0	2	1	1	India	Wellington
139	3	0	2	1	1	India	Christchurch
140	5	0	4	-1	1	India	Christchurch

Williamson's full set of Test match batting data is then presented visually in Figure 2.5, with each innings categorised by venue and team innings. The data for Williamson again illustrate the high innings-to-innings variation in batting scores, while also showing that players do not always bat twice in a match. The latter can be attributed to one of a number of reasons, such as the match coming to a premature conclusion due to poor weather, or because the player's second innings was not required as their team had already won the match.

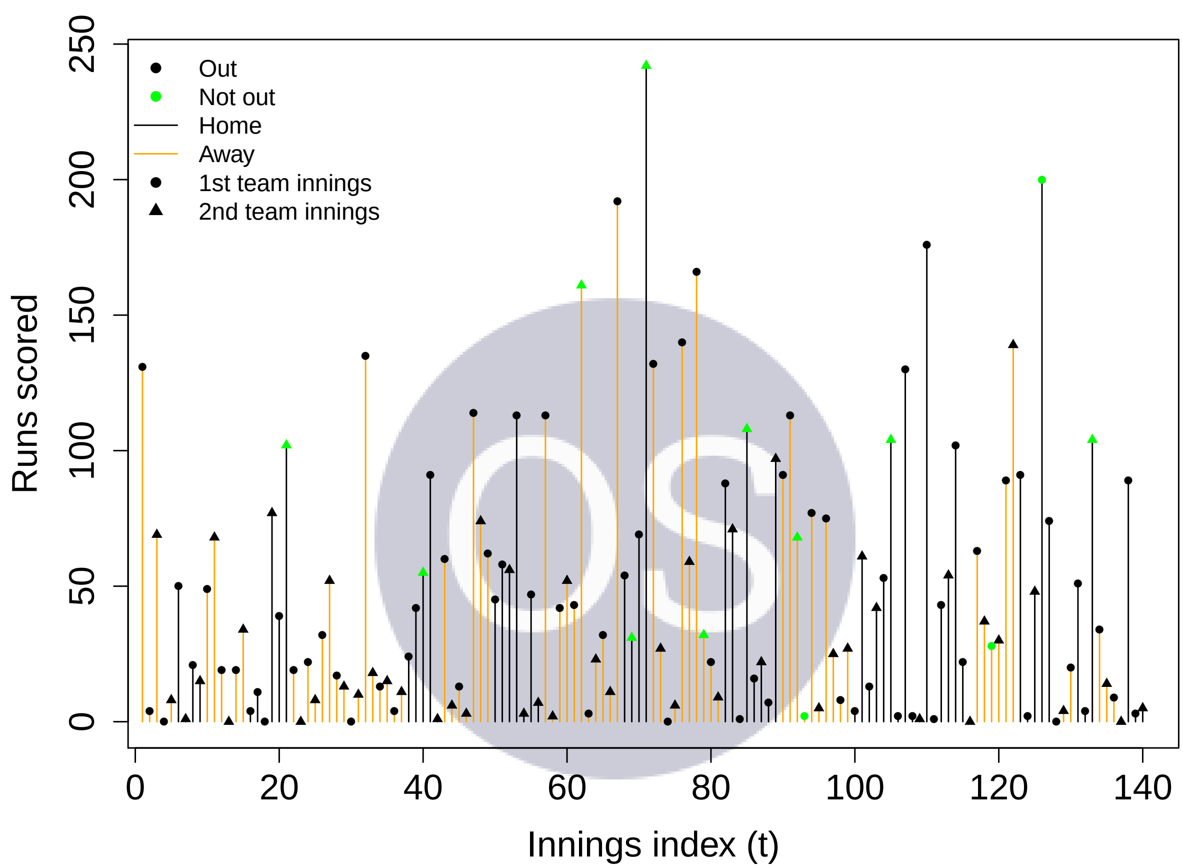


Figure 2.5. *Career batting data for Kane Williamson in Test matches, with innings split by venue and team innings # in match.*

As the model assumes a player's underlying ability is not influenced by the specific match scenario, it is best applied to longer form cricket, such as domestic first-class or international Test matches, where there is generally minimal external pressure on batsmen to score runs at a prescribed rate. However, it is plausible that the model could provide insights into one-day cricket, particularly for opening batsmen, whose role with the bat tends to be fairly consistent between matches.

Nested sampling

As discussed in Section 1.5.2, the batting career trajectory model is fitted using a C++ implementation of the nested sampling algorithm proposed by Skilling (2006). The output of the nested sampling algorithm provides posterior samples for each of the model parameters, as well as the marginal likelihood, which is used for model comparison. The effective sample size (ESS) of each nested sampling run is also computed using Shannon entropy (Shannon, 1948), to ensure the algorithm has effectively explored the parameter space. The results reported in Section 2.3 for each player are based on nested sampling runs initiated with 1,000 particles and use 1,000 MCMC steps per nested sampling iteration and were not sensitive to changes in these computational parameters, indicating the sampling was sufficient.

The run-time of the model varies significantly, from seconds to days, depending on the number of career innings in a player's career record, which modifies the number of parameters used to fit the model. As such, when applying the model to the careers of hundreds of players, the model fitting process can take weeks to complete. This can quickly become an issue when trying to maintain up-to-date results; in the height of summer Test matches are played on a near-weekly basis. Therefore, by the time the nested sampling algorithm has finished running on all players, another match may have been played, requiring the model to be re-fitted to the relevant players' updated career batting data. To deal with this problem, the model fitting process was implemented via parallel cloud computing, using the high performance computing facilities provided by the New Zealand eScience Infrastructure (NeSI). This allows the batting career trajectory model to be fitted to all players simultaneously, rather than sequentially, considerably increasing the computational efficiency of the entire procedure.

2.3 Results

2.3.1 Analysis of individual batsmen

When assessing the model output, the computation of the posterior predictive distribution for $\nu(t)$ is of primary concern and can be obtained by drawing a number of posterior parameter samples. Plotting the effective average, $\nu(t)$, against time, t, provides the batting career trajectory for an individual player, estimating how a player's underlying batting ability has varied over the course of their career to date, as well as providing a forecast of their future ability. Career trajectories of all 1,018 players analysed are available to view via an RShiny application at `www.oliverstevenson.co.nz/phd_cricket_visualisation`.

The career trajectory for Kane Williamson is presented in Figure 2.6, showcasing the evolution of his Test match batting ability. The posterior predictive estimate for $\nu(t)$ in red represents

Williamson's underlying batting ability in units of a batting average, assuming a neutral venue and his team's innings number in the match is unknown. The posterior predictive estimate in blue provides the estimate for $\nu(t)$, given the innings-specific variables, i_t and v_t, are known.

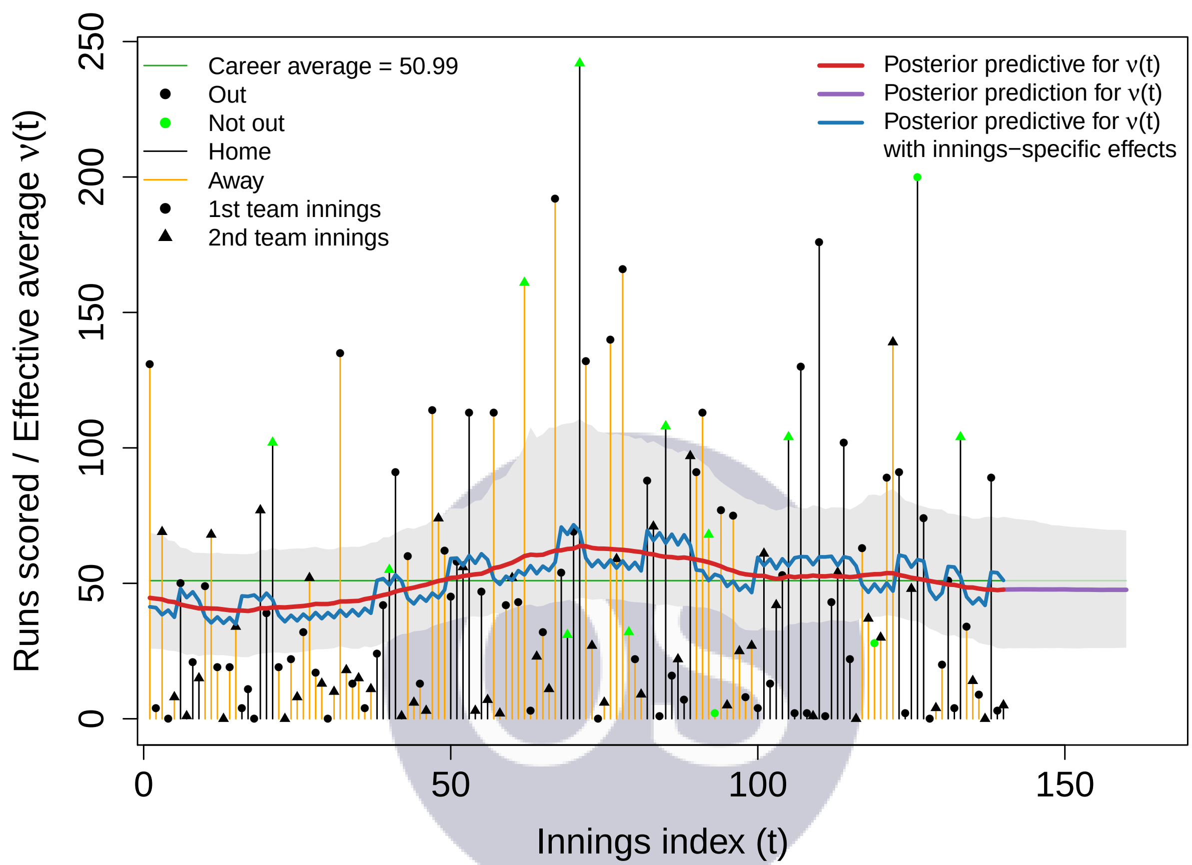


Figure 2.6. *Test match batting career trajectory (the posterior median of $\nu(t)$) for Kane Williamson, including the 95% credible interval (shaded region). A prediction for $\nu(t)$ is also made for 20 innings into the future (purple).*

Similar to the posterior distribution for the getting your eye in effective average function, $\mu(x)$, detailed in Stevenson & Brewer (2017), the posterior distribution for $\nu(t)$ is not necessarily symmetric and can have relatively heavy tails. As such, the posterior point estimate for $\nu(t)$ is computed using the posterior median, rather than the posterior mean. The posterior median also provides more accurate predictions of future scores compared with the posterior mean, which is discussed in more detail in Section 2.4. Posterior summaries for each of the model parameters are presented in Table 2.7.

A subset of 1,000 posterior samples used in the computation of the posterior predictive distribution for $\nu(t)$ are shown in Figure 2.7. Here, it is possible to observe the variety of feasible career trajectories, ranging from smoother functions that vary gradually over time, to ragged, more erratic looking processes that fluctuate significantly from innings to innings. Figure 2.7 also depicts the amount of uncertainty in the model's predictions of past, present, and future batting ability.

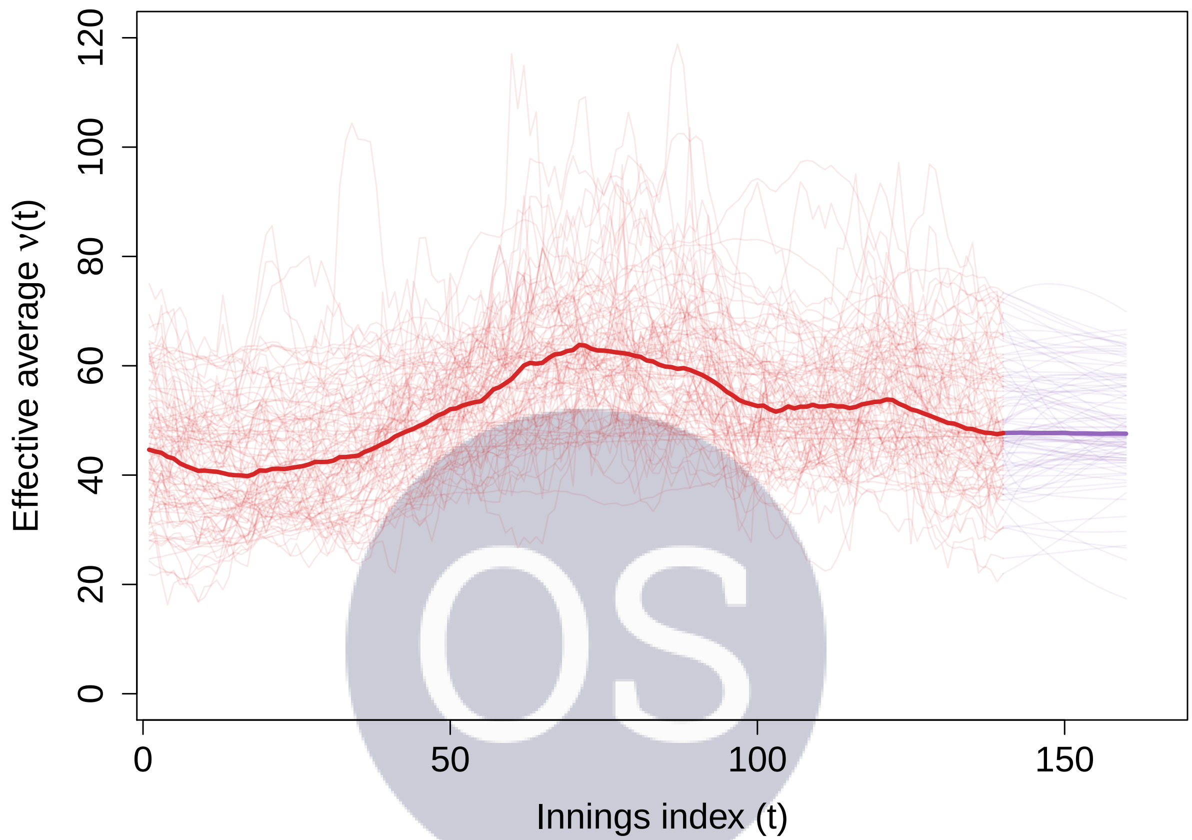


Figure 2.7. *A subset of 1,000 posterior samples for $\nu(t)$, the expected score given the parameters, for Kane Williamson. The purple lines represent predictions for $\nu(t)$ for 20 innings into the future. Due to the noisiness of batting scores, a range of career trajectories are compatible with the data. The posterior predictive estimate for $\nu(t)$ is overlaid to illustrate the moderate amount of uncertainty in the estimates.*

Individual-specific effects

As shown in Figure 2.6, the model estimates that Williamson's underlying batting ability in Test cricket has varied gradually over the course of his career. At the beginning of his career, the comparatively lower posterior predictive estimates for $\nu(t)$ imply that Williamson was not as good of a batsman as indicated by his current career average of 50.99. Since his first Test innings his ability appears to have improved, likely as a result of gaining experience in a variety of match conditions and being exposed to a range of world-class bowling attacks. In order to better understand what the model is suggesting about Williamson's career trajectory, one can consult the posterior distributions for each of the individual-specific Gaussian process parameters, which are presented in Figure 2.8.

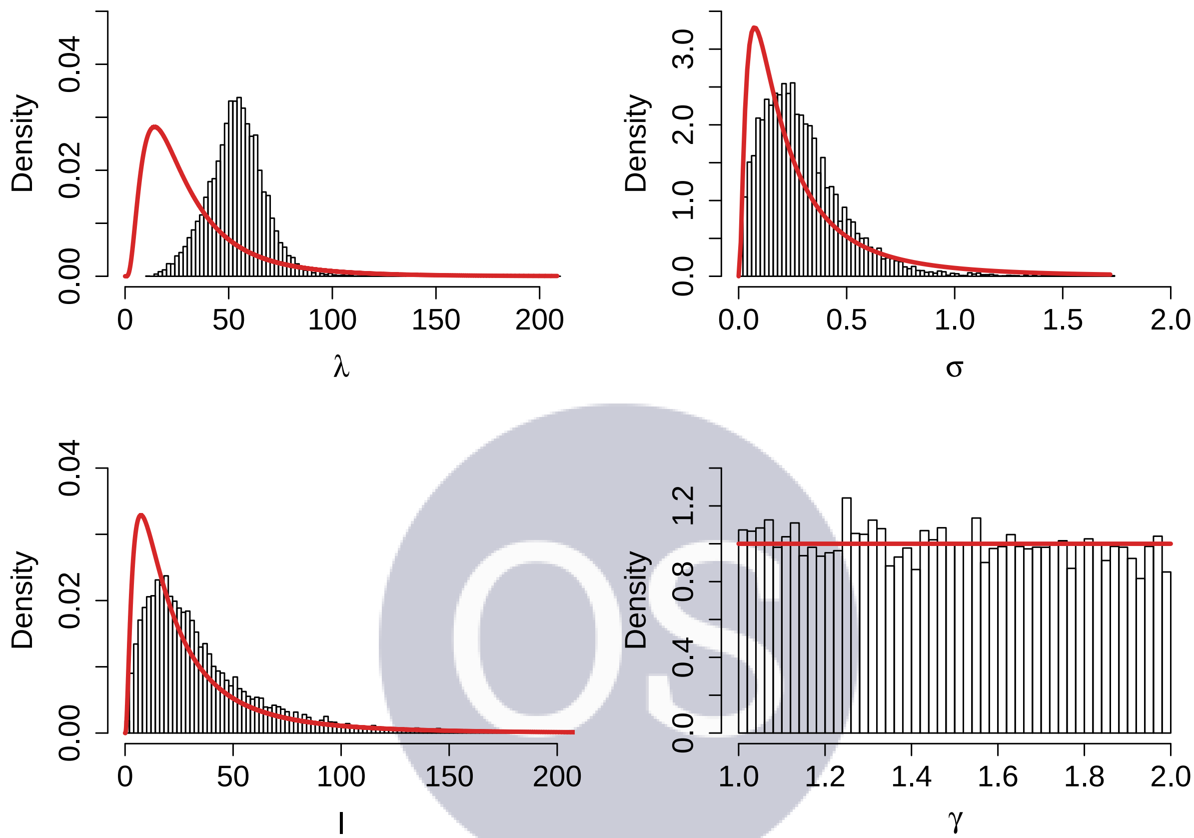


Figure 2.8. *Posterior distributions for each of the Gaussian process parameters, λ, σ, ℓ and γ. Red lines indicate the respective prior distributions. It appears as though Williamson's data are unable to modify the prior distributions for both ℓ and γ, suggesting the model is unable to distinguish between smooth or ragged underlying career trajectories, due to the noisy data.*

Clearly, a lot is learnt about λ, suggesting that the data are informative about Williamson's set of underlying eye-in abilities, $\{\mu_{2_t}\}$. In regards to σ, there is less posterior weight assigned to values of σ near zero. This result provides some evidence to support the presence of long-term variation in Williamson's underlying batting ability, as career trajectories with low σ values imply there is little variation in ability over an individual's career. However, the data have been unable to inform about smoothness parameters, ℓ and γ, which provides no evidence to either confirm or refute the presence of short-term, innings-to-innings variation in ability. That is, there is no reason to confirm or refute the notion that Williamson's underlying batting ability is affected by short-term form.

Table 2.7. *Posterior parameter summaries for Kane Williamson, including the 68% and 95% credible intervals.*

Parameter	Mean	Median	68% C.I.	95% C.I.
C	0.30	0.29	(0.21, 0.40)	(0.15, 0.54)
D	0.12	0.11	(0.06, 0.18)	(0.03, 0.30)
λ	53.9	53.8	(40.7, 66.1)	(26.8, 81.3)
σ	0.29	0.26	(0.11, 0.46)	(0.04, 0.75)
ℓ	38.9	28.2	(11.9, 62.5)	(4.2, 141.1)
γ	1.49	1.49	(1.15, 1.83)	(1.02, 1.97)
ϕ	1.03	1.03	(0.94, 1.12)	(0.86, 1.22)
ψ	1.11	1.11	(1.02, 1.21)	(0.93, 1.32)

Innings and venue-specific effects

In Figure 2.6, the posterior predictive estimate for $\nu(t)$ that includes the innings and venue-specific effects (blue), indicates that historically, Williamson has tended to perform better at home venues (black bars), compared with away venues (orange bars). Similarly, superior estimates for $\nu(t)$ are observed when Williamson is batting in his team's first innings of a match (circles), compared with his team's second innings (triangles), which is visible from the career trajectory's jagged behaviour when observing changes between his first and second innings in the same match. These results are fairly typical for many player's batting career trajectories.

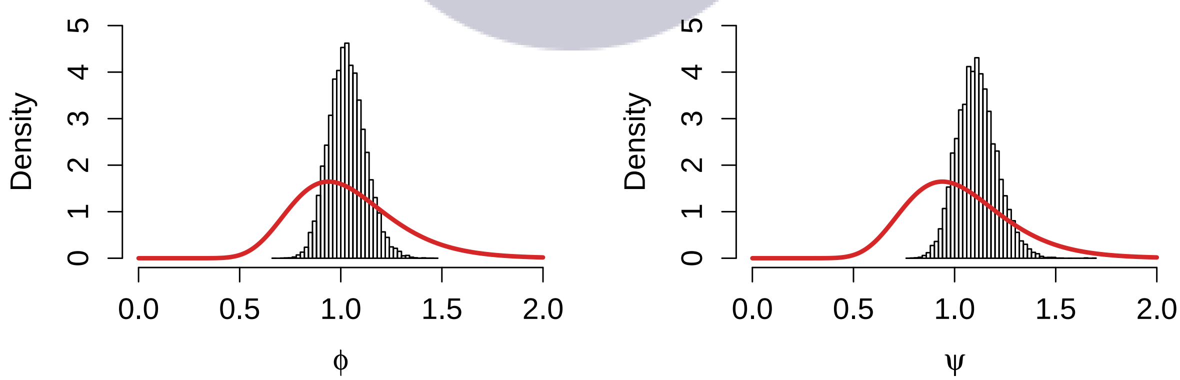


Figure 2.9. *Posterior distributions for the innings-specific parameters, ϕ and ψ. Red lines indicate the Log-normal(1, 0.25^2) prior. The data appear to provide moderate evidence that Williamson bats better at home venues, in his team's first innings of a match.*

Posterior distributions for the innings-specific effects, ϕ and ψ, are shown in Figure 2.9, suggesting the data have been somewhat informative with respect to both parameters. The parameter point estimates presented in Table 2.7 allow for the magnitude of these effects to be quantified. The point estimate for ψ suggests that one can expect Williamson to score 11% more runs when batting at a home venue, compared with a neutral venue. It is possible to compare performances at a home venue against an away venue by squaring the parameter estimate for ψ, giving an estimate of $\psi^2 = 1.24$, and corresponding 95% credible interval (0.87, 1.73). That is to say, Williamson's underlying ability is estimated to be 24% higher when batting at home venues, compared with away venues. Similarly, the multiplicative effect of batting in his team's first innings of a Test, compared with the second innings, can also be obtained by squaring, giving an estimate of $\phi^2 = 1.07$ and corresponding 95% credible interval (0.75, 1.48), indicating an expected 7% difference in runs scored between innings in the same match. It is worth noting that the relatively large uncertainties over ϕ and ψ suggest there is no definitive evidence at the individual level that Williamson necessarily performs better in home conditions, in his team's first innings of a match. However, when the data for all players are considered jointly in Section 2.3.2, the likely presence of such effects becomes more apparent.

Quantifying batting career progression

As seen in Figure 2.6, Williamson's underlying ability appears to have developed over the course of his Test career to date. Similar to the model estimates provided in Boys & Philipson (2019), the proposed batting career trajectory model is able to quantify batting ability, in units of a batting average, at the lowest and highest points of a player's career. This allows for the comparison of batsmen across eras, although unlike Boys & Philipson (2019), an adjustment to account for how difficult batting was in the era is not made. However, recalling that T is the total number of career innings a player has batted in, the batting career trajectory model is able to make a prediction for $\nu(T + 1)$, the quantity defining the number of runs an individual player is expected to score in their next Test innings. These predictions are typically made assuming a neutral venue and it is unknown whether the player is batting in their team's first or second innings, however, if the venue and innings-specific variables v_t and i_t are known, the information can be incorporated into the prediction.

Table 2.8. *Posterior point estimates for $\nu(t)$ at Kane Williamson's lowest and highest points of his career and prediction for his next career innings, $\nu(T + 1)$, including the 68% and 95% credible intervals.*

	Point estimate	**68% C.I.**	**95% C.I.**
Career low $\nu(t)$	33.2	(24.8, 43.5)	(18.2, 52.7)
Career high $\nu(t)$	73.8	(59.3, 98.1)	(50.4, 133.9)
$\nu(T + 1)$	47.7	(36.8, 58.9)	(26.1, 74.2)

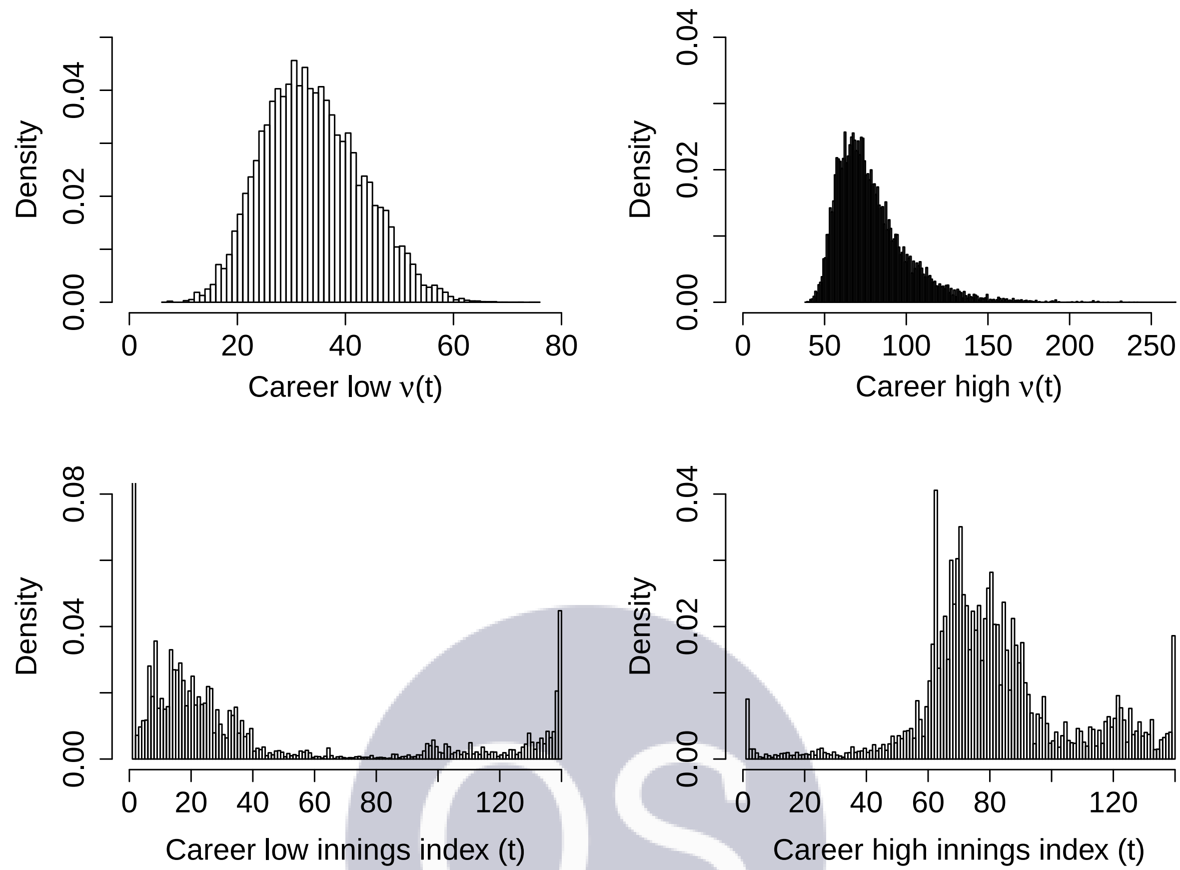


Figure 2.10. *Posterior distribution for $\nu(t)$ at the estimated lowest and highest points of Kane Williamson's Test career to date, and the career innings index, t, at which Williamson is estimated to have experienced the lowest and highest points of his career.*

The posterior distribution for $\nu(t)$ at Kane Williamson's lowest and highest points of his career to date are presented in Figure 2.10 and are summarised in Table 2.8. At the highest point of Williamson's career, his underlying ability corresponded to an expected average of 73.8, which incidentally was the highest estimate among all players globally at the time. At the lowest point of his career, the model estimates Williamson to have had an expected average of 33.2 — over a 50% difference in estimated ability compared with his career peak.

Furthermore, the posterior distribution estimating when Williamson experienced these points of his career are also shown in Figure 2.10. Perhaps unsurprisingly, the estimates in Figure 2.8 suggest that Williamson most likely experienced the lowest point of his career relatively early on. This finding is consistent with the cricketing concept of *finding your feet*, whereby players are unlikely to begin their Test careers playing to the best of their ability, which was a major focus of Stevenson & Brewer (2018, 2021). Rather, it takes time and experience for players to adjust to the demands of international cricket and can take players a number of innings to reach their peak ability. In the case of Williamson, Figure 2.10 suggests that he most likely experienced the peak of his career sometime after his 60$^{\text{th}}$ career innings, or he is still yet to experience it. Given Williamson has only recently turned 30 years of age, he likely has a number of years left ahead of him where his underlying batting ability will be close to its peak.

2.3.2 Hierarchical analysis of batsmen

In order to generalise the results across the entire group of 1,018 players analysed, a hierarchical analysis was performed for the set of Gaussian process parameters, $\{\lambda, \sigma, \ell, \gamma\}$, as well as the set of innings and venue-specific parameters, $\{\phi, \psi\}$. Defining a set of hyperparameters, $\boldsymbol{\eta}$, for each relevant model parameter and implementing a hierarchical model structure, allows for the quantification of the typical values that each parameter is clustered around, without having to analyse the data jointly. This is achieved by obtaining posterior samples for the appropriate subset of model parameters, $\boldsymbol{\theta} = \{\phi, \psi, \lambda, \sigma, \ell, \gamma\}$, for all players, then post-processing the results using MCMC (Hastings, 1970), to construct what the hierarchical model would have produced.

The analysis assumes that the typical values for the subset of parameters, $\{\lambda, \sigma, \ell, \phi, \psi\}$, are approximately log-normally distributed, while γ is assumed to follow a normal distribution, truncated at $[1, 2]$. Based on these assumptions, the hierarchical model structure for the model parameters that define the Gaussian process takes the following form, conditional on the set of hyperparameters, $\boldsymbol{\eta} = \{\mu_\lambda, \xi_\lambda, \mu_\ell, \xi_\ell, \mu_\sigma, \xi_\sigma, \mu_\gamma, \xi_\gamma, \mu_\phi, \xi_\phi, \mu_\psi, \xi_\psi\}$:

$$
\begin{aligned}
\log(\lambda) &\sim \text{Normal}(\log(\mu_\lambda), \xi_\lambda^2), \\
\log(\ell) &\sim \text{Normal}(\log(\mu_\ell), \xi_\ell^2), \\
\log(\sigma) &\sim \text{Normal}(\log(\mu_\sigma), \xi_\sigma^2), \\
\gamma &\sim \text{Normal}_{[1,2]}(\mu_\gamma, \xi_\gamma^2), \\
\log(\phi) &\sim \text{Normal}(\log(\mu_\phi), \xi_\phi^2), \\
\log(\psi) &\sim \text{Normal}(\log(\mu_\psi), \xi_\psi^2).
\end{aligned}
\tag{2.17}
$$

The hyperparameters, $\boldsymbol{\eta} = \{\mu_\lambda, \xi_\lambda, \mu_\ell, \xi_\ell, \mu_\sigma, \xi_\sigma, \mu_\gamma, \xi_\gamma, \mu_\phi, \xi_\phi, \mu_\psi, \xi_\psi\}$, are assigned the following prior distributions:

$$
\begin{aligned}
\mu_\lambda, \mu_\ell &\sim \text{Uniform}(0, 50), \\
\mu_\sigma, \mu_\gamma, \mu_\phi, \mu_\psi &\sim \text{Uniform}(0, 2), \\
\xi_\lambda, \xi_\ell, \xi_\sigma, \xi_\phi, \xi_\psi &\sim \text{Uniform}(0, 2), \\
\xi_\gamma &\sim \text{Uniform}(0, 1).
\end{aligned}
\tag{2.18}
$$

Given the data of all 1,018 analysed players, $\boldsymbol{D}$, and the hierarchical model structure defined in Equations 2.17 and 2.18, the marginal posterior distributions for the set of hyperparameters,

$\boldsymbol{\eta}$, may be written as per Equation 2.19.

$$P(\boldsymbol{\eta}|\boldsymbol{D}) \propto P(\boldsymbol{\eta}) \prod_{i=1}^{N} \mathbb{E}\left(\frac{f(\boldsymbol{\theta}_i|\boldsymbol{\eta})}{P(\boldsymbol{\theta}_i)}\right) \tag{2.19}$$

Here, $f(\boldsymbol{\theta}_i|\boldsymbol{\eta})$ is the log-normal or truncated normal prior distribution assigned to each of the model parameters in Equation 2.17, conditional on the relevant hyperparameters, applied to the data of the i^{th} player in the analysis. The expectation term inside the product can then be approximated by averaging over the posterior samples for each player. Each of $P(\boldsymbol{\theta}_i)$ and $P(\boldsymbol{\eta})$ relate to the prior distributions assigned to the model parameters and hyperparameters respectively, defined in Table 2.5 and Equation 2.18.

The MCMC algorithm was run for 100,000 iterations for each model parameter, to obtain the joint posterior distribution for the relevant hyperparameters using Equation 2.19. For example, to obtain the joint posterior for the hyperparameters relating to model parameter λ, the relative quantities in Equation 2.19 are defined as follows: $\boldsymbol{\eta} = \{\mu_\lambda, \xi_\lambda\}$, $\boldsymbol{\theta} = \lambda$.

The joint posterior distributions for the set of hyperparameters defining each of the Gaussian process parameters, are shown in Figure 2.11, while the posterior distributions associated with the hyperparameters defining the innings and venue-specific effects, are shown in Figure 2.12. As γ is assumed here to follow a truncated normal distribution, rather than the Uniform(1, 2) distribution defined in Table 2.5, the central value of the Uniform hyperpriors assigned to μ_γ and ξ_γ were chosen as the starting point of the algorithm, corresponding to values of 1.5 and 0.5 respectively. As each of the starting points appear to be fairly typical of the corresponding joint posterior distribution, no burn-in period was applied (Meyn & Tweedie, 1993).

In the case of the Gaussian process parameters, it appears as though the hierarchical model was somewhat informative in respect to λ and σ, but less so for ℓ and γ. The results for ℓ and γ are hardly surprising, as the model generally had difficulty in distinguishing between career trajectories with shorter and longer length-scales, for the majority of players analysed. One result of potential interest with respect to γ, is that the posterior parameter space tends to have lower density where $\mu_\gamma \to 1.0$ and $\mu_\gamma \to 2.0$. This suggests that for the typical batsman, the hierarchical analysis assigns less posterior weight to highly volatile batting career trajectories or overly smooth trajectories, similar to the Gaussian processes generated by the Matérn$_{\frac{1}{2}}$ and squared-exponential covariance functions in Figure 2.4. The main area of high density for λ supports the prior assumption that the typical player has an eye-in effective batting average close to 25.0 runs, although as expected, the relatively variable nature of values for μ_λ reflects the wide range of batting abilities seen in Test cricketers. For σ, values for μ_σ of, or near zero, appear to have low density, indicating that the average player is likely to exhibit some form of temporal variation in their batting ability over the course of their career.

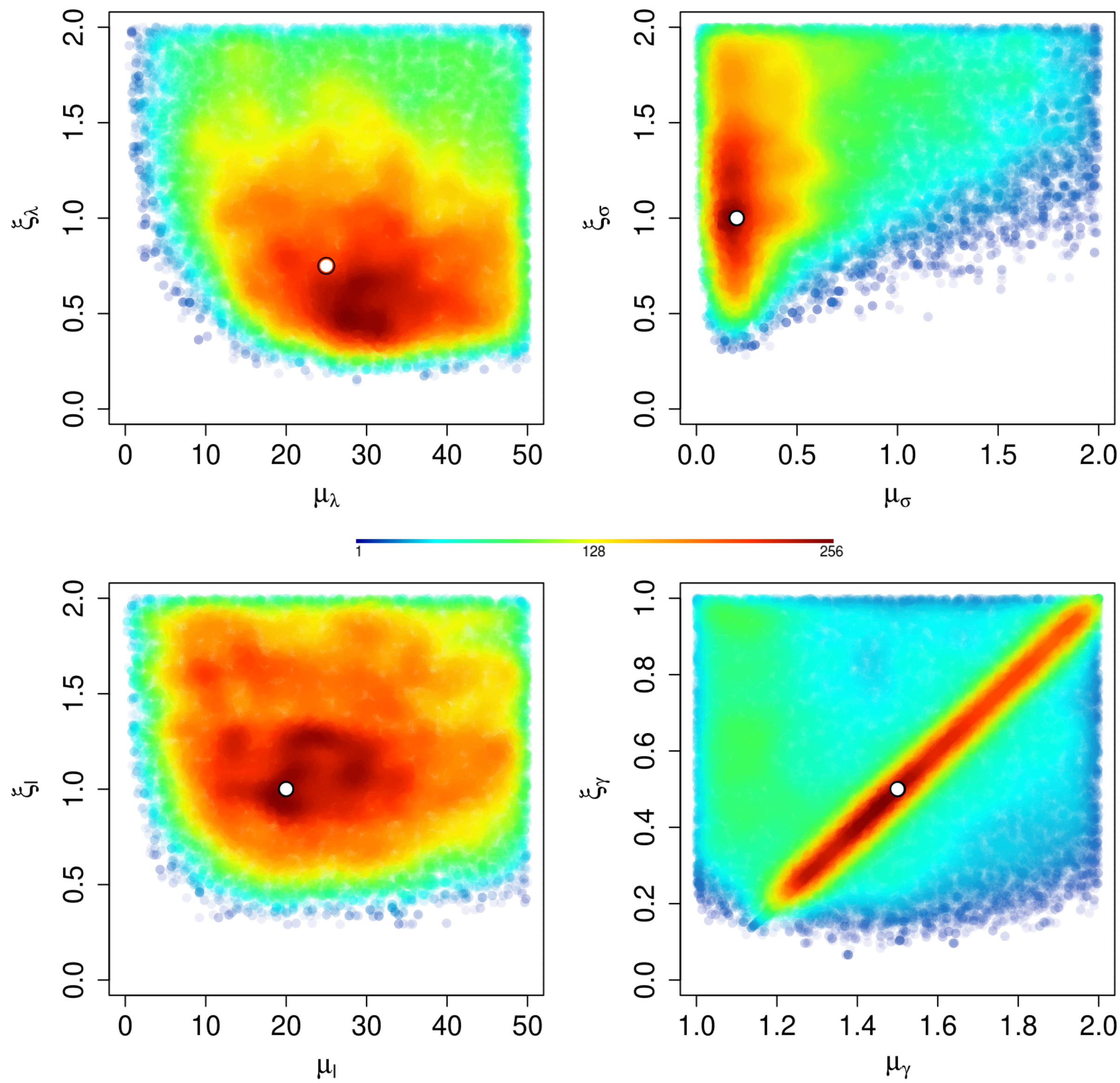


Figure 2.11. *Joint posterior distributions for the set of hyperparameters* $\{\mu_\lambda, \xi_\lambda, \mu_\ell, \xi_\ell, \mu_\sigma, \xi_\sigma, \mu_\gamma, \xi_\gamma\}$, *shown across the uniform prior parameter space. Dark red indicates areas of high density, while dark blue indicates areas of low density. The darkest red areas are 256 times more dense than the darkest blue areas as shown by the scale. The white circle indicates the starting point of the MCMC algorithm.*

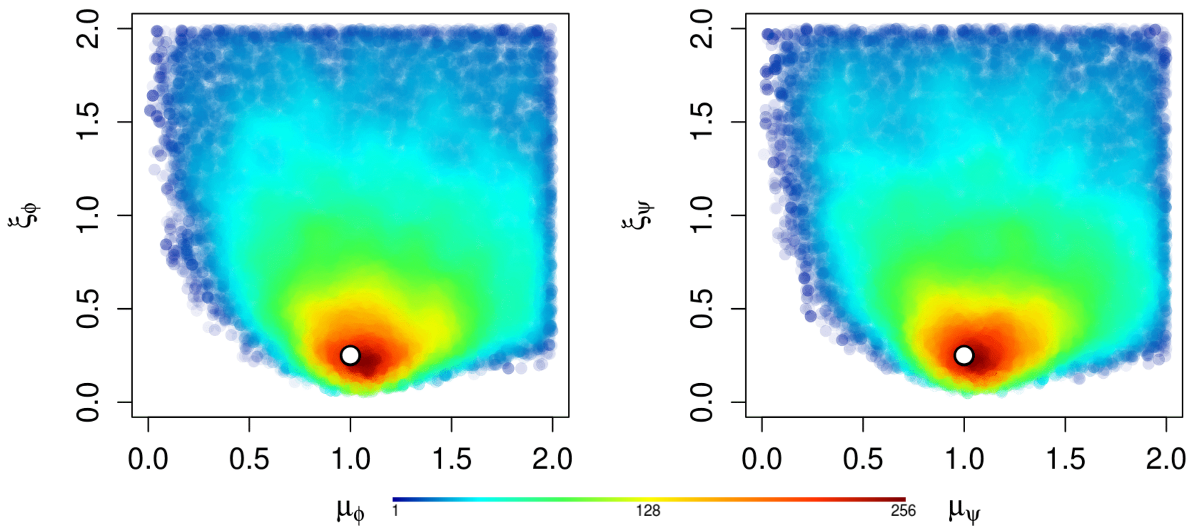


Figure 2.12. *Joint posterior distributions for the set of hyperparameters* $\{\mu_\phi, \xi_\phi, \mu_\psi, \xi_\psi\}$*, shown across the uniform prior parameter space. Red indicates areas of high density, while dark blue indicates areas of low density. The darkest red areas are 256 times more dense than the darkest blue areas as shown by the scale. The white circle indicates the starting point of the MCMC algorithm.*

In the case of the innings and venue-specific parameters, the majority of the posterior mass for both μ_ϕ and μ_ψ is centered on values greater than 1, as presented in Table 2.9. This result supports for the initial assumption; that the average player tends to perform better with the bat at home venues, in their team's first innings of a Test match, although the large amount of uncertainty in the parameter estimates must be acknowledged.

Table 2.9. *Posterior mean and credible intervals for hyperparameters* μ_ϕ *and* μ_ψ*.*

Hyperparameter	**Posterior mean**	**68% C.I.**	**95% C.I.**
μ_ϕ	1.11	(0.65, 1.32)	(0.28, 1.92)
μ_ψ	1.11	(0.66, 1.38)	(0.27, 1.92)

2.3.3 Comparison of batting career trajectories

To illustrate how the model can be used to directly compare the abilities and careers of multiple players, the batting career trajectories for the big four have been plotted in Figure 2.13. All four players appear to exhibit behaviour typical of finding your feet, taking a number of innings to reach their peak ability, with each player exhibiting some form of improvement during the early stages of their careers. Interestingly, Williamson appears to be the player who took the longest to fulfill his potential, which is perhaps unsurprising as he was just 20 years old when making his Test debut, compared with Smith (21), Kohli (22) and Root (21). As a result, Williamson was possibly a less developed batsmen when first entering the Test arena.

Figure 2.13. *Test match batting career trajectories (the posterior median for $\nu(t)$) for the big four: Steve Smith, Virat Kohli, Kane Williamson and Joe Root, including predictions of ability for the next 20 innings (dotted).*

The career trajectories presented in Figure 2.13 suggest there is an argument to be made that Williamson is the most improved batsmen among the big four. In the early stages of their respective careers, Williamson had the lowest estimated ability, however, at his peak, Williamson has the highest estimated ability. Nevertheless, unlike Steve Smith and to a lesser extent Virat Kohli, Williamson has been unable to sustain peak performance for an extended period of time. On the other hand, English batsman Joe Root experienced a successful start to his career, but has been unable to maintain the same level of continual improvement as the others. While still a world class batsmen, as suggested by his career average of 48.0, Root appears to be a model of consistency, without exhibiting the same stellar patches of form seen by Smith and Williamson.

As a counterpoint to the idea of finding your feet, seen in the career data of the big four and many other players, Figure 2.14 depicts the career trajectories of four players who made their Test debuts at a much older age, with a vast amount of domestic experience behind them. Each of these players made their international debuts in the latter half of their professional careers and did not exhibit the same behaviour typical of a player finding their feet. While no amount of domestic cricket can truly prepare a player for the challenges of Test cricket, this result does at least suggest there is some evidence to support the idea that more experienced players may reach their peak abilities much quicker than those who debut in their teens or early twenties, such as the big four.

