PERFORMANCE MODELLING IN SPORT

Stephen R. Clarke
BSc(Hons), DipEd, MA

A thesis submitted for
the degree of

Doctor of Philosophy

School of Mathematical Sciences
Swinburne University of Technology

December 1997

Acknowledgments

My thanks to the following for their contribution to this thesis.

John Norman for first arousing my interest in applications in sport, for his continued encouragement, support and collaboration over many years, and for his advice and assistance on this thesis.

Ray Stefani, a most knowledgeable person on sport, for his collaboration on football tipping and home advantage.

David Noble, for the interest he has shown over the years in sporting applications, for his collaboration on several projects, for conscientiously reading my drafts and offering his considered and helpful advice, and for being such a considerate supervisor.

Many anonymous referees, editors and other colleagues, who with their comments and advice have improved the original papers which form the bulk of this thesis.

Swinburne and in particular the School of Mathematical Sciences for providing encouragement for my research and leave to write up this thesis.

My mother, Connie Clarke, for providing the initial impetus by often asking when I was going to do a PhD, and who by taking her first parachute jump at age 83 demonstrated it was never too late to achieve an ambition.

Finally my wife Kaye, for bearing the brunt of late nights and untidy home offices, for listening to my half formed ideas, correcting my original drafts, helping to proof read this thesis, and assisting me in many other ways.

Candidates statement

This work has not previously been submitted by the candidate for any degree or similar award in a tertiary institution.

Except where acknowledged in the text the work is my own.

Stephen R. Clarke

Preface

This thesis is submitted under Swinburne University Policy for the Degree of Doctor of Philosophy, Regulation 4.2, Admission to candidature for thesis by publication. This allows for

> "thesis by publication on the basis of research which has been carried out prior to admission by candidature and which has been published, normally in texts or refereed publications. Only those publications not previously submitted by the applicant for a degree in any tertiary institution may be included in support of the application for candidature and in the Candidates thesis. In such cases the...Candidate must produce a significant body of work on an integrated theme that will comprise the thesis. This thesis would normally include a substantial introduction showing the relevance of the publications to development of the theme, plus a series of publications and any necessary linking commentary. All publications must be appropriately identified and referenced and the contributions of the Candidate to each publication must be clearly specified."

From my work in applications in sport over many years, I have selected for this thesis some of the papers relating to performance measurement in football (including Australian rules and English soccer), and cricket. Most of the football work relates to measurement of home advantage and forecasting match results. The football work has the advantage from an operational research viewpoint that it is heavily based on real data, and has also been implemented to the degree that the results have been regularly published in the popular media. The work on performance measurement and tactics in cricket makes up the second part of the thesis and uses the traditional OR technique of dynamic programming. I toyed with the idea of using my work on squash and badminton, which also falls in the general theme, but thought the thesis already long enough.

In this thesis each chapter usually comprises one publication. This has meant some repetition, as there is some commonality in literature surveys and game descriptions in various papers. However it has the advantage that each chapter is self contained and may be read in isolation. The content of each publication is as it originally appeared, but with style changes to create uniformity of presentation throughout the thesis. This necessitated some superficial changes. All headings, tables, figures and equations have been renumbered using the chapter number as a prefix. The style of referencing has been standardised, and the references consolidated in the bibliography at the end of the thesis. In a very few cases notation has been changed and tables and figures given a more descriptive caption.

A commentary section which details later developments or further work has been added at the end of some chapters.

The following lists the publications used in each chapter, states if they were refereed, and specifies the candidates input to co-authored papers.

Chapter I. Written specifically for the thesis.

Chapter II. The early part of this chapter is based loosely on Stefani, R. T., & Clarke, S. R. (1991), Australian rules football during the 1980s. *ASOR Bulletin*, 10(3), 11-15. The paper evolved from a 50% contribution by Professor Stefani and a 50% contribution from myself. In Chapter II, I have repeated all the analysis on 16 years of data.

The latter part of the chapter is from Clarke, S. R. (1997a), Home ground advantage in the Australian Football League, 1980-85, a paper presented at the APORS conference 1997, Melbourne.

Chapter III. Clarke, S. R., & Norman, J. M. (1995), Home ground advantage of individual clubs in English soccer. *The Statistician,* 44, 509-521. Refereed. The conception of this paper resulted from joint collaboration. Professor Norman assisted with the collection of data, which I computerised and analysed. I wrote the first draft of the paper, but Professor Norman assisted with following drafts and provided the necessary local knowledge.

The commentary material is based on Clarke, S. R. (1996b), Home advantages in balanced competitions - English soccer 1990-1996. In N. de Mestre (Ed.), *Mathematics and Computers in Sport* (pp. 111-116). Gold Coast, Qld.: Bond University.

Chapter IV. Clarke, S. R. (1993), Computer forecasting of Australian rules football for a daily newspaper. *Journal of the Operational Research Society*, 44(8), 753-759. Refereed.

Chapter V. Clarke, S. R. (1992), Computer and human tipping of AFL football - a comparison of 1991 results. In N. de Mestre (Ed.), *Mathematics and Computers in Sport* (pp. 81-93). Gold Coast, Qld: Bond University.

Chapter VI. Stefani, R. T., & Clarke, S. R. (1992), Predictions and home advantage for Australian rules football. *Journal of Applied Statistics*, 19(2), 251-261. Refereed. This paper resulted from joint discussions over several years on computer tipping and home advantage. I provided the data and my computer methods for the Australian rules analysis, Professor Stefani provided the least squares method as well as the data and results for the other sports.

Chapter VII. Clarke, S. R. (1996a), Calculating premiership odds by computer - an analysis of the AFL final eight play-off system. *Asia Pacific Journal of Operational Research*, 13(1), 89-104. Refereed. The commentary is based on my own work not yet submitted.

Chapter VIII. Clarke, S. R., & Norman, J. M. (1998b), When to rush a behind in Australian rules football: a dynamic programming approach. *Journal of the Operational Research Society*, in press. Refereed. In terms of input this paper is the complement of Chapter III. The conception of the paper resulted from joint collaboration. The development of the model was a joint exercise, but Professor Norman wrote the computer implementation and the first draft of the paper. I assisted with further development of the model and the paper, and provided the necessary local knowledge.

Chapter IX. Clarke, S. R. (1997b), Test Statistics. In J. Bennet (Ed.), *Statistics in Sport*. Edward Arnold (to appear). This is a slightly reduced version of the material to be published in the text. Several figures and tables extracted from published works and required for the text have been omitted here.

Chapter X. Clarke, S. R. (1986), Another look at the 1985/86 Sheffield Shield competition cricket results. *Sports Coach*, 10(3), 16-19.

Chapter XI. Clarke, S. R. (1988), Dynamic Programming in one day cricket - Optimal scoring rates. *Journal of the Operational Research Society*, 39(4), 331-337. Refereed.

The commentary is based on Johnston, M. I., Clarke, S. R., & Noble, D. H. (1992), An analysis of scoring policies in one day cricket. In N. de Mestre (Ed.), *Mathematics and Computers in Sport* (pp. 71-80). Gold Coast, Qld: Bond University. I wrote this paper based on joint development of my original idea by all authors. Mr Johnston wrote all the necessary computer programs.

Chapter XII. Clarke, S. R., & Norman, J. M. (1998a), Dynamic programming in cricket: Protecting the weaker batsman. *Asia Pacific Journal of Operational Research*, in press. Refereed. I developed the models and paper with the continued assistance of Professor Norman's help and advice.

Chapter XIII. Clarke, S. R., & Norman, J. M. (1997c), To run or not? Some Dynamic Programming models in cricket. In R. L. Jenson & I. R. Johnson (Eds.), *Proceedings of the Twenty-Sixth Annual Meeting of the Western Decision Sciences Institute* (pp. 744-746). Hawaii: Decision Sciences Institute. The full paper, which appears here, was refereed. The above is a shorter version published in the proceedings as is the practice at this conference. I developed Models 1 and 2 along with the necessary computer programs. Professor Norman developed Model 3 along with the necessary computer programs. I wrote the first draft of the paper. Both authors provided assistance and advice to the other in the development of the models and the paper.

Chapter XIV. Johnston, M. I., Clarke, S. R., & Noble, D. H. (1993), Assessing player performance in one-day Cricket using dynamic programming. *Asia-Pacific Journal of Operational Research*, 10, 45-55. Refereed. I wrote the paper based on joint development of my original idea by all authors. Mr Johnston wrote all the necessary computer programs.

Chapter XV. Written specifically for the thesis.

Table of Contents

List of Illustrations

List of Tables

Abstract

This thesis investigates problems of performance modelling in sport. Mathematical models are used to evaluate the performance of individuals, teams, and the competition rules under which they compete. The thesis comprises a collection of papers on applications of modelling to Australian rules football, soccer and cricket. Using variations of the model $w_{ij} = u_i + h_i - u_j + e_{ij}$ where w_{ij} is the home team winning margin when home team i plays away team j, u_i is a team rating, h_i is an individual ground effect and e_{ij} is random error, the evaluation of team home ground advantage effect (HA) is studied in detail. Data from the Australian Football League and English Association Football for 1980 to 1995 are investigated. The necessity of individual team HAs is demonstrated. The usual methods of calculating HA for competitions is shown to be inappropriate for individual teams. The existence of a spurious HA when home and away performances are compared is discussed. For a balanced competition, fitting the above model by least squares is equivalent to a simple calculator method using only data from the final ladder. A method of calculating HA by pairing matches is demonstrated. Tables of HA and paired HA in terms of points/game for each year are given. The resultant HAs for both Australian rules football and soccer are analysed. Clearly there is an isolation effect, where teams that are isolated geographically have large HAs. For English soccer, the paired HA is shown to be linearly related to the distance between club grounds. As an application of these methods, the development and implementation of a computer tipping program used to forecast Australian rules football by rating teams is described. The need for ground effects for each team and ground, and the use of heuristic methods to optimise the program is discussed. The accuracy of the prediction model and its implementation by publication in the media is discussed. International comparisons show prediction methods are limited by the data. Methods for evaluating the fairness of the League draw and the finals systems are given. The thesis also investigates the use of dynamic programming to optimise tactics in football and cricket. The thesis develops tables giving the optimal run rate and the expected score or probability of winning at any stage of a one-day cricket innings. They show a common strategy in one-day cricket to be non-optimal, and a heuristic is developed that is near optimal under a range of parameter variations. A range of dynamic programming models are presented, allowing for batsmen of different abilities and various objective functions. Their application to performance modelling are shown by developing a radically different performance measure for one day cricket, and applying it to a one-day series.

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

1.1. Introduction

Would Casperov have beaten Casablanka? Was John Coleman better than Gary Ablett? Is Greg Norman a better golfer than Nick Faldo? Is Dean Jones the best one-day cricketer? Which is the best golf hole? Is Essendon a better football team than Collingwood? Is the final 8 system better than the final 6? To answer these questions you need criteria.

In most sporting competitions there is a major criterion that teams or individuals are judged on - win the premiership, win the tournament, score as much as you can. However other statistics and measures of performance are also used as they presumably affect the likelihood of the desired outcome, or the way in which it is achieved. The weight of a boxer is worth documenting because it is known that this affects markedly his punching power, whereas his shoe size should be of little interest as it is of no relevance to his chance of winning or his method of winning. For a golfer, the average length of drive is recorded as it is thought to influence his chance of winning. However too often such statistics are measured in isolation of other factors (out of context). For example the putting ability of a golfer may be measured by the number of putts. However while a small number of putts could indicate the golfer is a good putter, it could also indicate a good chipper so his putts are generally from a shorter distance.

Similarly the average score for a one day cricketer is usually recorded. However a large score, scored at a slow rate, may be detrimental to the team. Thus not only should the size of scores be measured, but the rate of scoring. But again, while obviously a large score at a fast rate is optimal, what are the offsets. How do we compare a small score at a fast rate and a large score at a slow rate? How do these interact? Either may be appropriate at different stages of a game. Modelling the game can assist by providing a much more subtle measure of performance.

In many sporting competitions there are three levels at which decisions are taken that reflect on the outcome. The individual player, the team they play for and the competition in which the individual or team participates. Performance measurement should be driven by some model of the system, and reflect what the player or team or competition is trying to achieve. Thus, in team events, individual player performance measures should reflect the extent to which they contribute to team goals. The best interests of the team are then achieved by the player simply optimising his performance measure.

The performance of a set of rules for a competition can also be measured. A competition is generally designed to produce a winner or order of merit. It can be measured by the degree to which this order of merit corresponds to some desired order. For example the effects of different lengths of matches in racquet sports on the chances of the best player winning has been studied in the literature. Similarly, one day cricket matches are often decided by the application of rain interruption rules. The fairness of these rules in preserving a team's chances of winning can be measured through modelling. In multi-team competitions, while the order of merit a team achieves may be the primary measure of its success, there may be other factors that combine to produce the teams success. For example, in most team sports it is recognised that team ability and home advantage (HA) are important. Modelling is necessary to separate out the importance of these various effects to individual teams. The competition rules can then be judged on the degree to which these traits are rewarded or allowed for.

While one linking theme of my research is performance measurement through modelling, a second is measuring variability. Almost all current performance measures in sport are averages or sums. Yet it is widely recognised that consistency (or variability) is a major determinant of sporting success. However there are virtually no commonly kept statistics in major sports that measure variability. Once a modelling approach is taken, because variability is such an important determinant of success, many measures of variability will naturally arise. Thus HA is a natural measure to apply to football teams as it explains the variability in their performance from home to away.

1.2. Why sport?

When I first began research into sport, a perception existed that something so much fun could not be a serious pursuit. However the application area has become acceptable, and I no longer have to justify working in sport. A radio interviewer once accused me of wasting government money by indulging in such frivolous activity. Just in case there are still some doubters, it is perhaps worthwhile to reiterate some of the reasons why sport is a worthwhile area for the application of mathematical modelling.

- Sport is important to a large proportion of the population. Many people spend a great part of their lives playing, watching, discussing or thinking about sport. This interest is not limited to one socio-economic group but encompasses all income and education levels. A quick look at any daily newspaper will illustrate the extent to which sport dominates large sections of the media.
- Sport is big business. For example, the American Football league sold its 1990 rights to TV for $1.25 billion. Although still well behind the USA, with the introduction of sophisticated marketing methods, sports like cricket, football and tennis are becoming big money earners in Australia. Over the years Governments and industry have capitalised on the commercial aspects of sport so that it is now a multi billion dollar industry, and probably rivals traditional sectors. Its importance to economies and reputation is illustrated by the fervour with which cities compete for the honour of holding the Olympic games.
- Most importantly, sport abounds with problems to tackle. Administrators, selectors, coaches and players continually make decisions which could be assisted by analysis. Every sport has its golden rules which are often untested. Many sports collect reams of statistics, which are often left unanalysed. Every fan continually makes conjectures which are unproven. Many of these problems are amenable to analysis by statistical and operational research techniques.
- Sport has become an acceptable, and even encouraged, area of publication. Mottley (1954) first suggested the use of OR techniques in sport, and the publication in the mid 1970's of Machol, Ladany, & Morrison (1976) as a special issue of Management Science, and)Ladany & Machol (1977) based on previously published articles on strategy showed the many possibilities for quantitative research in sport. As researchers have tackled problems and published solutions, the realisation has grown on learned societies and journals that members and readers are interested in these applications, not only because of the techniques used but also because they share the general public's inherent interest in sport. The problem and its solution is of major interest, not just the solution technique.
- Many professional societies, and some journals, now have special sections devoted to applications in sport. To give three examples, the American Statistical Society has a special interest section in sport, The Royal Statistical Society has a regular section in its Series D journal focusing on sport, and the Australian Mathematical Society have a biennial conference on mathematics and computers in sport.
- Because many papers are problem oriented, the area is truly inter disciplinary. Papers on mathematical modelling in sport appear in mathematics, operational research, computing, statistics, psychology, engineering and even sports science

journals.

- Finally sport is a great source of problems for teaching. Because of their interest in the problems, because they have a stake in the answers, students will question techniques and methods used if the application is in sport. They see quantitative methods as being useful in solving real problems.

Since this thesis is a selection of my work in the area of sports modelling and performance measurement, I will give a review of that work. In doing so I will highlight some of the themes that have characterised my work and by association an overview of operational research work in related areas.

1.3. Scoring in racquet sports

Clarke & Norman (1979), my first paper, resulted from an investigation with John Norman on squash. This began my interest in applications in sport and the paper illustrated three important aspects of modelling in sport. The first was measurement based on a mathematical model. Players and commentators would often complain that the English scoring system resulted in longer games than the American system, but these observations were based on subjective feelings rather than any analysis. It would have been possible to tackle this problem by collecting data on actual matches. However Schutz (1970) had developed a mathematical model for tennis and applied it to various scoring systems in tennis. Schutz & Kinsey (1977) applied the method to squash, and assuming players had a constant probability of winning a rally, investigated the average length of a match under both scoring systems. However they were unable to solve the model for the mean and standard deviation of the length of the game and hence resorted to simulation. We were able to set up recurrence relations for the probability of winning and the mean and variance of the number of rallies left in the match, to solve the problem numerically. (In retrospect, the paper would have been better titled numerical results).

These results showed the importance of randomness and variation. The work to some extent quantified the importance of randomness or luck in squash, by giving tables that showed the probability the better player won the match. Randomness exists in all sporting contests but its effects are rarely quantified. In this case the effects were felt not only in which player won, but the length of the game. The mean length of the two games was much the same - it was the variance of the English game that was much larger. The perception of participants that games under the English scoring system were longer was really based not on the average length but the extremes. This is a second important theme in most of my work - the variation in sport and the effects of randomness.

The third issue was tactics. In English scoring a player makes a decision at 8-8 to play to 9 or 10. As a player should make the choice which maximises their probability of winning, we needed to solve this simple preliminary problem before tackling the main point of the paper. Our results were to some extent anti intuitive. The more likely a person was to lose the next point, the greater the advantage in choosing to play to 9 rather than 10. There are many opportunities in sport to analyse tactics, and some of my work has gone down this path. In Clarke (1979) I extended the model to allow for each player to have different probabilities of winning their serve, and looked in detail at tie point decisions in both English and American squash. Clarke & Norman (1978) put the results of the tie breaker research to the attention of squash players and administrators. It is important that results of such research be brought not only to other researchers through the professional journals, but also to players and followers of the game. Finally, as I stated above, examples from sport are great motivation for students. Clarke (1984, 1985a, 1985b) were the first of several of my papers designed to bring this idea to the attention of teachers.

Many researchers have subsequently investigated the effects of different scoring systems in racquet sports on the chances of players winning and the mean and variance of the length of matches, for example Croucher (1982c, 1986), Pollard (1983), Riddle (1988). The problem has now been solved analytically, as distinct from the numerical solutions we found. Miles (1984) related the probability of winning to the mean length of the match through the concept of efficiency. Tennis commentators often talk about important points, and this concept of importance was first given a mathematical definition by Morris (1977). Pollard (1986a) performed an extensive study of sports scoring systems, in particular investigating the concept of 'important points' and efficiency. Because a sports scoring system is attempting to find the better of two players, the results found more general application in hypothesis testing in Pollard (1992). On a different tack, Wright (1988) extended the model in squash to allow for different probabilities of winning a point depending on the side of the court to which the player is serving.

The research output of papers in this area also illustrates a problem perhaps more common to sport than other application areas. In most topic or application areas, researchers concentrate on a technique or application area and build up a knowledge of the literature and previous work over many years. Papers in sport are often 'one off'. Because of their own involvement in sport, a researcher will independently analyse a problem unaware of previous work. Since papers are published in a wide range of journals, a literature search solely in the researchers field will fail to bring up relevant papers. Because of this, problems often reappear in the literature with little recognition of previous work. For example, Alexander, McClements, & Simmons (1988), Brooks & Hughston (1988) and Simmons (1989) all revisit tie point strategy without recognition of much of the previous work.

1.4. Dynamic programming in cricket

In analysing squash and tennis scoring systems, the method involved defining the state of the match as the score. By finding a relationship between the value of the required measure (probability of winning, mean and variance of number of rallies in the remainder of the match etc.) in the state before a rally was played, to its value in surrounding states reached after the point was played, it was possible to start at the end of the match when the values were known and work backwards to the beginning or any intermediate score. This is only dynamic programming (DP) without the decisions, and it was a natural progression to use DP in sporting applications. Its suitability in sporting applications had been proposed by Bellman (1977) and there are now many examples where a DP formulation has the potential to assist the sports person with decision making. Norman (1995) in giving one example of an application of DP in each sport lists 10 papers. Sphicas & Ladany (1976) use it to extend the static strategy developed in an earlier paper, Sphicas, Humes, & Ladany (1975), by allowing the optimal aiming line in the long jump to depend on the jumps already taken. Norman (1985) used DP in a simple decision problem on whether to serve hard or soft at tennis, and Hayes & Norman (1984) is an interesting study on optimal route choice in orienteering. Recently Clarke & Norman (1997b, 1998b) have applied the technique to Australian rules football for the first time (see Chapter VIII). However the method had surprisingly not been applied to cricket, although Thomas (1978) had used cricket to illustrate a Markov problem in a student exercise (see also White (1993, p177-179).

In spite of the interest in cricket, and the availability of extensive statistics, there has been surprisingly little research on cricket (see Chapter IX for a literature review). Elderton (1945) and Wood (1945) in two of the earliest articles using quantitative methods in sport, investigate the measurement of consistency and the distribution of cricket scores. Elderton notes that if a score is equally likely to advance by one whatever the score, cricket scores should follow a geometric distribution. However he then denies the inherent variability this implies by defining a consistent batsman as one with a zero variance. Wood more correctly in my view, suggests that as the scores of consistent batsmen should follow the geometric distribution, a perfectly consistent batsmen should have a coefficient of variation of scores of 100. Although the record of meeting shows a lively discussion, the issue was not taken up in the literature. In Clarke (1991b, 1994c) I use an improved model that assumes for each *ball* a constant probability of dismissal. This idea is at the heart of most of my cricket models, and produces a geometric distribution for the number of balls faced rather than the score. A distribution of scores off each ball then produces coefficients of variation greater than 100, so that different batsmen, each perfectly consistent but with different distribution of scores off each ball, can have different coefficients of variation. These papers also suggest several other measures of performance for batsmen.

Croucher (1979, 1982b) is typical of the occasional papers on cricket in statistics journals. These often show up interesting points by standard statistical analysis of results. However the OR literature is strangely devoid of cricket applications. Willis (1994), Willis & Armstrong (1993) and Wright (1991, 1992) are more recent papers but apply to scheduling matches and umpires, and are not concerned with the game per se.