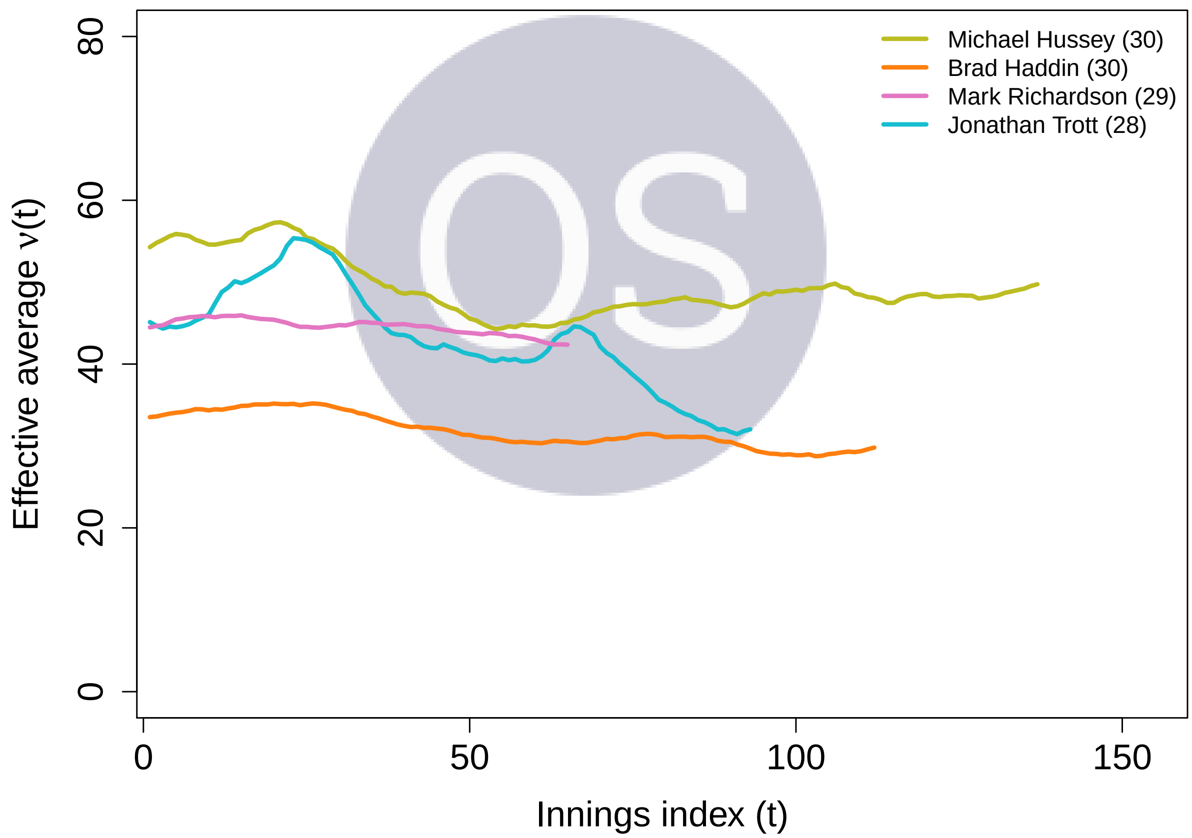


Figure 2.14. *Test match batting career trajectories (the posterior median for $\nu(t)$) for a group of players who reached their peaks very early in their Test careers. The age at which each player made their Test debut is also provided.*

2.3.4 Player batting rankings

The model results can also be used to gain an understanding of who the best batsmen in the world are at present. Each individual player has an estimate for $\nu(T+1)$, where $T+1$ is the innings index for the player's *next* career innings. The current[1] top 20 batsmen in the world according the model are presented in Table 2.10, where players are ranked by the number of runs they are expected to score in their next innings. Players must have batted in a minimum of 15 Test innings to be ranked. An up-to-date list of the top 100 Test batsmen is maintained at `www.oliverstevenson.co.nz/#research`. Note the predictions for $\nu(T+1)$ assume the player's next innings is played at a neutral venue and it is unknown whether the player is batting in their team's first or second innings of a match. For comparison, each player's ICC rating and world ranking are also provided. The rankings between methods are compared in Figure 2.15 are generally similar, for example, Steve Smith is unanimously agreed to be the top ranked batsmen in the world, however, there are several notable differences.

Firstly, the proposed batting career trajectory model ranks Indian batsman Rohit Sharma 8th in the world, while he is ranked 16th according the the ICC ratings. Looking closely at Sharma's batting career record indicates that he has an unusually high number of not out scores greater than 50. This suggests that there a number of occasions where Sharma has overcome the difficult getting your eye in process, but has not had the opportunity to turn these scores into large scores of 100 or more. As discussed in Section 2.2.1, for not out scores, the proposed model assumes the player would have gone on to score some unobserved score, conditional on $\mu(x,t)$. Meanwhile, the ICC ratings provide not out innings with an undefined bonus. In this manner, the proposed model supports the conclusion in Boys & Philipson (2019), that the ICC does not appropriately adjust for not out scores and provides a bonus that, at least for large scores, is too low.

Secondly, the ICC ratings method appears to weight recent performances more heavily than the batting career trajectory model. This is evident from the disparity in model and ICC rankings for certain players who have experienced a recent run of good form. For example, English all-rounder Ben Stokes is presently ranked 8th by the ICC, but has a ranking of 17th under the proposed model. In his most recent 20 Test innings, Stokes has averaged 48.6 runs per dismissal, increasing to an average of 58.0 in his most recent 10 innings. This is well above his career average of 37.8, which suggests the difference in Stokes' rankings is potentially due to the ICC ratings placing more emphasis on recent innings. Similarly, English wicket-keeper Jos Buttler is ranked 21st according to the ICC ratings, but 40th by the batting career trajectory model. Buttler has averaged 47.1 in his most recent 10 Test innings, considerably higher than his career average of 33.9. Regrettably, the ICC rating methodology is not public, making it

[1] as of 1st December 2020

Table 2.10. *Current (as of 1st December 2020) top 20 Test match batsmen ranked by expected number of runs scored in their next career innings, $\nu(T+1)$, including the 68% credible interval. ICC Test batting ratings and world rankings (#) are shown for comparison.*

Rank	Player		Innings	Career average	$\nu(T+1)$	ICC rating	(#)
1.	S. Smith	(AUS)	131	62.8	57.9 (47.9, 68.8)	911	(1)
2.	M. Labuschagne	(AUS)	23	63.4	55.8 (43.0, 75.2)	827	(3)
3.	B. Azam	(PAK)	53	45.4	53.5 (40.8, 75.8)	797	(5)
4.	V. Kohli	(IND)	145	53.6	51.4 (43.4, 59.8)	886	(2)
5.	D. Warner	(AUS)	155	48.9	48.0 (41.1, 56.8)	793	(6)
6.	K. Williamson	(NZ)	140	51.0	47.7 (36.8, 58.9)	812	(4)
7.	A. Mathews	(SL)	154	45.3	47.6 (40.4, 59.3)	658	(17)
8.	R. Sharma	(IND)	53	46.5	46.6 (37.7, 58.9)	674	(16)
9.	M. Agarwal	(IND)	17	57.3	46.2 (34.2, 63.5)	714	(11)
10.	J. Root	(ENG)	177	48.0	45.7 (39.4, 52.3)	738	(9)
11.	R. Taylor	(NZ)	178	46.1	43.6 (36.9, 50.6)	677	(15)
12.	C. Pujara	(IND)	128	48.7	43.4 (36.6, 50.6)	766	(6)
13.	A. Ali	(PAK)	152	42.9	42.8 (35.5, 52.1)	627	(23)
14.	A. Rahane	(IND)	109	42.9	41.6 (34.3, 50.5)	726	(10)
15.	T. Latham	(NZ)	92	42.3	40.8 (33.9, 48.8)	710	(12)
16.	M. Rahim	(BAN)	130	36.8	40.5 (33.8, 54.6)	654	(18)
17.	B. Stokes	(ENG)	122	37.8	39.4 (33.3, 47.7)	760	(8)
18.	D. Chandimal	(SL)	103	40.8	39.3 (31.3, 47.2)	563	(28)
19.	T. Head	(AUS)	28	42.0	38.1 (29.5, 49.2)	643	(20)
20.	BJ. Watling	(NZ)	110	38.5	38.1 (31.1, 45.4)	621	(25)

difficult to assess and compare how the ICC estimates the effects of recent performances and short-term form on current and future ability. An exponential weighted average is used, however the length scale is not public knowledge.

Thirdly, it is worth recalling that the expected number of runs to be scored in the next innings, $\nu(T+1)$, assumes a neutral venue. Under this assumption, the model estimates South African batsmen Quinton de Kock and Aiden Markram to score 35.0 and 25.3 runs in their next innings respectively. Each of these predictions are below de Kock and Markram's career averages of 39.1 and 38.5, however, the estimates are not lower due to a run of recent poor form. Rather, de Kock and Markram both excel when batting in home conditions but struggle heavily when batting at venues outside of South Africa. Therefore, when considering a prediction at a neutral venue, both players are penalised by the proposed model, as they are yet to prove themselves in conditions outside of their home country. As visible in Figure 2.15, this results in markedly different rankings between the batting career trajectory model and ICC rating method (37th versus 13^{th} for de Kock; 75^{th} versus 26^{th} for Markram).

Finally, the ICC method applies a decay to the ratings of players who miss selection in their side's recent Test matches. While well-intentioned to penalise players who are not consistently active in the Test scene, this approach has the unintended effect of disproportionately punishing players from countries who generally play less Test cricket. Missing a series due to injury or personal reasons can have a significant impact on batting ratings for players from smaller Test playing sides such as New Zealand or the West Indies, who may only play in a handful of Test series each year. Conversely, players from Australia, England and India are able to miss the odd game with less impact on their batting rating; these nations tend to have the luxury of competing in four and five-match series on a much more frequent basis and consequently players are able to recoup their ratings much faster.

While both methods provide an indication of overall player ability, the batting career trajectory model has the advantage of quantifying batting ability in units of a batting average, which can be easily understood by all viewers of cricket. As such, the proposed model can quantify differences in ability between players in a far more meaningful manner. Rather than concluding, 'Steve Smith is 99 rating points better than Kane Williamson', the model can make more useful probabilistic statements by computing $P(\nu_{Smith}(T_{Smith}+1) > \nu_{Williamson}(T_{Williamson}+1))$. In this case, one can conclude 'Steve Smith has a 56.4% chance of outscoring Williamson in their next respective innings', or, 'Steve Smith is estimated to outscore Williamson by 10.1 runs in their next respective innings'.

Such probabilistic statements can be particularly useful when comparing players with similar career averages, for example, Pakistani batsmen Babar Azam (45.4) and Azhar Ali (42.9). Objectively, based solely on the batting average, one would assume that these two players are relatively similar in terms of their underlying ability. However, the predictions obtained from the batting career trajectory model suggest that while both players are presently among the best batsmen in the world, Azam is more likely the superior batsmen, given his higher prediction for $\nu(T+1)$ of 53.5, compared with 42.8 for Ali. Comparing the relevant posterior distributions for $\nu(T+1)$, one can conclude 'Azam is predicted to outscore Ali by an average of 10.7 runs in their next respective innings', or, 'Azam has a 55.4% posterior probability of outscoring Ali'. Having access to such insights would provide coaches and selectors with a more meaningful measure of predicting and quantifying the risks and rewards of selecting one player over another. As usual, these predictions ignore any innings or venue-specific effects, which can easily be taken into account if this information is known. Additionally, the predictions assume that players in question will be facing the same quality of bowling during their next innings, which is reasonable when comparing two batsmen from the same team, or even when batsmen between teams that have bowling attacks of similar strength.

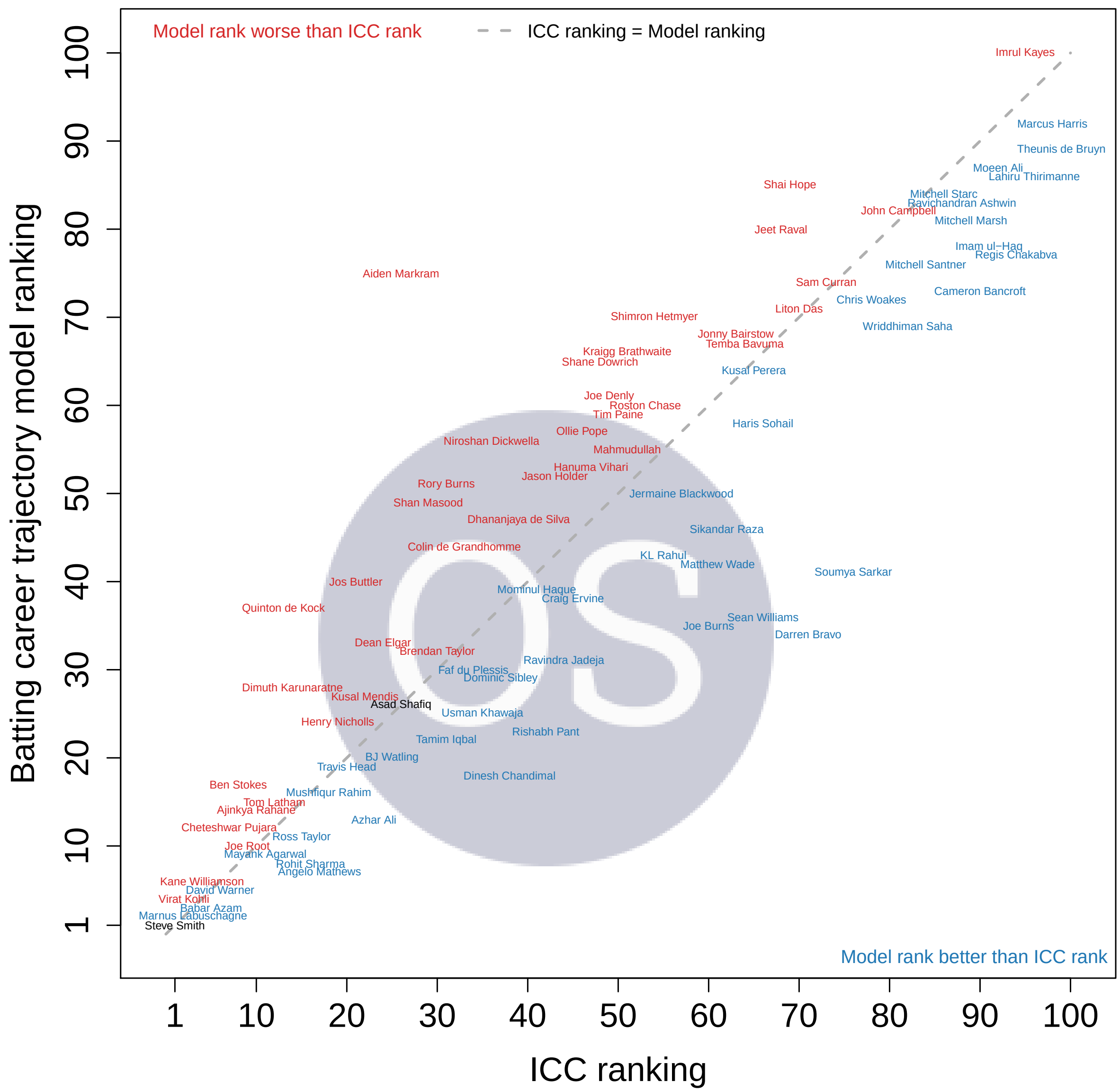


Figure 2.15. *Comparison of current (as of 1st December 2020) world rankings between the batting career trajectory model and the established ICC ratings. The proposed model considers players in red to be overvalued by the ICC method, while players in blue are considered to be undervalued. Players in black represent cases where there is consensus in rankings between the two methods.*

2.4 Model diagnostics

2.4.1 Model prediction

The model's ability to describe variations in a player's underlying ability, via the career trajectory, is a secondary feature and can only be considered useful if the model is able to predict future player performance more accurately than other known methods. Therefore, the primary means of assessing model performance is to compute the relative prediction errors for future innings.

When making a prediction for a player's next career score, x_{T+1}, one can assume the model has access to all previous scores, $\{x_1, ..., x_T\}$. However, x_{T+1} is a score yet to be observed. Instead, the predictive capabilities of the model are assessed by obtaining a prediction for x_T, and comparing this estimate to the actual observed value of x_T. The subsequent prediction error is then computed using the mean squared error (MSE). This is achieved using leave-out-out cross-validation (Sammut & Webb, 2010), whereby each player's most recent observed score, x_T, is removed from the data and a prediction for x_T is realised by fitting the model using the remaining data, $\{x_1, ..., x_{T-1}\}$. To avoid the complications that arise when predicting not out scores, the most recent out score is predicted for each player.

To provide a means of comparison with the batting career trajectory model, predictions and corresponding prediction errors have been computed for set of simple moving average (SMA) models of varying orders. The SMA models compute a prediction for a player's next career score, using the previous, 10%, 25%, 50% and 100% of a player's career data. For example, if a player has batted in 100 career innings, the SMA(10%) model computes a prediction for their next score by taking the average of their most recent 10 innings, while the SMA(50%) model uses the most recent 50 innings. Note that the SMA(100%) model is equivalent to a model that assumes a player has a batting ability equal to their career batting average, at any given point of their career.

As leave-one-out cross-validation requires a player's most recent innings where they were dismissed to be removed from the data, prediction errors cannot be predicted for players who have only batted in one career innings, or for players who have never been dismissed. Overall, prediction errors were able to be computed for 913 of the 1,018 players in the ESPNcricinfo data set. The performance of each model is assessed by computing the respective model predictions for all players, and taking the mean squared prediction error (MSE), which are presented in Table 2.11. The MSE is also shown for the subset of batsmen who have batted in a minimum of 10, 20 and 50 career innings.

When it comes to predicting the next score in a player's career, the batting career trajectory model outperforms all the SMA models. Additionally, the proposed model is able to provide a more accurate description of a player's career trajectory to date and is able to deal with the

Table 2.11. *Mean squared prediction errors using leave-one-out cross-validation. The batting career trajectory model outperforms all other models, while the SMA(10%) model tends to perform worst of all.*

	Minimum # of career innings			
Model	**No minimum**	**10 innings**	**20 innings**	**50 innings**
SMA(10%) model	633.1	751.3	684.2	857.7
SMA(25%) model	588.4	696.3	646.1	837.7
SMA(50%) model	608.2	704.2	661.7	859.0
SMA(100%) model	589.1	681.9	655.6	829.3
Batting career trajectory model	544.0	649.6	616.8	785.8

complexities of not out scores in the data. Of the fitted SMA models, the SMA(10%) model tends to perform worst of all, suggesting that only using recent form to predict player performance is inadvisable.

While the batting career trajectory model appears to provide the most accurate predictions of player performance, the best fitting SMA model appears to be the SMA(100%) model, which is equivalent to using a player's current batting average to predict future performance. This is a finding of particular interest and provides a stark warning to coaches and selectors; unless you are utilising modern analytical techniques to evaluate the effects of recent form on a player's underlying ability, it may be incredibly risky to use recent performances as an indicator of what is come in the future. Additionally, this may also cast some doubt on the accuracy of the ICC ratings if the exponential scale length is too small.

It is also worth noting that the same priors were used when fitting the batting career trajectory model to each player's career data. Doing so assumes that no prior information is known in regards to the batting abilities of individual players, which is often not the case. Instead, it would be plausible to apply different priors when analysing the careers of players who are considered specialist batsmen, all-rounders or bowlers, as there is a clear disparity in career batting averages, depending on a player's role in their team. It would also be reasonable to make use of domestic first-class career data when predicting how successful a player is going to be at the Test level. No doubt including such auxiliary information would further improve the predictive capabilities of the proposed batting career trajectory model.

2.4.2 Model comparison

As discussed in Section 1.5.2, a major advantage of using nested sampling to fit Bayesian models is the ability to compute the marginal likelihood at minimal extra cost, via Equation 1.9. This provides another means of assessing the fit of the batting career trajectory model, in addition to the computation of prediction errors in Section 2.4.1.

Table 2.12. *Marginal likelihood estimates for the top 20 Test match batsmen as ranked by the batting career trajectory model. The summation of marginal likelihoods for all players and the logarithm of the Bayes factor averaged across all players shows the data generally support the proposed model over the SMA(100%) model.*

Rank	**Player**		$\log(Z)$	$\log(Z_0)$	$\log(\frac{Z}{Z_0})$
1.	S. Smith	(AUS)	−592.0	595.0	3.0
2.	M. Labuschange	(AUS)	−120.5	−120.7	0.2
3.	B. Azam	(PAK)	−214.7	−219.1	4.4
4.	V. Kohli	(IND)	−669.6	−676.5	6.9
5.	D. Warner	(AUS)	−721.9	−727.9	6.0
6.	K. Williamson	(NZ)	−623.9	−630.2	6.3
7.	A. Mathews	(SL)	−642.4	−639.3	-3.1
8.	R. Sharma	(IND)	−219.3	−225.1	5.8
9.	M. Agarwal	(IND)	−86.8	−87.7	0.9
10.	J. Root	(ENG)	−797.9	−798.3	0.4
11.	R. Taylor	(NZ)	−756.7	−762.8	6.1
12.	C. Pujara	(IND)	−586.1	−589.9	3.8
13.	A. Ali	(PAK)	−672.8	−684.6	11.8
14.	A. Rahane	(IND)	−469.3	−469.7	0.4
15.	T. Latham	(NZ)	−416.9	−420.9	4.0
16.	M. Rahim	(BAN)	−553.4	−556.5	3.1
17.	B. Stokes	(ENG)	−541.6	−545.9	4.3
18.	D. Chandimal	(SL)	−451.5	−450.7	-0.8
19.	T. Head	(AUS)	−125.9	−125.0	-0.9
20.	BJ. Watling	(NZ)	−441.8	−445.2	3.4
	All players		−151,909.5	−153,473.4	1.5

In Table 2.12, the marginal likelihood is used to compare the support for the batting career trajectory model, Z, against the SMA(100%) model that assumes a player's ability remains constant throughout their career, Z_0. The marginal likelihoods are presented for each of the top 20 ranked Test batsmen identified in Section 2.3.4. The logarithm of the Bayes factor between the two models are also shown for each player, indicating the factor by which the proposed model is preferred over the SMA(100%) model. A positive value for this quantity implies that the batting career trajectory model is more likely to apply to a player's career data than the

SMA(100%) model; that is, there is a reasonable probability that a player's underlying ability has not remained constant throughout their career. The sum of marginal likelihoods over all players is also presented for each model, alongside the logarithm of the Bayes factor, averaged across all players.

In general, the batting career trajectory model is favoured over the SMA(100%) model for the majority of players. This provides further evidence to support the presence of temporal variation in batting ability during the careers of many players. It is worth noting that as the nested sampling algorithm is inherently a Monte Carlo process, the estimates for marginal likelihood are subject to sampling error. However, the algorithm was run with a large number of particles and MCMC iterations per nested sampling iteration, and this error is very small when compared with the differences in marginal likelihoods between the models.

2.5 Discussion

2.5.1 Limitations and further work

While the batting career trajectory model presented in this chapter attempts to account for temporal variation in ability that may exist on both short-term and long-term scales, due to a range of variables, it is worth acknowledging that a number of important variables have ignored. In regards to the getting your eye in process, variables such as balls faced or minutes batted may provide further information, in addition to runs scored.

Of particular note is the fact that opposition strength has been ignored. During the model development phase, it was difficult to establish a means of incorporating the quality of bowlers faced for each innings, without having a prior estimate for each player's bowling ability. Therefore, the assumption was made to treat all runs scored equally. As the data suggest there is far more variation in the abilities of batsmen bowled to for bowlers, compared with the variation in abilities of bowlers faced for batsmen, an adjustment for opposition strength was prioritised for the model that measures the career trajectories of bowlers, which is detailed in Chapter 3. However, including a measure of the strength of bowlers faced in each innings may improve the predictive capabilities of the batting career trajectory model. In the future, it may be possible to account for the strengths of bowlers faced during each batting innings, by taking the results and findings discussed in Chapter 3 and making post-hoc adjustments to the data that is fed into the model that evaluates batting career trajectories.

A worthwhile consideration in regards to the innings effect, ϕ, is whether the team batting in the first innings of a match has won the toss and chosen to do so, or whether the opposing team has sent them in to bat. Including such a variable may provide additional context when considering a player performance in their team's first innings of a match. In relation to the venue effect, ψ, it may be of value to group venues that tend to present similar batting conditions. For example, pitches in sub-continental countries, such as India, Bangladesh and Sri Lanka, tend to be analogous, while conditions in countries such as England and New Zealand are often observed to be alike. Treating the venue variable, v_t, on a geographical basis, rather than purely as home, away or neutral, may improve the accuracy of model prediction

Finally, while the hierarchical analysis across all players provides a general idea of the abilities of the average Test batsman, it is worth acknowledging some weaknesses of this two-stage approach, whereby all players are first analysed separately, then combining the output to summarise player performance across a wider group. Traditional hierarchical modelling would perform the individual inference by sharing information between players, allowing for the career trajectories of players with less available data to be informed by the data of other players. However, incorporating the data of all players in the ESPNcricinfo data set when analysing the careers of individual players would only increase the amount of processing required in an already computationally expensive process. Approaches exist that aim to improve efficiency in this regard — such as variational Bayesian methods — and computationally would be likely outperform the current nested sampling implementation. However, variational methods require a carefully designed family of target distributions used in the estimation of the posterior and can give poor results if these distributions are not well calibrated to the problem at hand. Conversely, MCMC provides accurate results if enough computing power is invested, but can be time inefficient. The present two-stage hierarchical approach, while somewhat computationally costly, is not so costly as to be impossible and provides a reasonable estimate of the abilities of a typical Test batsman. Developing a variational method that accurately approximates the posterior would likely provide a significant improvement on the computationally expensive nature of the current method, which is one of its limiting factors.

2.5.2 Concluding remarks

This chapter has proposed a novel method of measuring and predicting the past, present, and future batting abilities of individual players in Test match cricket. The results support a number of assumptions generally made about the batting performances of Test match cricketers: that the majority of players tend to score more runs when batting in their home country, in their team's first innings of a match (Figure 2.12). Additionally, there is evidence to support the idea that batting ability does not remain constant throughout a playing career, especially for those who have spent a number of years on the international circuit. Instead, batting ability appears to vary over time. For some players, this variation in ability exists on a shorter timescale as a result of recent performances. However, it is more common that this variation can be observed gradually, over the long-term, likely as a result of players gaining experience in a variety of match conditions around the world and being exposed to a range of world-class bowling attacks. In any case, the results indicate that it may be worthwhile considering the effects of form on an individual basis, with the abilities of some players more heavily influenced by recent performances than others.

In addition to supporting the concept of getting your eye in discussed in Stevenson & Brewer (2017), a number of players also exhibit behaviour typical of the concept of finding your feet, taking a number of innings to begin performing to the best of their abilities. A caveat to the idea of finding your feet may be applied to older players making their international debuts in the twilight of their careers, who already have a vast amount of domestic experience, with the results suggesting that there may be some correlation between a player's age and the speed at which they adjust to the demands of international cricket. While it is not always practical to provide every batsman with ample opportunity to prove themselves at the Test level, the results suggest that it may be beneficial to give younger players a few extra chances to showcase their talent. This recommendation is far from game-changing; one can assume such advice is already inherently followed by many coaches and selectors, given the clear advantages of unearthing a young superstar, as opposed to an aging one. However, it is always pleasing to have some statistical evidence to support the theory behind such decisions.

Although the predictions for batting ability can be associated with a reasonable amount of uncertainty as a result of the large amount of noise exhibited in many players' career data, the proposed model is shown to outperform a number of simple methods that are still routinely used to gauge player ability (Table 2.11), including the career batting average. Of practical interest is the result suggesting that estimating a player's current ability and future scores using only recent performances, is generally one of the least reliable methods. Often, both selectors and fans of the game will cite a player's recent form when it comes to justifying a new selection, or dropping an incumbent. These findings would suggest that doing so is inadvisable and may be a textbook case of falling for recency bias. Instead, dropping and selecting players on the basis of recent form may only be advisable if the players in question have shown a consistent tendency to string together numerous strong or poor performances in a row over the course of their career, which can be observed via their career trajectory and the posterior distributions for ℓ and γ.

The results have also been compared with more established means of measuring and ranking player ability, such as the ICC rating system. There is a reasonable amount of consensus between the two methods, however, by providing an intuitive cricketing interpretation of past, present, and future ability in units of a batting average, rather than arbitrary rating points, the model output can be easily understood by all followers of the game. This enables the differences in batting abilities between players to be quantified in more real terms, which allows for the results to be utilised by a wider audience. Moreover, as the model has been constructed within a Bayesian framework, it is simple to compare players and quantify differences in ability through the use of probabilistic statements, providing deeper insights pertaining to the risks and rewards of selecting one player over another. Consequently, the results may have have practical implications in high performance areas, such as talent identification and team selection policy.

Chapter 3

Estimating bowling career trajectories

3.1 Introduction

Bowling performances are typically measured in a similar manner to batting performances, using the bowling average: a metric defining the average number of runs a bowler concedes per wicket taken in their career (Equation 1.2). As with batting, bowling performances are typically summarised on an innings-by-innings basis. However, bowling performances are generally more difficult to objectively assess than batting performances, as more than one variable must be considered. Both the number of wickets taken and number of runs conceded are important factors when summarising a bowler's contribution during an innings. Consequently, there is no standardised means of visualising a player's career bowling performances, as there is for batting. The average number of runs conceded per over (economy rate) can also be worth considering, although this is far more important in one-day and T20 matches, where the batting side has a limited number of overs to score from.

As with the batting average, the same limitations that are discussed in Section 1.3.4 exist for the bowling average and its proposed alternatives, including the ICC ratings system. These are primarily (1) a lack of a clear cricketing interpretation; and (2) an inability to measure changes in ability that occur between individual innings and matches, as a result of players gaining experience and changes in player fitness and technique. There is far less discussion in regards to short-term variation in ability due to a getting your eye in effect for bowlers and as such, little research has been done in this area. As bowling is generally a more physically demanding task than batting (particularly pace bowling), it is entirely plausible, if not likely, that a bowler's underlying ability will deteriorate over the course of an innings, due to fatigue. It is a rare and highly impressive feat to see a non-spin bowler bowl for an entire session of play, which corresponds to a player bowling roughly up to 15 consecutive overs. On the contrary, multiple batsmen will often bat for consecutive sessions during a single match.

Given the difficulty of incorporating all available information, far fewer methods have been developed for assessing bowling performances, in comparison to the analysis of batting. Lemmer (2002) proposed the *combined bowling rate* (CBR) as an alternative to the bowling average for limited overs cricket. This metric combines the bowling average, economy rate, and strike rate in the evaluation of a bowling performance, rewarding not only the number of wickets taken, but also the rate at which both wickets are taken and runs are conceded. The CBR was further developed to apply to first-class and Test match cricket in Lemmer (2006) and included an adjustment to account for recent performances and the consistency of a bowler. However, the CBR ultimately lacks a meaningful interpretation, limiting its usage to the ranking and ordering of player performances.

Many of the potential reasons that players may exhibit variation in batting ability, which are discussed in Chapter 2, can also be applied to bowling and must be considered in the development of any proposed alternative to the bowling average. Similar innings and venue-specific effects are likely to be observed, as well as variations in ability that occur over the course of a playing career as a result of recent performances, experience, fitness, and improvements or deteriorations in technique. However, the difficulty in assessing the value of specific bowling performances is further compounded by the fact there can be a significant amount of variation in the abilities of batsmen bowled to.

When evaluating batting performances in Chapter 2, it is presumed that all runs scored by batsmen are equal. Of course, some teams will have stronger bowling attacks than others and this assumption has several limitations. However, as not every player is required to bowl, the variation in ability between a team's best and worst bowler in an innings is far smaller than the difference in ability between a team's best and worst batsman. Therefore, it is far more reasonable to treat runs scored by batsmen in a uniform manner than it is to make the assumption that all wickets taken by bowlers are equal. Two bowlers may have identical bowling figures in an innings, but context will determine who was the more valuable bowler. For example, multiple bowlers may have innings figures of 3/80 (read as three wickets for 80 runs conceded); if bowling to top-order batsmen, such a performance would usually be considered as highly valuable to the bowling team's effort, however, if a bowler took three wickets at the expense of 80 runs while bowling to the opposition's numbers 9, 10 and 11 batsmen, this would be seen as a rather costly performance.

Furthermore, the extent and timing of long-term variations in bowling ability may depend on the type of bowler in question, with bowlers broadly classified into one of two categories, each with multiple sub-categories. Firstly, there are *pace bowlers*, who aim to bowl the ball quickly, giving batsmen as little time as possible to react in order to induce a false shot. Pace bowlers who consistently bowl in excess of 140-145 km/h are often simply referred to as *fast* bowlers and use speed as their primary asset against opposition batsmen. Those with a slower average speed are often classified as *fast-medium* or *medium* bowlers, and often use a combination of speed and *swing* through the air to outfox their opponent. Pace bowling is considered the most physically demanding discipline in cricket and tends to result in the largest number of injuries. As such, pace bowlers do not usually have the same longevity as batsmen and will often peak in terms of speed earlier in their careers, when their bodies are at their physical peak — usually in their mid to late twenties.

Secondly, there are *spin bowlers*, who are much slower than pace bowlers, generally bowling at speeds below 100 km/h. As the name suggests, spin bowlers aim to generate movement off the pitch by imparting spin on the ball, which can be tricky to play, especially in conditions that promote significant deviation from a delivery's initial trajectory. Similar to pace bowlers employing the use of swing, spin bowlers aim to deceive the batsmen through the air through the use of *flight*, which is achieved by varying the pace and trajectory of individual deliveries. In terms of physicality, spin bowling is far less demanding than pace bowling. Instead, spin bowling is often referred to as one of the finer arts of cricket and is often associated with mind games and out-thinking the batsman, rather than using raw pace to gain the upper hand. As a result, spin bowlers can take longer to hone their craft, potentially peaking later in their careers and continuing to enjoy success well into their thirties.

In this chapter, the batting career trajectory model from Chapter 2 is extended to apply to the analysis of bowling performances. Before deriving the model, a means of visualising bowling performances is proposed and used throughout the chapter (Section 3.2.1). A method of adjusting raw bowling data to account for the quality of batsmen bowled to is then discussed in Section 3.2.2. The model likelihood and relevant parameters are specified in Section 3.3 before being fitted to the aforementioned adjusted data. The model output allows for construction of bowling career trajectories, allowing for the estimation of past, present, and future bowling abilities, including appropriate adjustments to account for the historic strength of batsmen that individual players have bowled to over their careers. The general findings are presented in Section 3.4 with player rankings compared with the ICC ratings method in Section 3.4.4. The validity of the model's fit and predictions of future performance are then assessed in Section 3.5.

3.2 Bowling data

3.2.1 Visualising bowling performances

As discussed in Section 3.1, bowling performances are typically presented on a per-innings basis and quantify two major variables: (1) the number of wickets taken, and (2) the number of runs conceded. To illustrate how a player's career bowling data are commonly recorded, the recent bowling performances for New Zealand pace bowler Neil Wagner are presented in Table 3.1. Here, it is possible to see the number of wickets taken and runs conceded in each innings, as well as the innings and venue-specific information for each performance. While the data can be readily understood when presented in table format, it is difficult to visualise multiple bowling performances over time, as this would require the data to be plotted in three dimensions: time, wickets taken, and runs conceded.

Table 3.1. *A summary of Neil Wagner's innings-by-innings bowling performances in his three most recent Test matches.*

Performance index	Overs	Balls	Runs conceded	Wickets	Innings	Team innings	Venue	Opposition	Ground
85	38	0	83	4	1	1	-1	Australia	Melbourne
86	17	2	50	3	3	-1	-1	Australia	Melbourne
87	33	1	66	3	1	1	-1	Australia	Sydney
88	9	0	37	0	3	-1	-1	Australia	Sydney
89	10	0	29	1	1	1	1	India	Christchurch
90	8	0	18	1	3	-1	1	India	Christchurch

Therefore, rather than considering bowling performances innings-by-innings, it may be advantageous to consider an individual performance as the number of runs conceded before a wicket is taken. As the Cricsheet data set discussed in Section 1.4.1 includes ball-by-ball data for almost all Test matches since 2008, the computation of runs conceded between each wicket is relatively straightforward. Rather than summarising Wagner's first innings performance against Australia in Melbourne as 4/83, this performance can be split into 4 distinct observations, where Wagner conceded 8, 36, 26 and 9 runs respectively, between each wicket. However, astute readers will note that the above four observations sum up to 79 runs conceded, rather than 83 as expected. This is a result of Wagner conceding an additional 4 runs, between taking his fourth wicket and the end of the opposition's batting innings. In this sense, the additional 4 runs conceded can be treated as a supplementary fifth observation, making note that this performance has no wicket associated with it, similar to how not out innings are treated in the context of batting. Under this framework, each observation is referred to as a *bowling spell*, allowing the same data from Table 3.1, to be presented as in Table 3.2.

As with the bowling data summarised innings-by-innings, each observed bowling spell in Table 3.2 includes the number of runs conceded as well as the relevant innings and venue-specific information. Similar to the batting data, the not out dummy variable indicates whether the observation corresponds to a wicket-taking spell (a value of 0), or a non wicket-taking spell (a value of 1). In this format, an individual's performances can be visualised over time, as is

Table 3.2. *A summary of Neil Wagner's spell-by-spell bowling performances in his three most recent Test matches.*

Spell index	Overs	Balls	Runs conceded	Not out	Innings	Team innings	Venue	Batsman dismissed	Wicket type	Opposition	Ground
262	3	2	8	0	1	1	-1	DA Warner	caught	Australia	Melbourne
263	22	2	36	0	1	1	-1	SPD Smith	caught	Australia	Melbourne
264	8	3	26	0	1	1	-1	TD Paine	lbw	Australia	Melbourne
265	2	1	9	0	1	1	-1	TM Head	caught	Australia	Melbourne
266	1	4	4	1	1	1	-1			Australia	Melbourne
267	4	1	19	0	3	-1	-1	DA Warner	caught	Australia	Melbourne
268	3	3	11	0	3	-1	-1	SPD Smith	caught	Australia	Melbourne
269	9	4	20	0	3	-1	-1	TM Head	bowled	Australia	Melbourne
270	6	3	12	0	1	1	-1	DA Warner	caught	Australia	Sydney
271	23	1	47	0	1	1	-1	JL Pattinson	bowled	Australia	Sydney
272	3	3	7	0	1	1	-1	MA Starc	bowled	Australia	Sydney
273	9	0	37	1	3	-1	-1			Australia	Sydney
274	9	4	29	0	1	1	1	GH Vihari	caught	India	Christchurch
275	0	2	0	1	1	1	1			India	Christchurch
276	5	3	7	0	3	-1	1	AM Rahane	bowled	India	Christchurch
277	2	3	11	1	3	-1	1			India	Christchurch

illustrated in Figure 3.1 with Neil Wagner's Test match bowling data.

3.2.2 Incorporating batsman-specific information

While the data presented in Table 3.2 does allow for the visualisation of performances over time, it still ignores the essential consideration in the context of bowling performances discussed in Section 3.1: the abilities of batsmen bowled to during each observation. To do so requires an estimate of the batting abilities of each batsman bowled to, over the course of each bowling spell. Fortunately, this is exactly what the batting career trajectory model presented in Chapter 2 attempts to measure.

The standardised bowling average

Recall the effective batting average function, $\mu(x, t)$ (Equation 2.9), which estimates a batsman's batting ability on score x, in their t^{th} career innings, in units of a batting average. The computation of $\mu(x, t)$ for an individual player, across all values of x and t, allows for the estimation of that player's underlying batting ability, for any ball faced during any innings of their career. It is possible to then append these estimates for batting ability, $\mu(x, t)$, to the Cricsheet ball-by-ball Test data set to provide an indication of batsman quality for every single delivery. Once an estimate for batsman ability exists for all balls bowled, it is possible to make adjustments for the quality of batsmen bowled to, during each bowling spell.

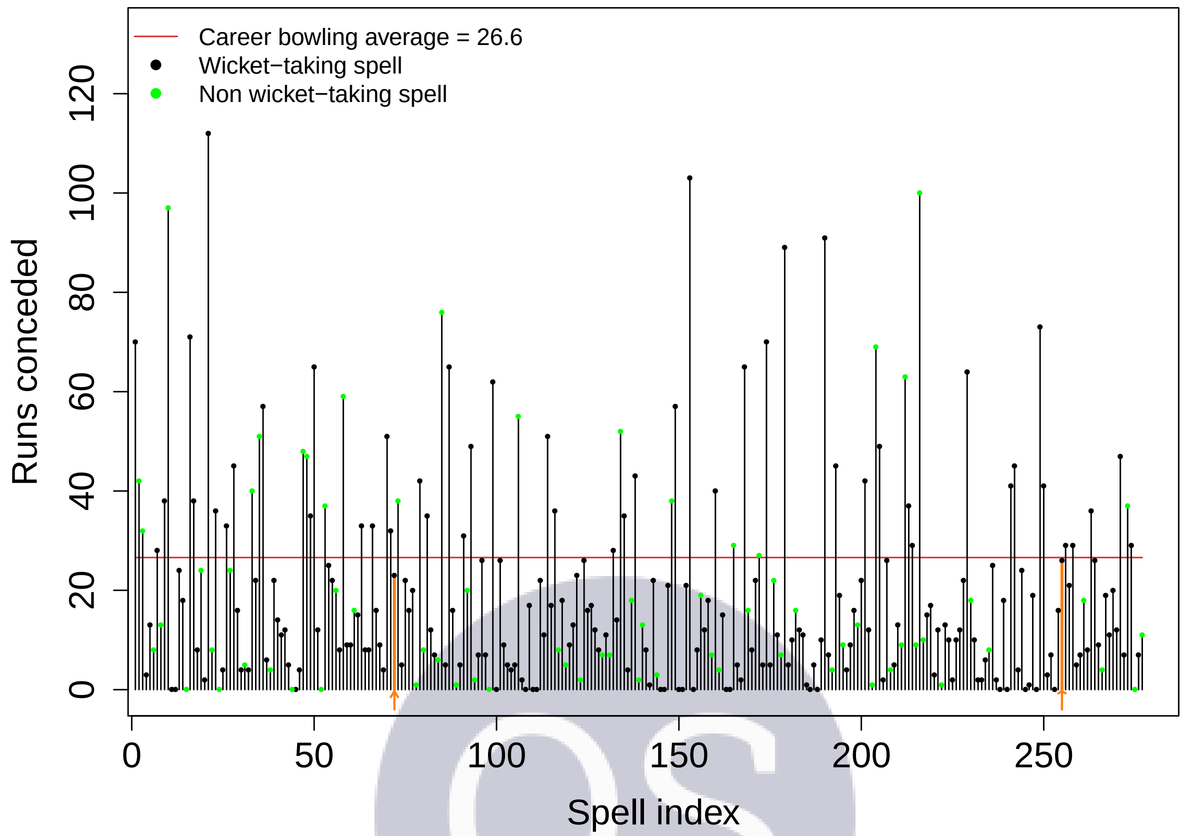


Figure 3.1. *Spell-by-spell bowling data for Neil Wagner in Test matches. Wagner's* 72^{nd} *and* 255^{th} *spells are highlighted in orange for comparison with Figure 3.2.*

This is achieved by introducing a new quantity, *standardised runs conceded*, denoted s_t, which represents the number of runs conceded during a bowler's t^{th} bowling spell, as a proportion of opposition batting ability, $\mu(x,t)$. For each ball bowled, the number of standardised runs conceded can be computed using Equation 3.1, noting that the estimate for $\mu(x,t)$ is the point estimate computed from the posterior predictive distribution. Taking the summation of standardised runs conceded during a bowling spell provides an estimate for s_t, allowing for a more direct means of comparing bowling spells, as the number of standardised runs conceded by a bowler accounts for the quality of batsmen bowled to.

$$\text{Standardised runs conceded} = \frac{\text{Runs conceded}}{\mu(x,t)} \tag{3.1}$$

For example, consider a batsman with an underlying batting ability estimate of $\mu(x,t) = 40.0$. If this batsman scored a boundary (four runs) from a single delivery, this would be counted as conceding $\frac{4}{40.0} = 0.1$ standardised runs. However, for a different batsman who has an underlying estimate of $\mu(x,t) = 10.0$, scoring a boundary would be counted as conceding $\frac{4}{10.0} = 0.4$ standardised runs. The batsman in the former example is estimated to have an underlying batting ability that is four times greater than the batsman in the latter example, which is reflected accordingly in the number of standardised runs conceded. Therefore, if a bowler were to concede a total of 0.5 standardised runs over the course of a bowling spell, before taking a wicket, one can assume the batsman bowled to scored a number of runs equivalent to half their estimated effective batting average, before being dismissed. Likewise, if a total of 2.0 standardised runs were conceded by a bowler before taking a wicket, one can assume the batsmen scored approximately twice as many runs as expected per their effective batting average, before being dismissed.

An example of the standardised Test match bowling data for Neil Wagner, which includes estimates for s_t, is presented in Table 3.3. Similar to Figure 3.1, the spell-by-spell standardised bowling data is presented in Figure 3.2 and provides a visualisation of performance over time, while also accounting for the strengths of batsmen bowled to, allowing for a more direct means of comparison between individual bowling spells.

Table 3.3. *A summary of Neil Wagner's spell-by-spell bowling performances in his three most recent Test matches, including estimates for the number of standardised runs conceded, s_t.*

Spell index	Overs	Balls	Runs conceded	Standardised runs conceded	Not out	Innings	Team innings	Venue	Batsman dismissed	Wicket type	Opposition	Ground
262	3	2	8	0.10	0	1	1	-1	DA Warner	caught	Australia	Melbourne
263	22	2	36	0.65	0	1	1	-1	SPD Smith	caught	Australia	Melbourne
264	8	3	26	0.63	0	1	1	-1	TD Paine	lbw	Australia	Melbourne
265	2	1	9	0.23	0	1	1	-1	TM Head	caught	Australia	Melbourne
266	1	4	4	0.17	1	1	1	-1			Australia	Melbourne
267	4	1	19	0.40	0	3	-1	-1	DA Warner	caught	Australia	Melbourne
268	3	3	11	0.25	0	3	-1	-1	SPD Smith	caught	Australia	Melbourne
269	9	4	20	0.50	0	3	-1	-1	TM Head	bowled	Australia	Melbourne
270	6	3	12	0.20	0	1	1	-1	DA Warner	caught	Australia	Sydney
271	23	1	47	0.89	0	1	1	-1	JL Pattinson	bowled	Australia	Sydney
272	3	3	7	0.26	0	1	1	-1	MA Starc	bowled	Australia	Sydney
273	9	0	37	0.64	1	3	-1	-1			Australia	Sydney
274	9	4	29	0.69	0	1	1	1	GH Vihari	caught	India	Christchurch
275	0	2	0	0.00	1	1	1	1			India	Christchurch
276	5	3	7	0.20	0	3	-1	1	AM Rahane	bowled	India	Christchurch
277	2	3	11	0.41	1	3	-1	1			India	Christchurch

Several interesting features of the standardised data presented in Figure 3.2 stand out immediately, in comparison to the data that measures each bowling spell in terms of raw runs conceded. Firstly, as observed in Figure 3.1, Wagner's most expensive performance was his 21$^{\text{st}}$ career spell, where he conceded 112 runs before taking a wicket. However, after adjusting for the strength of the opposition batsmen, it appears Wagner's most expensive performance was his 190$^{\text{th}}$ spell, where he conceded 3.04 standardised runs (90 runs), compared with 2.89 standardised runs (112 runs) in his 21$^{\text{st}}$ career spell.