The ball by ball nature of cricket lends itself to analysis by DP. In Clarke (1988b), I used a DP model to investigate optimal strategies for both first and second innings. In the first innings, the formulation maximised the expected number of runs in the remainder of the innings. This gave an optimal run rate and expected score in the remainder of the innings depending on the number of wickets in hand. The second innings is a final value problem, and maximises the probability of achieving more runs than the opposition. This gave the optimal run rate and the probability of achieving more runs than the opposition depending on the number of wickets in hand and the number of runs to go. While a relationship between probability of dismissal and scoring rate had to be assumed, Johnston (1992) later showed the results were valid over a wide range of relationships. This work also followed up a suggestion made in Clarke (1988b) and developed the first innings formulation into a radically different performance measure for individual players. This measure successfully combined the twin needs in one day cricket of scoring big and scoring fast by allocating to each player the excess scored each ball over the optimal given by the DP formulation. Details and some results are given in Johnston et al. (1992, 1993). The method ensures that a player's performance measure is consistent with team goals, and a player will maximise his measure by maximising his team's chance of winning.

The above models all assume batsmen are of equal ability. This restriction has been removed in several models which investigate a range of possible objective functions, some applicable to test cricket. Clarke & Norman (1995b, 1998a) investigate the problem of putting the weaker batsman on strike at the end of the over, so the better batsman is on strike at the beginning of the new over. The model is soluble analytically and could form a useful teaching example. However, while it produces optimal strategies to achieve the objective, the paper also shows this does not necessarily minimise the weaker batsman's exposure to the strike. Clarke & Norman (1997c) maximise the expected score in the remainder of the innings. In both these papers the decision is made after the ball is bowled and the batsman decides whether to take all the runs or not. The second also looks at a model where the batsman can also decide on his type of stroke. The applications of DP to cricket forms Chapters XI to XIV of this thesis.

1.5. Home advantage

In Clarke (1986a) , I looked at the problem of HA in cricket (see chapter X). That paper showed that in Australian cricket, most states won more than 50% of their matches at home, and all state teams performed better at home than away. However, because some grounds were more likely to produce outright results which carried more points than first inning victories, many more points were allocated for matches on some grounds than others. Thus the final year ladder was distorted by this unfair advantage carried by some states. Alternative methods of ranking the states were investigated.

The phenomena of HA has long been recognised by sports fans and has been the basis of considerable study since the seventies. In the first detailed study, Schwartz & Barsky (1977) found the percentage of matches won by the home team to be 53% in pro baseball, 60% in pro football, 64% in ice hockey and 64% in college basketball. They advanced three explanations for HA, learning factors, travel factors and crowd factors. Various studies have subsequently looked at the importance of these factors.

1. *Learning factors*: These cover ground familiarity etc. Pollard (1986b) found the percentage of wins for soccer teams with the smallest and largest playing surfaces, the team playing on artificial turf, and teams with large capacity grounds did not differ

significantly from other teams in the League. Barnett & Hilditch (1993), using more refined analysis, showed soccer teams playing on artificial surface enjoyed a higher HA than the remaining teams. This finding was supported by Clarke & Norman (1995).

2. *Travel factors*: These generally are a disadvantage to the visiting team and cover physical and mental fatigue, and disruption of routine due to travel. Both Schwartz & Barsky (1977) and Pollard (1986b) point out that as travel has become easier and progressed from train to air over a long period the HA has remained constant. Schwartz & Barsky (1977) argued that as the season progresses, travel effects should accumulate and HA increase. However they found no consistent evidence of this, and this has been repeated in later studies by Courneya & Carron (1991) and Pace & Carron (1992). In similar multiple regression studies in baseball and hockey, they found that travel factors such as distance, number of time zones crossed, direction of travel, time between games, number of successive games at home, and length of visitors road trip accounted for less than 1.5% of the variation in win-loss outcome. Snyder & Purdy (1985) found the HA was 64.7% when visiting basketball teams travelled less than 200 miles, but 84.6% when they travelled more than 200 miles. On the other hand Pollard (1986b) found in soccer that in both cases the HA was 64.3%.

3. *Crowd factors*: This includes social support for the home team and possible referee bias. Schwartz & Barsky (1977) attributed HA in the main to audience support. They claimed in baseball that increments in attendance can directly enhance the home team's chances of winning. Dowie (1982) argued that the HA is common across four soccer divisions where average crowd size varies by a factor of ten. Pollard (1986b) concurred, and also used the constancy of HA over the pre war and post war period when crowds halved as further evidence of a lack of influence of absolute crowd size on HA. Pollard (1986b) also used the constancy of HA across divisions to discount crowd density, which ranges from 20% in Division 4 to 70% in Division 1 as a factor in HA. Clarke & Norman (1994) support this finding although Bland & Bland (1996) disputed the claim. However Schwartz & Barsky (1977) found the home team's winning percentage increased with increasing crowd density, and Neville et al. (1996) found HA varying significantly across divisions in a manner related to mean attendance.

As can be seen from the above selection, most of this research is documented in the sports science and psychology literature, and uses in my view inappropriate measures. Courneya & Carron (1992) give a comprehensive review of this work. In the section where they survey the 'what' of HA, (the relationship between game location and outcome), 16 studies are listed, covering at various levels sports such as baseball, hockey, US football, basketball and soccer, and all but one with the home win percentage listed. Other measures such as points per game are not investigated in any of these studies. Pollard (1986b) is typical of the approach in that HA is measured as the number of games won by teams playing at home expressed as a percentage of all games played. However modelling shows the probability that the home team wins depends not only on the HA but the difference between team performance levels and the variability of results. Thus the percentage of home wins in a competition depends on the range of performance levels in the group as well as the HA. Snyder & Purdy (1985) show the limitations of the usual approach, when in looking at a university basketball competition they found division 2 teams won only 40% of their home matches against division 1 teams. This implies the quality of opposition effect overshadowed the HA effect. Gayton, Mathews, & Nickless (1987) also make this concession, when they exclude from their study teams that make a clean sweep of the finals because, "as Baumeister & Steinhilber (1984) note, the home court advantage is not likely to appear when one team is far superior to the other." Because the quality of teams differ, we must allow for differences in ability and measure HA by comparing a team's home and away performance. Thus HAs, particularly of individual clubs, can only be investigated properly through the use of models that incorporate the performance level of the teams as well as a HA.

Such work was being undertaken, but by researchers whose primary interest was in forecasting sporting contests. To do this successfully they had to measure HA. Although Harville (1980) gave estimates of just over two points per game with standard errors of about 0.4 of a point for the common HA in NFL for each year from 1971 to 1977, and Stefani (1980) quotes a three point HA for college and a two point HA for pro football, both these papers are ignored by Courneya & Carron (1992) in their literature review.

1.6. Computer forecasting of sport results

There are several models that can be used in modelling results of games between two teams. To facilitate discussion we give some that are common in the literature. Let w_{ij} be the winning margin when the home team i plays away team j in match k. Note that w_{ij} could be a 1,0 or similar win/loss variable, but it is more usual to use a goal or point margin. Let u_i be a rating for team i. This summarises a team's level of performance, their ability or form, and is modelled as either a fixed or random effect. Let e_{ij} be a random error, usually assumed to be zero mean.

Model 1. $$w_{ij} = u_i - u_j + e_{ij} \tag{1.1}$$

Since j and k are dependent through the match schedule, j is more correctly specified as $j(k)$, but I will normally not specify k in the models. This model allows for no HA. As such there is no need to use home and away to differentiate between teams, and some writers use winning and losing team. Since (1.1) is clearly only soluble to within an additive constant, an extra condition such as the average rating is 100 or zero is necessary.

Model 2. $$w_{ij} = u_i + h - u_j + e_{ij} \tag{1.2}$$

where h is a common HA and includes all that is advantageous for a team playing at home and disadvantageous for a team playing away. The team rating u_i now is interpreted as a measure of the performance of team i on a neutral ground. Since in some competitions not all matches are played on the home ground of one of the teams, h is often interpreted as 0, $+h$ or $-h$ depending on whether the kth match is played on a neutral ground, or the home ground of team i or team j. In some cases the approach taken is to pre calculate HA, and adjust the match results for HA. Model 1 can then be fitted to the adjusted results.

Most of the literature before work in which I was involved used models such as the above. One of the advances of the work in this thesis is the use of models that allow for different teams to have different HAs.

Model 3. $$w_{ij} = u_i + h_i - u_j + e_{ij} \tag{1.3}$$

where h_i includes all that is advantageous for team i playing at home and disadvantageous for any other team playing at i's home ground. This allows for different HAs for each team.

This model was suggested by Stefani & Clarke (1992) as a special case of

Model 4. $$w_{ij} = u_i + h_{ij} - u_j + e_{ij} \tag{1.4}$$

where h_{ij} is the paired HA between team i and j. This includes all that is advantageous to the home team and disadvantageous to the away team when team *i* plays team *j*.

In actually fitting these models to an *N* team competition, there will be *N* team ratings, and the models are usually fitted using a matrix of dummy variables. Let $\boldsymbol{w}$ be an *m* x 1 column vector of the *m* match results, $\boldsymbol{e}$ an *m* x 1 column vector of the *m* match errors, $\boldsymbol{r}$ a column vector of the *N* team ratings and a number of HAs depending on the model. $\boldsymbol{A}$ is a selection matrix, where $a_{k,i}$ is +1 if team *i* is the home team for the *k*th match, $a_{k,j}$ = -1 if team *j* is the away team and 0 otherwise, and similar selection coefficients for the h_is. We then have

$$\boldsymbol{w} = \boldsymbol{A}\,\boldsymbol{r} + \boldsymbol{e} \qquad (1.5)$$

Standard least squares theory shows the value of $\boldsymbol{r}$ that minimises the sums of squares of the errors $\boldsymbol{e}^{\boldsymbol{T}}\boldsymbol{e}$ satisfies the equations

$$(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{r} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{w} \qquad (1.6)$$

In this case, the matrix $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ is singular and we need to add the restriction $\Sigma u_i = 0$. This is accomplished using the Lagrange multiplier technique and minimising $\boldsymbol{e}^{\boldsymbol{T}}\boldsymbol{e} + \lambda\, \Sigma u_i$. This results in extra terms being added to the equations represented by (1.5). (In solving the above we are finding estimates of the true model values. I have not used the usual notation to differentiate parameters and their estimates, as it usually obvious from the context, and only complicates the notation.)

All of the above models can be used at the end of a season to rate a team's performance or calculate HAs. This is sometimes called a bit misleadingly prediction, as we can use it to predict (albeit an event that has already occurred) the match results based on the ratings. Perhaps it would be better called match fitting. Alternatively, the ratings can be calculated or estimated on an ongoing basis, and used to forecast future match results. Unfortunately this is also often called prediction.

While there had been sporting computer predictions and ratings published in the American media since the 1940s, their methods were proprietary. Leake (1976) gave details of a least squares rating system for college football, which he showed is analogous to an electrical circuit. The method takes no account of HA. The two pioneers in computer tipping are Stefani and Harville. Stefani began forecasting American college and professional football in 1970 and published details of his least squares method in Stefani (1977). He used least squares to obtain ratings (and hence predictions) on college and pro football teams and college basketball. He claimed 72% accuracy for college football, 68% for pro football and 69% for basketball. His method used (1.1) with *i* and *j* being the winning and losing team. The least squares solution of (1.5) is $\boldsymbol{r} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{w}$. He used the extra criterion average rating equals 100 to ensure $\boldsymbol{A}^T\boldsymbol{A}$ was non singular. In his case (120 teams) the matrix was too large to invert, but by investigating the form of $\boldsymbol{A}^T\boldsymbol{A}$ he showed that the least squares solution gives a team's rating as their average winning margin plus the average rating of their opponents. This can be solved iteratively provided a step is taken to avoid cycling. Note that with this method a team's rating will change, even though they do not play, due to the change in the rating of teams they previously played. For basketball, as the schedule often changes, a simpler method was developed where only ratings of teams that have played are changed. By making some approximations he arrived at the formula for updating ratings based on a match result of $r_{\text{new}} = r_{\text{old}} + \frac{1}{n+1}$ (error in previous forecast) , where n is the number of matches a team has played. This is similar to exponential smoothing with a smoothing parameter that is high at the beginning of the season and low at the end. In a study of half a pro football season, Stefani claimed "essentially the same accuracy" as the iterative method.

In Stefani (1980), he extends his method to include a common HA. By approximating for a large number of games, h becomes the average of home team points minus the away team points. The win margin is adjusted by removing the HA, and then the previous procedure carried out on the adjusted margins. Because he found his expected margin of victory greater than the actual he reduced the predicted margins by a factor L, so his prediction model is $w_{ij} = h + L*(r_i - r_j)$. Again L was chosen to minimise the sum of squares of errors and was about .67 for pro football and .75 for college football. In the case where the number of games used to determine each team's rating is the same, a team's rating will only change due to the outcome of its most recent game. The updating equation he derived is $r_{\text{new}} = r_{\text{old}} + \frac{m-1}{mn-1}$ (error in prediction) , where n is the number of games the team has played and m is the number of games its opponent has played. In this case the necessary calculations are simple enough to be processed on a calculator. Stefani claimed improved accuracy both in terms of correct winners and average absolute error. As it affects my approach, it is interesting to note that although much analysis is performed to find equations that minimise sums of squares of errors, some heuristic models still need to be used, and the accuracy of the predictions is judged by the mean absolute deviation, not the variance, of the errors.

Harville (1977, 1980) described using mixed linear models to predict scores in American National Football League (NFL). He used Model 2 but whereas the common HA is a fixed effect, the performance levels are random effects which change between seasons following a first order auto regressive process. Maximum likelihood procedures are used and the resultant equations solved using numerical iteration, Kalman filtering and smoothing algorithms. The method is applied to seven years of NFL data obtaining the correct winner 70.93% of the time. He also showed the method is slightly less accurate than the betting line as measured by number correct, average absolute error, and average squared error. Accuracy is generally better than Stefani, both in percentage correct and absolute average error. However the method obviously requires more computing power. Harville suggested the accuracy might be improved by allowing the weekly performance levels to be correlated. This would place increased emphasis on more recent games. However when this model was fitted he found no evidence of improvement. He also suggested that in using the method for rating, large margins should be truncated to avoid teams 'running up points'.

There are some problems with the above approaches as they apply to prediction of Australian rules football. The very large margins that can occur in Australian rules, and prediction errors of well over 100 points puts doubt on least squares as the optimising goal. In sport, proportion correct and mean absolute deviation are the usual measures of accuracy used, so it is preferable to optimise these. Although least squares theory can produce exact equations, simplifying assumptions often need to be made to enable solution. It is significant, I believe, that Stefani's methods were developed for regular publication, and from 1971 appeared in the Fort Wirth Star - Telegram. When tips need to be provided regularly, with little time between matches predicted and those already played on which the predictions are based, ease of computation is an important consideration. For these reasons I used a heuristic approach. In a subsequent review of his methods, Stefani (1987) suggested the simpler method of updating only the ratings of teams that play is not only computationally efficient but more accurate. With all ratings being adjusted, some untypical results tend to ripple through the ratings of all teams. In comparing his methods with those of Harville, Stefani finds little difference in the accuracy.

1.7. Application to Australian rules football

In 1980, following a student project, I became interested in computer tipping of Australian rules football. The development of this work is outlined in several papers and is discussed in Chapters IV to VI of the thesis. The results from the program have been published in the daily press from 1981 to 1986, and from 1990 to 1996. They are currently published in newspapers in both Victoria and South Australia, and broadcast weekly on TV in South Australia. This led to joint work with one of the pioneers of computer tipping, Ray Stefani.

In designing a prediction model for Australian rules to compete with human tipsters, my approach, outlined in Clarke (1981), followed the lead of Stefani in producing ratings that could be updated after each match rather than computed from scratch. Since his least squares methods ultimately ended up with updating equations similar to that used by Elo (1978) for rating chess players, it was decided to use exponential smoothing methods. Model 2 was used with a common HA. Because of the large margins possible in Australian rules, the model allowed truncation as suggested by Harville (1980). In Clarke (1988a, 1988c) an improved method, developed in 1986, was described. This used an extension of Model 3 which allowed a ground/team effect in addition to several other effects and used a power method to gradually reduce the effect of large margins and errors. The parameters were optimised by minimising the absolute average error. The methods and output are described in detail in Chapter IV of the thesis. This thesis shows that extremely simple exponential smoothing methods can be used in Australian rules football to provide team ratings and measures of HA that can be used to predict match results with an accuracy comparable to the expert tipster Stefani & Clarke (1992) found this method and Stefani's least squares method gave similar results when applied to 1446 Australian football matches from 1980 to 1989. The least squares method had slightly more correct predictions, while not surprisingly my method had a smaller absolute average error. Neither difference was statistically significant. (see Chapter VI).

1.8. Home advantage of individual clubs

In investigating the reasons for HA, researchers still tend to look at an overall competition. For example, to investigate the effects of travel on HA, two competitions would be investigated, one with a lot of travel and one with little. However this is very inefficient. Within a competition some teams would travel a lot, others little. If the individual HAs of teams could be calculated, then these could be related to the travel those individual teams have to undergo.

The researchers involved in computer tipping were also looking at HA. Stefani & Clarke (1991, 1992) used Model 3 and 4 to investigate individual HAs and paired HAs in Australian rules. This work is the first time in the literature individual HAs have been calculated. Clarke & Norman (1995a) used Model 3 and 4 to investigate individual and paired HAs in English soccer. Kuk (1995), Dixon & Coles (1996) and Dixon & Robinson (1996) are later papers that use models incorporating HA in soccer prediction models.

Harville & Smith (1994) were also investigating individual HAs. They defined home court advantage as the net effect of several factors that may have a (generally positive) effect on the play of the home team, and a (generally negative) effect on the play of the visiting team. They fitted models equivalent to 1, 2 and 3 to college basketball. They pointed out that Model 2 implies the expected difference between two teams on a neutral court is halfway between the expected differences on their own home courts. Model 3 implies the expected difference on a neutral court is the same as the difference in the expected differences between them and a common opponent on that team's home court. They defined home court advantage as the expected difference in score in a game played by a team on its home court minus the expected difference in the score played by the same team on a neutral court against the same opponent. In Model 1 this is zero, in Model 2 this is h and in Model 3 this is h_i. They discuss various estimable functions that reflect performance level, among them the expected overall performance level in relation to the average, $.5\,(2u_i + h_i - (1/n)(\Sigma(2u_i + h_i)) = u_i + .5h_i - (1/n)(\Sigma h_i)$. By looking at the marginal difference in the sums of squares between the models, they find strong evidence for a common HA, and some evidence for different HAs, but also that the practical difference is not great. Clarke (1997a) uses a similar method to show significant evidence of different HAs in Australian rules. Stefani & Clarke (1991, 1992) had already shown the size of these differences to be practically important. My contributions to quantifying HA is discussed in Chapters II, III and VI of the thesis.

1.9. Teaching applications

I have always believed the area of sport is excellent for motivating students at all levels and for making the study of mathematics, operational research and statistics more interesting. Several authors have taken up this theme, for example Croucher (1984,

1994) and Townend (1984) . In Clarke (1984, 1985a, 1985b, 1986b, 1988a, 1992b, 1994b, 1995), Tobin & Clarke (1993) and Clarke & Handley (1994), I have presented much of my material in a form suitable for teachers.

1.10. Thesis structure

This thesis investigates performance measurement in sport through mathematical modelling.

This chapter gives an overview of my and others contribution to modelling in sport. In Chapter II data from Australian rules football from 1980 to 1995 is investigated. When measuring a team's success, the necessity of modelling HA is demonstrated, and problems of HA in an unbalanced competition discussed. The existence of different HAs for different teams is demonstrated, and paired and individual HAs for Australian football teams tabulated and analysed. Chapter III develops special methods of calculating HA and team ratings for balanced competitions. Using 15 years of English soccer data, HAs for all clubs are calculated and analysed. Chapter IV demonstrates the applicability of the methods used by describing a computer tipping program used to rate teams and measure HA. Chapter V compares the accuracy of the computer with that of human tipsters. Chapter VI combines both HA and computer forecasting by comparing different forecasting methods and making some international comparisons of HA. In Chapter VII we begin to look at the performance of the competition structures within which teams compete. We first analyse the performance of a finals system, and then demonstrate how the results of the computer tipping program can be used to quantify the effects of administration decisions and evaluate the fairness of an unbalanced draw. Chapter VIII introduces the use of DP to analyse a tactical situation in football. Chapter IX begins the work on cricket by giving a complete literature review. Chapter X shows that many of the problems discussed in Chapters II to VII have their counterpart in cricket, and that competition rules can give an unfair HA. Chapter XI develops a DP formulation to optimise first and second innings strategy in one day cricket. Chapters XII and XIII extend this in various ways to allow for batsmen of differing abilities and alternative objective functions. Chapter XIV demonstrates how the DP models can be used to measure performance by developing a player rating. Chapter XV summarises the findings of the thesis.

CHAPTER II

HOME GROUND ADVANTAGE IN THE AUSTRALIAN FOOTBALL LEAGUE 1980-95

2.0. Abstract

In this chapter we raise some of the issues concerning home advantage (HA) in Australian rules football. We first look at traditional measures of HA as applied to whole competitions, such as percentage of games won, and alternative measures such as average margin of victory for the home team. We investigate the stability of these measures from year to year and throughout the season. We then look at individual HAs for each team. Two alternative methods of calculating these are investigated. The first is through paired matches, which operates independently of rating teams and does give more detailed information on HA, but is wasteful of data. The second uses all the data, and also gives team ratings as well as HAs. We then investigate the HAs obtained, and look at variations and possible causes or groupings of HA. Finally we look at the overall significance of various models, and show that the models are a significant improvement over previous models assuming a common HA.

2.1. Introduction

The major winter sport of the southern states of Australia is Australian rules football. The game is played with a rugby shaped ball between teams of 18 players on oval grounds of different sizes (the same grounds used for cricket during the summer). A match is played for four quarters, and when time on for play interruptions is added these each last for about 30 minutes. While the major method of moving the ball is by kicking, players can also punch the ball and run with it provided they bounce or touch it to the ground every 15 metres. Players running with the ball can be tackled by the opposition. For infringements of the rules, a free kick is awarded, which allows the recipient some time to dispose of the ball without being tackled by an opponent. A free kick is also given for a mark, awarded when a player catches a kick before the ball first touches the ground or another player. One of the great spectacles of the game is a player soaring over his opponents to take an overhead mark. Over the last 20 years, increasing use of handball and the option to play on after a mark or free kick combined with no off-side rule has meant the game has developed into a very fast flowing game requiring lots of running by players. The scoring region consists of four upright posts. Kicking the ball between the two centre posts scores a goal worth six points, while the region between either centre post and the corresponding outside post scores a behind worth one point. Draws are rare. A typical score might be 18 goals 12 behinds 120 points to 12 goals 15 behinds 87 points for a winning margin of 33 points. Ladder position is in order of premiership points (four for a win and two for a draw) with ties decided on percentage (100 x total points for / total points against). The top teams at the end of the home and away draw play off in a final series to determine the premier team.