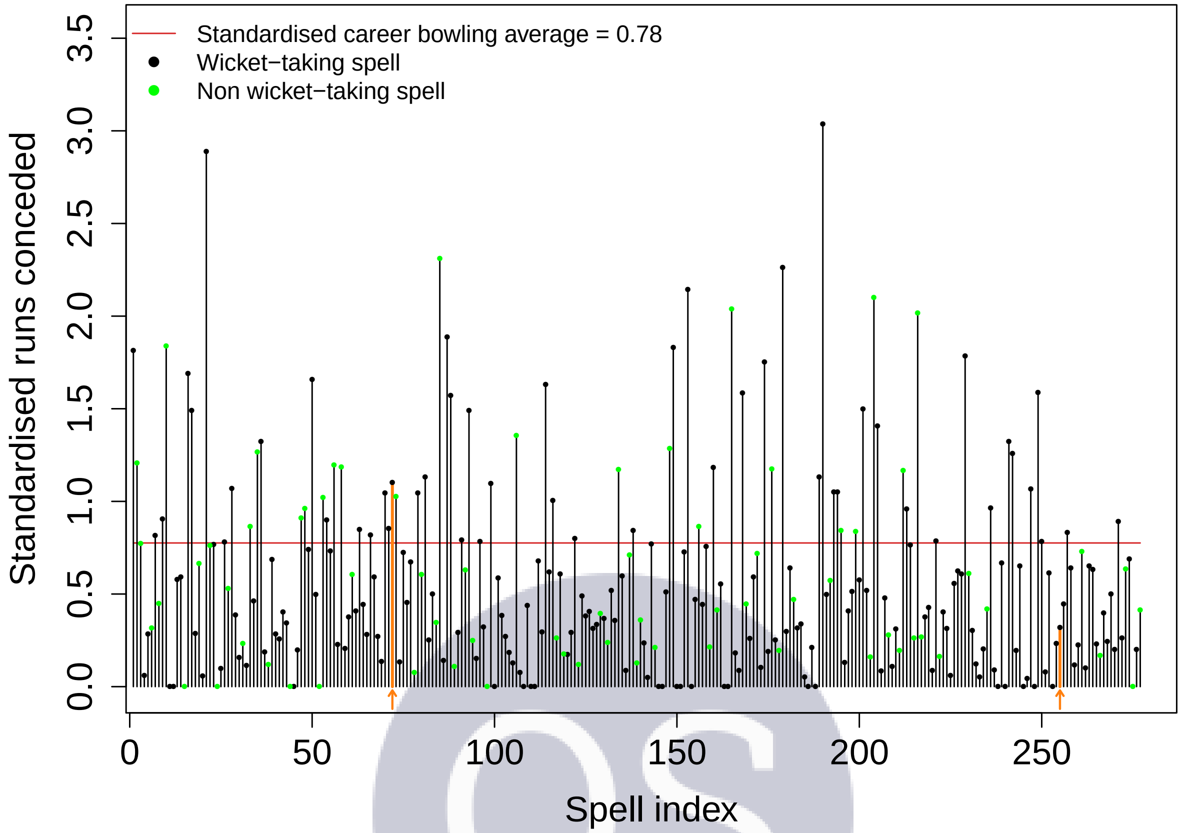


Figure 3.2. *Spell-by-spell bowling data for Neil Wagner in Test matches, in units of standardised runs conceded. Wagner's* 72^{nd} *and* 255^{th} *spells are highlighted in orange.*

Secondly, useful insights can be deduced by comparing bowling performances that appear similar in Figure 3.1, but very different in Figure 3.2. For example, in his 72^{nd} and 255^{th} career bowling spells, Wagner conceded 23 and 26 runs respectively — two objectively similar-looking performances. After adjusting for batsman ability, these performances correspond to Wagner conceding 1.10 and 0.32 standardised runs respectively, suggesting in relative terms, his 72^{nd} spell was almost four times more expensive than his 255^{th}. Upon closer inspection, it appears as though Wagner conceded a number of runs to India's lower order in his 72^{nd} spell before taking a wicket, while in his 255^{th} spell Wagner's main opponent was Australian Steve Smith, who in Chapter 2 was identified as one of the best batsmen in the world over the last several years. To illustrate the magnitude of this difference, Wagner's 72^{nd} and 255^{th} spells are highlighted in orange in Figure 3.2.

Finally, like the traditional career bowling average, it is possible to compute a standardised career bowling average, via Equation 3.2. This quantity represents the number of runs a bowler concedes per wicket, on average, as a proportion of a batsman's underlying batting ability,

$\mu(x,t)$, which is expressed in units of a batting average.

$$\text{Standardised bowling average} = \frac{\sum \text{Standardised runs conceded}}{\sum \text{Wickets taken}} \tag{3.2}$$

Using Equation 3.2, it is possible to compute Neil Wagner's standardised career bowling average of 0.78, which is included in Figure 3.2. This value indicates that on average, Wagner concedes runs equivalent to 78% of a batsman's estimated underlying effective batting average, $\mu(x,t)$, for every wicket taken.

In order to gain a deeper understanding of the relationship between runs and standardised runs, it is worth taking a look at the summary statistics of the original data, and the data after it has been adjusted to account for the estimated abilities of batsmen bowled to. Table 3.4 summarises the batting and bowling data obtained from the ball-by-ball Cricsheet data source. It is worth noting that extra runs, such as wides and no balls are counted as runs conceded by bowlers, but are not counted as runs scored by batsmen. In the case of these extra runs, the number of standardised runs conceded is calculated based on the estimate for the effective batting average, $\mu(x,t)$, of the batsman who was on strike.

Table 3.4. *Summary of batting and bowling data across all 532 Test matches in the Cricsheet data source. Summaries for the data split by both innings and venue are also provided.*

		Team innings		**Venue**		
Quantity	**Overall**	1$^{\text{st}}$ innings	2$^{\text{nd}}$ innings	Home	Away	Neutral
Runs scored	534,906	339,412	195,494	271,943	259,717	3,246
Standardised runs scored	14,979.7	9,013.2	5,966.5	7,093.0	7,776.5	110.2
Runs conceded	543,468	344,949	198,519	263,704	276,455	3,309
Standardised runs conceded	15,274.7	9,196.2	6,078.5	7,922.5	7,239.7	112.5
Bowler-credited wickets	16,512	9,769	6,743	8,760	7,615	137
Non bowler-credited wickets	458	265	193	234	220	4
Batting average	31.5	33.8	28.2	34.7	28.9	23.0
Standardised batting average	0.88	0.90	0.86	0.91	0.86	0.78
Bowling average	32.9	35.3	29.4	30.1	36.3	24.2
Standardised bowling average	0.93	0.94	0.90	0.90	0.95	0.82

The results from Table 3.4 indicate that the overall bowling average across all Test matches in the Cricsheet data is 32.9. This is slightly higher than the overall batting average of 31.5, as a result of extras conceded, as well as wickets such as run outs, not being credited to bowlers. These batting and bowling averages correspond to overall standardised batting and bowling averages of 0.88 and 0.93. Intuitively, one might expect the overall standardised bowling average to be value equal to, or very close to 1.0. However, as the estimates for batting ability, $\mu(x, t)$, correctly reward batsmen for remaining on not out scores, the estimates for a player's underlying batting ability tend to be slightly larger than their overall career average. Consequently, the overall standardised bowling average is slightly less than 1.0. With this in mind, it is possible to say that a player with a standardised bowling average less than 0.93 is better than average, while a player with a standardised bowling average greater than 0.93 is worse than average.

It is important to clarify that when computing the standardised bowling data presented in Table 3.4, the innings and venue-specific strengths and weaknesses of batsmen have been accounted for. Therefore, any difference in standardised bowling averages observed between innings or across difference venues, is the result of any residual effect of bowling in the first or second innings, or at a home or away venue, after adjusting for batting ability. For example, the raw bowling averages at home and away venues are 30.1 and 36.3 respectively, corresponding to an approximate 20% difference in the average number of runs conceded per wicket. However, the equivalent standardised bowling averages are 0.90 and 0.95, a difference of roughly 5%. That is to say, once the respective innings and venue-specific abilities of batsmen are taken into account, there is a residual unexplained 5% difference in standardised runs conceded per wicket by players bowling at home and away venues, which may be a result of bowler familiarity with local conditions or an effect due to playing in front of a supportive home crowd.

The career bowling averages and standardised bowling averages of all players in the Cricsheet data, who have taken at least 10 career wickets, are presented in Figure 3.3. As expected, there is a clear positive correlation between the two quantities, with the association becoming stronger as the total number of wickets a player has taken increases. This result is unsurprising; the law of large numbers would suggest that the overall quality of batsman bowled to over a player's career will converge to the average batsman, the more a player has bowled. However, for players with the same, or similar career bowling averages, there is still a meaningful amount of variation in standardised bowling averages, which provides further evidence that batsman quality is an important variable to consider when it comes to analysing bowling performances.

The adjusted bowling average

Although the standardised bowling average adjusts for the quality of opposition batsmen bowled to, it is a quantity that is unlikely to be readily understood by many of those within the cricketing community. This violates the essential criterion of *interpretability* that was identified in Section 1.2. Therefore, it is advisable to make a modification to the standardised bowling average to allow for a more meaningful and intuitive interpretation. This is achieved by introducing the *adjusted bowling average*, which represents a player's expected bowling average *if they were to bowl to the average Test batsman.*

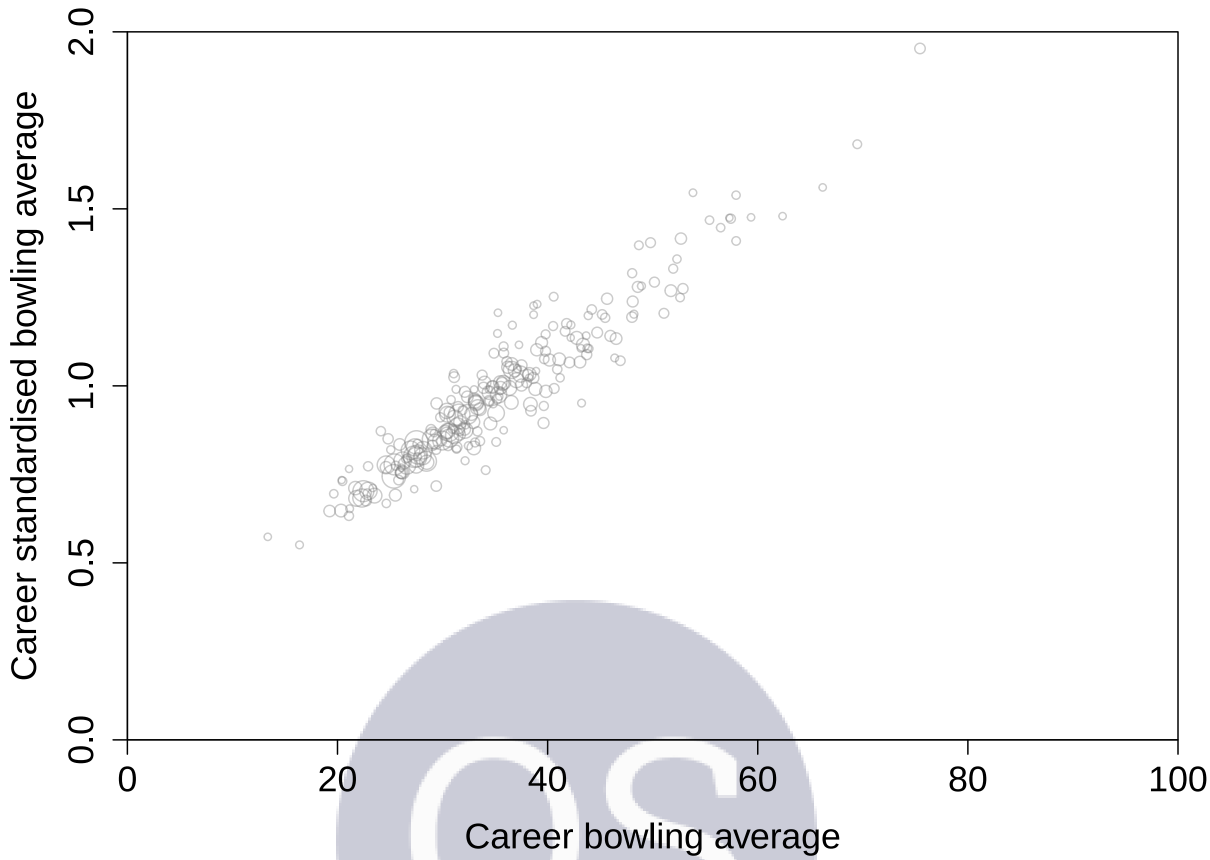


Figure 3.3. *Career bowling averages plotted against standardised career bowling averages for all players in the Cricsheet data set who have taken 10 or more wickets. The size of each point is representative of the total number of wickets taken by the player.*

In order to convert a player's standardised bowling average into an adjusted bowling average, an estimate for the batting ability of the average Test batsman must be computed. As the effective batting average, $\mu(x, t)$, expresses batting ability in units of a batting average, this quantity can be estimated easily be computing the mean effective batting average, denoted $\bar{\nu}$, across all Test matches, via Equation 3.3.

$$\bar{\nu} = \frac{\sum \text{Runs conceded}}{\sum \text{Standardised runs conceded}} \tag{3.3}$$

Using the quantities in Table 3.4, the average Test batsman has an estimated effective average $\bar{\nu} = \frac{543468}{15274.7} \approx 35.6$. It is then possible to obtain a player's adjusted career bowling average by multiplying their standardised career bowling average (Equation 3.2), by the the estimate for $\bar{\nu}$, as per Equation 3.4.

$$\text{Adjusted bowling average} = \text{Standardised bowling average} \times \bar{\nu} \tag{3.4}$$

Returning briefly to the bowling career of Neil Wagner to provide a practical example, it is possible to compute his adjusted career bowling average as: $0.78 \times \bar{\nu} = 27.6$. That is to say, if Wagner were to bowl to the average Test batsman, he would be expected to concede an average of 27.6 runs per wicket, over the course of his career. This is slightly higher than his career bowling average of 26.6. Therefore, one can infer that Wagner has bowled to slightly below average batsmen during his career.

It is possible to obtain an estimate for the average ability of batsmen bowled to during a particular player's career, by only including that player's bowling career data in the computation of $\bar{\nu}$ in Equation 3.3. For Wagner, $\bar{\nu}_{Wagner} = 34.3$, confirming the suspicion above, that the average batting ability of batsmen Wagner has conceded runs to over the course of his Test career, is slightly lower than the average Test batting ability, $\bar{\nu}$. The computation of $\bar{\nu}$ for individual players can be useful for comparing the relative strength of batsmen bowled to between players and indicates which players are going to see significant differences between their career bowling average and adjusted bowling average.

Table 3.5. *Career bowling averages, standardised career bowling averages, and adjusted career bowling averages for the current top 10 Test bowlers, as per the ICC ratings. An estimate for the average ability of batsmen bowled to, $\bar{\nu}_{player}$, is also provided for each player.*

Rank	Player		Wickets	Career bowling average	Standardised bowling average	Adjusted bowling average	$\bar{\nu}_{player}$	ICC rating
1.	P. Cummins	(AUS)	143	21.8	0.68	24.3	32.0	904
2.	S. Broad	(ENG)	513	27.5	0.84	29.9	32.7	845
3.	N. Wagner	(NZ)	206	26.6	0.78	27.6	34.3	843
4.	T. Southee	(NZ)	284	29.0	0.85	30.2	34.1	812
5.	J. Holder	(WI)	111	27.5	0.77	27.6	36.5	810
6.	K. Rabada	(SA)	197	22.9	0.70	25.0	32.6	802
7.	M. Starc	(AUS)	244	27.0	0.82	29.1	32.9	797
7.	J. Anderson[1]	(ENG)	538	25.4	0.74	26.5	34.1	781
8.	J. Bumrah	(IND)	68	20.3	0.65	23.0	31.4	779
10.	T. Boult	(NZ)	267	27.6	0.81	28.7	34.3	770

Thinking in terms of adjusted bowling averages allows for the direct comparison of players, as the historic strength of batsmen bowled to has been accounted for. For example, the career bowling averages and adjusted averages for the current top 10 Test bowlers, as ranked by the

[1] James Anderson's Test career began prior to 2008. Consequently, the Cricsheet data source does not contain bowling data for all 600 of his Test match wickets

ICC, are presented in Table 3.5. Interestingly, all of the 10 of the current top Test bowlers are pace bowlers, who frequently open the bowling. In every innings, opening bowlers are given the opportunity to bowl at two opening batsmen who are yet to get their eye in. That is to say, opening bowlers will always have a chance at bowling against two batsmen whose estimates of batting ability at the start of an innings, $\mu(0, t)$, are typically lower than of the average Test batting ability, $\bar{\nu}$. Consequently, adjusted bowling averages for all of the top 10 bowlers are larger than their career bowling averages, as the average ability of batsmen bowled to, is lower than that of the average Test batting ability. This is reflected in the individual estimates for $\bar{\nu}$, for each player. The player whose career bowling average and adjusted bowling averages are most similar is West Indian Jason Holder. When considering the context of Holder's role in the West Indian side, this result is unsurprising; Holder has split his time between being an opening bowler and a first-change bowler, with the latter not guaranteed the opportunity of bowling at two batsmen yet to get their eye in.

3.3 Model specification

3.3.1 Model likelihood

The derivation of the model likelihood in the context of bowling is closely related to the likelihood used in the analysis of batting performances in Chapter 2, derived in Stevenson & Brewer (2017, 2018, 2021). As discussed in Section 3.2.2, during a single bowling spell a player bowls and continues to concede runs until: (1) a wicket is taken, (2) the opposition's batting innings is concluded, or (3) the match is concluded. Let $R \in \{0, 1, 2, 3, ...\}$, denote the number of runs conceded in a given bowling spell. If $p \in [0, 1]$ defines the probability of a bowler taking a wicket without conceding any more runs in a given bowling spell, and bowling ability is assumed to be constant during a single bowling spell, then the probability mass function for R in Equation 3.5 can be expressed as a geometric distribution.

$$P(R = r) = p\,(1 - p)^r \tag{3.5}$$

However, in order to account for the strength of opposition batsmen bowled to during a given spell, bowling performances are measured in units of standardised runs conceded, as opposed to runs conceded. As a result, the data in question are continuous, rather than discrete. When moving from a discrete to a continuous response, a natural conversion is to move from a geometric distribution to an exponential distribution. Therefore, let $S \in \mathbb{R}_{\geq 0}$, represent the number of standardised runs conceded in a given bowling spell. By introducing the *standardised effective bowling average function*, $\omega(t) \in \mathbb{R}_{\geq 0}$, which defines a player's underlying bowling ability during their t^{th} career bowling spell, in units of a standardised bowling average, it is possible to express the exponential probability density function for S, as per Equation 3.6.

$$f_S(s) = \frac{1}{\omega(t)} \exp\left(\frac{-s}{\omega(t)}\right) \tag{3.6}$$

Under this specification, S has expectation: $\mathbb{E}[S_t] = \omega(t)$, and variance: $\text{Var}(S_t) = \omega(t)^2$. For any value of s, Equation 3.6 defines the probability of a bowler taking a wicket, without conceding any further runs and defines the likelihood function for a single bowling spell. However, as with not out scores in the batting model, there are a number of *non wicket-taking* observations, where a bowler's spell ends without a wicket being taken. In such cases, the likelihood is computed as $P(S \geq s)$, which assumes the bowler would have conceded a further number of standardised runs before taking a wicket, although this observation is not truly observed. Equation 3.7 defines $P(S \geq s)$, which is initially expressed as $1 - P(S \leq s)$ and is equivalent to $1 - F_S(s)$, where $F_S(s)$ represents the cumulative density function for S.

$$\begin{aligned}
P(S \geq s) &= 1 - P(S \leq s) \\
P(S \geq s) &= 1 - F_S(s) \\
P(S \geq s) &= 1 - \left(1 - \exp\left(\frac{-s}{\omega(t)}\right)\right) \\
P(S \geq s) &= \exp\left(\frac{-s}{\omega(t)}\right)
\end{aligned} \tag{3.7}$$

Therefore, if T is the total number of bowling spells in a player's career record, and N is the number of non wicket-taking spells, then the probability density of a set of conditionally independent wicket-taking observations, $\boldsymbol{x} = \{x_1, x_2, ..., x_{T-N}\}$, and not out wicket-taking observations, $\boldsymbol{y} = \{y_1, y_2, ..., y_N\}$, can be expressed as

$$f_S\left(\{\boldsymbol{x}, \boldsymbol{y}\}\right) = \prod_{t=1}^{T-N} \frac{1}{\omega(t)} \exp\left(\frac{-x_t}{\omega(t)}\right) \times \prod_{t=1}^{N} \exp\left(\frac{-y_t}{\omega(t)}\right). \tag{3.8}$$

For a set of known data, $\{\boldsymbol{x}, \boldsymbol{y}\}$, Equation 3.8 defines the likelihood function for any proposed value for $\omega(t)$. Therefore, conditional on the set of parameters, $\boldsymbol{\theta}$, defining the functional form of the standardised effective bowling average, $\omega(t; \boldsymbol{\theta})$, it is possible to derive the log-likelihood function, $\ell(\boldsymbol{\theta})$, from Equation 3.9.

$$\ell(\boldsymbol{\theta}) = -\left(\sum_{t=1}^{T-N} \log\big[\omega(t)\big] - \frac{x_t}{\omega(t)}\right) - \sum_{t=1}^{N} \frac{y_t}{\omega(t)} \tag{3.9}$$

The equations in Section 3.3.1 assume that bowling ability is constant during any given bowling spell, as there is little evidence — both anecdotal and empirical — to suggest that an effect similar to the concept of getting your eye in exists when it comes to analysing bowling performances. However, depending on the parameterisation of $\omega(t)$, bowling ability is free to vary over time, between spells and across a playing career, as discussed in Section 3.3.2.

3.3.2 Parameterising the effective bowling average function

Individual-specific effects

Conceptually, the standardised effective bowling average, $\omega(t)$, can be thought of as the rate at which a bowler will concede runs for every wicket taken, in their t^{th} career bowling spell. In order to parameterise $\omega(t)$ to allow for variation over time, a time-dependent parameter, w_t, is introduced. Here, w_t represents a player's underlying bowling ability in their t^{th} career bowling spell, in units of a standardised bowling average. While bowling performances do not exhibit the same amount of variation as seen in batting performances, the concept of form is still a worthwhile consideration on an individual level.

As with the batting career trajectory model, the time-dependent parameter used to measure the effects of short and long-term form, w_t, is modelled using a Gaussian process prior, with a mean value, λ, and covariance function $K(t_j, t_k)$, where t represents the index of a player's j^{th} and k^{th} career bowling spells (Rasmussen & Williams, 2006).

$$w_t \sim \text{GP}\left(\lambda, K(t_j, t_k)\right) \tag{3.10}$$

Given the success of the Gaussian process used to model batting career trajectories in Chapter 2, the same γ-exponential covariance function — presented again in Equation 3.11 — is used in the specification of the Gaussian process that governs the set of $\{w_t\}$ terms. The covariance function has three parameters, $\{\sigma, \ell, \gamma\}$, and allows for both immediate, short-term changes in w_t, and more gradual, long-term changes in w_t.

$$K(t_j, t_k) = \sigma^2 \exp\left(-\frac{|j-k|^\gamma}{\ell^\gamma}\right) \tag{3.11}$$

Innings and venue-specific effects

As with batting performances, innings and venue-specific effects are likely to exist when analysing bowling performances, as bowlers tend to get more assistance from a pitch that has deteriorated over several days and will generally prefer to bowl in familiar, home conditions. Several exceptions to this assumption will exist, for example, in New Zealand, rather than deteriorating from the first ball, pitches can often be at their most difficult to bat on during day one, before flattening out and being ideal for batting later in the match. Regarding venue, spin bowlers from countries where conditions are not typically spin-friendly (such as England and New Zealand), may prefer bowling in the sub-continent where pitches tend to be more favourable toward spin bowling. Similarly, pace bowlers from the sub-continent may relish the opportunity to play overseas, where there is often more pace and bounce to be exploited than in their local, home environment.

For each bowling spell, indicator variables i_t and v_t, provide the innings and venue-specific information, with ϕ and ψ representing the innings and venue-specific effects respectively.

$$i_t = \begin{cases} 1, & \text{if batting team's first innings of a match} \\ -1, & \text{if batting team's second innings of a match} \end{cases}$$

$$v_t = \begin{cases} 1, & \text{if bowling at a home venue} \\ 0, & \text{if bowling at a neutral venue} \\ -1, & \text{if bowling at an away venue} \end{cases}$$

Therefore, the standardised effective bowling average function, $\omega(t)$, represents a player's underlying bowling ability in their t^{th} career bowling spell, in units of a standardised bowling average, conditional on whether the opposition is batting in their first or second innings and the venue. The functional form for $\omega(t)$ is defined in Equation 3.12 as follows:

$$\omega(t) = w_t \times \phi^{i_t} \times \psi^{v_t}. \tag{3.12}$$

As with the effective batting average, the posterior predictive estimate for $\omega(t)$ can be plotted over time to provide a bowling career trajectory, illustrating how the model estimates underlying bowling ability to have varied over the course of a player's career. However, as noted in Section 3.2.2, the adjusted bowling average is a more intuitive quantity to interpret, compared with the standardised bowling average. Therefore the *adjusted effective bowling average function*, $\alpha(t)$, is introduced in Equation 3.13, allowing the model output to be presented and interpreted in units of an adjusted bowling average.

$$\alpha(t) = \omega(t) \times \bar{\nu} \tag{3.13}$$

3.3.3 Prior distributions

The full set of model parameters, $\boldsymbol{\theta}$, defined in Section 3.3.2, are as follows:

$$\boldsymbol{\theta} = \{\phi, \psi, \{w_t\}, \lambda, \sigma, \ell, \gamma\}.$$

As a constant ability is assumed during each bowling spell, the model has fewer parameters when analysing bowling performances, compared with the model in Chapter 2 used to analyse batting performances. Each of the relevant quantities, parameters, and functions are provided in Table 3.6, with prior distributions defined where necessary. Similar to the batting career trajectory model, the prior distributions are generally informative, but conservative, reflecting a general understanding of the distribution of bowling abilities of players in Test cricket.

Table 3.6. *The bowling career trajectory model parameters, data, and effective average functions, including the prior distribution for each quantity where relevant.*

Quantity	Interpretation	Prior
Data		
t	Career bowling spell index (time)	
o_t	Wicket-taking/non wicket-taking flag in t^{th} career bowling spell	
i_t	Batting team innings # in t^{th} career bowling spell	
v_t	Venue in t^{th} career bowling spell	
s_t	Standardised runs conceded in t^{th} career bowling spell	Likelihood function given in Equation 3.9
$\bar{\nu}$	Average Test match batting ability, in units of an effective batting average	Quantity computed via Equation 3.3
Innings and venue-specific effects		
ϕ	Team innings # effect	$\log(\phi) \sim \text{Normal}(\log(1), 0.25^2)$
ψ	Venue effect	$\log(\psi) \sim \text{Normal}(\log(1), 0.25^2)$
Gaussian process parameters		
$\{w_t\}$	Underlying bowling ability in t^{th} career bowling spell	$\log(\{w_t\}) \sim \text{GP}(\lambda, K(t_j, t_k; \sigma, \ell, \gamma)$
λ	Mean value of Gaussian process	$\log(\lambda) \sim \text{Normal}(\log(1), 0.5^2)$
σ	Scale parameter of covariance function, $K(t_j, t_k)$	$\log(\sigma) \sim \text{Normal}(\log(0.2), 1^2)$
ℓ	Length parameter of covariance function, $K(t_j, t_k)$	$\log(\ell) \sim \text{Normal}(\log(20), 1^2)$
γ	Smoothing parameter of covariance function, $K(t_j, t_k)$	$\gamma \sim \text{Uniform}(1, 2)$
Covariance and effective average functions		
$K(t_j, t_k)$	Covariance function for Gaussian process	Functional form given in Equation 3.11
$\omega(t)$	Bowling ability in t^{th} career bowling spell, in units of a standardised bowling average	Functional form given in Equation 3.12
$\alpha(t)$	Bowling ability in t^{th} career bowling spell, in units of an adjusted bowling average	Functional form given in Equation 3.13

As discussed in Section 3.3.2, the prior for the set of $\{w_t\}$ terms is specified by a Gaussian process, with an underlying mean, λ, and γ-exponential covariance function $K(t_j, t_k; \sigma, \ell, \gamma)$. However, bowling ability — regardless of whether it is measured in units of a bowling average, standardised bowling average or adjusted bowling average — must be positive. As such, it is the set of $\log\{w_t\}$ terms that are modelled by the Gaussian process prior, which are then back-transformed accordingly to ensure positivity in the estimates.

The log-normal priors over the innings and venue-specific effects, ϕ and ψ, are centred on one, implying the model assumes a player is equally capable at bowling at both home and away venues, across all innings of a match, unless the data suggest otherwise. Likewise, the log-normal prior for λ suggests that the median player will have a bowling ability, w_t, close to a value of 1.0, in units of a standardised bowling average. As seen from the summary data in Table 3.4, this is a reasonable specification. The conservative log-normal priors for the parameters controlling the flexibility of the Gaussian process prior, ℓ and γ, allow for a range of plausible career trajectories to be fitted to the data.

The most restrictive prior is again the log-normal prior over σ, which assumes the median player's bowling ability will vary by approximately plus or minus 20% from their underlying average bowling ability, λ, over the course of their career. A prior that entertains the possibility of larger vales for σ can lead to proposed career trajectories that suggest a player's underlying bowling ability fluctuates between bowling spells in a manner that is entirely unrealistic in the context of sporting ability.

3.3.4 Model fitting

Data

The model defined in Section 3.3 has been applied to all players for whom ball-by-ball bowling data exists in the Cricsheet data set. This corresponds to a total of 522 players, from 12 different countries, who have taken a total of 16,512 wickets, and have bowled in a combined total of 25,477 bowling spells.

The model is fitted to each player's standardised career data, to ensure the historic strength of opposition batsmen bowled to is accounted for. This enables the model output to then be viewed in terms of a standardised bowling average, or, post-processed to be viewed in units of an adjusted bowling average, which may be interpreted more easily by members of the cricketing community. As seen in the career bowling data of Neil Wagner, presented in Figures 3.1 and 3.2, there can be a reasonable amount of variation between observations, although perhaps less so than the batting data presented in Chapter 2. Nevertheless, such statistical noise may make it difficult for the model to distinguish between smooth career trajectory functions, which imply a gradual change in bowling ability over the long-term, or more jagged career trajectories, implying a bowler's ability is more closely tied with recent performances.

As the model assumes a player's underlying bowling ability is not heavily influenced by the current state of the match, the most appropriate application is in the analysis of bowling performances in longer form cricket. In one-day, and particularly T20 matches, a batsman's aggressiveness and rate at which they attempt to score runs can fluctuate profoundly. Therefore, when a wicket is taken in such situations it can be difficult to determine whether the wicket was a result of quality bowling, or because of a poorly executed attacking shot from the batsman. One could speculate that bowlers who are able to consistently elicit false shots from the batsman, regardless of their aggressiveness, should be rewarded and there is merit to this argument. However, in the closing stages of an opposing team's batting innings (often referred to as the *death* of an innings), batsmen are often trying to execute a strategy that maximises their team's final total, which is usually an approach that sees players attempt to hit the ball out of the ground on the majority of deliveries. Such scenarios tend to result in a higher than average proportion of attacking shots being played and consequently, a higher proportion of false shots. This is not to say that all players are able to regularly take wickets at the death of an innings, rather, it is more difficult to distinguish the abilities of bowlers from one another, compared with Test match cricket, where batsmen are generally free to score at their own pace and have minimal external pressure to do so. With this in mind, there is probably more merit in applying the proposed bowling model to one-day and T20 cricket than there is for the batting model, but for the sake of consistency, the results presented in Section 3.4 will focus solely on Test cricket.

Nested sampling

As discussed in Section 1.5.2, the bowling model is fitted using a C++ implementation of the nested sampling algorithm proposed by Skilling (2006). The output of the nested sampling algorithm provides posterior samples for each of the model parameters, as well as the marginal likelihood, which is used for model comparison. The effective sample size (ESS) of each nested sampling run is also computed using Shannon entropy (Shannon, 1948), to ensure the algorithm has effectively explored the parameter space. The results reported in Section 3.4 for each player are based on nested sampling runs initiated with 1,000 particles and use 1,000 MCMC steps per nested sampling iteration. The results were not sensitive to these adjustable tuning parameters, indicating the sampling was sufficient.

As with the batting career trajectory model, the run-time of the model in a bowling context varies considerably, depending on the number of career bowling spells a player has bowled in, which determines the number of parameters used to fit the model. Once again, the model fitting process was implemented using parallel cloud computing via the high performance computing facilities provided by the New Zealand eScience Infrastructure. This allows the model to be fitted to the career bowling data of many players simultaneously, significantly improving the computational efficiency of the process.

3.4 Results

3.4.1 Analysis of individual bowlers

Of primary interest when evaluating the model output for an individual player is the posterior predictive distribution for the effective bowling average functions, $\omega(t)$ and $\alpha(t)$, which allow for the results to be interpreted in units of both a standardised bowling average, or an adjusted bowling average. Plotting the effective bowling average over time, t, which in the context of bowling is represented in terms of bowling spells, gives a player's bowling career trajectory. This provides an indication of a player's past, present, and future bowling abilities, adjusted for the relative strength of opposition batsmen bowled to. The bowling career trajectories of all 522 players analysed are available to view via the same RShiny application that hosts the batting career trajectories at `www.oliverstevenson.co.nz/phd_cricket_visualisation`.

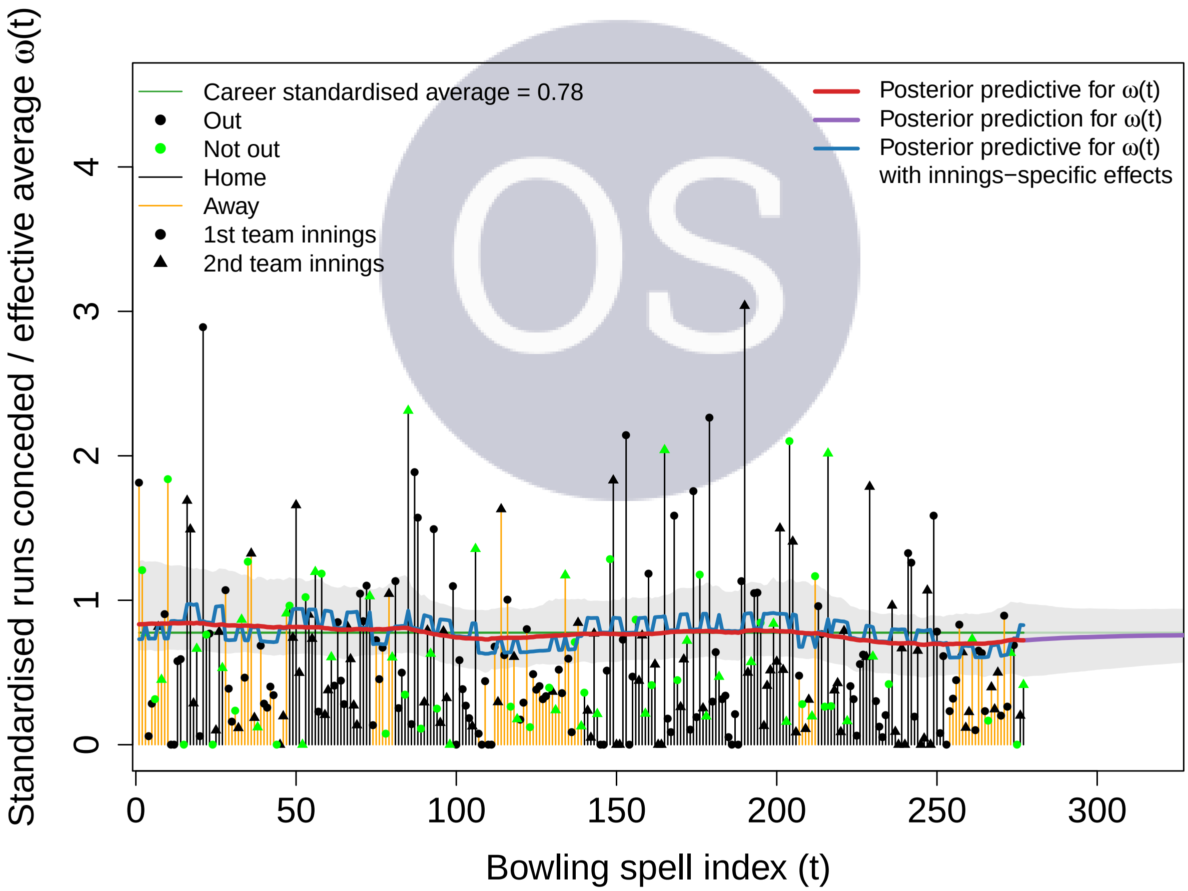


Figure 3.4. *Test match bowling career trajectory (the posterior median of $\omega(t)$) for Neil Wagner, in units of a standardised bowling average. The 95% credible interval is also provided (shaded region). A prediction for the standardised effective bowling average, $\omega(t)$, is also provided for 50 bowling spells into the future (purple).*

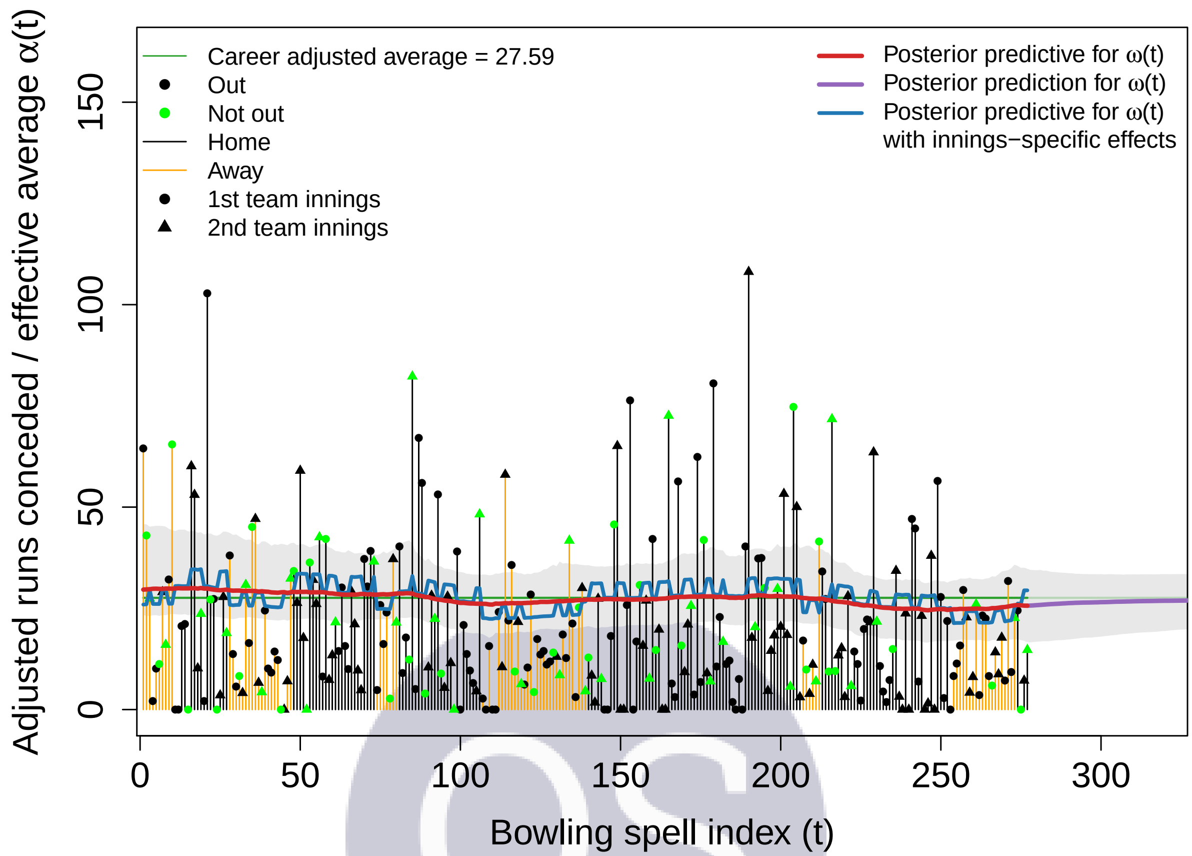


Figure 3.5. *Test match bowling career trajectory (the posterior median of $\alpha(t)$) for Neil Wagner, in units of an adjusted bowling average. The 95% credible interval is also provided (shaded region). A prediction for the adjusted effective bowling average, $\alpha(t)$, is also provided for 50 bowling spells into the future (purple).*

To demonstrate how individual player bowling career trajectories can be analysed, the bowling career trajectory for Neil Wagner is presented in Figures 3.4 and 3.5. Figure 3.4 presents the results in units of a standardised bowling average, while Figure 3.5 is presented in terms of an adjusted bowling average. A significant difference between the batting and bowling career trajectories, which is made apparent in Figures 3.4 and 3.5, is the estimated uncertainties of the effective average, which are comparatively smaller for bowlers. This observation holds true across the majority of players analysed, suggesting that after adjusting for the quality of batsmen bowled to, underlying bowling ability tends to vary less during a career than batting ability. It is also plausible that perhaps bowling is a more difficult skill to noticeably improve on in terms of on-field results, when measuring performance purely on runs conceded and wickets taken.

An important feature of Wagner's bowling career trajectory is the sustained improvement observed since his Test debut, illustrated by the decreasing trend in the posterior predictive

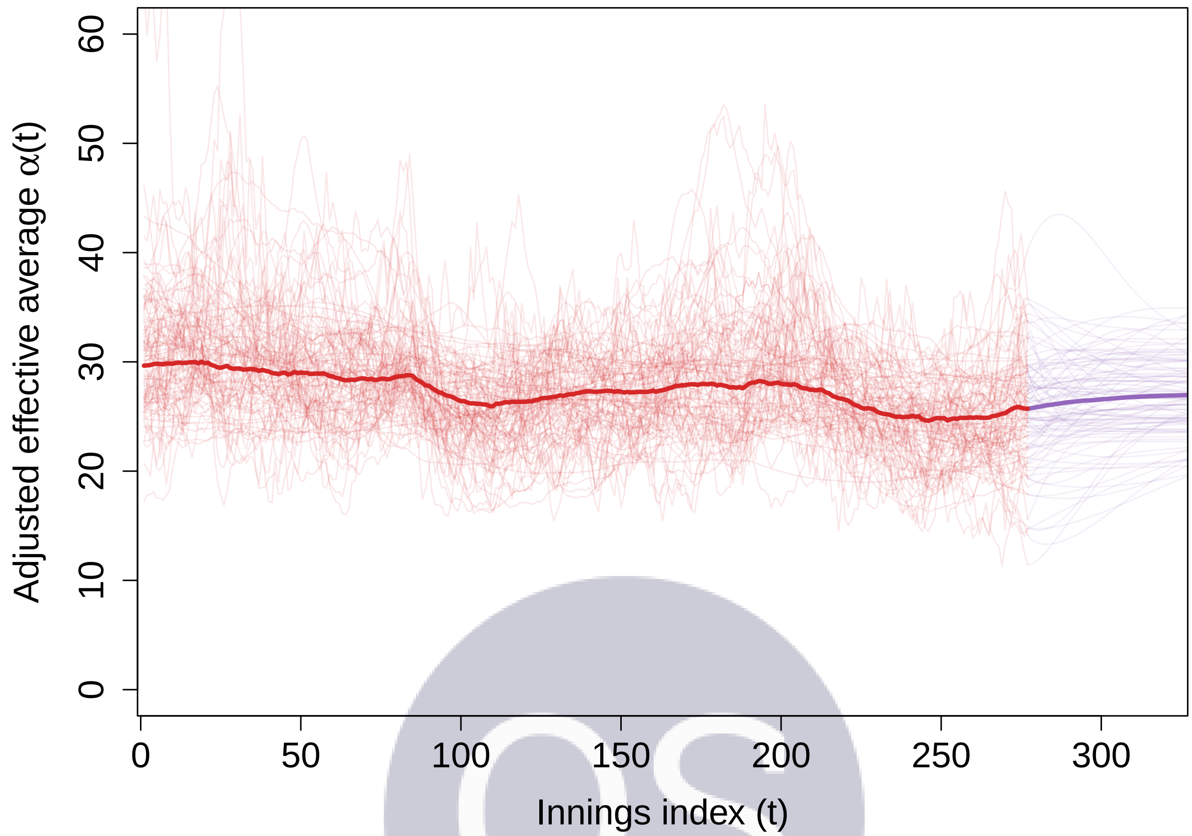


Figure 3.6. *A subset of 1,000 posterior samples for $\alpha(t)$, the expected adjusted bowling average given the parameters, for Neil Wagner. The purple lines represent predictions for $\alpha(t)$ for 50 bowling spells into the future. The posterior predictive estimate for $\alpha(t)$ is overlaid to illustrate the moderate amount of uncertainty in the estimates.*

estimate for his underlying bowling ability, represented by the red lines in Figures 3.4 and 3.5. The posterior predictive estimates for $\omega(t)$ and $\alpha(t)$ that include the innings and venue-specific information for each performance, i_t and v_t, are also provided (blue). As the posterior distributions for $\omega(t)$ and $\alpha(t)$ are not always symmetric and can have heavy tails, the posterior predictive estimates are computed as the posterior median, rather than the posterior mean.