The major competition in Australian rules, organised by the Victorian Football League (VFL), began in 1897 with 14 rounds between eight Victorian based clubs. By 1925 the competition consisted of 12 Victorian clubs, and it was still in that form in 1980 when the data for this study begins. With the exception of Geelong (Go Cats) all clubs were based in metropolitan Melbourne. The competition began to go national in 1982 with the movement of the South Melbourne club to Sydney. The administration of the competition was transferred to the Australian Football League (AFL) in 1990, and by 1995 the competition had grown to 16 teams including five interstate clubs. For the 1996 season, the entrance of another interstate club and planned mergers between Victorian clubs transformed the competition yet again. The competition receives a huge following both in terms of spectator and media interest. For example, each Friday, one Melbourne daily paper, The Sun News Pictorial, has a 12 page centre lift out in addition to several back pages devoted to AFL pre match coverage. Interstate matches are covered live on free to air TV, with replays of Melbourne matches. One of the cable pay TV operators has a channel expressly devoted to AFL coverage. Important matches are regularly attended by over 60,000 spectators.

The ultimate aim of a season of football is to win the premiership, or at least finish as high on the ladder as possible. Teams measure their success by ladder position. However final ladder position as a measure of team ability is tempered by the degree to which the draw is unfair. Russel (1980) discusses the problem of carry over effects due to a team continually playing the previous opponent of the same team. The problem arose from one team in the VFL draw having the same carry over effect in 21 of its 22 matches. However the unfairness of the draw goes much deeper. The VFL and AFL competitions have traditionally been unbalanced in quality of opposition. For instance, from 1926 through to 1967 (with the exception of some disruption in the forties due to the war) the draw consisted of 18 rounds between 12 clubs. Thus each team played some teams twice, and others once. From 1970 to 1986 there were 22 rounds and each team played each other twice, but with new teams in 1987 the opponent balance was again lost. Thus it may well be that some teams are playing against much stronger average competition than others. While this is recognised in general by fans and administrators, it is never quantified.

A major determinant of the difficulty of a draw for a particular team is the grounds that matches are played on. It is accepted by sports followers that most teams enjoy a HA. Matches on home grounds are more likely to be won than matches on away grounds. It is important for the draw to be balanced in terms of home and away games. While traditionally each team has played half its matches away and half at home, to maximise crowds it has become common to share grounds and move matches to large capacity grounds. While non-Victorian teams currently play half their matches on their home ground, the Victorian sides do not. We will show there is no semblance of ground balance in the current competition.

Since the draw is not fair on all teams, it is important that we look at measures of team success other than ladder position. As a major determinant of fairness, we will also look in detail at HA. All football followers recognise the importance of a home ground advantage, but never before this work has the actual advantage of individual teams in points been published. In particular, while it is recognised that individual teams have larger HAs than others, these effects are rarely quantified. One reason is probably the difficulty to assess HA because the draw is not balanced for either ability of opponent nor home and away matches.

In this chapter we look first at the HA of the competition as a whole using traditional measures. We then investigate the HA of individual clubs and the joint HA of pairs of clubs, and investigate whether these more complicated models are justified.

Data have been collected on an on going basis for all AFL football matches from 1980 onwards. The data consist of year, round number, home team, away team, ground, home score in goals and behinds and away score in goals and behinds. The data were originally collected on a weekly basis from daily newspapers and football records for the purpose of forecasting match results. A subset of home and away matches from 1980 to 1995 inclusive was used for the analysis in this paper. Rodgers & Browne (1996) was used as the authority for results and grounds.

2.2. HAs of the nominal home team

In the league draw, irrespective of where the match is played, the first named team is nominated as the home team. The usual measure of HA used in the literature is percentage of wins by the home team. Table 2.1 shows the percentage of losses draws and wins by the nominal home team. It has possibly shown some increase over the years. Over the 16 years 2361 matches resulted in 1345 wins for the home side with another 18 matches drawn. Counting a draw as half a win this gives 57.3 % wins for the home side. The table also shows how the number of matches has varied with the addition of new clubs and changes to the number of rounds played each year.

TABLE 2.1. Match results and HA in points ratio for the nominal home team for each year 1980 - 1995

Year	Total games	Win %	Draw %	Loss %	HA in points	Total Points	Ratio of total points to HA
80	132	54.5	2.3	43.2	1.3	210.0	161.2
81	132	54.5	0.0	45.5	8.8	199.8	22.7
82	132	54.5	0.8	44.7	10.6	224.2	21.2
83	132	53.8	0.0	46.2	5.7	212.2	37.3
84	132	55.3	0.0	44.7	5.1	205.8	40.6
85	132	54.5	0.8	44.7	6.0	211.0	35.4
86	132	58.3	0.0	41.7	11.0	203.5	18.5
87	154	60.4	1.3	38.3	12.7	209.3	16.5
88	154	60.4	0.6	39.0	8.8	195.1	22.1
89	154	61.0	0.6	38.3	11.9	189.1	15.9
90	154	59.1	0.0	40.9	8.1	200.1	24.9
91	165	56.4	1.2	42.4	8.4	205.2	24.6
92	165	55.2	1.2	43.6	8.1	207.3	25.7
93	150	56.7	0.7	42.7	8.7	210.2	24.3
94	165	60.6	0.6	38.8	11.0	188.9	17.2
95	176	55.7	1.7	42.6	4.1	188.8	46.0
80-95	2361	57.1	0.8	42.2	8.2	203.2	24.8

Since the percentage of home wins depends on the variation in the performance level of the teams as well as their HA, it is not a good measure to compare HAs between competitions, or even seasons. An alternative measure is the average margin of victory by the home team. To make comparisons across competitions and sports this can be standardised by comparing it to the total number of points scored in a match. Table 2.1 also gives the average winning margin of the nominal home team (HA in points), the average total points scored in a match and the ratio (the average number or points scored for every point attributable to HA). The table shows that HA is quite variable from year to year, but that over 16 years it averaged 8.2 points a game or about one point in every 25.

There are probably two competing effects at work here. One is the introduction of interstate clubs; we will see later they tend to have high HAs. On the other hand, a smaller proportion of games are actually played on a home ground. In Melbourne many grounds are shared, so what is nominally a home match for a particular club may in fact be on a neutral ground.

2.3. Changes in ground usage due to ground rationalisation

One of the reasons given for HA is ground familiarity. This can be obtained by training at a ground or playing at the ground. The current names of the training grounds of the clubs for the period 1980 to 1995 are listed in Table 2.2. We have used the current names where the names of the venues have changed, but the actual venues have not changed. In the same way we use Sydney to refer to South Melbourne, and AFL to refer to VFL.

However, in the AFL competition, teams do not necessarily play home matches on their training grounds. Over the years the AFL has sought to maximise crowd attendance by moving clubs from small capacity grounds to sharing larger grounds. The League also built their own ground, Waverley Park. It became available for regular use from 1970, and the League required all clubs to play some home matches there. The league has also attempted to maximise the use of the MCG. This has meant a steady erosion of the traditional home ground where a team plays and trains.

TABLE 2.2. Training grounds and home grounds of all clubs for period 1980-1995

Team	Training ground	Home ground
Adelaide	Football Park, 91-95	Football Park, 91-95
Carlton	Optus Oval	Optus Oval
Collingwood	Victoria Park	Victoria Park 80-93, MCG 94-95
Essendon	Windy Hill	Windy Hill 80-91, MCG 92-95
Fitzroy	Junction Oval and many others	Junction Oval 80-84, Victoria Park 85-86, Optus Oval 87-93, Whitten Oval, 94-95
Footscray	Whitten Oval	Whitten Oval
Geelong	Kardinia Park	Kardinia Park
Hawthorn	Glenferrie Oval	Optus Oval 80-91, Waverley Park 92-95
Melbourne	MCG	MCG
Nth Melbourne	Arden St.	Arden St 80-84, MCG 85-95
Richmond	Richmond CG	MCG
Sydney	Lakeside Oval 80-81, Sydney CG 82-95	Lakeside Oval 80-81, Sydney CG 82-95
StKilda	Moorabbin Oval 80-92, Waverley Park 93-95	Moorabbin, 80-92 Waverley Park 93-95
Brisbane	Carrara 87-92, Brisbane CG 93-95	Carrara Gold Coast 87-92, Brisbane CG 93-95
West Coast	Subiaco Oval, 87-95	Subiaco Oval, 87-95
Fremantle	Fremantle, 95	Subiaco Oval, 95

The changes over the years are best illustrated by comparing the first and last years of the period under study. For a typical year in the early sixties, prior to the introduction of Waverley Park and the beginnings of ground sharing, each team played all its home matches at its training ground, so each ground was used by only one team for home matches. In 1980, the first year in the period we are studying, the training grounds of Hawthorn (7) and Richmond (10) were not used for matches. These teams used Optus Oval and the MCG respectively as their home ground. However Waverley Park was used by all 12 teams for some home matches. Apart from Waverley, Optus Oval and the MCG were the only shared grounds, with each shared by two clubs. With the exception of the matches at Waverley Park, each team played all its home matches on its home ground.

By 1995 the pattern had become much more confusing. Training grounds no longer used for home matches included those of Essendon, Hawthorn, Nth Melbourne, Richmond and Fitzroy. Five clubs (Melbourne, Richmond, Nth Melbourne, Essendon and Collingwood) played the majority of their home matches at the MCG, but another three clubs played some home matches there. This resulted in other matches being moved away, usually to Optus Oval. Most of the MCG tenants had to play some home matches at Optus Oval, so even traditional owners Melbourne, who train at the MCG, were now one of six clubs to play some home matches at Optus Oval. Fitzroy played home matches at four different grounds, including the first ever match at Canberra. West Coast and Fremantle shared the WACA and Subiaco. Hawthorn and StKilda shared Waverley, but two other clubs also played a home match there. Only Adelaide, Sydney and Brisbane had the traditional pattern of a unique non shared ground for all their home matches.

2.4. Actual home grounds

Clearly, in the AFL, it is arguable which is the home ground of some teams. Teams play their home matches on a variety of grounds, which may or may not be their training grounds. In other cases, because of ground sharing, teams may play away matches on grounds other than their home ground several times. For example in 1995 with five teams sharing the MCG, a team could play five away matches at the MCG. With such a program, a team could become very familiar with grounds other than their home ground.

For this study, the home ground of a team for a particular year was defined as the ground on which the team played the most ***home*** matches. For example, in 1995 Collingwood played three home matches at Victoria Park, their training ground, but eight at the MCG. The MCG was therefore defined as their home ground. In all cases this resulted in the same ground as that officially recognised by the AFL. The resulting home grounds are given in Table 2.2. Most clubs have been very stable, only moving to the MCG or Waverley. Fitzroy is the exception, with four home grounds.

The above definition does mean that some games which may carry a large HA are not classified as played on a home ground. One case would be the Collingwood games played at Victoria Park mentioned above. Another example is West coast. They play about four home matches each year at the WACA, which because of the travel involved for visiting clubs would carry a large HA. Yet these would not be classed as a home ground on the above criteria. These matches, played on neither teams 'home' ground nevertheless may give one side a significant HA. Similarly, many 'home' matches are actually played on neutral grounds. An example would be any match between MCG cotenants. To check on actual HA, a team was categorised as having a perceived advantage if the match was played on their home ground or training ground or other ground with which they could be expected to have a HA. Thus Collingwood have a perceived advantage in any matches played at Victoria Park in seasons when this was not their home ground. Similarly West Coast and Fremantle have a perceived advantage in any matches played at the WACA or Subiaco. Matches were then discarded as neutral if neither or both teams had a perceived advantage. Table *2.3* gives the results and average HA for the remaining matches. These give a better reflection of the actual HA that exists in Australian rules than Table *2.1*.

The percent of games won by the 'home' teams has generally increased and in almost all cases the ratio of points to margin has decreased. Most of this is probably due to the introduction of interstate teams. For example, counting a draw as *0.5* of a win, in the seven years prior to the introduction of two more interstate teams in *1987* the win percentage for home teams averaged *56.6%*. The nine years following averaged *60.5%*. This difference was significant at the 5% level (p=0.017). The drop in points scored at the end of *1993* is attributable to a change in the length of the quarters and time off definitions which reduced the effective playing time. While this would presumably also cause a slight decrease in the average home winning margins, this effect has been ignored in the analysis in this thesis.

To sum up in *16* years of Australian rules football, *1869* matches carried a perceived HA. Of these the advantaged team won *1094* and drew another 15. Counting a draw as half a win this amounts to *59.3%*. In terms of margins the average win for the home team was just under *10* points per game, or just under one point advantage in every *21* points scored.

TABLE 2.3. Match results and HA in points ratio for the team with the perceived HA for each year 1980 - 1995

Year	Total games	Win %	Draw %	Loss %	HA in points	Total points	Ratio of total points to HA
80	107	57.0	2.8	40.2	3.4	210.5	62.2
81	102	54.9	0.0	45.1	10.3	199.4	19.4
82	104	56.7	1.0	42.3	12.9	226.6	17.5
83	106	55.7	0.0	44.3	8.9	214.7	24.1
84	103	59.2	0.0	40.8	10.5	205.3	19.6
85	100	52.0	1.0	47.0	5.6	212.7	38.1
86	100	58.0	0.0	42.0	11.7	205.5	17.6
87	119	62.2	1.7	36.1	13.7	212.0	15.5
88	117	61.5	0.9	37.6	12.0	194.4	16.1
89	121	64.5	0.8	34.7	14.4	187.9	13.0
90	118	60.2	0.0	39.8	9.6	201.6	21.1
91	130	59.2	1.5	39.2	12.0	204.7	17.1
92	134	53.0	1.5	45.5	5.3	208.6	39.4
93	131	60.3	0.8	38.9	10.9	208.5	19.1
94	135	62.2	0.7	37.0	11.4	189.6	16.6
95	142	57.7	0.0	42.3	4.8	187.7	39.1
ALL	1869	58.5	0.8	40.7	9.8	203.7	20.8

2.5. HA throughout the season

We can also look at the percentage of wins and HA in points throughout the season. Table 2.4 reorganises the data in Table 2.3 by combining the 16 years of data in groups of four or five weeks, so the table shows the monthly progression of scoring and the HA. Also included is a column showing the Win/Draw percentage, where draws are counted as 0.5 of a win. Scoring is somewhat lower in the last three months of the season as the weather deteriorates. There is some evidence for a lower HA in the first five rounds of the season, at least in terms of point margin and ratio of margin to points scored, but there is no consistent pattern of an increase or decrease in HA through the remainder of the year. Thus it may be an advantage for teams to play away matches against important opponents at the beginning of the season.

TABLE 2.4. Match results and HA in points ratio for the team with the perceived HA for each stage of season

Round	Total games	Win %	Draw %	Loss %	Win/ Draw %	HA in points	Total points	Ratio of total points to HA
1-5	423	57.2	1.2	41.6	57.8	7.6	211.8	27.9
6-9	331	58.0	0.9	41.1	58.5	9.6	205.3	20.2
10-13	335	61.8	0.6	37.6	62.1	11.2	199.5	17.7
14-17	331	57.1	0.9	42.0	57.6	9.6	199.2	20.7
18plus	449	58.8	0.4	40.8	59.0	10.7	201.5	18.8
Total	1869	58.5	0.8	40.7	58.9	9.8	203.7	20.8

2.6. Paired HA

A team may win more of its home matches simply because it is better than the teams it plays there. When we investigate HA of individual clubs, we need to balance for ability. One way to do this is to consider matches in pairs. We have seen that about 80% of the games involve a perceived home team situation. Most of those involve a return match with the same team, and about 80% of those also involve a perceived HA. The two results together give us a measure of the joint HA between the two teams. This can be estimated without the need for estimating performance levels by adding the two results. For Model 1.4 we had

$$w_{ij} = u_i + h_{ij} - u_j + e_{ij}$$

For the return match

$$w_{ji} = u_j + h_{ji} - u_i + e_{ji}$$

and so adding we obtain

$$w_{ij} + w_{ji} = h_{ij} + h_{ji} + error$$

where $h_{ij} + h_{ji}$ can be thought of as a paired HA. (Note that if we revert to model 2, where $h_{ij} = hi$, the paired HA becomes $h_i + hj$, the sum of the individual HAs of the two teams). By assuming this remains constant over several years, we can average over several years to reduce the error. Table 2.5 gives the number of pairs for each pair of teams (top half) and the average paired HA (bottom half). Note that in a perfectly balanced competition such as English soccer, each pair would occur 16 times. Here some clubs have many more pairs than others. For example, Sydney (team 11), which has never shared a ground, and as an interstate club was not required to share Waverley Park, generally has more pairs than others. Some values need to be treated with caution, as they are based on only a few pairs. For example, the largest paired advantage in the table is Geelong/Adelaide at 147 points. However this is based on only two pairs. Nevertheless it is interesting to note that in 1996, Geelong lost by 64 points in Adelaide but won by 35 points at Geelong - a turn around of 99 points. We should treat Fremantle (team 15) figures with extreme caution as they are only based on a single pair.

Some inferences can be drawn from the table. Most of the interstate clubs have large paired HAs with other teams, and virtually all the values over 40 points involve Adelaide, Brisbane or West coast. There are a few negatives in the table, and nearly all involve Victorian clubs.

TABLE 2.5. Mean paired HA for each team and number of pairs for the years 1980 to 1995

Team	Adel	Car	Coll	Ess	Fitz	Foot	Geel	Haw	Melb	NthM	Rich	Syd	StK	Bris	WC	Frem
Adel		2	2	1	2	2	2	2	2	2	3	3	3	3	4	
Car	48.0		5	7	2	8	8	3	5	8	8	11	7	5	6	
Coll	24.5	-31.2		6	6	8	9	7	8	4	5	11	7	3	4	
Ess	72.0	32.3	24.2		9	9	7	6	8	7	8	10	7	2	6	1
Fitz	62.0	7.5	27.8	16.7		8	9	4	6	6	9	13	10	5	3	1
Foot	61.0	22.1	39.3	28.2	25.8		8	9	7	6	6	11	10	6	6	1
Geel	147.0	21.4	-1.4	3.7	31.2	39.9		8	9	8	7	13	9	5	5	1
Haw	-34.0	-23.3	21.9	0.8	14.8	22.7	2.6		7	10	9	9	7	5	4	
Melb	22.5	-4.6	-29.6	11.0	4.7	27.1	32.4	14.6		3		12	5	6	5	
NthM	44.5	23.6	17.5	11.4	-18.8	-14.5	-14.1	30.4	27.7		5	11	10	4	3	
Rich	-18.0	43.4	42.4	-10.5	20.6	33.5	-7.6	-33.6		1.4		12	6	6	5	1
Syd	-16.0	19.3	9.5	39.1	12.6	25.6	-5.5	19.8	24.3	-9.5	9.8		12	6	4	1
StK	68.7	-4.9	36.3	18.3	13.2	4.4	51.6	-2.4	9.4	8.9	31.2	23.8		5	4	
Bris	5.3	42.6	-9.3	18.5	50.4	-7.5	28.4	48.8	43.8	26.3	60.7	33.0	32.0			
WC	67.5	61.5	57.8	56.5	47.7	27.7	15.8	-2.0	40.4	30.3	11.0	74.5	66.8	53.2		
Frem				39.0	9.0	33.0	42.0				-2.0	41.0				

2.7. Individual HAs

There are several ways these paired HAs can be manipulated to produce measures of individual team HA. A simple averaging is illuminating although as we shall see needs to be treated with some caution. Table 2.6 gives the averages of the above figures, for each team. We have also given the average margin of each team in their home matches and the average margin in their away matches. These are simple measures the average supporter would understand, particularly soccer followers as they are analogous to home and away goal difference. Because we have included only those matches which can be paired, the average paired HA is the difference in the two columns. Note the high average paired HA of all the interstate teams except for Sydney.

TABLE 2.6. Average of the paired HA for each team from 1980 to 1995

Team	Number of pairs	Average home margin	Average away margin	Average Paired HA
Adelaide	33	15.0	-21.7	36.8
Carlton	85	28.8	8.5	20.4
Collingwood	85	13.8	-1.1	14.9
Essendon	94	26.8	6.6	20.2
Fitzroy	93	2.0	-17.4	19.4
Footscray	105	0.9	-21.8	22.7
Geelong	108	19.9	2.4	17.4
Hawthorn	90	28.4	19.4	8.9
Melbourne	83	-1.8	-18.4	16.5
Nth Melbourne	87	6.8	-1.1	7.9
Richmond	90	-4.2	-17.3	13.1
Sydney	139	0.8	-16.0	16.8
StKilda	102	-7.5	-29.2	21.7
Brisbane	66	-6.7	-39.9	33.1
West Coast	64	34.6	-8.7	43.3
Fremantle	6	24.0	-3.0	27.0
All	1330	9.7	-9.7	19.5

The above paired HAs include a component due to the other teams' HAs. This increases the apparent HA of each team by including a component due to the HA of all the other teams. This spurious HA can be explained by the following simplistic argument. Each pair of matches gives us an estimate of $h_i + hj$ for the two teams i, j. By averaging these pairs for team i, we obtain an estimate of $h_i + h$, where h is the average of the hj. We need to remove h to get back to our estimate of hi. Stefani & Clarke (1991) give one method using an iterative procedure. An alternative is to use regression. The individual h_i can be estimated using general linear methods to find a least squares fit. While this could be done with the 220 averages of Table 2.5, it is better to use the original 664 individual pairs. While the actual estimates do not differ much, the second method gives greater weight to those averages based on more matches, and tends to produce more significant results. The REG procedure from SAS/STAT was used to obtain the estimates given in Table 2.7. Perhaps not surprisingly given the comments above, the figures are quite similar to those obtained by subtracting from the average paired HAs given in Table 2.6 their overall average of 9.7.