To illustrate the range of plausible career trajectories that can be fitted to Wagner's career data, a subset of 1,000 posterior samples for $\alpha(t)$ are presented in Figure 3.6. The posterior predictive estimate for $\alpha(t)$ is overlaid, again illustrating the gradual improvement seen in Wagner's bowling performances over time. Given the relatively noisy nature of the data, plausible career trajectories include those that exhibit fluctuations in ability as a function of both short and long-term form. Posterior summaries for each of the model parameters are provided in Tables 3.7 and 3.8, and are discussed in more detail below.

Individual-specific effects

The posterior distributions for the each of Gaussian process parameters, $\{\lambda, \sigma, \ell, \gamma\}$, which govern the shape of Wagner's career trajectory and the relative impact of short and long-term form, are presented in Figure 3.7 and are summarised in Table 3.7. Given Wagner's relative success in Test matches to date, the data have been reasonably informative in regards to the mean value parameter, λ, and therefore the set of underlying bowling spell abilities, $\{w_t\}$. The 95% posterior credible interval for λ does not include a value of 1, indicating there is evidence to suggest Wagner is a bowler who is expected, on average, to concede fewer runs than an opposition batsman's effective batting average, $\mu(x, t)$, for each wicket taken.

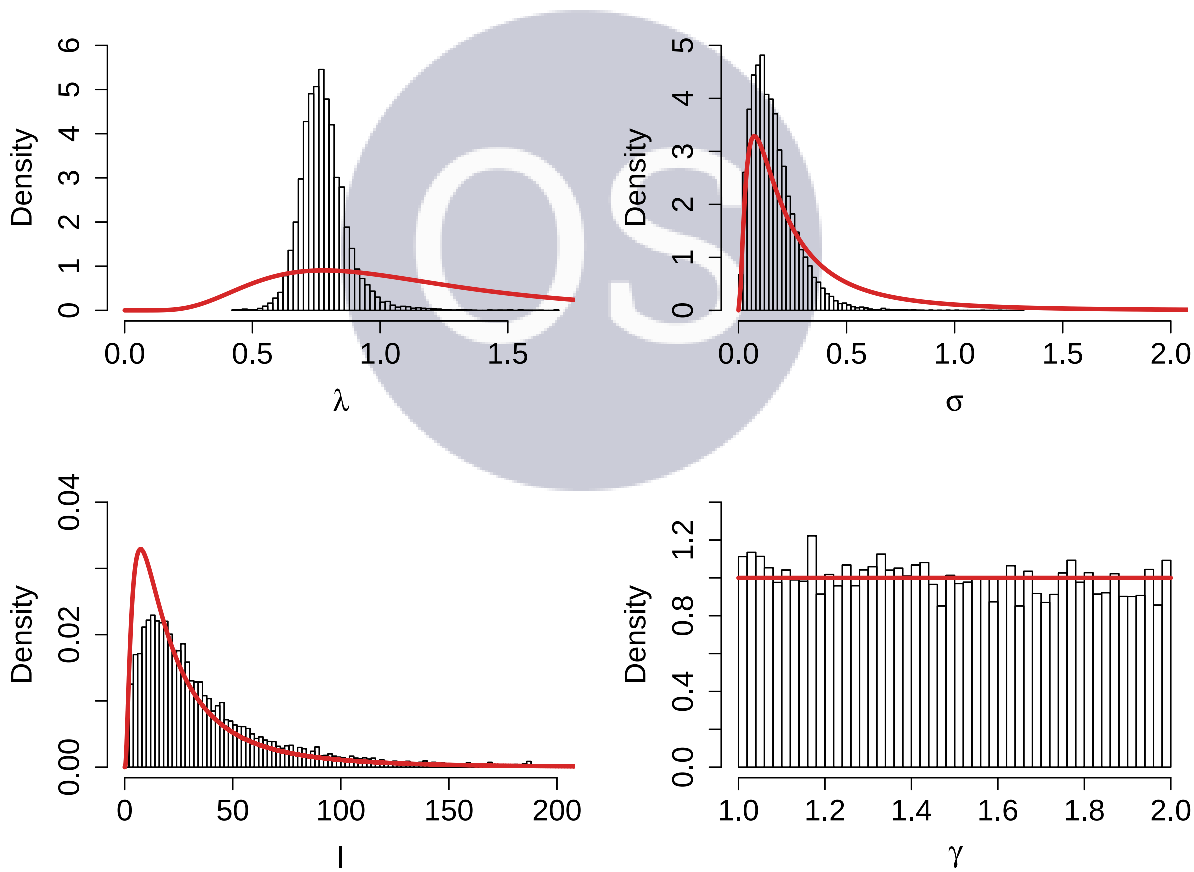


Figure 3.7. *Posterior distributions for each of the Gaussian process parameters, λ, σ, ℓ and γ. Red lines indicate the respective prior distributions. Note that λ is expressed here in units of standardised runs. It appears as though Wagner's data have a limited impact in modifying the prior distributions for both ℓ and γ, suggesting the model is unable to identify the smoothness of the best fitting underlying career trajectories, due to the noisy data.*

Table 3.7. *Posterior parameter summaries for the set of Gaussian process parameters for Neil Wagner, including the 68% and 95% credible intervals.*

Parameter	Mean	Median	68% C.I.	95% C.I.
λ	0.78	0.77	(0.70, 0.87)	(0.62, 0.99)
σ	0.16	0.14	(0.06, 0.26)	(0.03, 0.42)
ℓ	42.2	27.4	(10.8, 69.1)	(3.8, 179.2)
γ	1.49	1.48	(1.15, 1.83)	(1.02, 1.98)

With respect to σ, there is little posterior weight at values on or close to zero, providing some evidence to suggest that Wagner's bowling ability has not remained constant throughout his career to date. Little appears to have been learnt about each of the parameters controlling the smoothness of the Gaussian process, ℓ and γ, again indicating the model is unable to distinguish whether smooth or more jagged career trajectories provide the most appropriate fit to the data.

Innings and venue-specific effects

In Figures 3.4 and 3.5 the posterior predictive distribution for the standardised and adjusted effective bowling averages, $\omega(t)$ and $\alpha(t)$, which include the innings and venue-specific effects for each observation, are shown in blue. The results suggest that after adjusting for opposition batting strength, Wagner has tended to perform better at venues outside of New Zealand, as indicated by the comparatively lower estimates for away observations (orange bars). Likewise, Wagner appears to concede fewer runs per wicket taken when bowling against players who are batting in their team's first innings of a match. The posterior distributions for the innings and venue-specific effects, ϕ and ψ, are presented in Figure 3.8 and support these results.

The data have been informative with respect to both parameters, with the posterior parameter information summarised in Table 3.8, allowing for the magnitude of these effects to be quantified. The point estimate for ψ, suggests that after adjusting for the relative strengths of batsmen bowled to, Wagner concedes 8% more runs per wicket taken, in matches played in New Zealand, compared with matches played at a neutral venue. It is possible to compare the effect of bowling at a home venue versus an away venue, by squaring the estimates for ψ in Table 3.8, giving a value of $\psi^2 = 1.17$, suggesting that Wagner concedes an average of 17% more runs per wicket taken, in bowling performances that take place in New Zealand. In a similar manner, it is possible to compare Wagner's bowling performances against players who are batting in their team's first and second innings of a match by squaring the point estimate for ψ, giving an estimate of $\phi^2 = 0.89$. That is, on average Wagner concedes 11% fewer runs per wicket taken, when bowling against players batting in their team's first innings of a match.

Table 3.8. *Posterior parameter summaries for the innings and venue-specific effect parameters for Neil Wagner, including the 68% and 95% credible intervals.*

Parameter	**Mean**	**Median**	**68% C.I.**	**95% C.I.**
ϕ	0.94	0.94	(0.88, 1.01)	(0.82, 1.07)
ψ	1.08	1.08	(1.00, 1.16)	(0.93, 1.25)

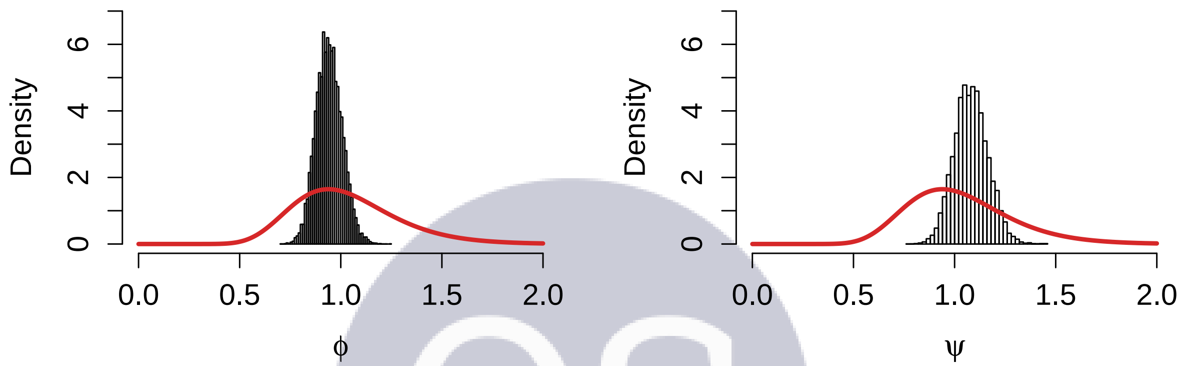


Figure 3.8. *Posterior distributions for the innings-specific parameters, ϕ and ψ. Red lines indicate the prior distribution. The data appear to provide some evidence to suggest that Wagner performs better at away venues, in the opposition batting team's first innings of a match.*

Intuitively, each of these results seem somewhat surprising on face value. As seen in Chapter 2, batsmen tend to perform at their peak when batting in home conditions, in their team's first innings of a match, yet, Wagner tends to excel in these exact conditions. Furthermore, New Zealand is widely regarded as country where playing conditions tend to suit pace bowlers, such as Wagner, more so than many other countries globally, which makes the point estimate for ψ particularly intriguing. However, digging a little deeper into the data reveals several plausible explanations for each of these findings.

Firstly, while New Zealand is generally considered a pace bowling haven, almost all of Wagner's bowling performances away from New Zealand, have taken place in either England, Australia or South Africa — countries where conditions also tend to suit pace bowling. Only four of Wagner's 48 Test appearances have been in sub-continental countries, where pace bowling can be more arduous and less rewarding. Secondly, as noted earlier in Section 3.3.2, modern New Zealand pitches have become somewhat known for offering some assistance to bowlers early in a match, before flattening out and becoming easier to bat on as a match progresses, before starting to deteriorate on day four or five. This does not align with the trend that is often seen in the majority of pitches around the world, where batting conditions are typically at their best on day one, and deteriorate over time. With this in mind and noting that 31 of Wagner's 48 Test matches have been played on New Zealand soil, it is perhaps unsurprising to see that Wagner has enjoyed more success when bowling against opposition players who are batting in their team's first innings of a match.

Quantifying bowling career progression

As observed in Figures 3.4 and 3.5, Neil Wagner's underlying bowling ability appears to have steadily improved over the course of his playing career to date. The posterior predictive estimates for $\omega(t)$ and $\alpha(t)$ provide an indication of a player's ability during any bowling spell of their career, allowing for an estimation of when a player's bowling ability was at its peak and at when it was at its worst. Additionally, the bowling career trajectory model has the capability to forecast a player's future ability, including an estimate for their next career bowling spell, $\alpha(T+1)$. The default prediction is made assuming a neutral venue and it is unknown what innings a player is bowling in, however, the innings and venue-specific information can be factored into the estimate if known.

The posterior distributions for Wagner's worst and best estimated career bowling abilities, in units of an adjusted bowling average, are provided in Figure 3.9 and are summarised in Table 3.9. Corresponding posterior distributions specifying when Wagner experienced each of these points, in terms of a career bowling spell index, are also provided in Figure 3.9.

Table 3.9. *Posterior point estimates for Neil Wagner's worst and best career bowling abilities, in units of an adjusted bowling average. The 68% and 95% credible intervals are provided, as well as a prediction of ability for his next career bowling spell,* $\alpha(T+1)$.

	Point estimate	**68% C.I.**	**95% C.I.**
Career worst $\alpha(t)$	35.2	(30.3, 43.3)	(27.0, 55.5)
Career best $\alpha(t)$	21.3	(17.3, 24.9)	(13.5, 27.9)
$\alpha(T+1)$	25.8	(21.5, 29.8)	(16.9, 34.6)

At his peak, Wagner is estimated to have had a bowling ability equivalent to an adjusted bowling average of 21.3, placing him among the best bowlers in the world at the time. As suggested in Figure 3.9, it is highly probable that Wagner is either at the peak of his career at present, or experienced his peak very recently. This result is unsurprising, given Wagner is presently ranked as the third best bowler in the world, based on the ICC bowling ratings. Conversely, Figure 3.9 suggests that Wagner's underlying bowling ability was at its lowest at some point near the beginning of his career, providing further evidence to support the notion that Wagner has been consistently improving since appearing on the Test scene.

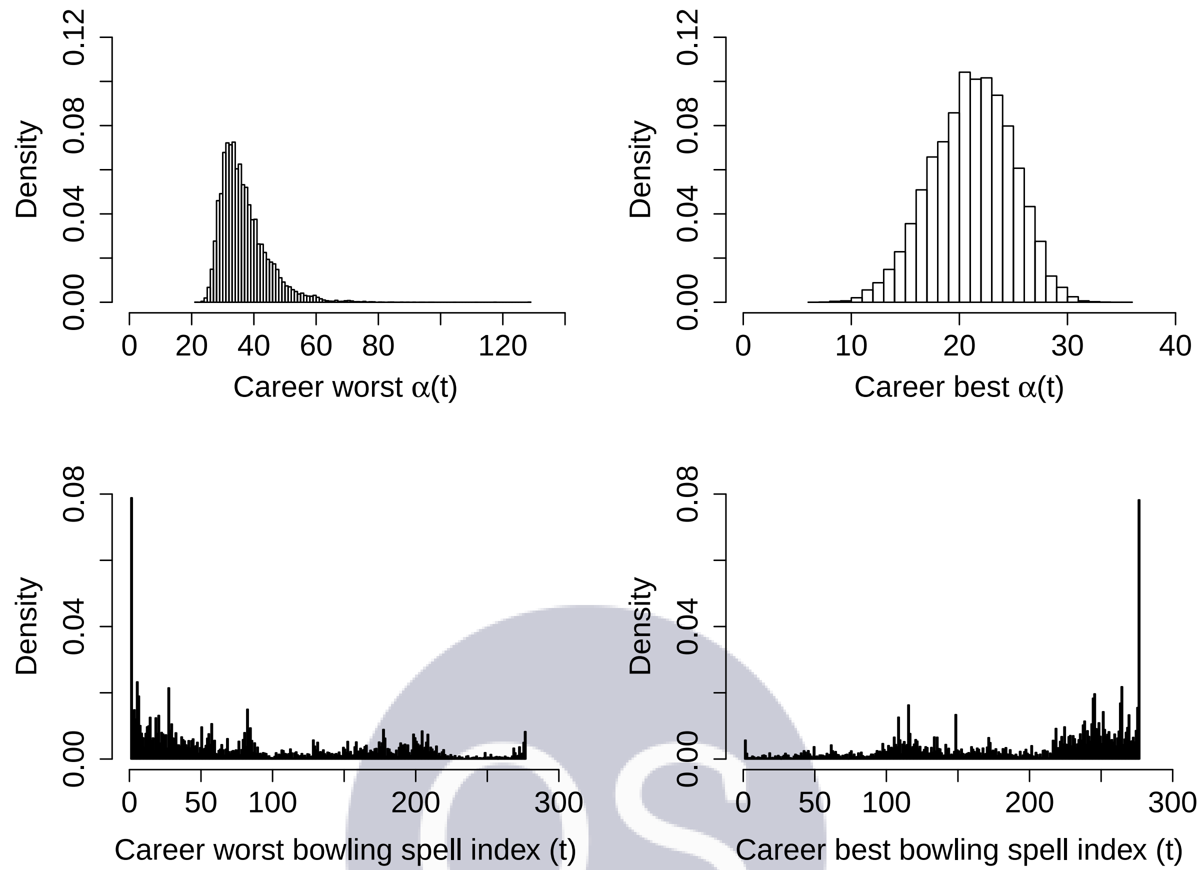


Figure 3.9. *Posterior distributions for Neil Wagner's worst and best career bowling abilities, $\alpha(t)$, in units of an adjusted bowling average. The estimated bowling spell index, t, for when Wagner experienced each of these points of his career are also provided.*

3.4.2 Hierarchical analysis of bowlers

As with the batting model presented in Chapter 2, a hierarchical analysis was performed to gain a deeper understanding of the typical Test match bowling career trajectory. Of particular interest was to determine whether short and long-term form appears to affect bowlers any differently to batsmen, which could have provide some insight as to how selectors should consider a player's recent bowling performances with a view to the future. The hierarchical analysis was performed across the set of Gaussian process parameters, $\{\lambda, \sigma, \ell, \gamma\}$, as well as the innings and venue-specific parameters, $\{\phi, \psi\}$.

For each model parameter, an underlying distribution is assumed and a set of relevant hyperparameters, $\boldsymbol{\eta}$, is introduced, allowing for the quantification of typical values each parameter is clustered near, without having to analyse the data jointly. The posterior estimates for set of model parameters $\{\lambda, \sigma, \ell, \gamma, \phi, \psi\}$, are obtained for each player and post-processed jointly using MCMC (Hastings, 1970), to construct what the hierarchical model would have produced, had the full data set of all 522 players been analysed together.

As with the hierarchical analysis performed on batting career trajectories in Chapter 2, the set of parameters, $\{\lambda,\ \sigma,\ \ell,\ \phi,\ \psi\}$, are each assumed to be well approximated by a log-normal distribution, while γ is assumed to loosely follow a Normal distribution, truncated at [1, 2]. Conditional on the set of hyperparameters, $\boldsymbol{\eta} = \{\mu_\lambda, \xi_\lambda, \mu_\ell, \xi_\ell, \mu_\sigma, \xi_\sigma, \mu_\gamma, \xi_\gamma, \mu_\phi, \xi_\phi, \mu_\psi, \xi_\psi\}$, the hierarchical model structure is as follows

$$\begin{aligned}
\log(\lambda) &\sim \text{Normal}(\log(\mu_\lambda), \xi_\lambda^2) \\
\log(\ell) &\sim \text{Normal}(\log(\mu_\ell), \xi_\ell^2) \\
\log(\sigma) &\sim \text{Normal}(\log(\mu_\sigma), \xi_\sigma^2) \\
\gamma &\sim \text{Normal}_{[1,2]}(\mu_\gamma, \xi_\gamma^2) \\
\log(\phi) &\sim \text{Normal}(\log(\mu_\phi), \xi_\phi^2) \\
\log(\psi) &\sim \text{Normal}(\log(\mu_\psi), \xi_\psi^2)
\end{aligned} \tag{3.14}$$

The hyperparameters, $\boldsymbol{\eta} = \{\mu_\lambda, \xi_\lambda, \mu_\ell, \xi_\ell, \mu_\sigma, \xi_\sigma, \mu_\gamma, \xi_\gamma, \mu_\phi, \xi_\phi, \mu_\psi, \xi_\psi\}$, are assigned prior distributions as per Equation 3.15 and sufficiently encompass the posterior parameter space that contains the bulk of the posterior mass. The marginal posterior distribution for each parameter can then be obtained via MCMC, using Equation 2.19.

$$\begin{aligned}
\mu_\ell &\sim \text{Uniform}(0, 50) \\
\mu_\lambda &\sim \text{Uniform}(0, 5) \\
\mu_\sigma, \mu_\gamma, \mu_\phi, \mu_\psi &\sim \text{Uniform}(0, 2) \\
\xi_\lambda, \xi_\ell, \xi_\sigma, \xi_\phi, \xi_\psi &\sim \text{Uniform}(0, 2) \\
\xi_\gamma &\sim \text{Uniform}(0, 1)
\end{aligned} \tag{3.15}$$

The MCMC algorithm was run for 100,000 iterations for each of the model parameters. The joint posterior distributions for the set of hyperparameters defining the Gaussian process parameters, $\{\mu_\lambda, \xi_\lambda, \mu_\ell, \xi_\ell, \mu_\sigma, \xi_\sigma, \mu_\gamma, \xi_\gamma\}$, are presented in Figure 3.10, while the joint posterior distributions associated with the hyperparameters defining the innings and venue-specific effects, $\{\mu_\phi, \xi_\phi, \mu_\psi, \xi_\psi\}$, are shown in Figure 3.11. As a truncated normal distribution is assumed for γ, the starting point for the algorithm was simply selected by taking the central value of each of the uniform hyperpriors defined in Equation 3.15, corresponding to starting values of: $\mu_\gamma = 1.5$, $\xi_\gamma = 0.5$. Once again, no burn-in period was deemed necessary, as each of the selected starting points for the MCMC algorithm appear to be typical of the corresponding joint posterior distribution (Meyn & Tweedie, 1993).

For the model parameters related to the Gaussian process prior for the set of $\{w_t\}$ terms, the

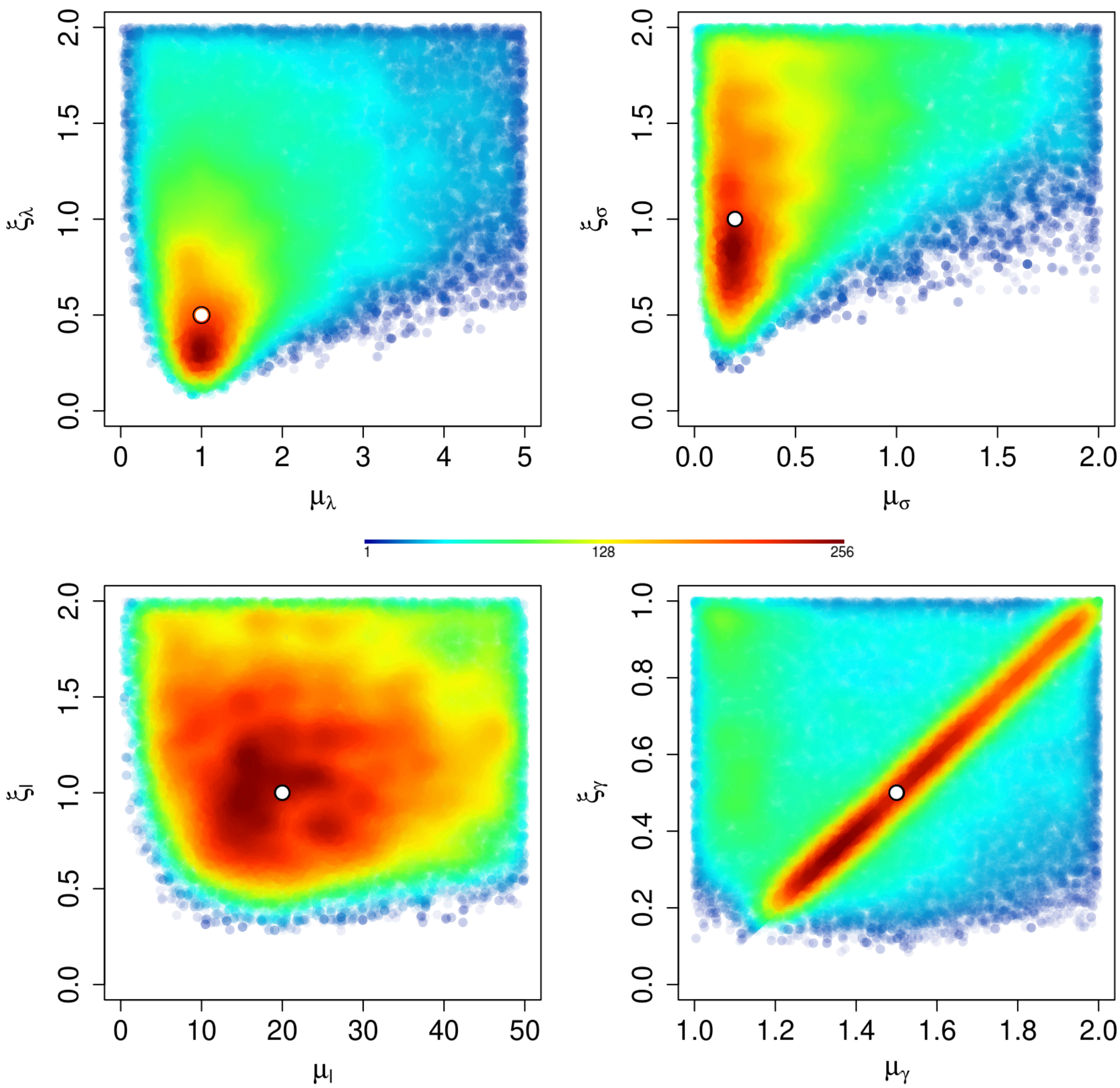


Figure 3.10. *Joint posterior distributions for the set of hyperparameters* $\{\mu_\lambda, \xi_\lambda, \mu_\ell, \xi_\ell, \mu_\sigma, \xi_\sigma, \mu_\gamma, \xi_\gamma\}$*, shown across the uniform prior parameter space. Red indicates areas of high density, while dark blue indicates areas of low density. The scale indicates that the darkest red areas are 256 times more dense than the darkest blue areas. The white circle indicates the starting point of the MCMC algorithm.*

hierarchical analysis was relatively informative with respect to λ and a little informative for σ. The areas of high density for λ confirm what was learnt from the summary data; standardised career bowling averages are typically clustered around values close to 1.0. Less variation is seen across values for μ_λ when performing the hierarchical analysis in a bowling context, compared with the analysis in Chapter 2. This observation provides some statistical evidence to support the assumption that batting ability tends to vary more than bowling ability, between players. With regards to σ, Figure 3.10 suggests that the hyperparameter μ_σ has little posterior weight assigned to values where $\mu_\sigma \to 0$, indicating that the average player will likely exhibit some form of variation in bowling ability, over the course of their Test career. Generally speaking the analysis has been less informative with respect to the parameters governing the effect of short and long-term form on a player's estimate career trajectory, ℓ and γ.

However, one result of potential interest with respect to ℓ, is that compared with the hierarchical analysis performed in the context of batting in Chapter 3, there is far more posterior mass where $\mu_\ell < 20$. This may suggest that in general, bowling performances may be more likely to be influenced by recent performances and effects due to short-term form, than batting performances. As such, there may be a valid argument that recent performances may be more indicative of current underlying ability for bowlers than they are for batsmen, which could have real-world implications when it comes to squad and team selection, for upcoming tours and matches.

As with the hierarchical analysis performed in Chapter 2, the posterior parameter space for γ tends to have higher density towards the centre of the parameter space for μ_γ, where $\mu_\gamma \to 1.5$. This result suggests that the proposed model tends to assign less posterior weight to highly erratic bowling career trajectories (implied by values of $\gamma \to 1$), and very smooth trajectories (implied by values of $\gamma \to 2$). As such, the Gaussian processes that are used to estimate bowling career trajectories tend to avoid following functions that are generated from a Matérn$_{\frac{1}{2}}$ or squared-exponential covariance function, as seen in Figure 2.4.

In Chapter 2, the hierarchical analysis for the innings and venue-specific effect parameters, ϕ and ψ, was relatively informative, with the bulk of the posterior mass for each of the hyperparameters μ_ϕ and μ_ψ, concentrated near values greater than 1. This provided further support for the widely accepted assumption, that most players score more runs when batting at a home venue, in their team's first innings of a match. In the context of bowling, the results of the hierarchical analysis are less clear-cut.

As shown in Figure 3.11, typical values for hyperparameters μ_ϕ and μ_ψ appear to be clustered around values of 1. Therefore, after the strengths of batsmen bowled to have been accounted for, the hierarchical analysis has not provided any definitive evidence to suggest that the typical Test bowler will tend to perform better when bowling at a home venue, against players who

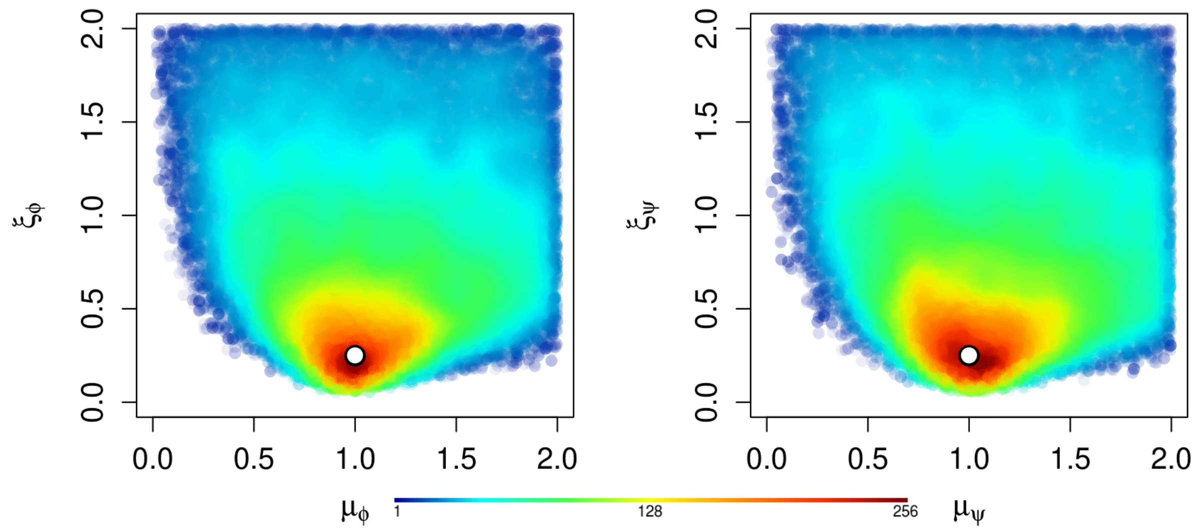


Figure 3.11. *Joint posterior distributions for the set of hyperparameters* $\{\mu_\phi, \xi_\phi, \mu_\psi, \xi_\psi\}$*, shown across the uniform prior parameter space. Red indicates areas of high density, while dark blue indicates areas of low density. The scale indicates that the darkest red areas are 256 times more dense than the darkest blue areas. The white circle indicates the starting point of the MCMC algorithm.*

are batting in their team's first innings of a match. Rather, observed differences in a player's career bowling averages between innings can generally be explained by the fact that batting usually becomes more difficult as a match progresses, which is accounted for in the data by the individual-specific estimates for the effective batting average, $\mu(x, t)$. A similar conclusion can be reached when considering the venue-specific effects on batting performances; players tend to find batting away from home a more difficult prospect. An outcome of some interest in regards to the venue-specific effect, ψ, can be observed when conditioning on player bowling type and considering the context in which pace and spin bowlers tend to succeed.

Spin bowlers and venue-specific effects: a case study

As discussed in Section 3.3.2, different countries around the world are home to conditions that can be more favourable to certain bowling types. Pitches and the amount of assistance they may offer both pace and spin bowlers can differ greatly, as a result of local climate, varieties of soil available and preparation techniques. Test playing nations located in the sub-continent such as Bangladesh, India, Pakistan, and Sri Lanka are renowned for their spin-friendly conditions, not only due to the pitches produced — which often favour spin bowling — but also as a function of the high average heat and humidity level, which can make it difficult for pace bowlers to bowl with sustained intensity, over extended periods of time. In contrast, pace bowlers tend to receive more assistance from the overhead conditions and pitches typically prepared in countries with more temperate climates, such as England and New Zealand. With this in mind, estimates for the venue-specific effect, ψ, may not depend so much on a player's affinity with their local, home conditions, rather their bowling type and country of origin.

To explore this point further, posterior estimates for ψ are presented in Table 3.10, for all spin bowlers in the Cricsheet data set who have taken 100 or more Test wickets. Of the 16 players satisfying this criteria, 10 are from the sub-continent (four from India and two from each of Bangladesh, Pakistan and Sri Lanka) and six are not (two from England and one from each of Australia, New Zealand, South Africa and West Indies).

Table 3.10. *Posterior point estimates for the venue-specific effect, ψ, for each of the spin bowlers in the Cricsheet data set who have taken 100 or more Test wickets. The 68% and 95% credible intervals are also provided, illustrating the preference of spin bowlers from the sub-continent to play at home venues, and those from outside the sub-continent to prefer playing away from home.*

Player		**Wickets**	ψ	**68% C.I.**	**95% C.I.**
R. Herath[2]	(SL)	397	0.91	(0.86, 0.96)	(0.82, 0.99)
N. Lyon	(AUS)	390	1.07	(1.02, 1.12)	(0.97, 1.19)
R. Ashwin	(IND)	365	0.91	(0.86, 0.97)	(0.81, 1.02)
G. Swann	(ENG)	255	1.02	(0.96, 1.09)	(0.90, 1.16)
Y. Shah	(PAK)	224	0.94	(0.87, 1.01)	(0.81, 1.10)
R. Jadeja	(IND)	213	0.87	(0.80, 0.94)	(0.74, 1.02)
S. al Hasan[2]	(BAN)	193	1.01	(0.93, 1.09)	(0.86, 1.16)
M. Ali	(ENG)	181	0.97	(0.89, 1.04)	(0.83, 1.12)
S. Ajmal	(PAK)	178	0.98	(0.90, 1.05)	(0.83, 1.13)
H. Singh[2]	(IND)	166	0.97	(0.90, 1.05)	(0.83, 1.13)
D. Perera	(SL)	156	0.88	(0.80, 0.96)	(0.73, 1.05)
D. Bishoo	(WI)	117	1.10	(1.00, 1.20)	(0.91, 1.32)
T. Islam	(BAN)	114	0.94	(0.85, 1.04)	(0.77, 1.13)
P. Ojha	(IND)	113	0.88	(0.76, 1.00)	(0.67, 1.14)
D. Vettori[2]	(NZ)	111	1.16	(1.05, 1.27)	(0.96, 1.40)
K. Maharaj	(SA)	110	1.13	(1.03, 1.24)	(0.94, 1.37)

For spin bowlers from the sub-continent, the point estimate for ψ is less than 1 for all 10

[2]The Test careers of Rangana Herath, Shakib al Hasan, Harbhajan Singh and Daniel Vettori began prior to 2008. Consequently, the Cricsheet data set does not contain all career bowling data for these players.

players, implying that all of these players have enjoyed more success when bowling at venues in their native country, compared with venues away from home. Furthermore, for eight of the 10 sub-continental players, the 68% credible interval provides evidence to support the presence of a venue-specific effect, although the 95% credible interval only provides support of such an effect for Sri Lankan Rangana Herath. Conversely, the posterior point estimate for ψ is greater than 1 for five of the six spin bowlers from outside the sub-continent, implying that the majority of these players tend to concede fewer runs per wicket when bowling outside of their respective home countries. Of these players, only Englishman Moeen Ali appears to prefer bowling on his home soil.

The results presented in Table 3.10 suggest that re-running the hierarchical analysis for the venue-specific effect, ψ, on two separate data sets of spin bowlers, conditional on each player's country of origin, would yield two different conclusions. Firstly, an analysis that only includes the data of spin bowlers from the sub-continent would likely indicate such bowlers tend to prefer bowling in home conditions. Secondly, running the analysis on the complementary set of spin bowlers originating from outside the sub-continent, would likely suggest that these players tend to enjoy more success when bowling at venues outside of their home country. It is entirely plausible, that similarly contrasting results would be obtained if separate hierarchical analyses were performed on unique groups of pace bowlers, where the inclusion of a player's data in an analysis was conditional on a player's country of origin.

3.4.3 Comparison of bowling career trajectories

It is possible to compare the bowling careers of multiple players by comparing their career trajectories, which are estimated using the posterior predictive distribution for the adjusted effective bowling average, $\alpha(t)$. The bowling career trajectories for the current top five Test bowlers, as ranked by the bowling career trajectory model, are provided in Figure 3.12. Varying levels of consistency and improvement in performance are shown between the careers of these players, illustrating the uncertainty and difficulty in estimating an individual's underlying bowling ability.

As suggested in Figure 3.12, the bowling performances of Indian pace bowler Jasprit Bumrah appear to be highly correlated with observations that are in close proximity of one another in the input space, t. That is to say, it appears as though Bumrah's career to date resembles that of a player whose underlying bowling ability is heavily affected by short-term form. However, it is likely that the relatively small sample size of Bumrah's career data is a contributing factor to the erratic function fitted to estimate his underlying bowling ability. One would expect Bumrah's career trajectory to become more smooth over time, unless the effect of short-term form truly has an impact of such a large magnitude on his bowling performances. While it is unlikely

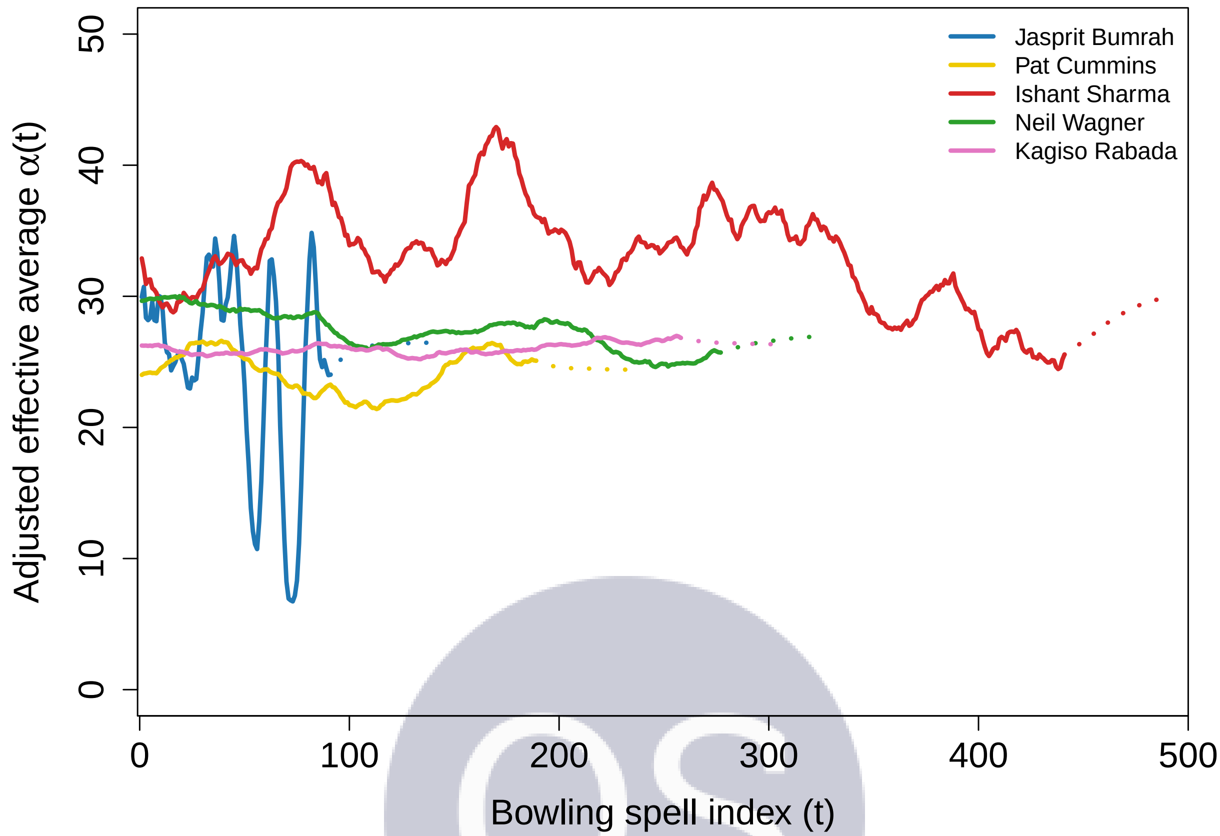


Figure 3.12. *Test match bowling career trajectories for current top five Test bowlers ranked by the bowling career trajectory model: Jasprit Bumrah, Pat Cummins, Ishant Sharma, Neil Wagner, and Kagiso Rabada. Predictions of bowling ability for the next 50 bowling spells are also included (dotted).*

that Bumrah's underlying bowling ability fluctuates as wildly as suggested by Figure 3.12, the results may at least suggest that Bumrah is the type of bowler who can experience moments of being *in the zone*, where he is unplayable during certain periods of a match, comparable to the concept of the hot hand in basketball, where players exhibit shooting performances where they are simply unable to miss. On the opposite end of the spectrum from Bumrah, exists the career trajectory for South African pace bowler Kagiso Rabada, whose career bowling performances have been highly consistent, but provide little evidence to suggest he has displayed any level of improvement or deterioration in bowling ability, over the course of his career to date.

In terms of consistency, the career trajectories estimated for Pat Cummins, Ishant Sharma and Neil Wagner, lie somewhere between that of Bumrah and Rabada. Like Bumrah — although on a less dramatic scale — Sharma's career performances appear to be somewhat related to one another in the short-term and it is possible to identify several points of Sharma's career where he was struggling with the ball: one period early in his career, approximately between bowling spells 65 and 95; and a second shortly after, approximately between spells 160 and 190. Given Sharma made his Test debut for India at the age of 19 and has been a mainstay in the side since 2008, it is unsurprising to see his career trajectory exhibit behaviour akin to the concept of finding your feet, discussed in Chapter 2, whereby he has taken some time to reach his peak bowling ability. Given Sharma has had over a decade to hone his craft to the point he has become the second highest Test wicket taker of all-time among Indian pace bowlers (sixth highest across all Indian bowlers), his overall improvement in recent years should not come as a surprise.

On the other hand, as discussed in Section 3.4.1, Neil Wagner has exhibited little else than constant improvement over his career to date and has certainly been able to find more consistency in his performances than Sharma. Wagner's bowling career trajectory exists as a great example of a player who made their debut towards the second half of their professional career, with a vast amount of domestic experience, and therefore has required little time to adjust to the demands of international Test cricket. Born and raised in South Africa, Wagner only moved to New Zealand in 2008 at the age of 22 and would have featured in national side sooner than 2012, but for the ICC's governing laws requiring players to reside in a country for four years before being eligible for selection. In the meantime, the 34 year old Wagner continues to improve and has shown few signs of slowing down.

3.4.4 Player bowling rankings

The model output can also be used to rank the top Test bowlers in the world at present by obtaining estimates for the adjusted effective bowling average, $\alpha(T+1)$, for all players, where $T+1$ is the bowling spell index for each player's next career bowling spell. The current[3] top 20 Test bowlers in the world are presented in Table 3.11, ranked by their expected adjusted bowling average, $\alpha(T+1)$, and provides an alternative means of quantifying and ranking the bowling abilities of cricket players, while maintaining a meaningful interpretation. To be eligible for ranking, each player must have participated in a Test match since the 1st January 2019 and have taken a minimum of 30 Test wickets, with an up-to-date list of the top 50 Test match bowlers maintained at `www.oliverstevenson.co.nz/#research`. It is worth noting that the predictions for $\alpha(T+1)$ assume a neutral venue and it is unknown whether the player is bowling against players who are batting in their team's first or second innings of a match. Corresponding ICC bowling ratings and world rankings are also provided for comparison. As with the batting rankings presented in Chapter 2, the bowling career trajectory model rankings and ICC rankings

[3] as of 1st December 2020

are relatively similar, although there are several notable differences. Respective career trajectory model and ICC rankings are presented in Figure 3.13, highlighting where the two methods are in agreement and where there is a lack of consensus.

Table 3.11. *Current (as of 1^{st} December 2020) top 20 Test match bowlers, ranked by expected adjusted average in their next career bowling spell, $\alpha(T+1)$, including the 68% credible interval. ICC Test bowling ratings and world rankings (#) are shown for comparison.*

Rank	Player		Wickets	Career bowling average	Career adjusted average	$\bar{\nu}_{player}$	$\alpha(T+1)$	ICC rating	(#)
1.	J .Bumrah	(IND)	68	20.3	23.0	31.4	24.2 (12.3, 35.7)	779	(9)
2.	P. Cummins	(AUS)	143	21.8	24.3	32.0	25.1 (19.2, 27.0)	904	(1)
3.	I. Sharma	(IND)	291	32.4	32.7	35.2	25.7 (19.6, 31.5)	729	(17)
4.	N. Wagner	(NZ)	206	26.6	27.6	34.3	25.8 (20.7, 28.6)	843	(3)
5.	K. Rabada	(SA)	197	22.9	25.0	32.6	26.8 (22.0, 27,9)	802	(6)
6.	J. Anderson	(ENG)	538	25.4	26.5	34.1	27.2 (24.2, 28.0)	781	(8)
7.	T. Southee	(NZ)	284	29.0	30.3	34.1	27.7 (21.0, 31.5)	812	(4)
8.	J. Hazlewood	(AUS)	195	26.2	28.1	33.2	28.0 (22.6, 29.8)	769	(11)
9.	J. Pattinson	(AUS)	81	26.2	26.9	34.6	28.4 (23.3, 33.4)	333	(43)
10.	S. Broad	(ENG)	513	27.5	29.9	32.7	28.5 (21.7, 30.3)	845	(2)
11.	M. Shami	(IND)	180	27.4	29.4	33.1	28.6 (23.2, 30.5)	749	(13=)
12.	R. Ashwin	(IND)	365	25.4	27.7	32.7	29.6 (24.6, 30.6)	756	(12)
13.	M. Abbas	(PAK)	80	21.7	25.3	30.5	29.7 (20.4, 34.9)	749	(13=)
14.	T. Boult	(NZ)	267	27.6	28.7	34.3	29.7 (25.7, 32.4)	770	(10)
15.	M. Starc	(AUS)	244	27.0	29.1	32.9	29.8 (24.1, 31.6)	797	(7)
16.	B. Stokes	(ENG)	158	31.4	31.6	35.4	30.1 (24.8, 34.6)	587	(24)
17.	C. de Grandhomme	(NZ)	47	31.6	30.7	36.7	31.1 (26.2, 39.9)	476	(32)
18.	S. al Hasan[4]	(BAN)	193	31.1	30.7	36.1	31.2 (27.8, 36.4)	N/A	(N/A)
19.	K. Jarvis	(ZIM)	46	29.4	33.8	31.0	31.3 (22.5, 33.4)	347	(41)
20.	K. Roach	(WI)	196	28.2	29.1	34.4	31.7 (26.6, 36.0)	744	(15)

Firstly, the bowling career trajectory model ranks Indian pace bowler Jasprit Bumrah as the current world number one, while the ICC ratings have him ranked 9^{th}. As shown in Figure 3.12 and discussed in Section 3.5.2, there is some evidence to suggest Bumrah is a player whose underlying bowling ability may be affected by short-term form and recent performances. As he has exhibited a number of strong performances in his most recent matches — his last 20 wickets have come at an average of 18.7 runs apiece — the bowling career trajectory model ranks him highly. A player who appears to be similarly impacted by short-term form, is West Indian pace bowler Jason Holder, whom is ranked 34^{th} by the proposed model, but 5^{th} by the ICC. Holder's last 10 wickets have come at an average of 30.1 runs, while his last five at an average of 51.8, which subsequently sees his ranking under the career trajectory model slip significantly. Comparable, although less dramatic, short-term form effects can be attributed to Indian pace bowler Ishant Sharma's model ranking of 3^{rd}, versus his ICC ranking of 17^{th}. It is interesting to

[4]Shakib al Hasan is currently serving a two-year ban for violating the ICC Anti-Corruption Code, which expires in October 2021. Consequently, he is unrated by the ICC at present.

note that these differences in rankings between methods appear to be due to the bowling career trajectory model weighting recent performances more heavily than the ICC ratings method, the opposite of what was observed for a number of batsmen in Chatper 2.

A second discrepancy between rankings is observed when considering the adjusted career bowling averages of players where there is significant disagreement between the two ranking methods. As noted in previous sections, players who have bowled against a higher quality of batsman, on average, are rewarded by the adjusted average, while those who have tended to bowl against weaker opposition are penalised. In Figure 3.13, all-rounders Ben Stokes and Colin de Grandhomme are among several players who are favoured more heavily by the proposed career trajectory model than the ICC ratings method. The adjusted career bowling averages and individual-specific estimates for the average batsman bowled to, $\bar{\nu}_{player}$, are presented in Table 3.11 for each player. Clearly, both Stokes and de Grandhomme are among those penalised less heavily by the adjusted average, as they appear to have bowled to a higher quality of batsman, on average, than many other players in the top 20. In fact, de Grandhomme is only one of two players (the other being Shakib al Hasan), whose adjusted average is lower than their career bowling average. Conversely, two players favoured by the ICC ratings over the bowling career trajectory model, Englishman Stuart Broad and Australian Mitchell Starc, are two players who, based on estimates for $\bar{\nu}_{Broad}$ and $\bar{\nu}_{Starc}$, appear to have bowled against a lower quality of batsman, on average, over the course of their Test careers. It is therefore possible to speculate that while the ICC rating method does make an attempt to account for opposition strength, the adjustment in some cases may not be enough. Again, due to the closed source nature of the ICC ratings formula, it is unknown how the adjustment for opposition strength is incorporated into their methodology.

A third consideration is to note that the player-specific prediction for current bowling ability, $\alpha(T+1)$, assumes a neutral venue. Under this assumption, Sri Lankan pace bowlers Suranga Lakmal and Lahiru Kumara, are predicted to have adjusted bowling averages of 38.5 and 43.0 respectively, for their next career bowling spells. Both predictions are in excess of their adjusted career bowling averages of 36.8 and 37.8, however, these inferior estimates for $\alpha(T+1)$ do not appear to be related to a poor run of recent form. In fact, a closer look at Lakmal and Kumara's career trajectories suggest they both appear to be improving and have been relatively successful in their most recent performances. Instead, the inflated estimates for $\alpha(T+1)$ are a result of both players performing strongly in home conditions, at Sri Lankan venues, but struggling heavily when bowling outside of their home country. Subsequently, when considering predictions of bowling ability for a neutral venue, both players are penalised by the model, which sees their respective rankings of 39^{th} and 44^{th} under the proposed model, pale in comparison to their ICC rankings of 22^{nd} and 29^{th}.

Finally, the significant differences observed between the bowling career trajectory rankings and ICC rankings for James Pattinson (9^{th} versus 43^{rd}) and Kyle Jarvis (19^{th} versus 41^{st}), can be attributed to the rating decay system implemented by the ICC. Pattinson has consistently delivered strong bowling performances during his nine year tenure in the Australian side, however, his career has been plagued by injury, limiting the number of appearances he has made for the national side. While Pattinson has featured in four Tests across 2019 and 2020, his most recent appearances prior to that were in 2016. The bowling career trajectory model estimates Pattinson to be a very strong bowler, based on his historical career data, while the ICC ratings method is less convinced about his quality, due to a lack of more recent performances. Meanwhile, Jarvis' rating is prone to decay, due to his native country, Zimbabwe, playing far fewer Tests than any other Test playing nation. Since the start of 2015, Zimbabwe have only featured in 13 Test matches, whereas countries such as Australia and India will typically play this many Tests in a calendar year.

Ultimately, both the bowling career trajectory model and ICC rating method attempt to quantify the relative bowling strengths of players around the world. Both methods make adjustments to account for recent performances and opposition batting strength, however, the proposed model's predictions have the advantage of providing a meaningful explanation of ability. While the interpretation of: *expected number of runs conceded before taking next career wicket, adjusted for opposition batsman strength*, is not quite as simple as the model interpretation in the context of batting, it is nevertheless a quantity that can be readily understood by members of the cricketing community and certainly has more meaning than the ICC's rating points.

Moreover, the proposed model is able to provide a more direct means of quantifying differences in bowling ability between players. For example, by computing the posterior probability $P(\alpha_{Cummins}(T_{Cummins}+1) < \alpha_{Wagner}(T_{Wagner}+1)) = 0.525$, the bowling career trajectory model can surmise 'there is a 52.5% chance that Pat Cummins concedes fewer runs than Neil Wagner before taking his next career wicket, assuming equal opposition batsman strength' — a much more meaningful conclusion than 'Pat Cummins is 61 rating points better than Neil Wagner'. Additionally, once known, the innings and venue-specific information can be incorporated into the bowling career trajectory predictions, providing more accurate estimates of upcoming player performance. These kinds of probabilistic statements are especially useful for comparing the abilities of players with similar career bowing averages, but different estimates for $\alpha(T+1)$, allowing for coaches and selectors to gain more insight in regards to the risks and rewards of selecting one player over another.

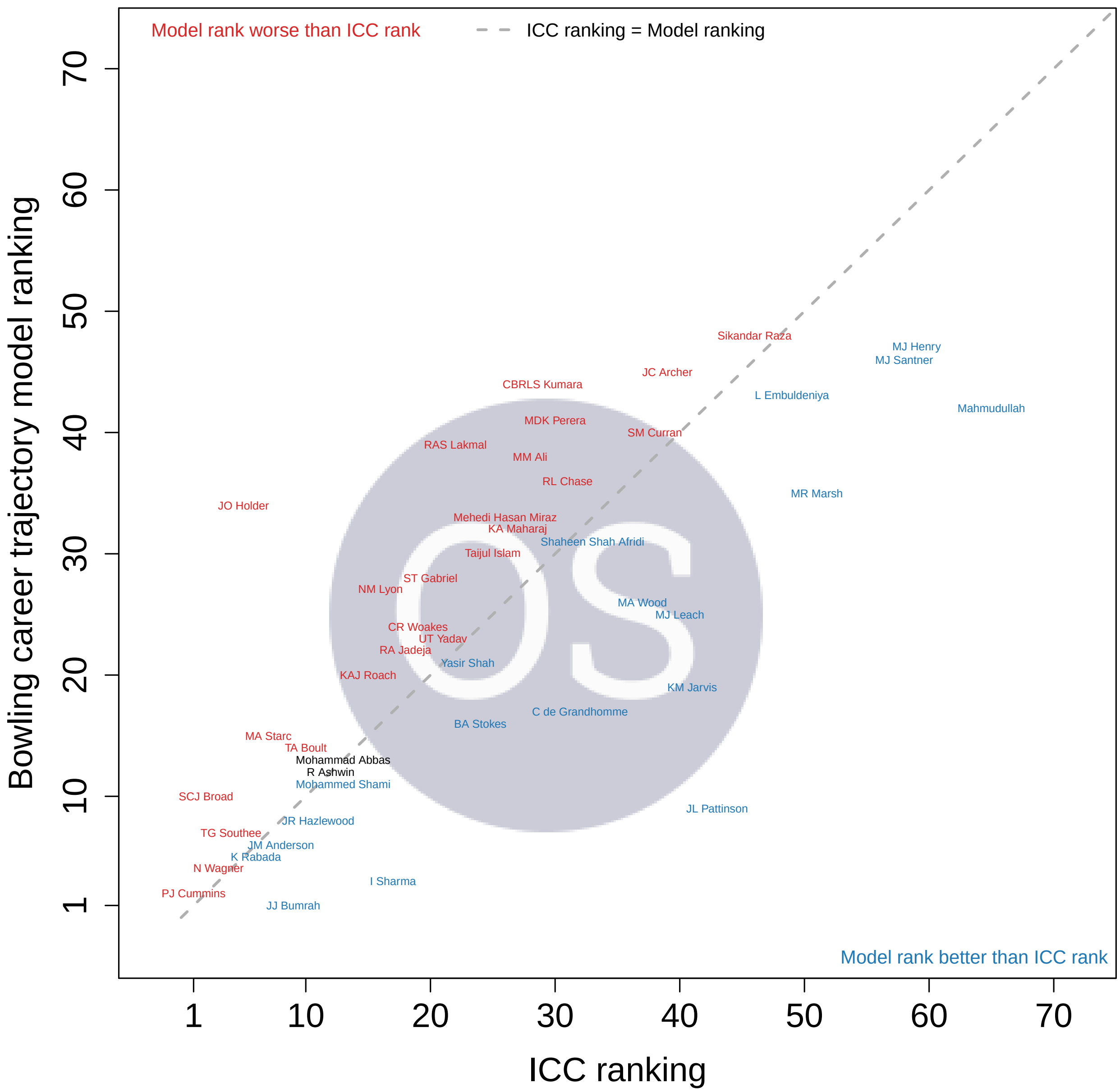


Figure 3.13. *Comparison of world rankings between the bowling career trajectory model and the established ICC ratings method. The model considers players in red to be overvalued by the ICC method, while players in blue are considered to be undervalued. Players in black represent cases where there is consensus in rankings between the two methods.*

3.5 Model diagnostics

3.5.1 Model prediction

The bowling career trajectory model attempts to describe how a player's bowling ability has varied over the course of their career, as well as providing a forecast for future estimates of ability. However, the proposed model can only be considered a useful tool if estimates of future performances are more accurate than predictions made using metrics such as the career bowling average. Therefore, an assessment of the proposed model's predictive capabilities are essential when it comes to validating the model, which is achieved by computing the relative prediction errors for future bowling performances. Note that the data for each observation fitted by the nested sampling algorithm is measured in units of standardised runs conceded, the mean squared prediction error is also reported in terms of standardised runs conceded.

The approach to computing prediction errors for the bowling career trajectory model is similar to the method used in Chapter 2. It is assumed that the model has access to all previous bowling performances, $\{s_1, s_2, ..., s_T\}$, when predicting the number of standardised runs to be conceded before taking a wicket in a player's next career bowling spell, s_{T+1}. However, as s_{T+1} is a future observation, yet to be observed, the prediction error is computed using leave-one-out cross-validation (Sammut & Webb, 2010), by forecasting a value for s_T, and evaluating how close this value is to the true observed value of s_T, using mean squared error (MSE). In order to avoid the complications of non-wicket taking bowling spells, a player's most recent wicket taking bowling spell, $\{s_T\}$, is used as the test data, while $\{s_1, ..., s_{T-1}\}$ is used as the training data.

To provide a means of comparing the predictive accuracy of the bowling career trajectory model, predictions of future bowling performance are also obtained using set of simple moving average (SMA) models of varying orders. The SMA models compute a prediction for s_T, using a player's most recent 10%, 25%, 50% and 100% of career bowling performances. For example, if a player has 100 observed bowling spells in their career bowling data, the SMA(10%) model makes a prediction for s_T using the most recent 10 bowling spells, while the SMA(100%) model would account for all 100 of the player's bowling spells. Again, the SMA(100%) model assumes that ability remains constant throughout a career and is equivalent to using a player's standardised career bowling average to predict future observations.

Under the conditions of leave-one-out cross-validation, a player's most recent wicket taking bowling spell is removed from the training data set and is treated as the test data, s_T. Therefore, where a player has only taken one or fewer wickets in their career, prediction errors are unable to be computed. Of the 522 players in the Cricsheet data set, prediction errors were able to be computed for 405 players, which are presented in Table 3.12. Given the hierarchical analysis presented in Section 3.4.2 provided little evidence to support the presence of both innings and venue-specific effects, after adjusting for opposition batting strength, the prediction error for the bowling career trajectory model that excludes these effects has been included.

Table 3.12. *Mean squared prediction errors using leave-one-out cross-validation. The bowling career trajectory model that excludes innings and venue-specific effects outperforms all other models across all players, while the SMA(10%) model tends to perform worst of all.*

	Minimum # of career bowling spells			
Model	**No minimum**	**10 spells**	**50 spells**	**100 spells**
SMA(10%) model	1.30	1.37	1.06	0.74
SMA(25%) model	0.97	0.94	0.66	0.52
SMA(50%) model	0.77	0.70	0.53	0.46
SMA(100%) model	0.68	0.66	0.48	0.43
Bowling career trajectory model (no innings/venue-specific effects)	0.59	0.59	0.53	0.46
Bowling career trajectory model (with innings/venue-specific effects)	0.62	0.62	0.58	0.53

In terms of which model provides the most accurate predictions of future performances, the results provided in Table 3.12 are a little variable. Across the entire data set of players analysed, the bowling career trajectory models appear to have the smallest prediction error, with the model that excludes the innings and venue-specific effects performing best of all. This result does not come as a huge surprise; after taking into account the strength of opposition batsmen bowled to, there was less clear-cut evidence to support the presence of such effects in a bowling context for many players, as shown by the hierarchical analysis conducted in Section 3.4.2. Applying the principle of Occam's razor (Myung & Pitt, 1997; Rasmussen & Ghahramani, 2000), it is possible to conclude that the most appropriate bowling career trajectory model is the one that does not include the innings and venue-specific effects.

Interestingly, when conditioning the data to only include players who have bowled in at least 50 career bowling spells, the SMA(100%) model appears to perform best of all. This result may imply that bowling ability tends to settle somewhere close to a player's true underlying ability, after bowling in a certain number of spells. This is supported by the observation made in Section 3.4.1, that after adjusting for the quality of batsmen bowled to, bowling performances tends to vary less during a career than batting performances. Or perhaps, bowling is not a skill that can be measured purely by runs scored and wickets taken.

Furthermore, the support for the SMA(100%) model when conditioning on players who have bowled in more than 50 spells, may imply that a major advantage of the proposed career trajectory model is the somewhat informative prior distribution imposed for the parameter controlling the mean value of the Gaussian process, λ. For example, a player who has taken two career wickets at an average of 100.0, is unlikely to have a true underlying bowling ability as suggested by their current career average. Instead, the prior over λ limits such extreme predictions of underlying ability for players who have bowled in a finite number of career bowling spells. This finding is supported by the results of the hierarchical analysis in Section 3.4.2, where the posterior distribution for hyperparameter μ_λ suggests there is far less variation in bowling ability between players, than there is for batting ability.

3.5.2 Model comparison

Again, as nested sampling has been employed as the Bayesian sampling scheme during the model fitting process, model comparison is effortless via the marginal likelihood or evidence, Z. In Table 3.13, the marginal likelihood for the bowling career trajectory model is presented for the current top 20 ranked bowlers identified in Section 3.4.4. The marginal likelihood for the SMA(100%) model, which assumes a player's bowling ability can be accurately estimated by their standardised or adjusted career bowling average, is also provided (Z_0). The logarithm of the Bayes factor between the two models is also given, providing the factor by which one model is favoured over the other for individual players. A negative value for this quantity indicates that the SMA(100%) model is more likely to apply to a player's bowling career data than the proposed career trajectory model. The sum of marginal likelihoods over all players in the data is also presented, as is the average logarithm of the Bayes factor.

Generally speaking, the constant ability is favoured across the majority of players who have bowled in a significant number of spells; Kyle Jarvis is the only player in Table 3.13 whose career data appears to be better approximated by the bowling career trajectory model. These results support the findings from Table 3.12; the SMA(100%) model tends to be more accurate at predicting current player bowling ability the more spells a player has bowled in.

It is worth noting that as the nested sampling algorithm employs an MCMC sampler to propose new parameter values, the estimates for marginal likelihood are subject to a degree of sampling error. Considering the size of the sampling errors and the small difference in likelihoods between models for many players, it is difficult to confidently conclude whether one model is more appropriate than the other when it comes to estimating player bowling ability. Once again the principle of Occam's razor can be applied (Myung & Pitt, 1997; Rasmussen & Ghahramani, 2000), suggesting it is perhaps not always worth the effort to model bowling performances using computationally demanding models, such as Gaussian processes. Instead, for the majority of players, it may be sufficient to account for the historic strength of batsmen bowled to over the course of their careers using standardised and adjusted runs. Current bowling ability can then be assumed to be close to a player's standardised or adjusted career bowling average.

Table 3.13. *Marginal likelihood estimates for the top 20 ranked Test match bowlers as ranked by the bowling career trajectory model. The summation of marginal likelihoods and the average logarithm of the Bayes factor over all players shows the data generally support the SMA(100%) model, over the proposed model.*

Rank	**Player**		$\log(Z)$	$\log(Z_0)$	$\log(\frac{Z}{Z_0})$
1.	J. Bumrah	(IND)	-40.2	40.2	0.0
2.	P. Cummins	(AUS)	-92.0	-90.4	-1.6
3.	I. Sharma	(IND)	-270.3	-268.8	-1.5
4.	N. Wagner	(NZ)	-157.3	-155.7	-1.6
5.	K. Rabada	(SA)	-132.9	-129.9	-3.0
6.	J. Anderson	(AUS)	-383.4	-381.9	-1.5
7.	T. Southee	(NZ)	-241.9	-240.1	-1.8
8.	J. Hazlewood	(AUS)	-154.1	-150.8	-3.3
9.	J. Pattinson	(AUS)	-61.6	-60.1	-1.5
10.	S. Broad	(ENG)	-428.0	-426.3	-1.7
11.	M. Shami	(IND)	-148.3	-147.7	-0.6
12.	R. Ashwin	(IND)	-278.7	-275.7	-3.0
13.	M. Abbas	(PAK)	-55.3	-54.4	-0.9
14.	T. Boult	(NZ)	-215.2	-211.7	-3.5
15.	M. Starc	(AUS)	-199.7	-197.1	-2.6
16.	B. Stokes	(END)	-143.1	-140.9	-2.2
17.	C. de Grandhomme	(NZ)	-42.0	-41.3	-0.8
18.	S. al Hasan	(BAN)	-169.3	-166.2	-3.1
19.	K. Jarvis	(ZIM)	-44.7	-45.0	0.3
20.	K. Roach	(WI)	-159.6	-158.6	-1.0
	All players		$-15,368.3$	$-15,260.4$	-0.2

This is still a significant step forward in the measurement of bowling ability and analysis of bowling performances. To date, there exists no method of accurately and fairly accounting for the quality of opposition batsmen bowled to. The introduction of measurement units such as standardised and adjusted runs conceded, allows for a more direct means of comparing bowling abilities and assessing the relative impact of individual bowling performances. As has been seen in sports such as baseball and basketball, fans are willing to engage with the complexities of advanced statistics and metrics. However, it is essential that such quantities can be easily understood and interpreted by viewers of the sport who have minimal formal statistical training.

3.6 Discussion

3.6.1 Limitations and further work

Employing a Gaussian process to fit the function used to describe a player's bowling career trajectory attempts to account for variation in bowling data that may exist on both short and long-term scales. However, as indicated in Section 3.5, once adjustments to the data are made to account for the batting abilities of opposition batsmen bowled to, there is an underwhelming amount of evidence to suggest that bowling ability varies by the same extent as batting performances, nor does it appear that there are considerable innings and venue-specific effects.

There are several plausible explanations for the results and findings discussed in this chapter. Firstly, it is worth acknowledging that the bowling career trajectory was fitted sequentially after the batting career trajectory model, rather than the other way around. As discussed, this decision was made as one can expect the variation in batting abilities across teams to be far greater than the variation in bowling abilities, due to the fact that in most games all players must bat, while the decision of who bowls is at the discretion of the bowling team's captain. If the bowling model was to be fitted prior to the batting model, it is likely that there would be observable innings and venue-specific effects.

Admittedly, under the current approach, the assumption is made that variation in batting performances between innings and venues is intrinsically tied to changes in a player's underlying batting ability, rather than the opposition's bowling ability. However, this is a necessary trade-off that must be made in order to provide a means of accounting for opposition batting strength when analysing bowling performances. In reality, such differences are likely a result of something in the middle. Some batsmen really do perform with different levels of ability when playing in and away from their home country, while some bowlers really are masters of exploiting deteriorating pitch conditions in the second innings. Likewise, some batsmen have shown time and time again that they rise to the occasion in their team's second innings of a match, while many spin bowlers have been shown to prefer bowling in the sub-continent, be that at a home or away venue. It is therefore implicit that treating all runs scored equally, rather than all runs conceded, is a more sensible approach.

Furthermore, while the model uses standardised runs conceded to adjust for the quality of batsmen bowled to, all wickets taken are treated equally. Instead, it would be advantageous to incorporate a means of rewarding players who are able to consistently dismiss world-class batsmen. This may be achieved by estimating the number of runs a bowler has potentially saved by dismissing a given player. For example, taking the wicket of a tail-end batsman on a low score is useful, but chances are, that player would not have gone on to score too many more runs anyway. On the other hand, dismissing a top-order batsman who already has their eye in, will be more significant in helping restrict an opposition batting team's final total. Such a consideration could also be applied to analysing the impact of individual player fielding performances. Dropping a tail-end batsman is less likely to have a significant bearing on a match, whereas dropping a top-order batsman who is currently on a large score may well be a costly mistake. Such an analysis could even be extended to quantify the effects of contentious or incorrect umpiring decisions.

Although complex variables such as opposition strength have been accounted for in the bowling career trajectory model, there are still a number of external factors that can affect performance that have been ignored. For example, it is well known that pace bowlers are generally more likely to take wickets at the start of an innings, partially due to the hard, new ball with which they bowl. While a ball is new, it is more responsive to overhead conditions and will generate more swing and bounce off the pitch, which usually translates to a more uncomfortable batting experience. As a ball ages, delivery trajectories becomes easier to predict and therefore easier for batsmen to safely negotiate. Accounting for an effect due to the age of the ball may provide additional insight as to the quality of specific bowling performances, particularly for players who are able to consistently take wickets throughout an innings, regardless of the ball state.

Finally, unlike the career trajectory model used to measure batting ability in Chapter 2, the present model applied in a bowling context requires a far richer data set in the form of ball-by-ball data — rather than simply a list of career performances — in order to be fitted. Such data is only now becoming readily available to the public via sources such as Cricsheet and ESPNcricinfo and even then, is only generally accessible for international matches. Furthermore, the state in which this data is available is not fit for immediate analysis. Instead, users must be be proficient in a computing language that can wrangle and clean the data into a more appropriate format before attempting to obtain meaningful insights. Given the sheer quantity of domestic T20 leagues that have taken over the annual cricketing schedule, it is difficult to see how significant progress can be made in the field of cricket analytics, until a more public-friendly means of accessing such data is made available.

3.6.2 Concluding remarks

The bowling career trajectory model proposed in this chapter has presented a framework for estimating the past, present, and future bowling abilities of professional cricket players. As discussed, the proposed model does not provide much evidence to suggest that bowling ability varies significantly over the course of a Test career. It is unclear whether this is simply true of the data, or is a shortcoming of the Gaussian process used to fit the model. Perhaps the implication is that measuring bowling performances using just runs conceded and wickets taken does not measure the full impact of a player's bowling efforts. In any case, this may at least suggest that there is some merit to simply using a player's standardised career bowling average or adjusted career bowling average to estimate their true underlying ability.

Rather than measuring bowling performances using multiple measures, runs conceded and wickets taken, the proposed method of splitting data into runs conceded per bowling spell, provides a method of visualising performances that previously did not exist. Furthermore, applying the results of the batting career trajectory model presented in Chapter 2 to the data provides a means of quantifying the relative strengths of batsmen bowled to, an important factor that cannot be accurately measured using the traditional bowling average. The introduction and development of the adjusted bowling average, to measure player bowling ability as if they were bowling to the average Test batsmen, allows for a fairer comparison of bowling performances between players. This provides context to certain performances and in some cases can differentiate between similar sounding bowling figures, which may have occurred in very different match scenarios.

The findings of the hierarchical analysis indicate that once the relative strengths of opposition batsmen have been accounted for, there are no clear innings or venue-specific effects that exist for many players, although an exception may exist here for spin bowlers. Ultimately, this result suggests that the best course of action is usually to select the strongest performing bowlers, rather than selecting bowlers that may be thought to suit certain conditions. Of course, this finding is simply a consideration that should be taken on board by coaches and selectors. All sports analytics should be used as an additional tool in the selection and strategy-forming process, rather than a hard and fast rule.

In terms of predicting future bowling performances, the proposed bowling career trajectory model was found to perform similarly to a constant ability model, which simply assumes a player's underlying bowling ability is equivalent to their career standardised average (Table 3.12). This result was supported by the lack of an observable difference between the marginal likelihoods of the two models presented in Table 3.13. A finding of some interest, consistent with the results from the batting career trajectory model discussed in Chapter 2, was that using only a player's most recent bowling efforts to predict future performance is the least reliable method of prediction. This again signals that selectors should be cautious in selecting players who have enjoyed a brief period of good form, but have not typically performed at a consistently high level.

Finally, the predictions of current player ability have been compared with the ICC ratings method to gain an understanding of how each method ranks the bowling abilities of individual players and where the significant differences lie. Generally speaking, there is a reasonable amount of agreement between the two methods. However, the bowling career trajectory model provides estimates of ability in units of an adjusted bowling average, which can be more easily understood than arbitrary rating points and allows for a more intuitive understanding of the estimated differences between players. Additionally, as the model has been developed within a Bayesian framework, it is easy to compare players and provide probability-based estimates relating to the likelihood of certain players outperforming one another. Once again, the model output lends itself to practical applications in areas such as team selection policy and talent identification.

Chapter 4

A simulation-based method of match outcome prediction

4.1 Introduction

The batting and bowling career trajectory models, presented in Chapters 2 and 3, have been shown to provide more accurate predictions of individual player ability that typically outperform traditional measures, such as the batting and bowling average, while maintaining an intuitive cricketing interpretation. The next natural step is to combine these predictions of player ability in a manner that allows for the outcome of an upcoming match to be predicted, given two proposed playing XIs. Regardless of the sport in question, a common outcome when mixing a data enthusiast's enjoyment of sport and fascination with numbers, is some form of statistical model that attempts to predict the outcome of a match. This is no different with cricket. Over the years there have been multiple attempts at identifying the key factors to cricketing success and numerous studies that have defined a framework for predicting the outcome of Test, one-day and T20 cricket.

Interestingly, the most popular method of predicting the outcome of a cricket match is one that does not market itself as a match predictor — the Duckworth-Lewis-Stern (DLS) method (Duckworth & Lewis, 1998; Stern, 2016) — used to determine a winner in interrupted one-day or T20 matches. Where the first innings of a match is prematurely cut short, usually due to uncontrollable factors such as adverse weather conditions, the DLS method attempts to predict the score the team batting first would have scored, conditional on the number of wickets and overs remaining. A somewhat controversial element of the DLS method, is the adjustment made to account for the fact that the batting team did not know their innings would come to a natural conclusion. This adjustment has seen multiple changes over the years in an attempt to reflect the more aggressive nature of the modern game. In the second innings of a match, the DLS method uses the number of resources remaining to the batting team (balls and wickets remaining), to provide the current total the chasing team needs to have scored, to be on track to win the game. This is equivalent to estimating the decision boundary in terms of runs scored, where the probability of the chasing team winning the match exceeds 50%. An important point of note is that the DLS method does not take into account the relative strengths of the teams playing in a match. While this may ultimately reduce the method's predictive accuracy, it is a necessary trade-off. The DLS method is used to resolve the outcome of almost all interrupted matches, from social village cricket to international cricket, therefore, assuming equal team strength is the only way to ensure fairness in the sense each side has an equal chance of winning a match, prior to its commencement.

A number of past studies have tended to emphasise outcome prediction in one-day and T20 matches, likely due to the well defined nature of match parameters, such as the number of overs permitted to be bowled by each side. A multiple linear regression model was employed by Bailey & Clarke (2006) to predict the outcome of ODI matches, once a game has commenced, utilising similar concepts as the DLS method, such as resources remaining to the batting team. Other covariates considered include a home ground advantage effect, as well as country-specific estimates of past performance at the particular venue and past history between the two competing nations. A question of interest that underpinned the study was how betting markets tend to react to significant match events, such as a wicket falling, with a preliminary finding suggesting that punters overreact to such events. In a similar study, a logistic regression model was used to identify factors that most affected a team's probability of winning an ODI match (Bandulasiri, 2008). Home ground advantage was found to be the most significant variable, although a finding of interest was that winning the coin toss only appears to provide a competitive advantage in day/night one-day matches, which are partially or wholly played under floodlights.

A more well-known match predictor that has enjoyed some commercial success and TV airtime, is the winning and score predictor, more commonly referred to as WASP. The WASP model was born from the PhD thesis of Scott Brooker at the University of Canterbury, New Zealand (Brooker, 2011) and a later paper concerning the estimation of batting conditions in ODI cricket (Brooker & Hogan, 2011). The resulting product, WASP, aims to predict the outcome of one-day and T20 matches by producing two major outputs: (1) in the first innings, the model estimates the score to be attained by the batting team; and (2) in the second innings, the model estimates the probability of the batting team successfully chasing the target score. The model calculations are based on dynamic programming and like the DLS method, consider variables such as the number of balls and wickets remaining to the batting team. Other factors, such as the historic scores made at the relevant venue and past performances of teams in similar match scenarios to the one at the time of prediction, are also accounted for. As with the DLS method, WASP does not account for the individual abilities of the players participating in a match, instead opting to assume equal strengths between teams. This again may weaken the model's predictive accuracy, particularly in cases where the batting team has opted to select an additional batsman or bowler in their lineup, which could significantly increase or reduce their overall batting strength. However, a benefit of not including such granular estimates of team and player ability is that the model is computationally undemanding, allowing for the output to be updated over-by-over, or at times, ball-by-ball, providing real-time updates in regards to the estimated state of a match. A second notable drawback of the model is the requirement for a subjective par score to be set prior to the commencement of a match. If this par score is considerably smaller or larger than the observed first innings total, the estimates of win probability in the second innings can tend to be a little extreme.

The WASP model made its TV debut in 2012, during a domestic T20 match in New Zealand and has continued to be used by broadcasters in both New Zealand and England for domestic and international matches. Unfortunately for those engaged in the field of research in cricket analytics, the WASP model has made the transition from the realm of academia to commercial product, limiting the amount of information available in terms of the current methodology. One can only speculate as to how the present iteration of WASP differs from the initial edition, however, for now WASP reigns as one of the leading — or at least most popular — methods of match outcome prediction for the short forms of the game.

Prior to the boom of T20 cricket, a number of attempts were made to apply statistical modelling to the prediction of match outcomes in Test cricket. Brooks et al. (2002) employed an ordered probit model to make such predictions, with the ordered responses being win, draw or loss for a given match. This study considered a number of common cricketing metrics, such as batting and bowling averages, strike rates and economy rates, to estimate a historic measure of batting and bowling performance for respective Test playing nations. These performance measures were then utilised in the model fitting and outcome prediction process. Based on a data set of 342 Test matches spanning from 1994 to 2000, the model was able to correctly predict the outcome for an impressive 71.1% of matches. It is worth noting that this rate of successful prediction is based purely on a training set, as such, the model may not be expected to perform as well under cross-validation.

More recently, Akhtar & Scarf (2012) applied a multinomial logistic regression model to predict the probability of a win, draw, or loss, occurring in a given Test match. Here, the relevant covariates were split into two categories, (1) *pre-match effects*, and (2) *in-game effects*. Pre-match effects include estimates for the respective team strengths, sourced using the ICC team ratings for Test playing nations; a venue effect, tied to the specific ground; a home ground advantage effect; and an effect for winning the toss. In-game effects are based on the current state of the match, such as the current score or lead held by a team, the number of overs remaining, the match run rate and the number of wickets remaining for each team across the match. Unsurprisingly, Akhtar & Scarf (2012) found that pre-match effects are more likely to be significant predictors of match outcome early in a match, while in-game effects become more significant as a match progresses. As with a number of previous studies, the estimated strength of each team is directly related to a team's past performances, rather than the individual players taking the field. As such, losing a key player to injury prior to a match would not impact a team's overall predicted probabilities of winning, losing or drawing.

The only Test match prediction model to be used in a public broadcast comes from CricViz, a UK based cricket analytics company, who have aimed to provide an answer to the age old cricketing question: 'who's winning?'. The aptly named WinViz model uses a Monte-Carlo simulator, which effectively simulates the outcome of a Test match a number of times, given two proposed playing XIs. However, as the CricViz product has become more popular globally, less and less information is publicly available in regards to its various models and how they work, including WinViz. An archived version of the CricViz website from April 2020 (`https://web.archive.org/web/20191203144418/http://cricviz.com/winviz/`) provides some information, which can only be assumed to be somewhat true of the current iteration of WinViz. Here, the WinViz model claims to account for numerous variables pertaining to a match, including the career records of all participants, historical data from the venue and country where a match is being played. Additionally, during the simulation process, WinViz makes adjustments based on individual players' career records against pace and spin bowling, a potentially important factor that has been overlooked by many past studies.

Of concern is the claim: *"The mathematics of a batsman's scores is fairly robust. They obey a common pattern known as a geometric distribution. If you have a good idea what his underlying average score will be in a certain innings then you can very accurately predict the probability of him making any other given score."* The geometric assumption first made by Elderton & Wood (1945) has since been disproved by multiple papers, in favour of the getting your eye in hypothesis (A. C. Kimber & Hansford, 1993; Brewer, 2008; Bracewell & Ruggiero, 2009; Stevenson & Brewer, 2017; Stevenson, 2017), which indicates that the developers of WinViz are not correctly estimating the underlying batting abilities of individual players during an innings.

While commercial match predictors, such as WASP and WinViz, claim to account for a multitude of factors, it is difficult to know exactly how each factor is considered in the model. This is the unfortunate reality of the disjointed nature of statistical applications and algorithms published in academia and those developed in the commercial business world. When it comes to promoting software and predictive algorithms, companies will often claim their product offers a certain functionality, but will not expose themselves to the peer-review process, at risk of losing their competitive advantage. Without the public availability of specific details, inner-workings, or methodology of a product, one must make a subjective judgement as to how much they can trust a company's word in regards to the quality of their advertised product. An ongoing example of this has been provided throughout this thesis with the ICC ratings system.

In this chapter, a simulation-based approach of predicting the outcome of a Test match is proposed, making use of the individual-specific estimates of batting and bowling ability, based on the career trajectory models detailed in Chapters 2 and 3. The simulator is built in R (Ihaka & Gentleman, 1996), as is the code that is used to post-process and analyse the results. As with the WinViz model, the proposed algorithm is simulation-based, drawing its estimates and projections of likely results from a number of hypothetical outcomes. Estimates of team batting and bowling strength are based on the individual players selected to play for each team, rather than the historic results of the team in question. This allows for an estimate of how a team's probability of winning a match is affected by the selection or non-selection of certain players. In this regard, the match simulator can provide an indication of a player's likely contribution towards their team's success.

The match simulation methodology is comprised of three main elements: (1) simulation initialisation (Section 4.2.1), (2) simulation processing (Section 4.2.2), and (3) simulation summary (Section 4.2.3). Each of these sections details the full process used to obtain the predictions outlining the most likely outcome of a match, from setting up the simulation process, through to a discussion of how the results can be visualised and used to obtain meaningful insights. Several real-world applications of the output are presented in Section 4.3, illustrating the match simulator's potential for practical use.

4.2 Methodology

4.2.1 Simulation initialisation

Prior to running the match simulator, a number of relevant quantities must be pre-computed to define the general framework within which Test match cricket is usually played. These specifications ensure that the simulated matches loosely mimic what may actually occur in reality. The match simulator is initialised via the `initialise_simulation()` function defined on the following page, which takes the following input arguments:

- `home_players` - a vector of player names, defining the home team.
- `away_players` - a vector of player names, defining the away team.
- `nsims` - the number of simulations to be computed.

```r
## Function that initialises objects required to simulate matches
initialise_simulation <- function(home_players,
                                  away_players,
                                  nsims = 1000)
{
  ## Get player information and store in a data.frame
  match_players <- match_player_info(home_players, away_players)

  ## Computes career quantities for the defined players
  ## - E.g. Batting/bowling averages, strike rates/economy rates
  averages <- compute_averages(match_players)

  ## Use historical data to compute empirical probabilities
  ## of any given event occurring on a given ball, E.g.:
  ## - Probability of a wicket falling/bowling a wide/hitting a 6
  prob_events <- compute_ball_events()

  ## Get estimates for player batting and bowling ability
  player_abilities <- compute_player_abilities(match_players, nsims)

  ## Compute bowler selection logic for the defined players
  bowling_logic <- ai_team_bowling(match_players)

  ## Initialise storage object for simulation results
  results <- vector("list", nsims)

  ## Return the objects in a list for access
  return(list(summary_data = list(prob_events = prob_events,
                                  averages = averages),
              match_players = match_players,
              player_abilities = player_abilities,
              bowling_logic = bowling_logic,
              results = results))
}
```

The `initialise_simulation()` function pre-computes a number of necessary quantities that are required in process of simulating a match and are stored in the object `summary_data`. These estimates include the player-specific values for metrics such as career batting and bowling averages; strike rates; and economy rates, which will define the tempo at which the match is likely to be played. Additionally, the empirical probabilities of certain events occurring on any given delivery are computed, based on historic ball-by-ball data, which in the case of Test match simulation, is obtained from the Cricsheet data source. As such, when referring to the data of 'all Test matches', this really means 'all Test matches since 2008'. The estimated batting and bowling abilities of the defined players are stored in the `player_abilities` object, while the logic of internal processes related to the likelihood of certain players bowling at various stages of an innings are defined in `bowling_logic`. The resulting output stores the relevant objects in a list, which can then be accessed and used during each simulation, including a `results` object, used to store each simulated result.

Proposed playing XIs

First and foremost, the simulator requires two proposed playing XIs to be defined, either assigning one team as the home team and the other as the away team; or, indicating the match is to be played at a neutral venue. To provide a running example of the simulation process throughout this chapter, a matchup between Australia and New Zealand is simulated using the proposed lineups defined in Tables 4.1. The hypothetical match is assumed to be taking placed in Australia, with each team's lineup based on common selections in their most recent Test matches. A number of players featured in both Chapters 2 and 3 appear in this matchup.

Once the team lineups are set, including a marker for the home and away team, predictions for player ability must be obtained. This is achieved by sampling from the posterior predictive distribution for both the effective batting average, $\mu(x,t)$, and standardised effective bowling average, $\omega(t)$, for each player. Each posterior sample will correspond to a player's assumed underlying batting and bowling ability for a single simulated match. Therefore, if 1,000 simulations are to be used to estimate the most likely outcome of a match, the algorithm requires 1,000 posterior samples for $\mu(x,t)$ and $\omega(t)$, for each player.

In some cases, such as when a player is making their Test debut, players will not have any Test match batting or bowling data to obtain predictions for batting and bowling ability. In such instances, there are several options available. One choice is to simply generate estimates for $\mu(x,t)$ and $\omega(t)$ using the prior parameter distributions defined in Chapters 2 and 3, which were shown to be fairly typical of the average player. However, if it is known that a certain player is coming into a match as a specialist batsman or bowler, these prior distributions can be altered accordingly, to better reflect their likely underlying abilities. Alternatively, where data exists

Table 4.1. *Proposed playing XIs for Australia and New Zealand, in batting order.*

Australia (home)	**New Zealand (away)**
1. David Warner	1. Tom Latham
2. Joe Burns	2. Tom Blundell
3. Marnus Labuschagne	3. Kane Williamson
4. Steve Smith	4. Ross Taylor
5. Matthew Wade	5. Henry Nicholls
6. Travis Head	6. Colin de Grandhomme
7. Tim Paine	7. BJ Watling
8. Pat Cummins	8. Mitchell Santner
9. Mitchell Starc	9. Tim Southee
10. Nathan Lyon	10. Neil Wagner
11. Josh Hazlewood	11. Trent Boult

for a player's domestic first-class record, estimates for current batting and bowling ability can be obtained, with a heuristic adjustment made to account for the jump in difficulty between first-class and Test cricket.

Bowler selection logic

Before a match can be simulated, the logic for several internal processes must be defined, namely, how the algorithm selects which players will bowl a particular over. This is an important feature of the match simulator, as it ensures that the proportion of simulated overs bowled by each player is consistent with historic performances. The proposed logic is defined via a two-step process, involving an estimate for the average quantity of overs to be bowled by a player and an estimate for when in an innings these overs are likely to be bowled.

Firstly, an estimate for each player's *bowling workload* must be computed, using the player's historic bowling career data. Bowling workload is defined as the probability that a player will bowl any given over. For example, a bowling workload of 20% implies that a player will bowl one fifth of their team's overs. The data used to estimate a player's bowling workload is simply every ball that a player has been involved with either as a fielder or bowler, during their Test career. Each ball is assigned an index, I, where $I = \{1, 2, 3, ..., n\}$, and a flag indicating whether or not the player in question was the bowler. This data is then modelled using backward elimination (Hocking, 1976), via three logistic regression models, presented in Equations 4.1, 4.2, and 4.3, to obtain a prediction for a player's bowling workload, p.

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \times I + \beta_2 \times I^2 \tag{4.1}$$

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \times I \tag{4.2}$$

$$\log\left(\frac{p}{1-p}\right) = \beta_0 \tag{4.3}$$

This procedure is effectively testing whether or not a player's bowling workload has remained constant over their career. Equation 4.1 tests whether there is an observable quadratic relationship between bowling workload and time, over the course of a player's career, while Equation 4.2 is simply testing whether or not bowling workload has remained constant. If neither Equation 4.1 or 4.2 are deemed suitable, a player's bowling workload is presumed have remained constant throughout their career, as implied by Equation 4.3.

A practical example is provided in Figure 4.1, where the bowling workload for Neil Wagner is modelled using Equations 4.1, 4.2, and 4.3. Wagner's primary role in the New Zealand side is as a bowler and as such, he has a high average workload of just over 20%, as indicated in Figure 4.1. Here, Equation 4.2 defines the preferred bowling workload model for Wagner, suggesting he has seen a significant linear increase in bowling workload, as his career has progressed.

Defining bowling workload alone will not quite achieve realistic simulated results. Certain bowlers and bowler types are known to be more or less likely to bowl at certain stages of an innings, for example, it is uncommon to see a spin bowler open the bowling. Therefore, estimates for *bowling workload timing*, defining when a player is most likely to bowl during an innings, must also be obtained. Computing the number of occasions a player has bowled the n^{th} over in an innings provides an understanding of the typical stages of an innings they are brought on to bowl. Wagner is generally used as a first change bowler by New Zealand, meaning he does not open the bowling, but is often the first person to bowl after the opening bowlers. This idea is supported in Figure 4.2, which shows that Wagner has a low probability of bowling in the first ten overs of an innings, but bowls a lot through the middle overs. The significant drop in Wagner's usage around the 80 over mark does not come as a surprise; in Test cricket the bowling team is given the opportunity to replace the current ball with a new ball every 80 overs. In most scenarios the new ball is immediately given to a team's opening bowlers, as pace bowlers tend to benefit more from a fresh ball than spin bowlers. Therefore, as a first-change bowler, Wagner is generally given a rest at this point of an innings, to ensure he is ready to bowl with the relatively new ball around the 90^{th}–100^{th} over mark, immediately following the opening bowlers.