The figures are consistent with those given in Stefani & Clarke (1992) which covered the years 1980-89, and their comments on the relative mix of travel, especially across time zones, crowd intimidation, and lack of familiarity with the playing conditions in regard to international comparisons apply here. It is logical that West Coast should benefit from the distances travelled by other teams, across two time zones. In addition, Western Australia and South Australia are traditional Australian rules states, and matches played in those states are in front of one sided capacity crowds. Note that all interstate clubs with the exception of Sydney have a large HA. The biggest four HAs all belong to interstate clubs. Fremantle is not significant due to there being only a few pairs, but the value is still large. Travel between Sydney and Melbourne involves no time zone changes. As Sydney is the one interstate team that was actually formed by relocating a Melbourne club to a traditionally non Australian rules city, the crowd support at matches in which Sydney plays is much less one sided than for other interstate teams. Some other teams have above average HAs, specifically Footscray, StKilda, Essendon and Carlton. While this is probably in line with most supporters' perceptions, other parts of the table are not. Geelong is in the lower half on the table, while Collingwood is thirteenth. The supposed intimidation effect of the Collingwood supporters at Victoria Park does not appear to show up in the figures.

The negative HA of North Melbourne is also surprising. Although not significantly different from zero, it is nevertheless interesting that over 16 years a team should average two points worse at home than away. Previously the only negative HAs reported were by Baumeister & Steinhilber (1984) in the context of deciding matches in finals series of baseball and basketball, but Gayton et al. (1987) failed to replicate in ice hockey and Benjafield & Liddel (1989) limited it to teams with an expectation of winning. The table lends weight to ground familiarity as a cause for HA. The bottom three teams have for most of the period all played on shared grounds on which they do not train. Thus they do not gain familiarity with the playing ground during training, and other teams gain more familiarity than usual by playing more away matches at the ground. Significantly the bottom five teams on the table all currently play on a shared ground.

TABLE 2.7. Individual HAs of all teams in points per game based on paired matches for the years 1980-1995, in order of decreasing HA

Team	HA Estimate	Standard Error	p value for HA = 0
West Coast	33.5	7.0	0.00
Adelaide	25.2	9.7	0.01
Brisbane	23.4	6.9	0.00
Fremantle	18.3	22.6	0.42
Footscray	12.9	5.5	0.02
StKilda	12.6	5.6	0.02
Essendon	10.9	5.8	0.06
Carlton	10.2	6.1	0.09
Fitzroy	9.9	5.9	0.09
Geelong	8.0	5.4	0.14
Sydney	7.9	4.8	0.10
Melbourne	5.6	6.2	0.37
Collingwood	5.2	6.1	0.40
Richmond	2.3	6.0	0.70
Hawthorn	-0.5	6.0	0.93
Nth Melbourne	-2.0	6.1	0.74

2.8. Using linear regression analysis on individual match results

The above method is quite wasteful of data. To allow for team ability we paired the data, so only those matches that can be paired are used, and this results in discarding 44% of the matches. This limits its use on a yearly basis, as many of the teams would only have a few pairs. An alternative is to fit a model with team and HA effects to the original match results for each year separately. We could use any of the Models 1, 2 or 3, or variations could be used. For example, you could allow for different hs for different grounds. Thus Collingwood could be allowed a different h for Victoria Park as against the MCG, or West coast a different value for the WACA and Subiaco. This could be used to possibly differentiate travel effects from ground effects. In all these cases, no data is wasted as even a match on a neutral ground contributes toward determining the us.

Figure 2.1 gives a sample regression output for 1995. In this case we have used Model 3 which allows for individual HAs, and used the actual home ground for that year as defined in section 2.4. Note the root mean square error is 39 points so football is very unpredictable. None of the HAs are significantly different from zero. This is due to the large error and the small number of matches played on each home ground. Again we need to look across seasons to obtain enough data to make reasonable observations about differences in HA. Nevertheless the us and hs are generally in line with that expected. The top teams have large us, the interstate teams generally large HAs. The correlation between the us and premiership points is 0.83, and between the us and percentage is 0.93. The better correlation with percentage is not surprising, as premiership points can be very dependent on a few close wins.

The SAS System

------------------------------ YEAR=95 ------------------------------

Model: MODEL1
NOTE: Restrictions have been applied to parameter estimates.
NOTE: No intercept in model. R-square is redefined.
Dependent Variable: MARGIN

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	31	150750.24391	4862.91109	3.092	0.0001
Error	145	228022.75609	1572.57073		
U Total	176	378773.00000			

Root MSE	39.65565	R-square	0.3980
Dep Mean	4.10795	Adj R-sq	0.2693
C.V.	965.33814		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for HO: Parameter=0	Prob > |T|
UO	1	-31.444878	12.06997367	-2.605	0.0101
U1	1	26.205001	11.28522369	2.322	0.0216
U2	1	-1.646862	12.97593121	-0.127	0.8992
U3	1	26.925728	13.79836118	1.951	0.0529
U4	1	-40.973561	10.41715016	-3.933	0.0001
U5	1	-2.340900	10.89383102	-0.215	0.8302
U6	1	26.167461	11.39850125	2.296	0.0231
U7	1	-8.320901	12.54937285	-0.663	0.5083
U8	1	3.829991	11.79737605	0.325	0.7459
U9	1	22.525549	11.91225347	1.891	0.0606
U10	1	0.912430	13.03292146	0.070	0.9443
U11	1	-4.379118	11.87188608	-0.369	0.7128
U12	1	-15.250602	12.33104761	-1.237	0.2182
U13	1	-11.118884	11.85595351	-0.938	0.3499
U14	1	6.591798	10.86629163	0.607	0.5450
U15	1	2.317748	10.83552132	0.214	0.8309
H0	1	28.698837	17.70174191	1.621	0.1071
H1	1	4.905900	17.93123963	0.274	0.7848
H2	1	5.599304	18.40599612	0.304	0.7614
H3	1	-7.534129	18.79107521	-0.401	0.6891
H4	1	-13.420776	18.48592833	-0.726	0.4690
H5	1	-5.950555	17.99247664	-0.331	0.7413
H6	1	-4.916018	17.60617708	-0.279	0.7805
H7	1	8.716675	17.88099084	0.487	0.6267
H8	1	-6.615524	18.11296561	-0.365	0.7155
H9	1	-22.089570	18.35510965	-1.203	0.2308
H10	1	6.985185	18.37912233	0.380	0.7045
H11	1	7.992999	17.49195170	0.457	0.6484
H12	1	-6.804360	17.59674637	-0.387	0.6996
H13	1	19.880955	17.49099143	1.137	0.2576
H14	1	26.939111	18.28337767	1.473	0.1428
H15	1	-14.528699	18.15355155	-0.800	0.4248
RESTRICT	-1	-2.55617E-14	.	.	.

Figure 2.1. **SAS** output obtained from regression procedure fitting a team rating and individual **HA** to 1995 margin results

Since teams have traditionally played half their matches at home we might use $u_i + 0.5\,h_i$ as a measure of a team's success through the year. This is in line with Harville & Smith (1994) who suggests an equivalent measure for a team's overall performance level in relation to the average performance level. This measure has a correlation of 0.90 with premiership points and 0.98 with percentage. Figure 2 shows a scatter plot of percentage against $u_i + 0.5\,h_i$ and demonstrates the extremely close fit. Thus the u_i and h_i together give a good measure of a team's overall success through the year, but separately give a measure of how much contribution the effects of team ability and HA made. It also suggests that percentage is a better measure of a team's average performance level than premiership points.

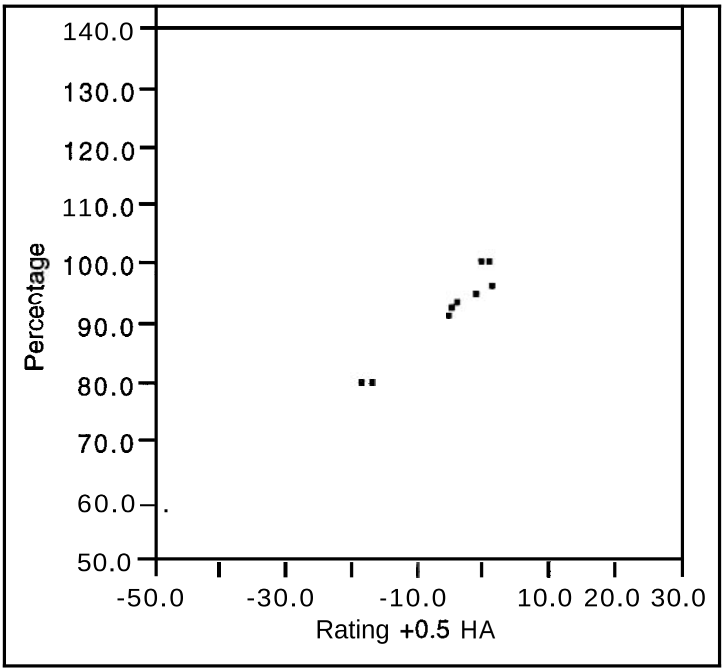


Figure 2.2. Percentage versus $u_i + 0.5\,h_i$ for 1995

The HAs for each club for each year obtained by this method are given in Table 2.8. Again the averages are roughly in line with those previously obtained, and most of the comments made in regard to the average paired HAs apply here. Three of the interstate teams have the largest average HAs. Sydney is the exception. Note in the early years after their move they had a high HA, but this has diminished, possibly as teams got used to travel. Note the very low HAs of teams that have shared the MCG for many years. Melbourne, Richmond and North Melbourne are all in the lower third of the table.

TABLE 2.8. Home advantages for all clubs in the AFL 1980-1995

	Year																
Team	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	80-95
Adel												46	34	44	29	29	36
Bris								-11	33	39	43	-5	45	48	30	20	27
WC								42	21	38	-2	17	17	12	-15	27	17
Foot	15	17	12	34	28	25	26	10	12	19	-50	29	7	20	33	-6	14
Car	13	5	39	33	-11	-3	12	17	32	-6	-6	20	16	-2	48	5	13
StK	-2	7	15	-16	8	6	20	38	17	33	25	13	11	8	-10	-7	10
Coll	-11	29	11	32	11	-33	6	-9	21	17	25	32	-2	8	13	6	10
Syd	36	5	17	9	5	15	6	28	-18	10	2	-4	4	34	-18	8	9
Ess	5	11	10	15	7	32	-2	31	8	17	-13	23	-19	-14	29	-8	8
Fitz	9	12	-3	35	6	-10	28	28	29	-8	15	9	-5	2	-10	-13	8
Geel	-3	28	27	-40	20	-5	22	17	-28	25	15	3	14	24	7	-5	8
Melb	-10	15	-9	12	-4	18	20	3	5	-15	21	-18	7	37	1	-7	5
Haw	25	-12	19	-28	9	30	-9	8	20	24	-24	10	-12	-12	1	9	4
Rich	-5	-4	-6	-10	14	-9	8	-0	14	12	17	-17	-13	-12	28	7	1
NthM	-17	2	21	28	23	-23	2	-10	12	-3	53	-3	-31	-35	18	-22	1
Frem																-15	-15
All	4	10	13	9	10	4	12	14	13	14	9	10	5	11	12	2	9.3

2.9. Further analysis of individual HAs

Do different teams have different HAs or are the above differences due to random variation? Table 2.8 provides data that can be analysed by normal statistical methods. The year to year variation in HA for individual teams is very large, and little faith could be placed on individual values. Nevertheless, some may be of interest to administrators and supporters. For example, why did Footscray, after a decade of consistently high positive HAs, have a huge -50 in 1990? StKilda, after its move to Waverley in 1993, had three consecutive HAs all less than they had enjoyed in any of the preceding nine years. Some aspects have been investigated in more detail below.

2.9.1. Team and year effects

The data were analysed in various ways using the general linear models. With HA as the dependant variable a model for all the data with a year and team effect was highly significant (p=0.036). The year effect was insignificant (p=0.83) but the team effect was highly significant (p=0.002). (Removing the single Fremantle observation resulted in virtually no change.) This is clear evidence that in Australian rules the HAs of all teams are not the same.