Once estimates of bowling workload and workload timing are obtained for each player, the probability of any given bowler in a team bowling the n^{th} over of a simulated match can be

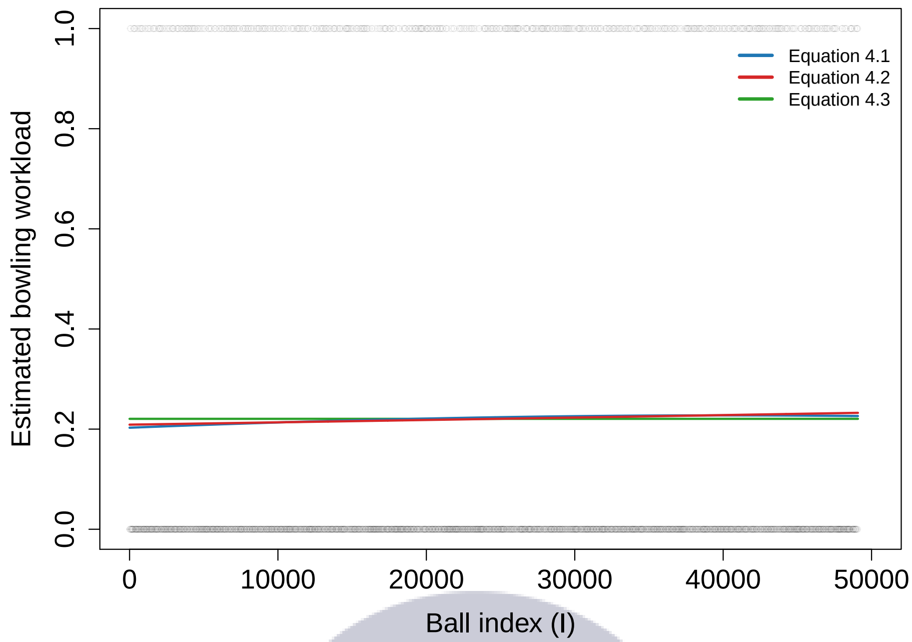


Figure 4.1. *Estimated bowling workload for Neil Wagner. Points plotted on the y-axis at values of 1 indicate balls that were bowled by Wagner, while points plotted on the y-axis at values of 0 indicate balls that were not bowled by Wagner. The model defined by Equation 4.2 (red) is deemed the most appropriate, indicating Wagner's general bowling workload has increased since the early stages of his career.*

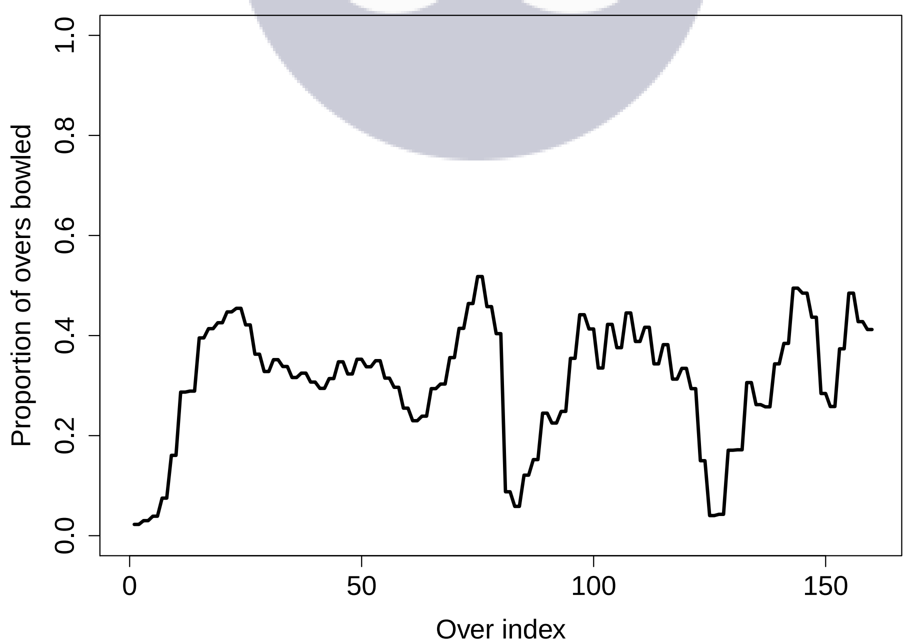


Figure 4.2. *The proportion of overs bowled by Neil Wagner, by over index, based on the empirical data. As a low proportion of innings last more than 120 overs, the data at larger over indexes is sparse and can be highly erratic.*

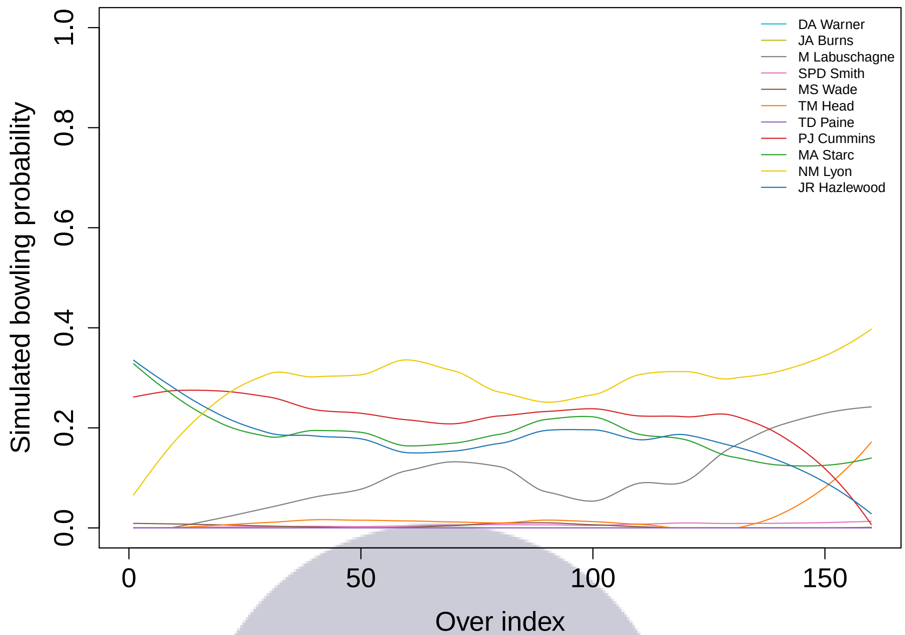


Figure 4.3. *Over-by-over simulated bowling probabilities for players in the proposed lineup for Australia.*

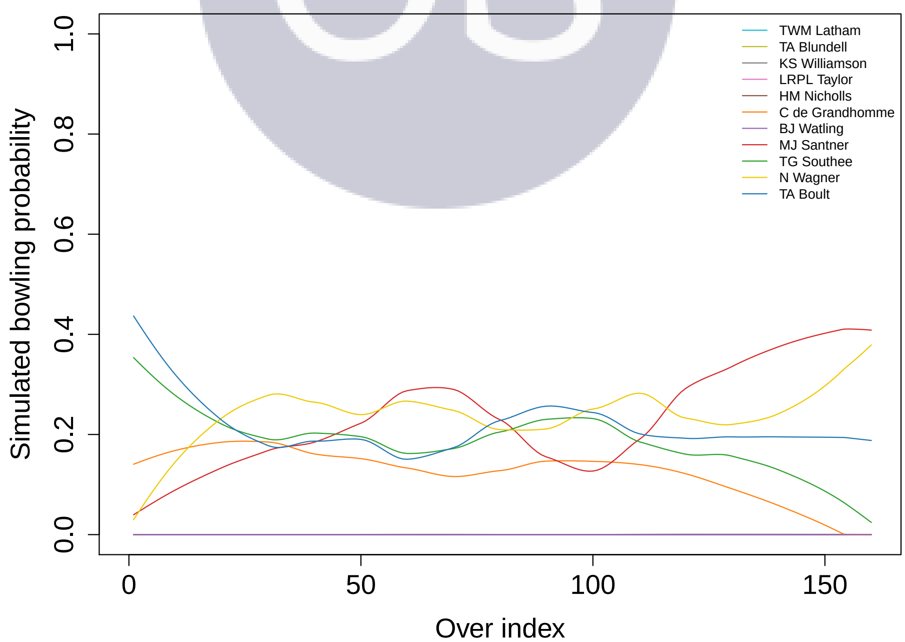


Figure 4.4. *Over-by-over simulated bowling probabilities for players in the proposed lineup for New Zealand.*

estimated. Figures 4.3 and 4.4 show the respective over-by-over bowling probabilities for the lineups proposed for Australia and New Zealand in Table 4.1. As the bowling workload timing can be highly erratic, a locally estimated scatterplot smoother is applied to the estimates to provide more reasonable values (Cleveland, 1979; Cleveland & Devlin, 1988). Although the plots are rather cluttered and noisy, several general trends are clear. Pace bowlers who typically open the bowling for their teams (Mitchell Starc and Josh Hazlewood for Australia; Trent Boult and Tim Southee for New Zealand) have higher bowling probabilities at the very beginning of an innings and around the 80^{th} over, when the new ball can be taken.

Conversely, bowling probabilities for spin bowlers Nathan Lyon (Australia) and Mitchell Santner (New Zealand), are low at the start of an innings, but increase as an innings progresses, until the 80^{th} over, where a sharp decrease can be observed, due to the new ball being taken. Finally, a close inspection of Figure 4.3 depicts the increasing probabilities of Australian part-time bowlers Marnus Labuschange and Travis Head being brought on to bowl, as an innings progresses. The majority of Test innings are concluded within 120 overs, therefore, where a team is still batting at this stage of an innings it can be an indication of the bowling team struggling. In such cases, part-time bowlers are often given the opportunity to bowl several overs to try and snag a wicket, while giving their strike bowlers a chance to rest.

4.2.2 Simulation process

Once estimates relating to necessary quantities such as player abilities and bowler selection probabilities have been pre-computed, the requisite framework exists to simulate a match and predict a possible pathway the proposed match could take. As discussed, a single Test match consists of up to four innings — two per side — with innings consisting of multiple overs and each over consisting of six legal balls or deliveries. Accordingly, the process of simulating a match is broken down into each of these sub-categories, which are discussed below.

Match simulation

Prior to running the simulation for a single match, there is a final step in the initialisation process, which requires the definition of several global match parameters, outlining the match conditions under which the simulation will be constrained. These are namely the total number overs to be bowled in a simulated match and the winner of the toss. In Test match cricket, 90 overs are scheduled to be bowled per day, so the total number of overs to simulate is initialised at 450. If the result of the toss is unknown then this variable is irrelevant, as teams should be assumed to have an equal probability of batting or bowling first, unless there is prior knowledge that one team is more likely to opt to bat or bowl than the other, should they win the toss. Each of these match parameters is defined in the `initialise_match()` function, which initialises an object to store the results of a simulated match. The function takes the following parameters as input arguments:

- `summary_data` - quantities pre-computed in `initialise_simulation()`, which come in two categories:

  1. `prob_events` - a list of empirical probabilities of any event occurring on a given delivery, based on the historical ball-by-ball data of all Test matches.
  2. `averages` - a list summarising the career batting averages, batting strike rates, bowling averages, bowling strike rates, and bowling economy rates, of all players in the match. The list also contains the mean batting average, batting strike rate, bowling average, bowling strike rate, and bowling economy, averaged across all Test matches.

- `iteration` - the simulation iteration number, which is used as an index to select the appropriate posterior estimates for player batting and bowling abilities that were pre-computed in `initialise_simulation()`.

Two of the more important objects within the object that is created by the `initialise_match()` function are: (1) `match_parameters`, and (2) `innings_data`. The `match_parameters` object is a list containing the information relating to the current state of the simulated match, such as the number of overs remaining, the team that has currently scored more runs (a positive value for `lead_deficit` indicates the home team is ahead, while a negative value indicates the away team is ahead), and an empty vector `result`, to store the result of the simulation once complete. Note the `innings_data` object is initialised as an empty list, which is populated as the innings within a match are simulated.

As innings are made up of overs, and overs are made up of six legal balls or deliveries, it is easier to understand how each of the ball, over and innings simulation processes are inter-related, by starting with the most repetitive and complex process — the simulation of a single ball or delivery — and working backwards. To demonstrate the match simulation procedure in its entirety, consider the proposed lineups for Australia and New Zealand in Table 4.1 and assume a match simulation object has been initialised using the `initialise_match()` function.

```
## Function that initialises the match simulation object
## Initialises the match parameters that store the match-specific
## information relating to the state of the match
## Computes the relevant batting and bowling abilities for all
## players participating in the match
initialise_match <- function(summary_data, match_players,
                              player_abilities, bowling_logic,
                              iteration)
{
  ## Extract player batting and bowling abilities
  ## for this iteration of the simulated match
  match_abilities <- compute_match_abilities(match_players,
                                             player_abilities,
                                             iteration)

  ## Initialise the match parameters
  ## toss() randomly assigns home/away team to bat/bowl first
  match_parameters <- list(total_overs = 450,
                           overs_bowled = 0,
                           balls_bowled = 0,
                           overs_remaining = 450,
                           toss = toss(),
                           lead_deficit = 0,
                           match_completed = FALSE,
                           result = character(1))

  ## Return relevant match data, including the match parameters
  return(list(summary_data = summary_data,
              match_players = match_players,
              match_abilities = match_abilities,
              bowling_logic = bowling_logic,
              match_parameters = match_parameters,
              innings_data = vector("list")))
}
```

Ball simulation

To simulate the result of a single delivery, or ball, the match simulator requires just three arguments: (1) a `batsman` object, (2) a `bowler` object, and (3) the pre-computed `summary_data` object. As shown in Tables 4.2 and 4.3, `batsman` and `bowler` objects contain various details about the players in question and store the player's current innings-specific performance.

Table 4.2. *Elements contained within a* `batsman` *object.*

Quantity	Type	Description	Initialised value
`$name`	Character	Player's name	`"DA Warner"`
`$position`	Numeric	Position in the batting order	`1`
`$runs`	Numeric	Runs scored in innings	`0`
`$balls_faced`	Numeric	Number of legal balls faced in innings	`0`
`$ball_count`	Numeric	Total number of balls faced in innings	`0`
`$fours`	Numeric	Number of fours scored in innings	`0`
`$sixes`	Numeric	Number of sixes scored in innings	`0`
`$mux`	Vector	$\mu(x)$ point estimates	`[1] 41.42246 [2] 58.43406 ...`
`$nut`	Numeric	$\nu(t)$ point estimate	`69.27811`
`$strike_rate`	Numeric	Career strike rate (runs scored per ball)	`0.72855`

Table 4.3. *Elements contained within a* `bowler` *object.*

Quantity	Type	Description	Initialised value
`$name`	Character	Player's name	`"TA Boult"`
`$ball_count`	Numeric	Total number of balls bowled in innings	`0`
`$overs`	Numeric	Number of overs bowled in innings	`0`
`$balls`	Numeric	Number of legal balls bowled in current over	`0`
`$maidens`	Numeric	Number of maiden overs bowled in innings	`0`
`$runs_conceded`	Numeric	Total runs conceded in innings	`0`
`$standardised_runs_conceded`	Numeric	Total standardised runs conceded in innings	`0`
`$wickets`	Numeric	Total wickets taken in innings	`0`
`$wt`	Numeric	$\omega(t)$ point estimate	`0.91231`
`$economy_rate`	Numeric	Career economy rate (runs conceded per ball)	`0.49617`

Here, the batsman on strike at the beginning of this match is Australian opener David Warner. In this particular simulation, Trent Boult has been selected to bowl the first over of the match for New Zealand, which is not a particularly surprising result, given the 46.0% probability assigned to him bowling the first over, as shown in Figure 4.4. It is possible to observe the simulation-specific estimates of player batting and bowling ability; in this case, Warner has been assigned an underlying effective batting average: $\nu(t) = 69.3$, with an estimate $\mu(0) = 41.4$, when on a score of 0. Meanwhile, for this simulation Boult has an underlying standardised effective bowling average: $\omega(t) = 0.91$.

The `batsman` and `bowler` objects are parsed to the `sim_ball()` function, which simulates the outcome of a single ball, based on the estimated abilities of the batsman and bowler. The process of simulating a ball can be split into two steps:

1. Simulate a `main_event`, which falls into one of three categories:

   (a) `no_wicket_event` — the batsman is not dismissed and runs may or may not be scored.

   (b) `bowler_wicket_event` — the batsman is dismissed by the bowler and no runs are scored.

   (c) `non_bowler_event` — a batsman is run out and runs may or may not be scored.

2. Given which of main events (a), (b), or (c) occurs, simulate a *sub-event*:

   (a) If main event (a) — simulate how many runs/extras are scored.

   (b) If main event (b) — simulate the type of dismissal.

   (c) If main event (c) — simulate which batsman is run out and how many runs/extras are scored.

The empirical probabilities of each main event and sub-event occurring are based on historic Test data and are provided on the following pages.

```
$main_events
              event      prob
  no_wicket         0.9838602
  bowler_wicket     0.0157042
  non_bowler_wicket 0.0004356
```

```
$no_wicket_event
  runs wicket_type extra_runs extra_type1 extra_type2      prob
     0                      0                         0.7252100
     1                      0                         0.1478854
     2                      0                         0.0369813
     3                      0                         0.0092685
     4                      0                         0.0600162
     5                      0                         0.0001634
     6                      0                         0.0039595
     7                      0                         0.0000029
```

```
$no_wicket_event (continued)
  runs wicket_type extra_runs extra_type1 extra_type2      prob
     0                      1        byes             0.0009386
     0                      2        byes             0.0004002
     0                      3        byes             0.0001083
     0                      4        byes             0.0014181
     0                      1     legbyes             0.0045560
     0                      2     legbyes             0.0007424
     0                      3     legbyes             0.0000754
     0                      4     legbyes             0.0009831
     0                      5     legbyes             0.0000068
     0                      1     noballs             0.0036163
     0                      2     noballs             0.0000232
     0                      3     noballs             0.0000106
     0                      5     noballs             0.0000367
     1                      1     noballs             0.0006699
     2                      1     noballs             0.0001798
     3                      1     noballs             0.0000454
     4                      1     noballs             0.0003461
     6                      1     noballs             0.0000126
     0                      1       wides             0.0020175
     0                      2       wides             0.0000377
     0                      3       wides             0.0000164
     0                      4       wides             0.0000019
     0                      5       wides             0.0002330
     0                      5     penalty             0.0000193
     1                      5     penalty             0.0000039
     0                      2        byes     noballs 0.0000010
     0                      5        byes     noballs 0.0000058
     0                      2     legbyes     noballs 0.0000039
     0                      5     legbyes     noballs 0.0000019
     0                      6     penalty       wides 0.0000010
```

```
$bowler_wicket_event
  runs        wicket_type extra_runs extra_type1 extra_type2      prob
     0             bowled          0                         0.1746609
     0            caught           0                         0.6101623
     0 caught and bowled          0                         0.0233769
     0         hit wicket          0                         0.0006056
     0                lbw          0                         0.1739947
     0            stumped          0                         0.0171996
```

```
$non_bowler_wicket_event
  runs        wicket_type extra_runs extra_type1 extra_type2      prob
     0            run out          0                         0.7598253
     1            run out          0                         0.1834061
     2            run out          0                         0.0414847
     3            run out          0                         0.0021834
     0            run out          1     legbyes             0.0043668
     0            run out          2     legbyes             0.0021834
     0            run out          1     noballs             0.0021834
     1            run out          1     noballs             0.0021834
     2            run out          1     noballs             0.0021834
```

In order to account for the relative batsman and bowler abilities for a given ball, several adjustments are made to the empirical main event and sub-event probabilities. The R code required to simulate the main and sub-event for a single delivery is rather lengthy. As such, the relevant code has been omitted here and is instead provided in Appendix A. However, the general procedure using the Warner/Boult matchup as an example is as follows:

1. Obtain an estimate for the *expected bowling average* (`expected_bowl_average`), given the abilities of the competing batsman and bowler. This is achieved using the estimates for batsman and bowler ability, $\mu(x)$ and $\omega(t)$, stored in the `batsman` and `bowler` objects. In this particular simulated match, Warner is estimated to have an underlying effective average, $\nu(t) = 69.27811$, while Trent Boult is estimated to have an underlying standardised effective bowling average $\omega(t) = 0.91231$. Note that the estimate for batting ability, $\mu(x)$, is conditional on the batsman's current score, which in this particular example is 0.

```
expected_bowl_average <- bowler$wt *
                         batsman$mux[batsman$current_score - 1]
```

Here, `expected_bowl_average` $\approx 0.91231 \times 41.42246 \approx 37.79033$, indicating Trent Boult is expected to concede approximately 37.8 runs to David Warner, on average, for every wicket he takes. This estimate assumes that Warner does not get his eye in any further and is updated each delivery as runs are scored.

2. Obtain an estimate for the bowler's *relative economy rate* (`relative_bowling_er`) by comparing the bowler's career economy rate with the historic Test economy rate.

```
relative_bowling_er <- bowler$economy_rate /
                       summary_data$averages$bowling_economy_rate
```

   Here, `relative_bowling_er` $\approx \frac{0.49617}{0.52057} \approx 0.95312$, indicating Trent Boult concedes fewer runs per ball on average, than the average runs conceded per ball across all Test matches.

3. Obtain an estimate for the batsman's *relative batting strike rate* (`relative_batting_sr`) by comparing the batsman's career batting strike rate with the historic Test batting strike rate.

```
relative_batting_sr <- batsman$strike_rate /
                       summary_data$averages$batting_strike_rate
```

   Here, `relative_batting_sr` $\approx \frac{0.72855}{0.50989} \approx 1.42886$, indicating David Warner scores more runs per ball on average, than the average runs scored per ball across all Test matches. This is indicative of Warner's aggressive batting style, which is one of his well-known traits.

4. Obtain estimates for the *expected number of runs conceded per ball* (`expected_rcpb`) and *expected number of runs scored per ball* (`expected_rspb`), given the specific batsman/bowler matchup. These quantities account for both the bowler's career economy rate and the batsman's career strike rate. Note that the expected runs conceded per ball includes extras such as wides and no balls conceded by the bowler, while the expected runs scored per ball only measures runs scored by the batsman and assumes that wides and no balls are bowled at a constant rate, regardless of the bowler.

```
## Expected runs conceded per ball by the bowler
expected_rcpb <- summary_data$averages$bowling_economy_rate *
                 relative_bowling_er * relative_batting_sr

## Expected runs scored per ball by the batsman
## Computing the difference between the historical runs
## conceded per ball and runs scored per ball implies that
## wides and no balls are bowled at the same rate,
## regardless of the bowler
expected_rspb <- expected_rcpb -
                 (summary_data$averages$runs_conceded_per_ball -
                  summary_data$averages$runs_scored_per_ball)
```

Here, `expected_rcpb` $\approx 0.52057 \times 0.95312 \times 1.42886 \approx 0.70895$, and `expected_rspb` ≈ 0.70083. This suggests Trent Boult is likely to concede more runs per ball than he has over the course of his career (0.49617), given Warner's generally aggressive approach to batting. However, this result also suggests Boult is likely to restrict Warner to a lower strike rate than his career strike rate (0.72855).

5. Obtain an estimate for the bowler's *expected bowling strike rate* (`expected_ bowler_sr`), given the specific batsman/bowler matchup, representing the average number of balls to be bowled, per wicket taken.

```
expected_bowler_sr <- expected_bowl_average / expected_rcpb
```

Here, `expected_bowler_sr` $\approx \frac{37.79033}{0.70895} \approx 53.30465$, indicating that Trent Boult is expected to dismiss David Warner every 53.3 deliveries, on average, assuming Warner does not get his eye any further.

6. Obtain an estimate for the *relative bowling strike rate* (`relative_bowler_sr`), given the specific batsman/bowler matchup, by comparing the expected bowling strike rate with the historic Test bowling strike rate.

```
relative_bowler_sr <- expected_bowler_sr /
                      summary_data$averages$bowling_strike_rate
```

Here, `relative_bowler_sr` $\approx \frac{53.30465}{63.2216} \approx 0.84314$, indicating that the expected bowling strike rate (`expected_bowler_sr`) for the proposed batsman/bowler matchup is lower than the historic Test bowling strike rate.

7. Adjust the empirical probabilities of each different main event occurring, to account for the relative abilities, strike rates, and economy rate of the specific batsman and bowler, using the quantities computed in (1) to (6).

   (a) Adjust the probability of a bowler-credited wicket event (`bowler_wicket_ event`) occurring, ensuring that wickets fall, on average, at the expected bowling strike rate (`expected_bowler_sr`) defined in (5). This is achieved by converting the empirical probability of a bowler-credited wicket occurring into odds, then, multiplying these odds by the inverse of the relative bowling strike rate (`relative_bowler_sr`). These odds are then backtransformed to obtain the adjusted probability of a bowler credited wicket occurring on a single ball, given the batsman/bowler-specific matchup. This has the effect of decreasing the likelihood of a bowler-credited wicket occurring if the relative bowling strike rate is greater than 1, and increasing the likelihood of a bowler-credited wicker occurring if the relative bowling strike rate is less than 1.

```r
## Initialise the probability of a bowler credited wicket
## occurring using the empirical data
bowler_wicket_prob <-
   subset(summary_data$prob_events$events,
          event == "bowler_wicket_event")$prob

## Convert probabilities into odds
bowler_wicket_odds <- bowler_wicket_prob / (1 - bowler_wicket_prob)
```

```r
## Adjust the odds using the relative bowling strike rate
bowler_wicket_odds <- bowler_wicket_odds * (1 / relative_bowler_sr)

## Convert odds back to probability to get the adjusted
## probability of a bowler-credited wicket event occurring
bowler_wicket_prob <- bowler_wicket_prob / (1 + bowler_wicket_odds)
```

The relative bowling strike rate computed in (6), `relative_bowler_sr` ≈ 0.84314, and the empirical probability of a bowler-credited wicket occurring on any given delivery is 0.0157, or 1.57%. After adjustment, the probability of a bowler-credited wicket event occurring is estimated to be equal to 0.0186, or 1.86%.

(b) After adjusting the probability of a bowler-credited event occurring, the total probability of all three different main events will no longer sum to 1. The assumption is made that the probability of a run out occurring (`non_bowler_wicket_event`) is constant, regardless of the batsman currently at the crease and the bowler in question. While it is possible that certain players may have an increased likelihood of being involved in a run out, the empirical probability of a run out occurring in the first place is so small that such effects are unlikely to have a meaningful impact on the outcome of a ball. Therefore, given the necessary condition defined in Equation 4.4, the probability of a non-bowler-credited wicket occurring is adjusted accordingly to satisfy this requirement.

$$P(\texttt{no\_wicket\_event}) + P(\texttt{bowler\_wicket\_event}) + P(\texttt{non\_bowler\_wicket\_event}) = 1 \tag{4.4}$$

```r
## Empirircal probability of a non-bowler-credited wicket
## event (run out) occurring
non_bowler_wicket_prob <-
   subset(summary_data$prob_events$events,
          event == "non_bowler_wicket_event")$prob
```

```r
## Adjust the probability of a non-wicket taking event
## occurring to ensure all main event probabilities
## sum to 1
no_wicket_prob <- bowler_wicket_prob - non_bowler_wicket_prob
```

Therefore, the adjusted probabilities of each of the main events `no_wicket`, `bowler_wicket`, and `non_bowler_wicket` occuring are 0.9810, 0.0186, and 0.0004 respectively, or, 98.10%, 1.86%, and 0.04%.

8. Adjust the probabilities of the run scoring sub-events contained in the `no_wicket_event` object.

   (a) Compute the `adjustment_factor`. This quantity represents the factor by which the probability of run scoring sub-events occurring must be multiplied, to ensure that on average: (i) the bowler will take wickets at an expected bowling average (`expected_bowl_avarage`) equivalent to the quantity defined in (1); (ii) runs will be conceded by the bowler at a rate equivalent to the expected number of runs conceded per ball (`expected_rcpb`), defined in (4); and (iii) runs will be scored by the batsman at a rate equivalent to the expected number of runs scored per ball (`expected_rspb`), defined in (4).

```r
## Before adjusting the run scoring sub-events, this many
## runs are expected to be scored per ball by the batsman
current_rspb <- sum(no_wicket_prob * no_wicket_event$runs *
                        no_wicket_event$prob) +
                sum(non_bowler_wicket_prob *
                        non_bowler_wicket_event$runs *
                        non_bowler_wicket_event$prob)

## Compute the adjustment factor
adjustment_factor <- expected_rspb / current_rspb
```

   To compute the `adjustment_factor`, one must first calculate the expected number of runs scored per ball by the batsman using the empirical sub-event probabilities and the adjusted main event probabilities, defined in (7). Assuming that the rate at which wides and no balls are bowled is constant, regardless of the bowler, the `adjustment_factor` is then equal to $\frac{\texttt{expected\_rspb}}{\texttt{current\_rspb}}$.
   Here, the `adjustment_factor` $\approx \frac{0.70083}{0.50724} \approx 1.38165$.

(b) Adjust the probability of run scoring sub-events accordingly.

```
## Use the adjustment factor to adjust probabilities of
## run  scoring sub-events occurring, conditional on
## batting strike rate and bowling economy rate
no_wicket_event[no_wicket_event$runs > 0, ]$prob <-
   no_wicket_event[no_wicket_event$runs > 0, ]$prob *
   adjustment_factor

## Normalise the sub-event probabilities to sum to 1
no_wicket[no_wicket$runs == 0, ]$prob <-
   no_wicket[no_wicket$runs == 0, ]$prob /
   sum(no_wicket[no_wicket$runs == 0, ]$prob) *
   (1 - sum(no_wicket[no_wicket$runs > 0, ]$prob))
```

Given the `adjustment_factor` ≈ 1.38, the likelihood each of the run scoring sub-events within the `no_wicket_event` object occuring are increased by a factor of 1.38 for this ball.

How this process accurately accounts for various factors, such as batting and bowling abilities; strike rates; and economy rates, is not immediately intuitive. Several assumptions are made, such as the constant rate at which wides and no balls are bowled, regardless of the bowler, and the rate at which runs out occur, regardless of the batsmen. Admittedly, a number of the adjustments made to the probabilities are heuristic in nature, however, it is possible to show via a simulation study that this approach, on average, does result in simulations that converge to expected results.

From this point, the simulation of a ball is straightforward; a main event and corresponding sub-event are simulated using the adjusted probabilities computed using the process defined in (1) to (8). The result of a simulated ball bowled by Trent Boult to David Warner is computed below using the `sim_ball()` function, defined in Appendix A. A logical, user-defined argument, `ball_output`, provides a user with the option to print the results of a the ball to be to screen. Here, the simulated ball bowled by Trent Boult to David Warner, has resulted in Warner scoring 1 run.

General logic checks are then performed to test for various scenarios. For example: Was the ball a legal delivery? Did the batsmen rotate the strike? If there was a run out, which batsman is out? The relevant output is then used to update the `batsman` and `bowler` objects accordingly, to prepare the simulation for the next simulated ball. The results of each delivery are ultimately uploaded to the respective batting and bowling `scorecard` objects, contained within the `innings_data` object. In a batting context, information such as runs scored, balls faced and strike rate is stored, while in a bowling context the number of overs bowled, runs conceded, standardised runs conceded and wickets taken is recorded.

```
## Simulate the first ball of the proposed Australia vs. NZ match,
## bowled by Trent Boult to David Warner
sim_ball(summary_data, innings_data, ball_output = TRUE)

TA Boult to DA Warner. 1 run.

$simulated_main_event
    event    prob
no_wicket 0.98099

$simulated_sub_event
  runs wicket_type extra_runs extra_type1 extra_type2    prob
     1                      0                         0.20433
```

```
match_data$innings_data[[1]]$batting$scorecard
      batsman runs BF    SR   out how_out bowler
    DA Warner    1  1 100.0 FALSE
     JA Burns    0  0   0.0 FALSE
M Labuschagne    0  0     0
    SPD Smith    0  0     0
      MS Wade    0  0     0
      TM Head    0  0     0
     TD Paine    0  0     0
   PJ Cummins    0  0     0
     MA Starc    0  0     0
      NM Lyon    0  0     0
 JR Hazlewood    0  0     0
       extras    0
        total    1
```

```
match_data$innings_data[[1]]$bowling$scorecard
  bowler overs balls maidens runs standardised_runs wickets
TA Boult     0     1       0    1           0.02414       0
```

Over simulation

With the simulation process of an individual ball or delivery now defined, the simulation of an over — consisting of six legal balls — is relatively easy. The ball simulation process is simply repeated until either: (1) six legal balls are bowled, (2) the bowling team takes 10 wickets, concluding the batting team's innings, or (3) the batting team reaches the target score, concluding the match. A single over is simulated using the `sim_over()` function (defined in Appendix A), which includes a user-defined argument that allows the output of a simulated over to be printed to screen, including the bowler's figures and batsman scores at the conclusion of the over.

Here, the first over of the simulated match between Australia and New Zealand has been rather eventful. David Warner was able to get off the mark immediately, Joe Burns was bowled out for a duck and Marnus Labuschange scored a boundary from his first delivery faced. The batting and bowling `scorecard` objects are then updated accordingly to reflect the events of the simulated over. Note that the bowling `scorecard` object records the number of runs conceded in both units of runs and standardised runs conceded, which allows for a fairer and more accurate means of comparing simulated bowling performances in the post hoc analysis of the entire simulation process.

```
## Simulate the first over of the proposed Australia vs. NZ match,
## to be bowled by Trent Boult
sim_over(match_data,
         match_parameters,
         innings_data,
         over_output = TRUE,
         ball_output = TRUE)

TA Boult to DA Warner. 1 run.
TA Boult to JA Burns. 0 runs + 2 leg byes.
TA Boult to JA Burns. 0 runs.
TA Boult to JA Burns. 0 runs.
TA Boult to JA Burns. 0 runs. OUT, bowled.
TA Boult to M Labuschange. 4 runs.
```

```
Overs bowled = 1. Runs scored from over = 7. Score = 7/1.
DA Warner 1, M Labuschagne 4.
TA Boult: 1/5 (1)
```

```
match_data$innings_data[[1]]$batting$scorecard
      batsman runs  BF     SR   out  how_out    bowler
    DA Warner    1   1  100.0 FALSE
     JA Burns    0   4    0.0  TRUE   bowled  TA Boult
M Labuschagne    4   1  400.0 FALSE
    SPD Smith    0   0      0
      MS Wade    0   0      0
      TM Head    0   0      0
     TD Paine    0   0      0
   PJ Cummins    0   0      0
     MA Starc    0   0      0
      NM Lyon    0   0      0
 JR Hazlewood    0   0      0
       extras    2
        total    7
```

```
match_data$innings_data[[1]]$bowling$scorecard
  bowler overs balls maidens runs standardised_runs wickets
TA Boult     1     0       0    5           0.09845       1
```

Innings simulation

Now, with the simulation process of both individual balls and overs being clearly defined, the simulation of an entire innings is not difficult. Overs are simulated until: (1) the bowling team has taken 10 wickets, concluding the batting team's innings, (2) the batting team reaches the target score, concluding the match, or (3) the total number of overs to be bowled in the match is reached, concluding the match in a draw.

There is one further scenario in which a team's innings can be concluded. In Test cricket, it is occasionally advantageous to the batting team to end their innings prematurely — known as *declaring* an innings — in order to achieve victory before a match is deemed drawn after 5 days without a result (or in the case of the match simulator, 450 overs). To mimic the concept of declaring an innings, some logic has been built in to the innings simulation process, allowing for declarations in the third innings in cases where the batting team would set a significant and presumably unattainable target to win the match, given the number of overs remaining. This prevents the batting team in the simulated match from continuing to bat and wasting time that could otherwise be used to bowl the opposition out and win the match.

The method used to implement declarations is fairly rudimentary. Given the number of overs remaining in a match, subjective estimates were made in regards to the minimum lead a team batting in the third innings would require before considering a declaration. Such estimates were made conditional on 10, 45, 90, 135, 180, 225 and 270 overs remaining in a match, with the lower and upper limits for these subjective estimates presented in Figure 4.5. The estimates are based on both the author's extensive experience of spectating Test cricket and historical Test records. For example, the highest target ever set is 742, imposed by England against Australia in 1928, therefore, the upper limit for the lead a team would require before considering a declaration with 270 overs remaining (3 days), is set at 750. Similarly, as the highest target successfully chased in the fourth innings is 418 by the West Indies versus Australia in 2003, the lower limit with 270 overs remaining is estimated to be a lead of 450 runs, which history would suggest is a large enough lead to guarantee victory.

In order to provide a framework to the logic governing declarations, the linear equation defined in Equation 4.5 is fitted to define the decision boundary for a third innings declaration, conditional on the current lead and number of overs remaining in the match. This function errs on the side of conservatism and rather crude, however, it does at least result in somewhat sensible declarations in simulated matches where one team is particularly dominant.

$$P(\text{Declaration}) = \begin{cases} 0, & \text{if current lead} < 150 + 2 \times \text{overs remaining} \\ 1, & \text{if current lead} \geq 150 + 2 \times \text{overs remaining} \end{cases} \tag{4.5}$$

In reality, a captain's decision to declare an innings is based on a far greater number of external factors beyond the number of overs remaining in a match. The amount of assistance the pitch is likely to offer bowlers and the number of runs scored in the match thus far, are both of worthwhile consideration, as is the forecasted weather, which can be unpredictable at the best of times. If the outcome is being predicted for a match that has already commenced, it is possible to recode the declaration logic to account for the state of the pitch and upcoming weather. Often, commentators and seasoned spectators alike will be able to get a feel for when a declaration is coming and will generally be in the ballpark in terms of the current lead and number of overs remaining at the time of a declaration.

Considering the various conditions that can lead to the conclusion of a simulated innings, logical checks are made at the end of each over to ensure none of these conditions have been met. A new bowler is then selected to bowl the next over, using the bowler selection logic defined when the simulation was initialised. This method of bowler selection can be a little unrealistic compared to what generally occurs in the Test cricket arena. Given the way the bowling logic has been defined, bowlers will rarely bowl long spells of seven or eight consecutive overs. While this does not have a significant effect on the likely overall outcome of a match, it does detract from the realism of the simulation. There may be a slight bias introduced as a result of the

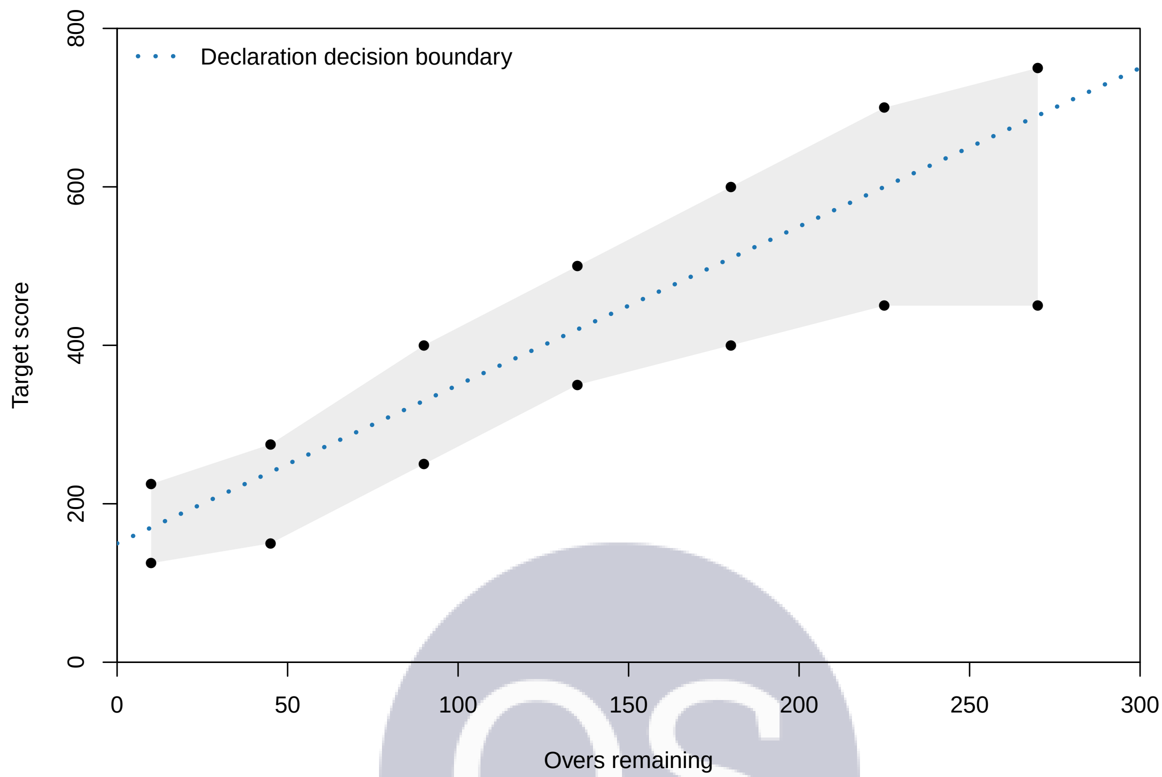


Figure 4.5. *Subjective estimates for the minimum lead required before the team batting in the third innings would consider a declaration. The grey area depicts the range of values in which team's may consider declaring their third innings, conditional on their current lead and the number of overs remaining. The blue line represents the conservative declaration decision boundary that loosely fits these subjective assessments.*

bowler selection method, as there is no means of recording bowler fatigue. It is entirely possible that pace bowlers will be simulated to bowl unrealistically long spells of more than 10 overs, which rarely occurs in reality. However, while this may happen on the odd occasion, it should not have a particularly large impact across a large number of simulated matches.

An innings can be simulated using the `sim_innings()` function, which again has a user-defined option to print the results of the simulated innings to screen. Finally, an entire match can then be simulated using the `sim_match()` function, which employs the use of the `sim_ball()`, `sim_over()`, and `sim_innings()` functions that have been defined in Appendix A. The result of the match is then stored inside a simulation object, `sim`, containing the `scorecard` objects and match results for all simulated matches. The resulting output from the four simulated innings for the proposed matchup between Australia and New Zealand is provided below, along with the relevant batting and bowling `scorecard` objects.

```r
## Simulate the proposed Australia vs. NZ match
sim_match(summary_data, match_players, bowling_logic, iteration = 1,
          match_output = TRUE, innings_output = TRUE,
          over_output = FALSE, ball_output = FALSE)
First innings.
Batting team: Australia. Score: 300/10 (86.2)

Second innings.
Batting team: New Zealand. Score: 298/10 (106.4)

Third innings.
Batting team: Australia. Score: 331/10 (107.0)

Fourth innings.
Batting team: New Zealand. Score: 121/10 (40.1)

Result: Australia win by 212 runs.
```

```
## First innings batting scorecard
match_data$innings_data[[1]]$batting$scorecard
       batsman runs  BF    SR   out  how_out      bowler
     DA Warner    5   4 125.0  TRUE   bowled  TG Southee
      JA Burns    0   4   0.0  TRUE   bowled    TA Boult
M Labuschagne   65 126  51.6  TRUE   caught    TA Boult
     SPD Smith  126 210  60.0  TRUE  run out
       MS Wade    6   9  66.7  TRUE   caught    TA Boult
       TM Head    0  10   0.0  TRUE   caught  TG Southee
      TD Paine    0   6   0.0  TRUE   caught  MJ Santner
    PJ Cummins    4   6  66.7  TRUE      lbw  TG Southee
      MA Starc   56  95  58.9  TRUE      lbw    TA Boult
       NM Lyon    2   4  50.0  TRUE   caught    N Wagner
  JR Hazlewood   28  49  57.1 FALSE
        extras    8
         total  300
```

```
## First innings bowling scorecard
match_data$innings_data[[1]]$bowling$scorecard
          bowler overs balls maidens runs standardised_runs wickets
        TA Boult    22     2       3   74           2.40981       4
      TG Southee    16     0       3   49           1.29413       3
C de Grandhomme     14     0       1   47           0.78477       0
        N Wagner    19     0       1   69           1.89104       1
      MJ Santner    15     0       3   54           1.41593       1
```

In the first innings, Australia is simulated to be bowled out for 300, after 86.2 overs. Only four Australians managed to make it to double figures, with a top score of 126 by Steve Smith and handy contributions down the order from Mitchell Starc and Josh Hazlewood. Pace bowlers Trent Boult and Tim Southee are the star performers here for New Zealand, finishing with figures of 4/74 and 3/49 respectively.

```
## Second innings batting scorecard
match_data$innings_data[[2]]$batting$scorecard
         batsman runs  BF   SR   out  how_out       bowler
      TWM Latham    1   8 12.5  TRUE   caught     MA Starc
     TA Blundell   11  30 36.7  TRUE      lbw  JR Hazlewood
   KS Williamson   82 143 57.3  TRUE   caught     MA Starc
     LRPL Taylor    0   5  0.0  TRUE   bowled     MA Starc
     HM Nicholls    1   5 20.0  TRUE      lbw      NM Lyon
C de Grandhomme    32  37 86.5  TRUE   caught  JR Hazlewood
      BJ Watling   77 252 30.6  TRUE   caught      NM Lyon
      MJ Santner   29  99 29.3  TRUE   caught  JR Hazlewood
      TG Southee    4   5 80.0  TRUE   caught  JR Hazlewood
        N Wagner   25  29 86.2  TRUE   caught     MA Starc
        TA Boult   12  29 41.4 FALSE
        extras     24
         total    298
```

```
## Second innings bowling scorecard
match_data$innings_data[[2]]$bowling$scorecard
       bowler overs balls maidens runs standardised_runs wickets
 JR Hazlewood    22     0       6   67           3.75268       4
     MA Starc    21     0       2   79           2.79865       4
   PJ Cummins    25     0       6   62           2.58381       0
      NM Lyon    19     4       9   41           1.83603       2
M Labuschagne     8     0       2   23           0.73933       0
    SPD Smith     1     0       0    9           0.44601       0
```

In the second innings, New Zealand are simulated to reply with a score of 298 all out after 106.4 overs, behind scores of 82 and 77 from Kane Williamson and BJ Watling, leaving them trailing by just two runs at the halfway point of the match. Opening bowlers Josh Hazlewood and Mitchell Starc have done most of the damage in this innings for Australia, each capturing four wickets.

```r
## Third innings batting scorecard
match_data$innings_data[[3]]$batting$scorecard
```

```
       batsman runs  BF    SR   out  how_out          bowler
     DA Warner   18  28  64.3  TRUE   caught C de Grandhomme
      JA Burns  113 204  55.4  TRUE   caught C de Grandhomme
M Labuschagne   17  38  44.7  TRUE   bowled        N Wagner
     SPD Smith   37 100  37.0  TRUE   caught      TG Southee
       MS Wade   42  81  51.9  TRUE  run out
       TM Head    1  10  10.0  TRUE      lbw        N Wagner
      TD Paine   23  39  59.0  TRUE   caught      MJ Santner
    PJ Cummins   34  83  41.0  TRUE   bowled C de Grandhomme
      MA Starc   23  21 109.5  TRUE   caught      MJ Santner
       NM Lyon    2  28   7.1  TRUE   caught      MJ Santner
  JR Hazlewood    7  13  53.8 FALSE
        extras   14
         total  331
```

```r
## Third innings bowling scorecard
match_data$innings_data[[3]]$bowling$scorecard
```

```
          bowler overs balls maidens runs standardised_runs wickets
        TA Boult    21     0       3   61           3.52618       0
C de Grandhomme    15     0       6   25           0.68892       3
      TG Southee    23     0       4   73           1.91672       1
        N Wagner    23     0       2   92           3.37100       2
   KS Williamson     1     0       0    6           0.19223       0
      MJ Santner    24     0       4   64           2.48146       3
```

A solid third innings total of 331 is simulated for Australia, leaving New Zealand a significant, but attainable target of 334 runs to win the match. The Australian innings is underpinned by a century from opening batsman Joe Burns, who bounces back after being dismissed for a duck in the first innings. Colin de Grandhomme and Mitchell Santner star with the ball here, each picking up three wickets.

```
## Fourth innings batting scorecard
match_data$innings_data[[4]]$batting$scorecard
         batsman runs  BF    SR   out  how_out      bowler
      TWM Latham   25  55  45.5  TRUE   caught     NM Lyon
     TA Blundell    7  18  38.9  TRUE   caught JR Hazlewood
   KS Williamson    0   2   0.0  TRUE   caught    MA Starc
     LRPL Taylor   34  32 106.3  TRUE   caught JR Hazlewood
     HM Nicholls   16  47  34.0  TRUE      lbw JR Hazlewood
C de Grandhomme     2   5  40.0  TRUE   bowled    MA Starc
      BJ Watling   19  43  44.2  TRUE   caught    MA Starc
      MJ Santner    5  20  25.0  TRUE      lbw     NM Lyon
      TG Southee    0   1   0.0  TRUE      lbw    MA Starc
        N Wagner    7  18  38.9 FALSE
        TA Boult    0   1   0.0  TRUE   caught  PJ Cummins
        extras    6
         total  121
```

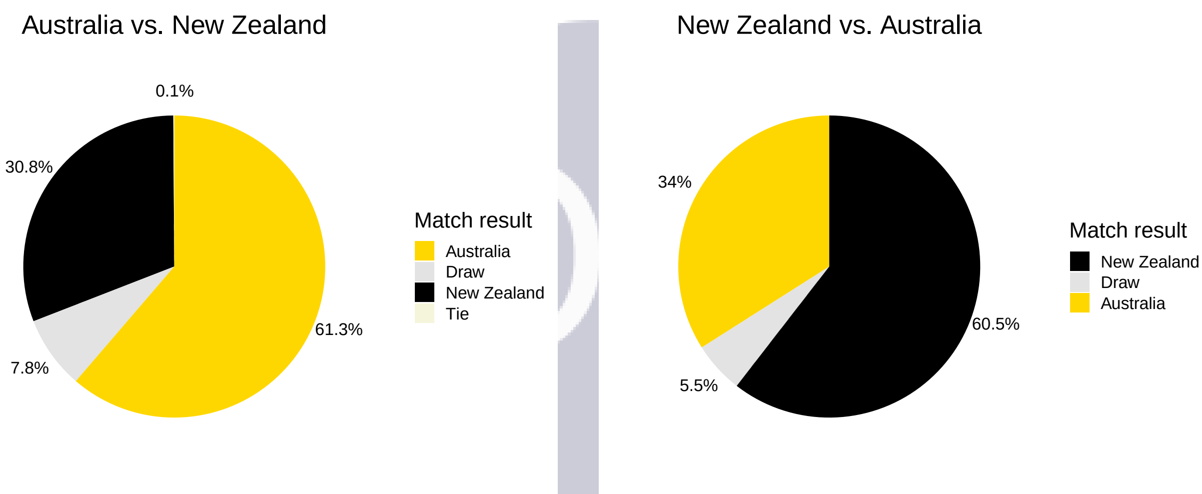
```
## Fourth innings bowling scorecard
match_data$innings_data[[4]]$bowling$scorecard
       bowler overs balls maidens runs standardised_runs wickets
 JR Hazlewood    10     0       3   20           1.22214       3
   PJ Cummins    13     0       2   39           1.43100       1
     MA Starc    10     0       2   34           2.94788       4
      NM Lyon     7     1       2   24           1.12086       2
```

The fourth innings sees New Zealand collapse for 121, losing the match by 212 runs. There is certainly no denying that this simulated match has produced an objectively realistic looking result. In fact, despite being a simulated match, for patriotic New Zealand fans this result is eerily similar to a number of recent real-life performances against trans-Tasman rivals Australia, where New Zealand appears to be competitive for the majority of the match, only to collapse spectacularly at the final hurdle.

4.2.3 Simulation summary

The procedure of simulating a single match, detailed in Section 4.2.2, is repeated multiple times in order to gain a deeper understanding of possible match outcomes. The results of each of these simulations are saved in a simulation object, including the performances of individual players in each of the simulated matches. It is then possible to post-process the results of the simulated matches, to gain a deeper understanding of the likely performances of each player. Here, the simulated results for the proposed match between Australia and New Zealand are discussed. In order to estimate the effect home ground advantage may have on the matchup, the simulation is run twice, with each nation given the opportunity to play as the home and away team.

Of primary interest is the most likely outcome of the match, given the two proposed playing XIs. Here, the matchup is simulated 1,000 times with Australia as the home nation, and 1,000 times with New Zealand as the home nation. The summary of the simulated results is provided in Figure 4.6 and suggests that home ground advantage is indeed a significant factor in this matchup. It goes without saying that these estimates are subject to some uncertainty, given they are obtained from a simulation-based process. However, re-running the simulator with 1,000 iterations typically yields estimates within approximately 2% of one another.



(a) **(b)**

Figure 4.6. *(a) Summary of simulated match outcomes with Australia as the home nation. (b) Summary of simulated match outcomes with New Zealand as the home nation.*

Overall, Figure 4.6 suggests Australia is the slightly stronger team, given the marginally higher win probability estimates across the two scenarios, although on the whole, this appears to be a fairly even contest. Interestingly, one of the 1,000 simulated matches with Australia as the home nation ended in a tie, an incredibly rare result, which has only occurred twice since the first Test was played in 1877. It is worth noting that the estimated probability of a draw occurring is likely lower than the actual likelihood of a draw, as these estimates ignore variables such as the weather, which is usually the main contributing factor in cases where Test matches end without a victor.

Table 4.4. *Summary of simulated batting performances for each player in the proposed Australia versus New Zealand matchup. The first section of results summarises simulated batting performances where the home team is Australia, while the second section of results summaries simulated batting performances where the home team is New Zealand. The expected batting average for each player in the proposed matchup provides an indication of how well each player is expected to perform with the bat, given the bowlers they are likely to face, while the simulated average summarises each player's actual batting performances across all simulated matches and illustrates the amount of noise present in the simulation process. The posterior predictive estimate for each player's effective batting average, $\nu(t)$, excluding any innings and venue-specific effects, is also provided for reference.*

		AUS vs. NZ		NZ vs. AUS	
Batsman	$\boldsymbol{\nu(t)}$	**Expected**	**Simulated**	**Expected**	**Simulated**
David Warner	47.9	49.2	48.1	29.2	29.9
Joe Burns	35.0	26.4	26.1	25.2	24.9
Marnus Labuschange	55.8	50.5	48.0	40.5	37.9
Steve Smith	57.9	47.7	49.5	42.7	42.9
Matthew Wade	34.1	28.3	27.9	25.3	24.9
Travis Head	38.1	33.7	32.7	24.9	24.6
Tim Paine	30.6	23.1	22.9	23.0	23.6
Pat Cummins	16.3	13.2	12.9	9.7	9.6
Mitchell Starc	23.3	17.9	17.6	16.6	15.8
Nathan Lyon	13.0	11.2	10.8	9.6	9.5
Josh Hazlewood	13.5	9.9	10.0	8.7	8.2
Tom Latham	40.8	26.8	26.4	34.0	33.5
Tom Blundell	43.6	35.0	34.0	34.9	34.3
Kane Williamson	47.8	31.5	31.8	37.1	37.3
Ross Taylor	43.5	30.1	29.6	38.0	37.7
Henry Nicholls	37.7	24.6	24.5	31.8	31.1
Colin de Grandhomme	34.0	22.5	22.5	34.8	35.5
BJ Watling	38.1	27.3	26.3	27.5	26.7
Mitchell Santner	25.4	17.1	17.4	23.1	22.4
Tim Southee	17.8	13.4	13.6	16.1	15.8
Neil Wagner	12.7	8.9	9.1	10.7	10.7
Trent Boult	18.4	10.5	10.5	15.6	15.2

Table 4.5. *Summary of simulated bowling performances for players in the proposed Australia versus New Zealand matchup who are expected to have a bowling workload $\geq 2\%$. The first section of results summarises simulated bowling performances where the home team is Australia, while the second section of results summaries simulated bowling performances where the home team is New Zealand. The expected bowling average for each player in the proposed matchup provides an indication of how well each player is expected to perform with the ball, given the batsmen they are likely to bowl to, while the simulated average summarises each player's actual bowling performances across all simulated matches and illustrates the amount of noise present in the simulation process. The posterior predictive estimate for each player's adjusted effective bowling average, $\alpha(t)$, excluding any innings and venue-specific effects, is also provided for reference.*

		AUS vs. NZ		**NZ vs. AUS**	
Bowler	$\boldsymbol{\alpha(t)}$	**Expected**	**Simulated**	**Expected**	**Simulated**
Marnus Labuschange	45.2	35.3	34.8	38.5	39.5
Steve Smith	53.9	48.9	59.7	41.5	52.6
Travis Head	55.9	38.5	44.4	50.3	54.3
Pat Cummins	25.1	20.2	20.0	25.9	25.9
Mitchell Starc	29.8	23.1	23.6	29.9	29.3
Nathan Lyon	33.8	29.1	28.9	33.2	33.6
Josh Hazlewood	28.0	21.1	21.0	27.1	25.7
Colin de Grandhomme	31.1	30.4	31.5	29.2	29.6
Mitchell Santner	44.4	53.2	50.6	34.3	33.1
Tim Southee	27.7	28.2	27.5	22.1	22.3
Neil Wagner	25.8	25.3	24.7	22.7	22.3
Trent Boult	29.7	31.5	31.1	23.6	22.8

As discussed in Section 4.2.2, the simulation process saves the individual player performances in every simulated match. Averaging these performances across all simulations provides an indication of the batting and bowling performances one should expect from each player, which are presented in Tables 4.4 and 4.5. The expected average signifies how well the player is expected to perform in the match, on average, while the simulated average summarises the player's simulated performances. Comparing the differences between expected and simulated averages provides a rough indication of the amount of variation present in the simulation process. Over a large enough sample of simulations, the expected and simulated averages should converge with one another.

It is clear from the results in Tables 4.4 and 4.5 that the majority of players in the proposed matchup prefer playing in a home environment. This appears to be particularly true for a number of Australian batsmen and is universal across all bowlers, with the possible exception of Steve Smith and his part-time leg spin bowling. The results appear to suggest that Australia are particularly reliant on batting performances from the trio of David Warner, Marnus Labuschagne, and Steve Smith — all players who were ranked in the current world top 10 Test batsmen in Chapter 2 — whereas New Zealand appears to have a more balanced batting lineup and may be more likely to see meaningful batting contributions from a larger number of players.

Furthermore, it is interesting to note that the majority of the predicted expected batting averages are lower than the posterior predictive estimates for a player's effective batting average, $\nu(t)$. However, after considering the quality of bowlers participating in the match, this result becomes less surprising. Each side boasts three world-class pace bowlers, who in Chapter 3 were ranked among top 20 bowlers in the world at present. As a result, the simulation results indicate — perhaps rightly so — that this matchup is likely to be dominated by the bowlers.

4.3 Practical applications

While the simulation results presented in Section 4.2.3 may be of interest to the general cricketing community, it is worthwhile to consider how this type of output can add value to one's cricket viewing experience, or, used to gain a competitive advantage over an opposition. As such, there are two main areas where the match simulator has been identified to have practical implications: (1) a broadcasting or public interest perspective, and (2) a high performance or private interest perspective, which are discussed in Sections 4.3.1 and 4.3.2 respectively.

4.3.1 'Who's winning?'

At some point of their lives, many fans of Test cricket will have been asked the exasperating question: 'who's winning?', by an earnest bystander attempting to feign interest in a sporting contest that can drag on for five days, with the players breaking for a drink and meal every couple of hours. Such a question often results in the inquisitor being chastised with a drawn-out sermon, detailing the various intricacies of Test cricket, hypothetical scenarios, and why there is no simple way of determining which team is going to win this early on in a match. In this regard, the author speaks from a position of great experience and authority. The output obtained from the match simulator attempts to provide some clarity to this question and to save onlookers the embarrassment of making such an egregious query.

As discussed in Section 4.2, the match simulator can be used to provide a pre-match prediction in an attempt to gain an understanding of what the most likely outcome of a match may be. In Section 4.2.3, Australia were predicted to have a pre-match probability of winning of roughly

61% in matches that were to be played in Australia, as shown in Figure 4.6. Already, a less knowledgeable observer can infer that Australia are the favourites to win the match, given their playing XI, the opposition team and their home ground advantage.

Now, consider the particular simulation discussed in Section 4.2.2. Here, after one day (90 overs) of simulated play, Australia had been bowled out for 300 and New Zealand were 12/1 in reply. At the close of play on day two of the simulated match, New Zealand were on a score of 240/8, trailing Australia by 60 runs. By the end of day three, Australia had reached a score of 236/6, leading by 296 runs. The match was then won by Australia in the latter stages of day four.

It is possible to re-run the simulator at any given point of a match that has commenced in order to obtain up-to-date estimates of the most likely outcome, given the current state of play. This was done at the end of each day's play in the example simulation presented in Section 4.2.2. The results shown in Figure 4.7 provide a visualisation of how the likelihood of each potential match outcome can vary over the course of a match. At the close of play on day one, Australia's probability of winning was marginally higher than the pre-match prediction, suggesting they had ever so slightly outperformed New Zealand. By the end of day two, New Zealand had improved their odds of winning since the start of the day, indicating they had taken the honours for the day's play. However, day three belonged to Australia, who managed to build a lead of 296 runs, corresponding to an 86% predicted probability of winning the match. The match was then won by Australia on day four.

Such simulations can be run and re-run throughout the course of a match, at the end of each session or day's play, and after significant match events, such as wickets of key batsmen. By incorporating this type of analysis into the TV broadcast of a Test match, viewers are provided with an idea of which team has the momentum, how the match is currently poised, and at least a partial answer to the age-old question of 'who's winning?'. This is in essence, exactly what is starting to happen with broadcasters across the globe, who are beginning to partner with cricket analytics companies, such as CricViz, to provide data-driven insights during a Test match.

As the match simulator also provides an estimate of the expected performance for each player, broadcasters can also use the output to generate talking points among commentators. For example, when a player walks out to bat, or comes on to bowl, their career record is often displayed on screen. Including information such as their current estimated effective batting and bowling averages, or expected performance in the match allows for the identification of players who are in or out of form. While Chapters 2 and 3 have shown that it is very difficult to predict a player's performance in a match, due to the significant amount of noise associated with batting and bowling performances, such auxiliary information may at least provide a storyline around a player's match performance, which may add to the overall viewing experience.

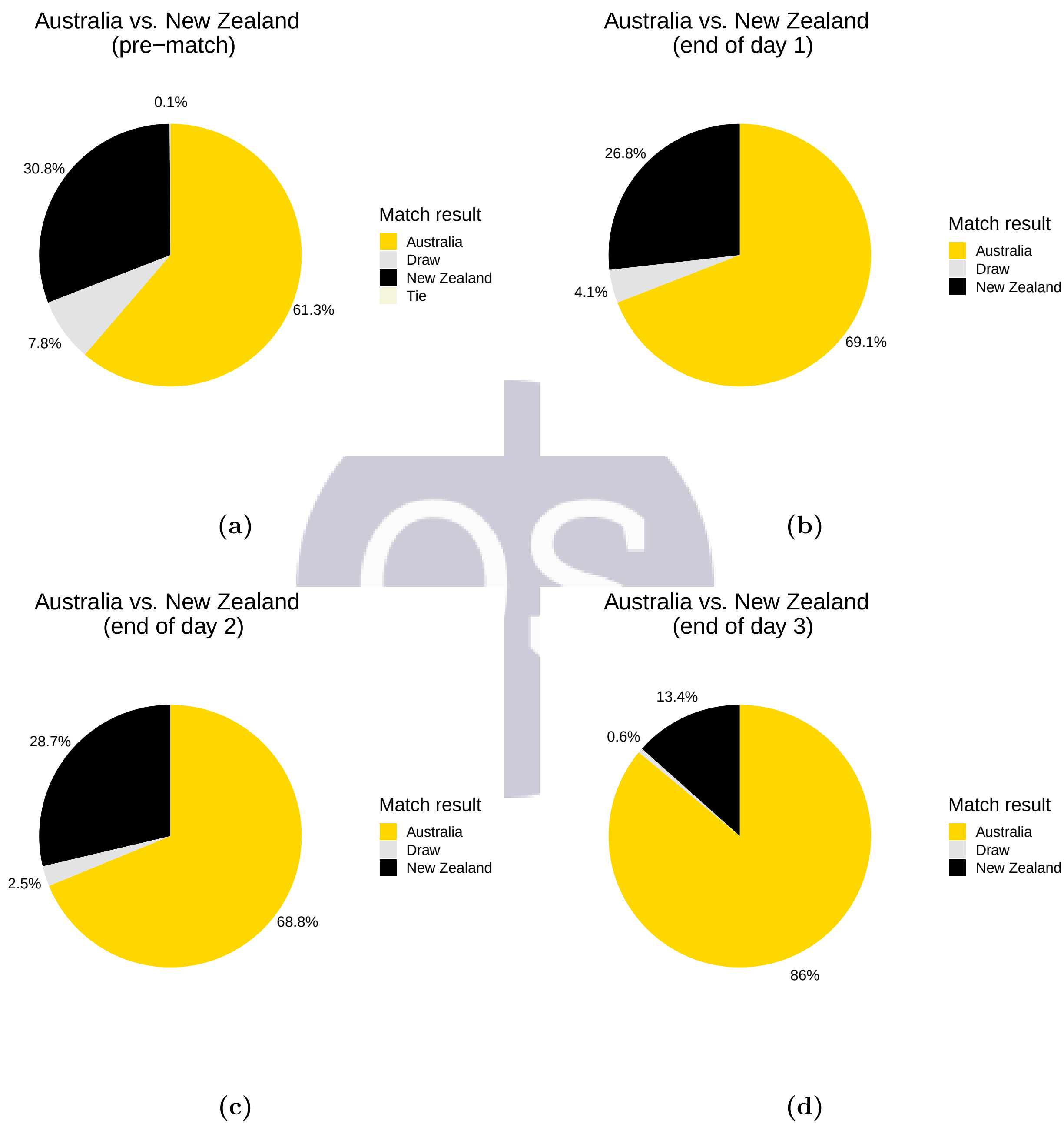


Figure 4.7. *(a) Pre-match predictions. (b) End of day one predictions. (c) End of day two predictions. (d) End of day three predictions.*

4.3.2 Gaining a competitive advantage

While knowing the potential outcome of a match may provide an interesting talking point, such results are of little use to professional teams, unless the findings can be used to gain a competitive advantage over an opposition. Therefore, the simulation output detailing the expected performances of individual players may be of interest to high performance units of both domestic and national teams, allowing for coaches and selectors to gain a deeper understanding of how well certain players are expected to play, given a proposed matchup. Such results may also provide teams with a means of quantifying the risks and rewards of selecting certain players over others, which may be particularly useful in situations where there are several viable players to choose from in an upcoming match.

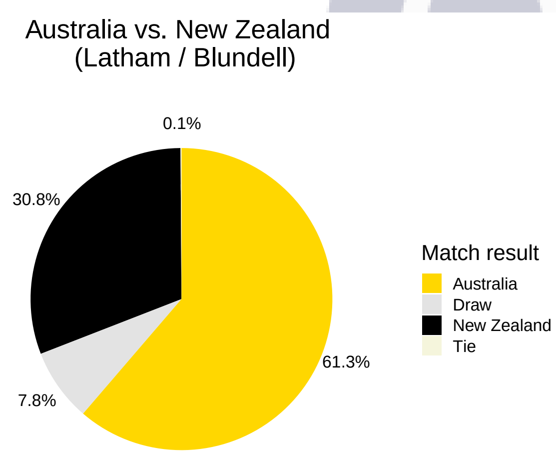
For example, consider the proposed playing XI for New Zealand introduced in Table 4.1, which was used in the simulation process throughout Section 4.2. Of the 11 selected players, nine have been fairly regular selections over the last few years. However, there are two spots within the New Zealand side that remain up for grabs, as they have done for several years, namely (1) the role of Tom Latham's opening batting partner, and (2) the role of primary spin bowler. Numerous individuals have attempted to stake their claim in each of these positions, although no player is yet to provide any conviction that they are the long term solution.

Even in 2017, New Zealand's ongoing issues at the top of the batting order were well documented, with several unflattering predictions made regarding the potential batting abilities of the next opening candidate (Stevenson & Brewer, 2017; Stevenson, 2017). At the time of publication, Jeet Raval had made promising start to his Test career, was outperforming all projections of ability made in Stevenson & Brewer (2017) and Stevenson (2017), and appeared to be a potential long-term solution to partner Latham. Unfortunately, as noted in Chapter 2, Raval has since experienced a change in fortunes, with his performances regressing to what was forecast at the beginning of his career; a bittersweet moment for the authors — the model has accurately predicted Raval's underlying Test match batting ability — but at what cost? The question of who will open the batting with Tom Latham has returned. At present, Tom Blundell has taken up the mantle as Latham's newest opening partner and despite making a successful start to his career, only time will tell if he is the solution to New Zealand's opening woes.

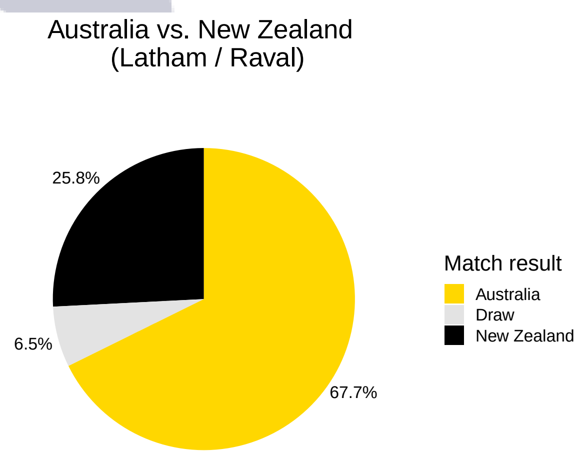
In a similar vein, New Zealand has been unable to find a suitable candidate to fulfil the role of primary spin bowler since the retirement of Daniel Vettori in 2014, who had been the incumbent since 1997. Again, while multiple players have attempted to stake their claim, the question of who is New Zealand's premier spin bowler, remains as pertinent as ever. For a period of time, Mitchell Santner provided some hope as a consistent replacement for Vettori, however, in the last 24 months, no fewer than five players have taken the field in a Test match as New Zealand's spin bowler: Mitchell Santner, Ish Sodhi, Todd Astle, Will Somerville and Ajaz Patel.

Wherever a selection dilemma appears, it is possible to make use of the match simulator to gain an deeper understanding of what the risks and rewards of selecting certain players over others may be. Simulating the same proposed matchup, using slightly different starting lineups can provide a rough indication of the impact a certain player may have on the result. It must be accepted that these simulations are not going to necessarily mimic what may happen in reality, but the results at least provide a foundation on which selectors can base their decisions.

In Figure 4.8, the simulated match outcomes are presented, conditional on New Zealand's opening partnership. One simulation is based on the Latham/Blundell opening partnership, as used throughout Section 4.2, with the other based on an the alternative Latham/Raval combination. The results suggest Blundell is expected to average 34.0 with the bat in the matchup, while Raval is predicted to average a mere 14.9. By observing the difference in win probabilities, one can infer that replacing Blundell with Raval may decrease New Zealand's probability of winning the match by up to 5%. It is again worth acknowledging that these estimates can be subject to some uncertainty, given the simulation-based nature of the process. However, this is still a fairly significant result, considering the difference in lineups is just a single player.



(a)



(b)

Figure 4.8. *(a) Match outcome predictions with Tom Latham and Tom Blundell opening the batting. (b) Match outcome predictions with Tom Latham and Jeet Raval opening the batting.*

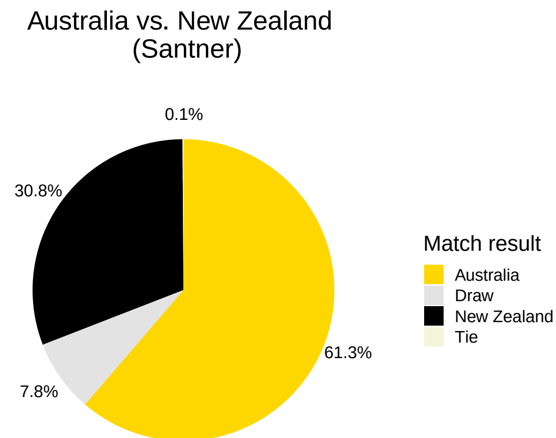
A similar analysis has been conducted for the previously identified spin bowlers, with the exception of Todd Astle, who has since retired from Test cricket. Simulations were run with Santner replaced in turn by each of Ish Sodhi, Will Somerville, and Ajaz Patel, with the corresponding match outcome predictions provided in Figure 4.9. The expected bowling and batting averages for each player are presented in Table 4.6. Interestingly, the results do not suggest that the player with the best expected bowling average, Ajaz Patel, is the player who provides New Zealand with the highest probability of winning. Instead, it appears as though Will Somerville would be the best choice in this matchup. Undoubtedly, this result may be partially due to randomness and uncertainty inherent in the simulation process, however, in this scenario it is also worthwhile considering the expected batting averages of each player, as such performances can factor in to the ultimate outcome of the match. Here, the simulation output implies that while Patel is expected to perform better than Somerville within a bowling context, the utility that Somerville offers with the bat more than makes up for his slightly inferior expected bowling average.

Table 4.6. *Summary of expected bowling and batting performances for New Zealand spin bowlers Mitchell Santner, Ish Sodhi, Will Somerville and Ajaz Patel, in the Australia versus New Zealand matchup.*

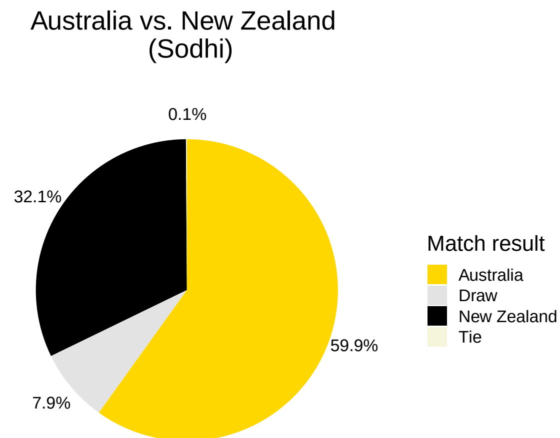
Player	**Expected bowling average**	**Expected batting average**	**Predicted NZ win %**
Mitchell Santner	53.2	17.1	30.8%
Ish Sodhi	43.2	15.0	32.1%
Will Somerville	35.3	15.1	34.3%
Ajaz Patel	32.7	9.3	32.5%

Ultimately, a solid argument could be made for any of these four players to be selected. After all, the model has not accounted for variables such as recent domestic performances or player form in other match formats, such as T20 and one-day cricket, which may translate to improved Test match performance. Furthermore, before every match, subjective assessments are made by both players and coaches, as to which types of bowler are likely to be more or less effective, given the opposition lineup and the pitch that has been prepared for the match. Santner and Patel are both left arm orthodox bowlers, Ish Sodhi is a more aggressive leg spin bowler, and Will Somerville is a right arm off spin bowler. As such, when selecting the most appropriate player for the match, coaches and selectors must consider far more variables than those employed by the match simulator.

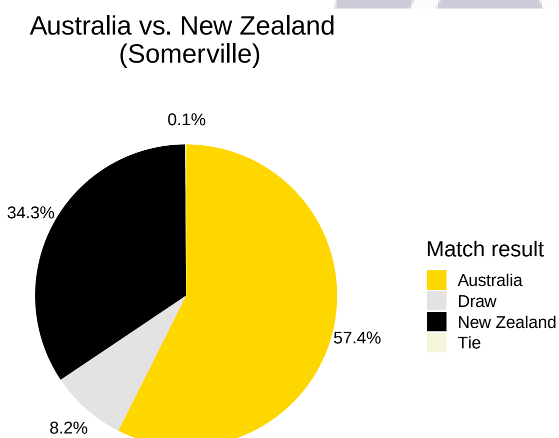
On face value, the estimated differences in win probabilities when replacing one player with another can be minimal and at times, it can be difficult to discern the noise from the signal. However, sports analytics can be all about identifying areas in which marginal gains can be achieved. Often such areas are unbeknownst to coaches and are not visible to the naked eye. As long as coaches and selectors continue rely on purely on the eye test and gut-feel when



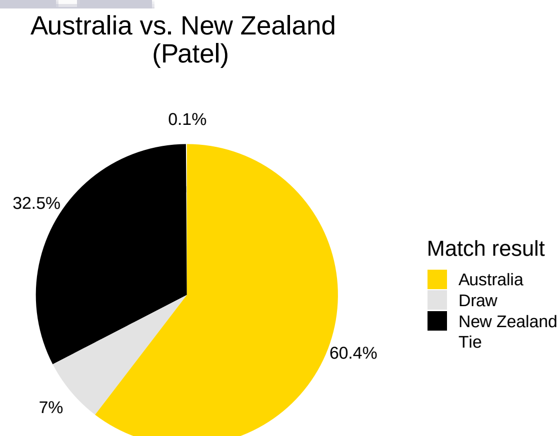
(a)



(b)



(c)



(d)

Figure 4.9. *(a) Match outcome predictions with Mitchell Santner playing as the spin bowler. (b) Match outcome predictions with Ish Sodhi playing as the spin bowler. (c) Match outcome predictions with Will Somerville playing as the spin bowler. (d) Match outcome predictions with Ajaz Patel playing as the spin bowler.*

making subjective judgements in regards to the abilities of players, it is inevitable finer details will be missed that can be explained and exploited by data. Undoubtedly, coaches are coaches for the very reason that their intuitions are generally fairly accurate and they tend to make the correct decision in the majority of cases. Therefore, the utilisation of analytical tools such as the proposed match simulator, should be done in a manner which informs, challenges and supplements a subjective opinion, rather than providing the underlying basis and primary rationale for one's argument.

4.4 Discussion

4.4.1 Limitations and future work

The proposed simulation-based method of predicting the most likely outcome of a match attempts to mimic plausible courses a match could take. Making use of the empirical data from a large sample of previous Test matches ensures that the general pace of a simulated match is similar to what has been observed historically. As shown, the general results and scorecards that are simulated tend to be indistinguishable from reality, however, it is important to acknowledge the method is far from perfect and there are several shortcomings and areas for potential improvement.

When estimating the abilities of individual players, the match simulator makes use of the results obtained from the batting and bowling career trajectory models, detailed in Chapters 2 and 3. Although these estimates have been proven to provide more accurate predictions of player ability than traditional metrics, such as batting and bowling averages, there is still a certain level of uncertainty associated with these forecasts. As these predictions only consider past performances in Test cricket, the accuracy of such estimates may be called into question for players who have only played in a handful of Test matches. For players new to the Test scene, it may be advantageous to develop a method of incorporating domestic first-class data into the predictions of player ability.

During the simulation of a given ball, the simulation process accounts for the respective abilities of the batsman and bowler in question. However, the specific strengths and weakness of players versus certain batting and bowling types have been ignored. Inevitably, some players have a preference of bowling to left or right-handed batsmen. Similarly, some players clearly perform better when facing spin bowling, rather than pace bowling, or vice versa. Regrettably, the Cricsheet data source does not provide this level of information at the individual level. However, such player-specific characteristics are available in the data source provided by New Zealand Cricket and NV Play. As such, when simulating the potential match outcomes for domestic first-class matches in New Zealand, the match simulator is able to make adjustments where certain batsmen have a specific weakness against particular bowling types, or where bowlers have a significant preference in bowling to right or left-handed batsmen. Unfortunately, the process of manually defining batting handedness and bowling type in the Cricsheet data source is an arduous process that cannot be automated. Nevertheless, this an obvious area in which the match simulator can be improved and is something which must be addressed eventually. Incorporating this level of information in the methodology is necessary for obtaining match outcome predictions that are truly reflective of reality. Doing so will provide coaches and selectors with more accurate quantifications of the pros and cons of selecting certain bowlers or batsmen in a matchup, given the opposition lineup, which will ultimately make the match simulation tool more valuable than it is at present.

One known advantage the proposed match simulator has over other more well-known methods, such as the WinViz model, is the inclusion of effects such as getting your eye in for individual players. By assuming a non-constant batting ability during an innings, the simulated results are going to have a higher predictive accuracy than methods that incorrectly assume batting scores can be modelled using a geometric distribution. However, one variable that is held constant in the match simulator is that of batting strike rate. In reality, it is likely that a batsman's strike rate is also influenced by the process of getting your eye in, much like their underlying batting ability. As a result, it is possible that the match simulator overestimates the scoring rate for batsmen on low scores.

On a similar note, the simulator does not have any built-in logic to manipulate the general tempo, or scoring rate of a match. Every now and then, there are occasions in a Test match where the batting team opts to take more risks to score quick runs and advance the state of a game, in order to avoid the possibility of drawing a match. Conversely, when teams are set an unattainable target score, run scoring becomes irrelevant as the batting team is simply trying to survive in order to secure a draw. At present, the proposed simulation method does not account for such scenarios.

Finally, within a bowling context, there is no adjustment made to account for the state of the ball. It is well-known that a new cricket ball will offer a lot more assistance to bowlers than an old, worn out ball. As such, it may be worthwhile to consider a method of quantifying how much more difficult batting is against a new ball. This is a fairly complex and challenging problem, as not only are different bowlers able to exploit the new ball with varying degrees of success, the state of the pitch — which is rarely identical between any two matches — will have a significant effect on how quickly the ball deteriorates.

On the whole, given the countless complexities and intricacies of the game of cricket, there are a multitude of factors that have not been considered within the proposed method of match simulation. This is acknowledged, however, care has been taken to at least address factors which can have a significant impact on the final outcome of a match.

4.4.2 Concluding remarks

The match simulator discussed in this chapter provides a novel means of estimating the most likely outcome of a Test match, given two proposed lineups. Unlike a number of past studies, the proposed method considers the strengths of each team on an individual player basis, rather than on the past performances of the team as a whole. The simulations take the estimates of player ability obtained from Chapters 2 and 3 as inputs and marginalise over them appropriately. This provides levels of detail in the estimates that many other methods are simply unable to compete with. Consequently, the match simulator may have significant real-world utility in both a public and private capacity.

The use of empirical data and built-in logic regarding several complexities of Test cricket, such as bowler selection, guarantees the simulation process to produce results which are generally in the realm of possibility in terms of what could really happen on the pitch. Adjusting the likelihood of various run scoring and wicket taking events occurring on a given delivery to account for the respective batsman and bowler abilities is an approach that is, in an academic sense, yet to be adopted or documented. Unfortunately, due to the element of secrecy and confidentiality surrounding commercial products that attempt to provide similar predictions and outputs, it is difficult to compare the predictive accuracy of the proposed method with competing models. Nevertheless, it is not unreasonable to believe that the proposed method could be among one of the most accurate predictors of match outcome in Test cricket, given the player ability estimates have been shown to be more accurate than measures such as batting and bowling averages. The confidence one can have in this statement will only become stronger, as the shortcomings identified in Section 4.4.1 are addressed.

Examples of how the results of simulated matches could be used from both a public and private standpoint have also been discussed. The simulation output can be used to provide up-to-date predictions for which team is most likely to win a match and can help answer the age-old question of: ‘who’s winning?’. Furthermore, the expected performances of each player can be used to generate and facilitate discussions among commentators as to who the key performers are likely to be in a match, as well identifying performances where players have significantly under or overachieved.

From the perspective of coaches and selectors, the match simulation tool provides a means of quantifying the expected performances of individual players and the impact they may have on their team’s chances of wining a match, which may help in fine-tuning the balance of a team. As shown, it is not always the player with the best career batting or bowling average who is necessarily the best selection for a given role within a side. Of course, the results from a simulated match are certainly not intended to replace the human element of the team selection process. Instead, the output can be used as another analytical tool in times of uncertainty, to provide discussion points and a secondary viewpoint that may not otherwise be available. At present, there is a significant element of guesswork when weighing up the pros and cons of selecting certain players to fill a role in a side; the results from the match simulator provide a more formal means of quantifying this uncertainty.

OS

Chapter 5

Conclusion

The research presented in this thesis has primarily focused on the development of statistical models that can be used to analyse player performance in the sport of cricket. At present, traditional cricketing statistics, such as the batting and bowling average, are commonly used to quantify the playing abilities of professional players. However, such metrics do not account for a multitude of factors that may impact a player's underlying abilities and subsequent performances, including recent form, opposition strength, and innings and venue-specific effects. The proposed models provide a means of adjusting for such variables and allow for the estimation of a player's batting (Chapter 2) and bowling (Chapter 3) career trajectories, which describe how playing abilities have varied over the course of a career to date and provide predictions of past, present, and future ability. Estimates relating to the level of uncertainty associated with these predictions are also readily available. Generally speaking, these predictions have been shown to provide more accurate forecasts of future performances than quantities such as the batting and bowling average. Furthermore, a simulation-based method of predicting the outcome of an upcoming match has been derived, which employs the player-specific estimates of batting and bowling ability and marginalises over the uncertainties (Chapter 4). Potential applications of the simulation output have been identified in both a public and private capacity, providing useful insights related to the likely performances of individual players, as well as quantifying the risks and rewards of selecting certain players in a given matchup.

5.1 Batting career trajectories

In cricket, recent performance, or form, is frequently cited as being a major reason for selecting or dropping specific players within a team, particularly in the context of batting. Therefore, if form is truly a valid criteria for selectors to use when picking a side, then recent performances should be a significant predictor of future batting scores. The batting career trajectory model developed in Chapter 2 aims to detect variations in batting ability that may be observed over the course of a playing career, due to the likes of recent form, as well as general improvements or deteriorations in technique, fitness or eyesight. In previous works by the author (Stevenson & Brewer, 2017; Stevenson, 2017), several models have been derived that allow for the identification of changes in batting ability that may occur within an innings, due to the cricketing concept of getting your eye in. These earlier models are used as a theoretical foundation, which is built upon and extended by the batting career trajectory model, allowing for the detection of variations in batting ability that may occur between innings.

Time dependence between individual batting performances is measured using a Gaussian process, a machine learning algorithm that allows for a range of plausible functions to be fitted to batting data, to describe a player's career trajectory. When evaluating each performance, several innings-specific variables are considered, namely, the specific innings of a match, and the venue — either home or away — in which the performance took place. As pitches in Test cricket are used for up to five days, it is often inevitable that batting conditions will deteriorate as a match progresses. This tends to result in batting being more difficult in the latter innings of a Test match. Furthermore, home ground advantage has been identified as a performance enhancing factor across a number of sports. Cricket is no different, however, it is plausible that in Test cricket, such an advantage manifests itself due to familiarity with the local pitch and weather conditions, which have a major bearing on how a match is played out. The findings support each of these pre-conceived hunches, with both the innings and venue-specific effects found to be significant when conducting a hierarchical analysis across the entire set of players analysed.

Interestingly, for the vast majority of players, little evidence was found to support the idea that recent form is a significant predictor of future performance. Instead, variations in ability are often observed over the long-term, suggesting that underlying batting capabilities develop and deteriorate slowly during a playing career, potentially due to any or all of the reasons previously identified. These results provide some support for the idea of *finding your feet*, whereby players — particularly those who make their international debut at a young age — can take some time to adjust to the demands of Test cricket and reach their peak ability. Comparing the predictive accuracy of the batting career trajectory model with other methods, such as the batting average and a range of simple moving average models, found the proposed model provides the most accurate estimates of future batting scores. Moreover, these diagnostic checks indicate that using only recent batting efforts to predict upcoming performance is actually the least accurate method of those tested, a finding of major interest with potential real-world applications. In fact, after the career trajectory model, the next best method of predicting future scores was the simple batting average. In this regard, there is perhaps some evidence to suggest that using form as a main criteria to drop and select certain players is not recommended, unless other analyses have provided evidence to the contrary. These findings give rise to the oft-touted phrase *"form is temporary, class is permanent"*, implying that while many players go through good and bad patches of form during their career, those who are truly talented will often bounce back to their best over the long term.

Finally, a comparison of the batting career trajectory model with the official ICC batting ratings has identified and exposed several problems with the ICC's method of ranking players. In particular, the ICC ratings are unable to provide a natural cricketing interpretation, making it difficult to properly quantify the differences in abilities between players. On the other hand, the batting career trajectory model is expressed in units of a batting average, a metric which all members of the cricketing community are familiar with and lends itself to the direct comparison of players. Furthermore, working within a Bayesian framework allows for such comparisons to be made via probabilistic statements, which provide clear insights as to the pros and cons of selecting certain players over others.

5.2 Bowling career trajectories

Although bowling data is typically less noisy than batting data, the concept of bowling form is still regularly discussed as being an indicator of a player's current bowling ability. However, unlike batting, bowling performances are usually summarised using multiple variables, making it difficult to easily compare individual bowling efforts both within and between different innings. Additionally, given the wide ranging nature of batting abilities between players, it is advantageous to consider the strengths of the individual batsmen bowled to during a performance. Rather than treating all runs conceded as equal, those conceded to lower quality batsmen should be weighted more heavily than those conceded to world-class players. Each of these challenges must be addressed before a career trajectory model can be implemented within a bowling context.

Firstly, rather than considering bowling efforts on a per-innings basis, each performance is split into separate *bowling spells*, defining the number of runs conceded between individual wickets. Under this specification, each bowling spell is simply measured by the number of runs conceded, allowing for performances to be visualised over time. One limitation of this method is the requirement of ball-by-ball data, which can be difficult to obtain. Secondly, to address the issue related to the variable strengths of opposition batsmen, the *standardised bowling average* quantity has been derived. This measure allows for the number of runs conceded by bowlers in a given spell to be adjusted accordingly, depending on the estimated batting abilities of the opposition, which are obtained using the results of the batting career trajectory model presented in Chapter 2. Then, by estimating the average batting ability of all players across modern Test cricket, the *adjusted bowling average* quantity has been derived, expressing a player's bowling average in terms of the number of runs they are expected to concede per wicket taken, if they were bowling to the average Test batsman.

The derivation of these new measures of performance are among the most important contributions of this thesis to the field of cricket analytics and should not be overlooked. When estimating and predicting the bowling abilities of individual players, it is often discussed and acknowledged that opposition batting quality is a variable which must be considered. However, to date there is still no commonly accepted method of doing so, or, those which claim to are not publicly available. While the usage of standardised and adjusted bowling averages may not catch on across the wider cricketing community, they are a step in the right direction and may promote further work to find a more permanent solution to this problem.

Using these newly derived measures of performance, a bowling career trajectory model has been developed to measure and predict past, present, and future bowling ability. Like the batting career trajectory model, a Gaussian process is employed to detect any time dependence that may exist between performances. Again, innings and venue-specific effects are considered to account for the varying conditions players may find themselves bowling in. Generally speaking, bowling ability appears to vary over shorter time frames, when compared with the results obtained from the batting career trajectory model. Meanwhile, for many players there was less support to imply bowling abilities vary and fluctuate nearly as much their batting counterparts. Additionally, there was minimal evidence to suggest that the innings and venue-specific factors have a significant impact on bowling bowling ability. However, it is important to note that the estimated batting abilities of opposition players already account for these external variables. As such, it may simply be a case that the proposed method assigns all innings and venue-specific variation in ability to batsmen, when in fact some of the explanatory power should be attributed to bowlers.

Once again, the results obtained from the bowling career trajectory model were compared with the ICC bowling ratings, leading to similar conclusions regarding the limitations of the ICC methodology. While there tends to be a reasonable amount of overlap between the methods, the proposed model provides an intuitive cricketing interpretation and allows for differences in abilities to be easily quantified. Furthermore, the bowling career trajectory model was found to provide the most accurate estimates of future performance, when compared with predictions obtained using the bowling average and a set of simple moving average models. Similar to the findings discovered in a batting context, the least accurate method of predicting future bowling performance was that which only considers a player's most recent efforts. These findings should act as another stark warning to coaches and selectors; unless you are using other advanced analytical tools to gauge player bowling ability, there is real potential of falling victim to recency bias if selecting bowlers based on short-term form.

5.3 Match simulation

The career trajectory models derived in Chapters 2 and 3 provide estimates of batting and bowling ability that are more accurate than traditional metrics such as the batting and bowling average. While these predictions are inherently useful to players, coaches and selectors alike, it is even more valuable to have a means of estimating the effect an individual is going to have on the outcome of a match. This is achieved by using the results obtained from the batting and bowling career trajectory models as inputs to a match simulation engine, which provides team and individual-specific predictions pertaining to the most likely outcome of a match, given two proposed playing XIs.

The match simulator uses empirical data from all Test matches from 2008 onward to drive the underlying output of every single delivery. This ensures that the tempo of all simulated matches loosely mimics what has previously occurred. Additionally, some logic has been built into the match engine in order to address several of the finer complexities of Test cricket, such as dealing with declarations and bowler selection during an innings. Again, this aims to ensure the resulting output is as realistic as possible and provides estimates that accurately reproduce what may really happen out on the pitch.

Several potential real-world applications of the match simulator have been identified and discussed. Firstly, the results may be of interest to public broadcasters of cricket, wishing to provide additional match analysis from a data-driven perspective to supplement the overall viewing experience. The simulation output provides up-to-date estimates regarding the current state of a match, attempting to provide a quantifiable answer to the question of 'who's winning?'. Additionally, the output can be used to identify which players are in and out of form coming into a match and provides an indication of which players are likely to be the key performers in a given matchup. This may provide commentators with various talking points and player-specific storylines which can be discussed prior to and throughout a match. While often considered over the top and hyperbolic, if there is one thing sports broadcasting in the US can teach the world, it is that pre-match narratives — regardless of their plausibility — can increase viewership and interest in a match.

Secondly, the match simulator may be utilised by coaches and selectors of professional teams to gain a competitive advantage over their opponents. Similar to the potential use in a public perspective, the output attempts to provide quantifications regarding questions that are usually based on subjective opinions. By predicting the outcome of a match using a range of potential playing XIs, teams can work out which balance of players are most likely to achieve a positive result. It is worth reiterating that the match simulator is not intended to replace the human element of the team selection process, rather to challenge and support opinions about various players and ask questions that may otherwise be overlooked. While the simulations ignore a number of contextual variables that are certainly important when selecting a team for match day, they at least provide a data-driven foundation for the optimal playing XI around which coaches and selectors can build their side.

5.4 Concluding remarks

As more and more industries, including sport, rush to keep up with the latest trends in data science and find a use for modern analytical techniques with exciting labels, such as machine learning and artificial intelligence, it is important not to lose sight of the ultimate goal when conducting data analysis. Analytics for analytics' sake is like a fisherman casting a rod into an empty pond, knowing full well there are no fish at the bottom; without a purpose and knowledge of what one wants to achieve from their analytical pursuit, the exercise is fruitless.

The development of the proposed methods and models in this thesis have undergone careful consultation with experts in the field of cricket analytics and an end goal in mind. This has ensured while the models are somewhat complex and employ the use of machine learning algorithms, the resulting output can be readily understood by the wider cricketing community, not just those engaged in the field of sports analytics. It has been pleasing to see the results support a number of pre-conceived cricketing hunches from a statistical point of view, although it is important to note that complex mathematical functions, such as Gaussian processes, were only employed where more traditional statistical tools were unable to perform with the same accuracy. Extending the methodology to apply to other formats, such as one-day and T20 cricket, is in the scope of future work, allowing for additional cricket-related hypotheses to be tested, and for the models to be implemented across a wider range of matches.

The present research does not come without its own limitations, which have been well documented throughout. In many cases, the results do not provide drastically difference estimates to currently used and accepted methods, which are generally far more time efficient than those that have been proposed. However, a primary application of analytics in sport is to identify areas in which teams can move the needle fractionally, by effecting small changes to their playing strategies and decision making processes. Ultimately professional sport is a results-driven industry. A 1% increase in win probability may not sound significant, but over the long term can be worth a lot in terms of both on-field reward and off-field gain.

The aim of this thesis has not been to reinvent the wheel when it comes to the general analysis of cricket. In the majority of instances, batting and bowling averages will continue to suffice as very reasonable estimates of player ability. Instead, the results and findings can be used to supplement one's understanding of the sport and to tailor exceptions in relation to concepts such as recent form. What this research has shown is that cricket is not a sport that can simply be boiled down to several equations outlining the most optimal way to play; however, it has been a fulfilling exercise to understand a little more about the numbers behind the game.

References

Akhtar, S., & Scarf, P. (2012). Forecasting Test cricket match outcomes in play. *International Journal of Forecasting*, *28*(3), 632–643.

Arlott, J. (1979). *An eye for cricket.* Hodder Stoughton.

Association of Cricket Statisticians and Historians. (1981). *A Guide to Important Cricket Matches Played in the British Isles 1707-1863 (1985).* ACS.

Bailey, M., & Clarke, S. R. (2006). Predicting the match outcome in one day international cricket matches, while the game is in progress. *Journal of Sports Science & Medicine*, *5*(4), 480.

Bandulasiri, A. (2008). Predicting the winner in one day international cricket. *Journal of Mathematical Sciences & Mathematics Education*, *3*(1), 6–17.

Boys, R. J., & Philipson, P. M. (2019). On the ranking of test match batsmen. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *68*(1), 161–179.

Bracewell, P. J., & Ruggiero, K. (2009). A parametric control chart for monitoring individual batting performances in cricket. *Journal of Quantitative Analysis in Sports*, *5*(3).

Brewer, B. J. (2008). Getting your eye in: a Bayesian analysis of early dismissals in cricket. *arXiv preprint arXiv:0801.4408*.

Brooker, S. (2011). *An economic analysis of ability, strategy and fairness in ODI cricket* (Unpublished doctoral thesis). University of Canterbury.

Brooker, S., & Hogan, S. (2011). *A method for inferring batting conditions in ODI cricket from historical data.* Working Papers in Economics 11/44, University of Canterbury, Department of Economics and Finance.

Brooks, R. D., Faff, R. W., & Sokulsky, D. (2002). An ordered response model of Test cricket performance. *Applied Economics*, *34*(18), 2353–2365.

Cai, T., Hyndman, R. J., & Wand, M. (2002). Mixed model-based hazard estimation. *Journal of Computational and Graphical Statistics*, *11*(4), 784–798.

Carter, M., & Guthrie, G. (2004). Cricket interruptus: fairness and incentive in limited overs cricket matches. *Journal of the Operational Research Society*, *55*(8), 822–829.

Clarke, S. R., & Norman, J. M. (1998). Dynamic programming in cricket: protecting the weaker batsman. *Asia Pacific Journal of Operational Research*, *15*, 93–108.

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

Clarke, S. R., & Norman, J. M. (2003). Dynamic programming in cricket: Choosing a night watchman. *Journal of the Operational Research Society*, *54*(8), 838–845.

Cleveland, W. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, *74*(368), 829–836.

Cleveland, W., & Devlin, S. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, *83*(403), 596–610.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281.

Csapo, P., Avugos, S., Raab, M., & Bar-Eli, M. (2015). The effect of perceived streakiness on the shot-taking behaviour of basketball players. *European Journal of Sport Science*, *15*(7), 647–654.

Davis, J., Perera, H., & Swartz, T. B. (2015). A simulator for Twenty20 cricket. *Australian & New Zealand Journal of Statistics*, *57*(1), 55–71.

Duckworth, F., & Lewis, A. (1998). A fair method for resetting the target in interrupted one-day cricket matches. *Journal of the Operational Research Society*, *49*(3), 220–227.

Duckworth, F., Lewis, A., & Stern, S. (2019). Cricket's raining champion: Two decades of Duckworth–Lewis (and Stern). *Significance*, *16*(3), 30–35.

Duff & Phelps. (2017). *Duff & Phelps 2017 IPL report.* Retrieved 02-07-2020, from `https://www.duffandphelps.com/about-us/news/duff-and-phelps-2017-ipl-report`

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

Duvenaud, D. K., Nickisch, H., & Rasmussen, C. E. (2011). Additive gaussian processes. In *Advances in neural information processing systems* (pp. 226–234).

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

ESPNcricinfo. (2018). *ESPNcricinfo Smart Stats - a new way to understand T20 cricket.* Retrieved 06-04-2018, from `https://www.espncricinfo.com/story/espncricinfo-smart-stats-a-new-way-to-understand-t20-cricket-1142569`

ESPNcricinfo. (2019). *Explainer: how Smart Stats works.* Retrieved 21-03-2019, from `https://www.espncricinfo.com/story/_/id/26312449/how-smart-stats-works`

ESPNcricinfo. (2020). *Rahul Dravid: 'Remember, a batsman who averages 50 has failed a lot more than he has succeeded'.* Retrieved 23-07-2020, from `https://www.espncricinfo.com/story/rahul-dravid-in-cricket-a-batsman-who-averages-50-has-failed-a-lot-more-than-he-has-succeeded-1227551`

Forbes. (2008). *By the numbers: MLB team valuations.* Retrieved 30-06-2020, from `https://www.forbes.com/2008/04/16/mlb-team-values-biz-sports-baseball08-cx_mo_kb_0416values_slide.html`

Forbes. (2020). *Forbes releases 23rd annual Major League Baseball valuations.* Retrieved 30-06-2020, from `https://www.forbes.com/sites/forbespr/2020/04/09/forbes-releases-23rd-annual-major-league-baseball-valuations/#15b273207eeb`

Franks, A. M., D'Amour, A., Cervone, D., & Bornn, L. (2016). Meta-analytics: tools for understanding the statistical properties of sports metrics. *Journal of Quantitative Analysis in Sports*, *12*(4), 151–165.

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*(6), 721–741.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, *17*(3), 295–314.

Hastings, W. K. (1970, 04). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*(1), 97-109.

Hocking, R. R. (1976). The analysis and selection of variables in linear regression. *Biometrics*, *32*(1), 1–49.

Ian, P., & Thomas, J. (2002). Rain rules for limited overs cricket and probabilities of victory. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *51*(2), 189–202.

Ihaka, R., & Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, *5*(3), 299–314.

James, B., & Henzler, J. (2002). *Win shares.* STATS Pub.

Jayadevan, V. (2002). A new method for the computation of target scores in interrupted, limited-over cricket matches. *Current Science*, *83*(5), 577–586.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795.

Kimber, A. C., & Hansford, A. R. (1993). A statistical analysis of batting in cricket. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 443–455.

Kimber, J. (2017). *My summer watching the big four.* Retrieved 23-07-2020, from `https://www.espncricinfo.com/story/jarrod-kimber-my-summer-watching-the-big-four-1107557`

Kishore, S. (2020). *Jeet Raval: 'I started to tense up, thinking way too much. I froze'.* Retrieved 20-08-2020, from `https://www.espncricinfo.com/story/_/id/29359196/started-tense-thinking-way-too-much-froze`

Knapp, B. (1963). *Skill in sport: the attainment of proficiency.* Routledge & K. Paul.

Koulis, T., Muthukumarana, S., & Briercliffe, C. D. (2014). A Bayesian stochastic model for batting performance evaluation in one-day cricket. *Journal of Quantitative Analysis in Sports*, *10*(1), 1–13.

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

Lemmer, H. H. (2006). A measure of the current bowling performance in cricket. *South African Journal for Research in Sport, Physical Education and Recreation*, *28*(2), 91–103.

Leuty, R. (2018). *49ers hope partnership with software giant will help sell more $10 beers and improve game day for fans.* Retrieved 12-06-2020, from `https://www.bizjournals.com/sanfrancisco/news/2018/10/22/san-francisco-49ers-data-fan-experience-sap-nfl.html`

Lewis, A. (2005). Towards fairer measures of player performance in one-day cricket. *Journal of the Operational Research Society*, *56*(7), 804–815.

Lewis, A. (2008). Extending the range of player-performance measures in one-day cricket. *Journal of the Operational Research Society*, *59*(6), 729–742.

Lewis, M. (2003). *Moneyball: The art of winning an unfair game.* WW Norton & Company.

MacKay, D. J. (2003). *Information theory, inference and learning algorithms.* Cambridge University press.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, *21*(6), 1087–1092.

Meyn, S. P., & Tweedie, R. L. (1993). Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, 518–548.

Moonda, F. (2017). *The making of a Kiwi.* Retrieved 20-08-2020, from `https://www.espncricinfo.com/story/firdose-moonda-on-jeet-raval-s-journey-to-opening-for-new-zealand-1088577`

Morley, B., & Thomas, D. (2005). An investigation of home advantage and other factors affecting outcomes in English one-day cricket matches. *Journal of Sports Sciences*, *23*(3), 261–268.

Myung, I. J., & Pitt, M. A. (1997). Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin & Review*, *4*(1), 79–95.

Neal, R. M. (2003). Slice sampling. *Annals of Statistics*, 705–741.

Neal, R. M. (2011). MCMC using Hamiltonian dynamics. *Handbook of Markov chain Monte Carlo*, *2*(11), 2.

Nevill, A. M., & Holder, R. L. (1999). Home advantage in sport. *Sports Medicine*, *28*(4), 221–236.

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

O'Hagan, A., & Forster, J. J. (2004). *Kendall's advanced theory of statistics, volume 2B: Bayesian inference* (Vol. 2). Arnold.

Oliver, D. (2004). *Basketball on paper: rules and tools for performance analysis.* Potomac Books, Inc.

Ötting, M., Langrock, R., Deutscher, C., & Leos-Barajas, V. (2020). The hot hand in professional darts. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *183*(2), 565–580.

Perera, H., Davis, J., & Swartz, T. (2016). Optimal lineups in Twenty20 cricket. *Journal of Statistical Computation and Simulation*, *86*(14), 2888–2900.

Perera, H., & Swartz, T. (2013). Resource estimation in T20 cricket. *IMA Journal of Management Mathematics*, *24*(3), 337–347.

Pollard, R. (1986). Home advantage in soccer: A retrospective analysis. *Journal of Sports Sciences*, *4*(3), 237–248.

Poulton, E. (1957). On prediction in skilled movements. *Psychological Bulletin*, *54*(6), 467.

Preston, I., & Thomas, J. (2000). Batting strategy in limited overs cricket. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *49*(1), 95–106.

R Core Team. (2020). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Rasmussen, C., & Ghahramani, Z. (2000). Occam's razor. *Advances in Neural Information Processing Systems*, *13*, 294–300.

Rasmussen, C., & Williams, C. K. I. (2006). *Gaussian processes for machine learning.* MIT Press.

Sahni, M., & Bhogal, G. (2017). Anxiety, depression and perceived sporting performance among professional cricket players. *British Journal of Sports Medicine*. Retrieved from `https://bjsm.bmj.com/content/early/2017/05/10/bjsports-2017-097827.5`

Sammut, C., & Webb, G. I. (2010). Leave-one-out cross-validation. In *Encyclopedia of machine learning* (pp. 600–601). Boston, MA: Springer US.

Scarf, P., Shi, X., & Akhtar, S. (2011). On the distribution of runs scored and batting strategy in Test cricket. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *174*(2), 471–497.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423.

Shea, S. (2014). In support of a hot hand in professional basketball and baseball. *PsyCh Journal*, *3*(2), 159–164.

Skilling, J. (2006). Nested sampling for general Bayesian computation. *Bayesian Analysis*, *1*(4), 833–859.

Soares, R. (2020). *Most frequent NBA shot locations 1997/98 - 2019/20.* Retrieved 18-06-2020, from `https://twitter.com/ryandsoares/status/1271891604736028673`

Stern, S. E. (2016). The Duckworth-Lewis-Stern method: extending the Duckworth-Lewis methodology to deal with modern scoring rates. *Journal of the Operational Research Society*, *67*(12), 1469–1480.

Stevenson, O. G. (2017). *The nervous 90s: A Bayesian analysis of batting in Test cricket* (Unpublished master's thesis). University of Auckland.

Stevenson, O. G., & Brewer, B. J. (2017). Bayesian survival analysis of batsmen in Test cricket. *Journal of Quantitative Analysis in Sports*, *13*(1), 25–36.

Stevenson, O. G., & Brewer, B. J. (2018). Modelling career trajectories of cricket players using Gaussian processes. In *Bayesian Statistics and New Generations* (pp. 165–173).

Stevenson, O. G., & Brewer, B. J. (2021). Finding your feet: A Gaussian process model for estimating the abilities of batsmen in Test cricket. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *70*(2), 481–505.

Stroustrup, B. (2013). *The C++ programming language.* Pearson Education.

Swartz, T. B., Gill, P. S., Beaudoin, D., et al. (2006). Optimal batting orders in one-day cricket. *Computers & Operations Research*, *33*(7), 1939–1950.

Totterdell, P. (1999). Mood scores: Mood and performance in professional cricketers. *British Journal of Psychology*, *90*(3), 317–332.

Tversky, A., & Gilovich, T. (1989). The cold facts about the "hot hand" in basketball. *Chance*, *2*(1), 16–21.

Uhlenbeck, G. E., & Ornstein, L. S. (1930). On the theory of the Brownian motion. *Physical Review*, *36*(5), 823.

Wetzels, R., Tutschkow, D., Dolan, C., van der Sluis, S., Dutilh, G., & Wagenmakers, E.-J. (2016). A Bayesian test for the hot hand phenomenon. *Journal of Mathematical Psychology*, *72*, 200–209.

Whiting, H. T. A. (1969). *Acquiring ball skill: A psychological interpretation.* Bell.

Young, S. (2019). *The NBA's 3-point revolution continues to take over.* Retrieved 11-12-2020, from https://www.forbes.com/sites/shaneyoung/2019/11/30/the-nbas-three-point-revolution-continues-to-take-over/?sh=66f72f5656b8

Appendix A

The following R functions are referred to and used throughout Chapter 4 when discussing the simulation of Test matches.

`sim_match` — a function that simulates a Test match from start to finish, assuming a maximum of 450 overs of play (90 per day).

```
## A function that simulates a Test cricket match
sim_match <- function(summary_data,
                      match_players, player_abilities,
                      bowling_logic, iteration,
                      match_output = TRUE, innings_output = TRUE,
                      over_output = TRUE, ball_output = TRUE)
{
  ## Initialise the match
  match_data <- initialise_match(summary_data = summary_data,
                                 match_players = match_players,
                                 bowling_logic = bowling_logic,
                                 player_abilities = player_abilities,
                                 iteration = iteration)

  ##### Innings 1 #####
  ## Simulate the 1st innings
  first_innings <- sim_innings(match_data = match_data, innings = 1,
                               innings_output = innings_output,
                               over_output = over_output,
                               ball_output = ball_output)

  ## Update the match_data object
  match_data$innings_data[[1]] <- first_innings$innings_data
  match_data$match_parameters <- first_innings$match_parameters
  ##### End of innings 1 #####
```

```
##### Innings 2 #####
## Simulate the 2nd innings
second_innings <- sim_innings(match_data = match_data, innings = 2,
                              innings_output = innings_output,
                              over_output = over_output,
                              ball_output = ball_output)

## Update the match_data object
match_data$innings_data[[2]] <- second_innings$innings_data
match_data$match_parameters <- second_innings$match_parameters
##### End of innings 2 #####

##### Innings 3 #####
## Simulate the 3rd innings
third_innings <- sim_innings(match_data = match_data, innings = 3,
                             innings_output = innings_output,
                             over_output = over_output,
                             ball_output = ball_output)

## Update the match_data object
match_data$innings_data[[3]] <- third_innings$innings_data
match_data$match_parameters <- third_innings$match_parameters
##### End of innings 3 #####

##### Innings 4 #####
## Check if 4th innings is necessary
## I.e. has either team won by an innings?
if((match_data$match_parameters$toss == "home" &
    match_data$match_parameters$lead_deficit < 0) |
   (match_data$match_parameters$toss == "away" &
    match_data$match_parameters$lead_deficit > 0))
{
  ## End the match
  match_data$match_parameters$match_completed <- TRUE

  ## Determine the results
  match_data$match_parameters$result <-
     ifelse(match_data$match_parameters$lead_deficit > 0, "home", "away")
}else{
```

```
    ## Simulate the 4th innings
    fourth_innings <- sim_innings(match_data = match_data,
                                  innings = 4,
                                  innings_output = innings_output,
                                  over_output = over_output,
                                  ball_output = ball_output)

    ## Update the match_data object
    match_data$innings_data[[4]] <- fourth_innings$innings_data
    match_data$match_parameters <- fourth_innings$match_parameters
  }
  ##### End of innings 4 #####

  ## Check how the match ended
  ## 1. If there are overs remaining, determine the result
  if(match_data$match_parameters$overs_remaining > 0)
  {
    if(match_data$match_parameters$lead_deficit > 0)
      match_data$match_parameters$result <- "home"
    if(match_data$match_parameters$lead_deficit < 0)
      match_data$match_parameters$result <- "away"
    if(match_data$match_parameters$lead_deficit == 0)
      match_data$match_parameters$result <- "tie"
  }

  ## 2. If no overs are remaining, the match was a draw
  if(match_data$match_parameters$overs_remaining == 0)
    match_data$match_parameters$result <- "draw"

  ## Output
  if(isTRUE(match_output))
    cat("Result: ", match_data$match_parameters$result, "\n", sep = "")

  ## Return the match parameters and innings data
  return(list(match_parameters = match_data$match_parameters,
              match_data = match_data$innings_data))
}
```

`sim_innings` — a function that simulates a team's innings.

```
## A function that simulates an innings
sim_innings <- function(match_data, innings,
                        innings_output = TRUE, over_output = TRUE, ball_output = TRUE)
{
  ## Initialise the innings and match parameter objects
  innings_data <- initialise_innings(match_data = match_data, innings = innings)
  match_parameters <- match_data$match_parameters

  ## Continue to simulate overs until 10 wickets are lost,
  ## or until the batting team declares,
  ## or until the batting team has chased down the target,
  ## or until there are no overs remaining in the match
  while(innings_data$wickets < 10 &
        isFALSE(innings_data$declare) &
        isFALSE(innings_data$innings_completed) &
        isFALSE(match_parameters$match_completed) &
        match_parameters$overs_remaining > 0)
  {
    ## Simulate an over
    over_data <- sim_over(match_data = match_data,
                          match_parameters = match_parameters,
                          innings_data = innings_data,
                          over_output = over_output,
                          ball_output = ball_output)

    ## At the end of each over:
    ## 1. Update the innings data
    innings_data <- over_data$innings_data

    ## 2. Update the match parameters
    match_parameters <- over_data$match_parameters

    ## 3. If in the 3rd innings of a match, should the innings be declared?
    if((innings == 3 & innings_data$batting_team == "home" &
        match_parameters$lead_deficit > (2 * match_parameters$overs_remaining + 150)) |
       (innings == 3 & innings_data$batting_team == "away" &
        match_parameters$lead_deficit < -(2 * match_parameters$overs_remaining + 150)))
    {
      ## Delcare and end the innings
      innings_data$declare <- TRUE
      innings_data$innings_completed <- TRUE
    }
```

```
  ## 4. Select the bowler to bowl the next over
  if(isFALSE(innings_data$innings_completed))
  {
    ## Current bowler (who cannot bowl the next over)
    current_bowler <- innings_data$bowling$bowler$name

    ## Overs bowled
    ## If innings has exceeded 160 overs, use the 160th over bowling workload
    ## estimates
    overs_bowled <- ifelse(innings_data$overs_bowled > 160,
                           160, innings_data$overs_bowled)

    ## Get the list of bowlers and relevant bowling workloads
    bowlers <- subset(match_data$bowling_logic,
                      team %in% innings_data$bowling_team)[ , c(1, 3 + overs_bowled)]

    ## Exclude current bowler so they do not bowl two consecutive overs
    bowlers <- subset(bowlers, name != current_bowler)

    ## "BOWLERS name?!"
    bowlers_name <- sample(bowlers$name, size = 1, prob = bowlers[ , 2], replace = TRUE)

    ## Extract the bowler's data from the list of bowler objects
    innings_data$bowling$bowler <- innings_data$bowling$players[[bowlers_name]]

    ## Check if bowler has already bowled in the innings, if not,
    ## add to the bowling scorecard
    if(!(innings_data$bowling$bowler$name %in% innings_data$bowling$scorecard$bowler))
    {
      ## Number of bowlers who have bowled so far in the innings
      nbowlers <- length(innings_data$bowling$scorecard$bowler)

      ## Append bowler to the bowling scorecard
      innings_data$bowling$scorecard <-
         rbind(innings_data$bowling$scorecard,
               data.frame(innings_index = innings,
                          bowler = innings_data$bowling$bowler$name,
                          wt = innings_data$bowling$bowler$bowl_ability,
                          overs = 0, balls = 0, maidens = 0,
                          runs_conceded = 0, standardised_runs_conceded = 0,
                          wickets = 0, RPO = 0, stringsAsFactors = FALSE))
    }
  }
}
```

```
  ## If innings ended midway through an over adjust overs remaining
  ## No need to do this if currently in the 4th innings of a match
  if(match_parameters$balls_bowled > 0 & innings_data$innings_index != 4)
  {
    ## Take an over off the overs remaining
    match_parameters$overs_bowled <- match_parameters$overs_bowled + 1
    match_parameters$overs_remaining <- match_parameters$total_overs -
                                        match_parameters$overs_bowled
    match_parameters$balls_bowled <- 0
  }

  ## End innings and match if there are no overs remaining
  if(match_parameters$overs_remaining < 1)
  {
    ## End the innings
    innings_data$innings_completed <- TRUE

    ## End the match
    match_parameters$match_completed <- TRUE
  }

  ## Output the innings
  if(isTRUE(innings_output))
    cat("Batting team: ", innings_data$batting_team, ". Score: ",
        innings_data$total_scored, "/", innings_data$wickets,
        " (", innings_data$overs_bowled, ".", innings_data$balls_bowled,
        ")", "\n", sep = "")

  ## Return the innings data
  return(list(match_parameters = match_parameters,
              innings_data = innings_data))
}
```

`sim_over` — a function that simulates an over.

```
## A function that simulates an over
sim_over <- function(match_data, match_parameters, innings_data,
                     over_output = TRUE, ball_output = TRUE)
{
  ## Initialise a data.frame to store the results for each ball
  over_data <- data.frame(runs = numeric(1),
                          wicket_type = character(1),
                          extra_runs = numeric(1),
                          extra_type1 = character(1),
                          extra_type2 = character(1),
                          legal_delivery = logical(1),
                          rotate_strike = logical(1),
                          bat_ability = numeric(1),
                          bowl_ability = numeric(1),
                          stringsAsFactors = FALSE)

  ## Separate out the elements of our match data object
  match_players <- match_data$match_players
  match_abilities <- match_data$match_abilities

  ## Initialise over parameters
  ball_count <- 1

  ## Bowl until 6 legitimate balls have been bowled,
  ## or until the batting side has lost 10 wickets,
  ## or until the batting side has chased down the target,
  ## or until there are no overs remaining in the match
  while(innings_data$bowling$bowler$balls < 6 & innings_data$wickets < 10)
  {
    ## Simulate a ball
    ball <- sim_ball(summary_data = match_data$summary_data,
                     innings_data = innings_data,
                     ball_output = ball_output)

    ## Save the ball event
    over_data[ball_count, ] <- ball$event

    ## Update the innings data
    innings_data <- ball$innings_data
```

```r
    ## Update match parameters:
    ## Lead/deficit
    match_parameters$lead_deficit <-
       ifelse(innings_data$batting_team == "home",
              match_parameters$lead_deficit + ball$event$runs + ball$event$extra_runs,
              match_parameters$lead_deficit - ball$event$runs - ball$event$extra_runs)

    ## Number of balls bowled
    match_parameters$balls_bowled <-
       ifelse(isTRUE(ball$event$legal_delivery),
                     match_parameters$balls_bowled + 1,
                     match_parameters$balls_bowled)

    ## If in the 4th innings, check if the batting team chased down the target
    if(innings_data$innings_index == 4 &
       ((innings_data$batting_team == "home" & match_parameters$lead_deficit > 0) |
        (innings_data$batting_team == "away" & match_parameters$lead_deficit < 0)))
    {
      ## Update innings_completed and match_completed parameters
      innings_data$innings_completed <- TRUE
      match_parameters$match_completed <- TRUE

      ## End the over, innings, and match
      return(list(match_parameters = match_parameters,
                  innings_data = innings_data,
                  over_data = over_data))
    }

    ## Was there a wicket?
    if(ball$event$wicket_type != "")
    {
      ## If 10 wickets have fallen, conclude the over
      if(innings_data$wickets == 10)
      {
        ## Update the innings_completed parameter
        innings_data$innings_completed <- TRUE

        ## If that was the 6th legal delivery of the over,
        ## update bowling figures accordingly
        if(innings_data$balls_bowled == 6)
        {
          ## Update bowler object
          innings_data$bowling$bowler$overs <- 1 + innings_data$bowling$bowler$overs
          innings_data$bowling$bowler$balls <- 0
```

```
        ## Update bowling scorecard
        subset(innings_data$bowling$scorecard,
               bowler == innings_data$bowling$bowler$name)
               [ , "overs"] <- innings_data$bowling$bowler$overs
        subset(innings_data$bowling$scorecard,
               bowler == innings_data$bowling$bowler$name)
               [ , "balls"] <- innings_data$bowling$bowler$balls
      }

      ## If 10 wickets have fallen to end the 4th innings, conclude the match
      if(innings_data$innings_index == 4)
        match_parameters$match_completed <- TRUE

      ## End the over and innings
      return(list(match_parameters = match_parameters,
                  innings_data = innings_data,
                  over_data = over_data))
    }else{
      ## Otherwise, bring in the next batsman
      next_batsman <- innings_data$batting$players[innings_data$wickets + 2]

      ## Add next batsmen to the innings data object
      innings_data$batting$batsmen[[innings_data$batting$striker]] <-
         initialise_batsman(name = next_batsman,
                            innings = innings_data$innings_index,
                            match_players = match_data$match_players,
                            match_abilities = match_data$match_abilities)

      ## Determine batsman's position in the batting order
      innings_data$batting$batsmen[[innings_data$batting$striker]]$position <-
         innings_data$wickets + 2
    }
  }
}

## If the over ended after 6 balls, perform usual checks:
## Batsmen switch ends at the end of an over
if(isFALSE(ball$event$rotate_strike))
  swap(innings_data$batting$striker, innings_data$batting$non_striker)
```

```
## Was over a maiden?
if(sum(over_data$runs) == 0 &
   all(over_data$extra_type1 != "wides") &
   all(over_data$extra_type1 != "noballs"))
{
  ## Record maiden in bowler object
  innings_data$bowling$bowler$maidens <-
     innings_data$bowling$bowler$maidens + 1

  ## Record maiden in bowling scorecard
      subset(innings_data$bowling$scorecard,
             bowler = innings_data$bowling$bowler$name)
             [ , "maidens"] <- innings_data$bowling$bowler$maidens
}

## Update innings data
innings_data$overs_bowled <- innings_data$overs_bowled + 1
innings_data$balls_bowled <- 0

## Update bowling scorecard and bowler data
innings_data$bowling$bowler$overs <- innings_data$bowling$bowler$overs + 1
innings_data$bowling$bowler$balls <- 0
subset(innings_data$bowling$scorecard,
       bowler == innings_data$bowling$bowler$name)
       [ , "overs"] <- innings_data$bowling$bowler$overs
subset(innings_data$bowling$scorecard,
       bowler == innings_data$bowling$bowler$name)
       [ , "balls"] <- innings_data$bowling$bowler$balls

## Update list of bowlers
innings_data$bowling$players[[innings_data$bowling$bowler$name]] <-
                                    innings_data$bowling$bowler

## Update match parameters
match_parameters$overs_bowled <- match_parameters$overs_bowled + 1
match_parameters$balls_bowled <- 0
match_parameters$overs_remaining <- match_parameters$total_overs -
                                    match_parameters$overs_bowled
```

```
  ## Output the results of the over
  if(isTRUE(over_output))
    cat("Overs bowled = ", innings_data$overs_bowled, ". ",
        "Runs scored from over = ",
        sum(over_data$runs) + sum(over_data$extra_runs), ". ",
        "Score = ", innings_data$total_scored, "/", innings_data$wickets, ".\n",
        innings_data$batting$batsmen[[1]]$name, " ",
        innings_data$batting$batsmen[[1]]$runs_scored, ", ",
        innings_data$batting$batsmen[[2]]$name, " ",
        innings_data$batting$batsmen[[2]]$runs_scored, ".\n",
        innings_data$bowling$bowler$name, ": ",
        innings_data$bowling$bowler$wickets, "/",
        innings_data$bowling$bowler$runs_conceded,
        " (", innings_data$bowling$bowler$overs, ")",
        "\n", sep = "")

  ## Return the over data, updated innings data and match parameters
  return(list(match_parameters = match_parameters,
              innings_data = innings_data,
              over_data = over_data))
}
```

`sim_ball` — a function that simulates a single ball.

```
## A function that simulates a ball
sim_ball <- function(summary_data, innings_data, ball_output = TRUE)
{
  ## Initialise ball parameters
  runs_conceded <- 0

  ## Extract the batsman/bowler objects from the innings data
  batsman <- innings_data$batting$batsmen[[innings_data$batting$striker]]
  bowler <- innings_data$bowling$bowler

  ## Simulate the ball event, conditional on the batsman/bowler
  ball <- sim_event(summary_data, batsman, bowler)

  ## Include information about the batsman/bowler in the ball
  bat_ability <- ball$event$bat_ability
  bowl_ability <- ball$event$bowl_ability

  ## Resolve the ball event
  wicket <- ifelse(ball$event$wicket_type != "", TRUE, FALSE)
  wicket_type <- ifelse(ball$event$wicket_type == "", NA, ball$event$wicket_type)
  boundary4 <- ifelse(ball$event$runs == 4, TRUE, FALSE)
  boundary6 <- ifelse(ball$event$runs == 6, TRUE, FALSE)
  legal_delivery <- ball$event$legal_delivery
  rotate_strike <- ball$event$rotate_strike

  ## Determine number of runs conceded by the bowler
  runs_conceded <- ball$event$runs

  ## Deal with extras
  if(ball$event$extra_type1 == "wides")
    runs_conceded <- ball$event$runs + ball$event$extra_runs
  if(ball$event$extra_type1 == "noballs")
    runs_conceded <- ball$event$runs + 1

  ## Compute the number of standardised runs conceded
  standardised_runs_conceded <- runs_conceded / bat_ability

  ## Update the innings scoreboard
  innings_data$runs_scored <- innings_data$runs_scored + ball$event$runs
  innings_data$extras_scored <- innings_data$extras_scored + ball$event$extra_runs
  innings_data$total_scored <- innings_data$runs_scored + innings_data$extras_scored
```

```
## Update the innings parameters
innings_data$balls_bowled <- ifelse(isTRUE(legal_delivery),
                                    innings_data$balls_bowled + 1,
                                    innings_data$balls_bowled)
innings_data$wickets <- ifelse(isTRUE(wicket),
                               innings_data$wickets + 1,
                               innings_data$wickets)

## Update the batsman object
batsman$runs_scored <- batsman$runs_scored + ball$event$runs
batsman$balls_faced <- ifelse(ball$event$extra_type1 != "wides",
                              batsman$balls_faced + 1,
                              batsman$balls_faced)
batsman$out <- ifelse(isTRUE(wicket), TRUE, batsman$out)
batsman$how_out <- ifelse(isTRUE(wicket), wicket_type, batsman$how_out)

## Update the batsman object
if(ball$event$runs %in% c(4, 6))
{
  batsman$fours <- ifelse(ball$event$runs == 4, batsman$fours + 1, batsman$fours)
  batsman$sixes <- ifelse(ball$event$runs == 6, batsman$sixes + 1, batsman$sixes)
}

## Update the bowler object
bowler$balls <- ifelse(isTRUE(legal_delivery), bowler$balls + 1, bowler$balls)
bowler$ball_count <- bowler$ball_count + 1
bowler$runs_conceded <- bowler$runs_conceded + runs_conceded
bowler$standardised_runs_conceded <- bowler$standardised_runs_conceded +
                                     standardised_runs_conceded
bowler$wickets <- ifelse(wicket_type %in%
                         c("bowled", "caught", "caught and bowled",
                           "hit wicket", "lbw", "stumped"),
                         bowler$wickets + 1, bowler$wickets)

## Update the batting scorecard
innings_data$batting$scorecard$runs_scored[batsman$position] <- batsman$runs_scored
innings_data$batting$scorecard$balls_faced[batsman$position] <- batsman$balls_faced
innings_data$batting$scorecard$strike_rate[batsman$position] <-
   innings_data$batting$scorecard$runs_scored[batsman$position] /
   innings_data$batting$scorecard$balls_faced[batsman$position]
innings_data$batting$scorecard$out[batsman$position] <- wicket
innings_data$batting$scorecard$how_out[batsman$position] <- ifelse(isTRUE(wicket),
                                                                    wicket_type, "")
```

```
  ## Extras and total runs scored
  innings_data$batting$scorecard$runs_scored[12] <- innings_data$extras_scored
  innings_data$batting$scorecard$runs_scored[13] <- innings_data$total_scored

  ## Update the bowling scorecard
  bowl_position <- which(innings_data$bowling$scorecard$bowler ==
                       innings_data$bowling$bowler$name)
  innings_data$bowling$scorecard$balls[bowl_position] <- bowler$balls
  innings_data$bowling$scorecard$runs_conceded[bowl_position] <- bowler$runs_conceded
  innings_data$bowling$scorecard$standardised_runs_conceded[bowl_position] <-
    bowler$standardised_runs_conceded
  innings_data$bowling$scorecard$wickets[bowl_position] <- bowler$wickets
  innings_data$bowling$scorecard$RPO[bowl_position] <-
    ifelse(innings_data$bowling$scorecard$runs_conceded[bowl_position] == 0,
           0,
           innings_data$bowling$scorecard$runs_conceded[bowl_position] /
           (6 * innings_data$bowling$scorecard$overs[bowl_position] +
           innings_data$bowling$scorecard$balls[bowl_position]))

  ## Update the batsman/bowler objects in the innings data
  innings_data$batting$batsmen[[innings_data$batting$striker]] <- batsman
  innings_data$bowling$bowler <- bowler

  ## Update the list of bowlers
  innings_data$bowling$players[[bowler$name]] <- bowler

  ## Did batsmen rotate strike?
  if(isTRUE(rotate_strike))
    swap(innings_data$batting$striker, innings_data$batting$non_striker)

  ## Output
  if(isTRUE(ball_output))
  {
    output_ball(batsman = innings_data$batting$batsmen[[innings_data$batting$striker]],
                bowler = innings_data$bowling$bowler,
                runs_scored =  ball$event$runs,
                boundary4 = boundary4, boundary6 = boundary6,
                wicket = wicket, wicket_type = wicket_type,
                extras_scored = ball$event$extra_runs,
                extra_type1 = ball$event$extra_type1,
                extra_type2 = ball$event$extra_type2)
  }
  ## Return the updated innings data
  return(list(event = ball$event, innings_data = innings_data))
}
```

`output_ball` — a function that prints the result of the `sim_ball` function to screen.

```
## A function that summarises the results of a delivery
output_ball <- function(batsman, bowler,
                        runs_scored, boundary4, boundary6,
                        wicket, wicket_type,
                        extras_scored, extra_type1, extra_type2)
{
  ## Output the result of the delivery
  cat(bowler$name, " to ", batsman$name, ". ", sep = "")

  ## Number of euns scored
  if(runs_scored == 1)
  {
    cat(runs_scored, " run", sep = "")
  }else{
    cat(runs_scored, " runs", sep = "")
  }

  ## Extra runs scored
  if(extras_scored == 0)
  {
    cat(". ")
  }else{
    if(extra_type2 == "")
    {
      cat(" + ", extras_scored, " ", extra_type1, sep = "")
    }else{
      cat(" + ", extras_scored, " ", extra_type1, " & ", extra_type2, sep = "")
    }
  }

  ## Wicket
  if(isTRUE(wicket))
    cat("OUT, ", wicket_type, ".", sep = "")

  ## End line
  cat("\n")
}
```

`sim_event` — a function that simulates the sub-event that occurs in the `sim_ball` function.

```
## A function that simulates the event that occurs on a given ball
sim_event <- function(summary_data, batsman, bowler)
{
  ## Get the event summary information for the given ball
  event_summary <- event_probability(summary_data, batsman, bowler)

  ## no wicket/bowler wicket/non-bowler wicket
  event <- sample(event_summary$prob_events$event, size = 1,
                  prob = event_summary$prob_events$prob)

  ## Compute the relevant sub-event
  sub_event_index <- sample(1:length(event_summary[[event]][ , 1]), size = 1,
                            prob = event_summary[[event]]$prob)
  sub_event <- event_summary[[event]][sub_event_index, ]

  ## Create a data.frame containining all relevant main event and sub-event
  ## information pertaining to the ball
  event_output <- cbind(sub_event[ , c("runs", "wicket_type", "extra_runs",
                                       "extra_type1", "extra_type2",
                                       "legal_delivery", "rotate_strike")],
                        "bat_ability" = event_summary$bat_ability,
                        "bowl_ability" = event_summary$bowl_ability)

  ## Output the relevant quantities (runs, wicket, wicket type, extra, etc.)
  return(list(event = event_output,
              wicket_prob = event_summary$bowler_wicket_prob +
                            event_summary$non_bowler_wicket_prob,
              bowler_wicket_prob = event_summary$bowler_wicket_prob,
              bat_ability = event_summary$bat_ability,
              bowl_ability = event_summary$bowl_ability,
              expected_rspb = event_summary$expected_rspb,
              expected_rcpb = event_summary$expected_rcpb))
}
```

event_probability — a function that computes the probability of each main event and sub-event occurring in the sim_event and sim_ball functions, given the specific bowler/batsman matchup.

```
## A function that computes the probability of each main event and sub-event
## occurring on any given delivery adjusted for the batsman/bowler matchup
event_probability <- function(summary_data, batsman, bowler)
{
  ## Extract the batsman's ability given the score and innings #
  ## Effective batting average mu(x)
  if(batsman$runs_scored > 400)
  {
    bat_ability <- batsman$mux[401]
  }else{
    bat_ability <- batsman$mux[batsman$runs_scored + 1]
  }

  ## Compute the bowler's ability given the innings #
  bowl_ability <- bowler$bowl_ability

  ## Compute the expected bowling average for the batsman/bowler matchup
  expected_bowl_average <- bowl_ability * bat_ability

  ## Compute the relative bowling economy and batsman strike rates
  relative_bowling_er <- bowler$economy_rate / summary_data$averages$bowling_economy_rate
  relative_batting_sr <- batsman$strike_rate / summary_data$averages$batting_strike_rate

  ## Compute the expected runs scored/conceded per ball
  ## Assume the number of wides/no balls conceded remains constant regardless
  ## of bowler ability
  expected_rcpb <- summary_data$averages$bowling_economy_rate *
                   relative_bowling_er * relative_batting_sr
  expected_rspb <- expected_rcpb -
                   (summary_data$averages$runs_conceded_per_ball -
                    summary_data$averages$runs_scored_per_ball)

  ## Relative runs scored/conceded per ball compared with the historic
  ## Test average
  relative_rcpb <- relative_bowling_er * relative_batting_sr
  relative_rspb <- expected_rspb / summary_data$averages$runs_scored_per_ball

  ## Compute the bowler's expected strike rate
  expected_bowler_sr <- expected_bowl_average / expected_rcpb
  relative_bowler_sr <- expected_bowler_sr / summary_data$averages$bowling_strike_rate
```

```r
## The expected bowler strike rate is therefore inverse of the updated
## bowler-credited wicket probability
## The next step is to adjust the run scoring events to ensure the batsman
## scores runs and the bowler takes wickets at the appropriate rate

## Initialise the probability of a bowler wicket using the
## historical Test average
bowler_wicket_prob <-  subset(summary_data$prob_events$events,
                              event == "bowler_wicket")$prob

## Convert probabilities into odds
bowler_wicket_odds <- bowler_wicket_prob / (1 - bowler_wicket_prob)

## Adjust odds using the relative estimates
adjusted_wicket_odds <- bowler_wicket_odds * (1 / relative_bowler_sr)

## Convert bowler wicket odds back to probability
adjusted_wicket_prob <- adjusted_wicket_odds / (1 + adjusted_wicket_odds)

## Normalise the event probabilities using the adjusted wicket probability
## The probability of a non-bowler credited wicket should remain constant,
## regardless of the batsmen at the crease
## The only thing we want to change is the probability of a bowler-credited
## wicket, or no wicket occurring
prob_events <- summary_data$prob_event$events
non_bowler_wicket_prob <- subset(prob_events,
                                 event == "non_bowler_wicket")$prob
subset(prob_events, event == "bowler_wicket")$prob <- adjusted_wicket_prob
subset(prob_events, event == "no_wicket")$prob <- 1 - adjusted_wicket_prob -
                                                  non_bowler_wicket_prob
no_wicket_prob <- subset(prob_events, event == "no_wicket")$prob

## Default sub-event probabilities
no_wicket <- summary_data$prob_events$no_wicket
run_scoring <- subset(no_wicket, runs > 0)
bowler_wicket <- summary_data$prob_events$bowler_wicket
non_bowler_wicket <- summary_data$prob_events$non_bowler_wicket

## Without adjustment, there are this many runs conceded per ball
current_rspb <- sum(prob_events$prob[1] * run_scoring$runs * run_scoring$prob) +
                sum(prob_events$prob[3] * non_bowler_wicket$runs * non_bowler_wicket$prob)
```

```
## After adjusting the wicket probabilities, we also need to adjust the runs
## scored on run scoring deliveries by computing the adjustment factor
## Assume rate of no balls/wides is the same regardless of bowler
adjustment_factor <- expected_rspb / current_rspb

## Adjust run scoring probabilities using the adjustment factor.
## This accounts for batting strike rate and bowling economy rate
subset(no_wicket, runs > 0)$prob <- adjustment_factor * subset(no_wicket, runs > 0)$prob

## Normalise by adjusting the non-run scoring probability
## In rare instances where the batsman's ability is super low, this can lead
## to negative probabilities
## In such cases, simply make the probability of a non-run scoring event = 0
## This will lead to cases where expected_rspb != adjusted_rspb but such cases
## are so rare that this should not make any practical difference
if(sum(no_wicket[no_wicket$runs > 0, ]$prob) > 1)
{
  ## Set probability of all non-scoring events to 0
  no_wicket[no_wicket$runs == 0, ]$prob <- 0

  ## Normalise the probabilities
  no_wicket$prob <- no_wicket$prob / sum(no_wicket$prob)
}else{
  ## Normalise by adjusting the non-run scoring probability as usual
  susbet(no_wicket, runs == 0)$prob <-
     subset(no_wicket, runs == 0)$prob /
     sum(subset(no_wicket, runs == 0)$prob *
     (1 - sum(subset(no_wicket, runs > 0)$prob))
}

## Expected runs scored per ball after adjustment
adjusted_expected_rspb <- sum(no_wicket$runs * no_wicket$prob * no_wicket_prob) +
                          sum(non_bowler_wicket$runs *
                              non_bowler_wicket$prob *
                              non_bowler_wicket_prob)

## Expected runs conceded per ball after adjustment
adjusted_expected_rcpb <- adjusted_expected_rspb +
        sum(subset(no_wicket, extra_type1 == "noballs")$prob * no_wicket_prob) +
        sum(subset(no_wicket, extra_type1 == "wides")$prob *
            subset(no_wicket, extra_type1 == "wides")$extra_runs * no_wicket_prob)
```

```
  ## Return a list of data.frames containing the main event and sub-event
  ## probabilities, adjusted for the batsman/bowler matchup
  return(list(prob_events = prob_events,
              no_wicket = no_wicket,
              bowler_wicket = bowler_wicket,
              non_bowler_wicket = non_bowler_wicket,
              bowler_wicket_prob = adjusted_wicket_prob,
              non_bowler_wicket_prob = non_bowler_wicket_prob,
              bat_ability = bat_ability,
              bowl_ability = bowl_ability,
              expected_rspb = adjusted_expected_rspb,
              expected_rcpb = adjusted_expected_rcpb))
}
```