2.9.2. Interstate teams

The cause of this difference can be traced to the interstate clubs. If the same model is fitted to the years 1980 to 1986 (before the introduction of Brisbane and West coast) the model as a whole is not significant, and the team effect has a p value of 0.57. However when fitted to the years 1987 to 1995, the overall model was significant at p = 0.04 and the team effects at p=0.01. When the data were split into two groups, Victorian and interstate, the 178 Victorian values had a mean HA of 7.7 with standard deviation of 17.6. The 38 interstate values had a mean of 17.5 with a standard deviation of 19.8. The differences in standard deviation were statistically insignificant, allowing an equal variances t test which showed the differences in the mean for the two groups was significant with p=0.003. Note that the differences are not just statistically significant, but of practical significance. The average HA of the interstate teams is more than double the Victorian clubs.

2.9.3. MCG teams

One reason advanced for HA is ground familiarity. Since the MCG is played on by many clubs, other teams will also be familiar with the ground, and the home team's advantage will be reduced. To investigate if teams sharing the MCG had a different HA than other teams the interstate teams were removed. Since they have a higher HA than average, and none use the MCG, their inclusion would invalidate the results. This analysis was therefore restricted to the Victorian clubs only. The 49 HA values for clubs in the seasons they used the MCG as a home had a mean of 1.4, the other 129 values for Victorian clubs had mean of 10.0. The difference was significant at p=0.002.

2.9.4. New Ground effect

When clubs change ground, they would be less familiar with their new ground and hence should have a lower HA. Table 2.1 shows that Victorian clubs changed grounds to another Victorian venue eight times. The eight HAs for the first season at the new grounds had a mean -2.9. The other 170 seasons by Victorian clubs had a mean of 8.2. The difference was marginally significant (p=0.08). This effect could be confounded with the effect shown in the previous section, as three of the moves were to the MCG.

2.10. Significance of various models

While the above analysis suggests that teams have different HAs, they may occur in groups. For example, we may only need one HA for interstate teams, one for MCG teams and one for other teams. Harville & Smith (1994) test the significance of Models 1 through **3** by successively fitting the models and testing if the incremental improvement is significant. They find that while including a common HA in Model 2 is highly significant, the gains made by the extra complication of including individual HAs gained some improvement but this was not significant. One of the improvements they suggested was to group teams. We adopt this approach here, and also test two incremental models between models 2 and **3.** In the first of these extra models (model 2a) we allow a different but common HA for all interstate teams (hi) to the Victorian teams (hv), while in the second (model 2b) we allow a common HA for the MCG teams (hm). Because we are interested primarily in the interstate effect, in this section we only consider the years 91 to 95, and all games at Subiaco and the WACA are classed as home games. The process is explained in detail for 1995, and then summary results are given for all years 1991-1995. In keeping with Harville, the notation $S_{i|j}$ represents the difference between the residual sum of squares obtained by fitting model **i** and **j** respectively - i.e. the extra sum of squares explained by fitting Model **j** over that obtained by fitting Model **i.** Table 2.9 shows the marginal sums of squares explained by progressively fitting the models. An F ratio can be formed to test if the model is a significant improvement over the previous model using the final residual mean square (these are the values given in the table). More correctly, the marginal sums of squares can be totalled to test any required hypothesis. Suppose we wish to test whether Model 2 is a significant improvement over Model 1. The improvement in the sum of squares gained by fitting the extra parameter in model 2 is 2634.1. The residual sum of squares is 10547.8+11.2+10794.9+223567.5 = 244921.4 with 1+1+1+13+145 = 150 degrees of freedom. This gives an F statistic of 1.6, in this case not significant. To test if Model 3 is an improvement over Model 2, the extra 15 parameters of Model 3 contribute 10547.8+11.2+10794.9=21353.9 for a mean square of 1423.6. Compared to the error mean square of 1541.8 this gives an F value of 0.92, clearly insignificant.

The various hypotheses are relatively simple to implement with PROC REG in SAS. We simply start of with the most complicated Model 3 and progressively restrict groups of the hs to be equal, and request the relevant tests of hypothesis. For example to test Model 3 against Model 2 we request the hypothesis test $h_0=h_1=h_2=h_3=h_4=h_5=h_6=h_7=h_8=h_9=h_{10}=h_{11}=h_{12}=h_{13}=h_{14}=h_{15}$. To test Model 2 against Model 1 we put a restriction $h_0=h_1=h_2=h_3=h_4=h_5=h_6=h_7=h_8=h_9=h_{10}=h_{11}=h_{12}=h_{13}=h_{14}=h_{15}$ and test the hypothesis $h_0=0$.

On the basis of 1995 only, there is strong evidence for Model 1 and 2a, somewhat marginal evidence for Model 2 and no evidence for the other models. While it is surprising that Model 2 is not significant, note from Table 2.3 that 1995 had one of the lowest HAs in terms of points for all years. Other years may yield a different result. The results of the analysis for each model for the years 91 to 95 is now discussed in detail.

TABLE 2.9. Marginal significance of various models for the year 1995

Source	Degrees of freedom	Marginal Sum of squares	Mean square	F	p
Model 1	15	$SS_1 = 131217.4$	8747.8	5.7	.00
Model 2| Model 1	1	$SS_{2\|1} = 2634.1$	2634.1	1.7	.19
Model 2a| Model 2	1	$SS_{2a\|2} = 10547.8$	10547.8	6.8	.01
Model 2b| Model 2a	1	$SS_{2b\|2a} = 11.2$	11.2	.01	.92
Model 3| Model 2b	13	$SS_{3\|2b} = 10794.9$	830.4	.5	.92
Residual	145	223567.5	1541.8		
Total	176	378773.0			

Model 1: Model 1 alone is highly significant, with a p value < 0.0001 each year. There is no doubt there are differences in the mean level of team performances.

Model 2: Model 2 proved to be significant in most years. Table 2.10 gives the significance of the improvement for Model 2 over Model 1 (ie for the hypothesis test h=0), the R-square value (i.e. the percentage of variation in the margins explained by the model) and the estimated value of the common HA for each year and its standard error.

TABLE 2.10. Model 2 results for the years 1991-95

Year	p for H_0: h=0	R^2	h	se(h)
1991	.002	.44	11.2	3.6
1992	.074	.42	6.5	3.6
1993	.001	.41	11.8	3.4
1994	.004	.37	12.2	3.6
1995	.191	.35	4.4	3.4

Clearly Model 2 is significant and its place as the standard model is justified by these results. Note the estimated values for the common HA are all within one standard error of those given in Table 2.3, which suggests that the simple methods used there do give a reasonable estimate of a common HA. The model generally explains about 40% of the variation in results, which again illustrates the large variation present in Australian rules.

Model 2a: The inclusion of a different HA for interstate teams is generally significant. Table 2.11 gives the significance of the improvement for Model 2a over Model 2 (ie for the hypothesis test hi = hv), the R-square value and the estimated value of the common HAs for interstate and Victorian teams. With three of the five years significant there is strong evidence for the interstate teams having a different HA to Victorian clubs. However there are still large errors in the estimates and averages over several years need to be taken to obtain accurate estimates. 1994 is atypical with the estimate for Victorian clubs slightly higher than interstate clubs. Nevertheless the table as a whole is conclusive evidence that some clubs do have higher HAs than others, and that a more complicated model than a common HA is justified.

TABLE 2.11. Model 2a results for the years 1991-95

Year	p for H_0: $hi=hv$	R^2	hi	se(hi)	hv	se(hv)
1991	0.296	0.45	19.4	8.6	7.2	5.1
1992	0.006	0.45	27.4	8.2	-3.6	5.0
1993	0.016	0.44	31.1	8.6	3.6	4.8
1994	0.759	0.37	10.0	8.0	13.2	4.9
1995	0.008	0.38	21.2	7.1	-5.3	4.9

Model 2b: The evidence for a different HA for the MCG and other Victorian clubs is somewhat inconclusive. The p values for 1991 through to 1995 for the improvement in Model 2b over Model 2a were 0.05, 0.11, 0.50, 0.77, and 0.93, which suggests that while a difference may have existed it has disappeared in recent years.

Model 3: There is weak evidence for an improvement in Model 3 over Model 2. The respective p values are 0.51, 0.29, 0.11, 0.25 and 0.54. While none individually are significant, taken as a group they provide some evidence for improvement. However all of this improvement can be put down to the gains made by the simpler Model 2a. (The p values for the improvement of Model 3 over 2a are 0.52, 0.77, 0.30, 0.20, and 0.93, which show no tendency towards significance). However not a great deal is lost by using Model **3** in place of Model 2. The adjusted R^2 value, which adjusts R^2 making allowances for the number of parameters in the model, is generally slightly higher for Model **3.**

Summary: The above results suggest that based on the results in 1991 to 1995, while a common HA or separate HAs for interstate and Victorian teams is justified, further increasing the number of distinct HAs is not. Other groupings may be possible. For example it may be advantageous to remove Sydney from the group of interstate teams, as they are much less isolated than the others in the group. There may also be justification for singling out Victorian sides such as Footscray, whose calculated HAs appear to confirm their long held reputation for a large HA. Of course, one way of deciding possible candidates for groupings is to fit a model with unique HAs for each team and further investigate their similarities and differences. For this reason alone it is worth persevering with Model 3.

2.11. Conclusion

The AFL competition is not balanced with respect to quality of opposition nor HA. The added complication of ground sharing makes it difficult to calculate team ratings and HA. All measures of HA vary greatly from year to year. Over the period 1980 to 1995 the team with a perceived HA won approximately 59% of the matches. This is made up of two distinct periods. Prior to 1987 the home win percentage was 56%, but this increased to 60% after the introduction of new interstate teams in 1987. However a better measure is the average winning margin of the home team, which was just short of 10 points. Although this is not adjusted for ability of opposition, it gives similar yearly values to fitting a regression model with a common HA. Quality of opposition can also be allowed for by looking at the paired HA, but several years data is necessary to obtain reasonable averages. These clearly show an isolation factor, with virtually all paired HAs over 40 points involving Adelaide, Brisbane or West Coast. There are various methods for extracting individual HAs from the paired values, but care must be taken to allow for a spurious HA caused by the HA of the opposition. However when this is done an ordering of the clubs by HA clearly shows an isolation effect. West Coast, Adelaide, Brisbane and Fremantle head the table with HAs over 20 points. Melbourne, Collingwood, Richmond, Hawthorn and Nth Melbourne, all inner city Melbourne clubs, bring up the rear. Over the 16 years, Nth Melbourne had a negative HA (-2 points). Regression analysis was used to calculate individual HAs for each year. Detailed analysis of these showed that the team effect was highly significant. There is strong evidence that this is due to the interstate teams having a different HA to the others. There was also evidence for MCG teams and teams playing for the first season on a new ground having a lower than average HA. By investigating models of varying complexity, it is clearly shown that the use of models more detailed than those incorporating only a common HA is justified. While unique HAs for all clubs may not be necessary, in the AFL competition they are at least as accurate as using a common HA. The optimum appears to be somewhere in between, with perhaps a different HA for interstate teams from the others.

The analysis has clearly shown that different clubs have different HAs. This allows an investigation of the effects of differences in travel, crowd and familiarity factors. This could previously only be done by comparing competitions between different sports or grades.

CHAPTER III

HOME GROUND ADVANTAGE OF INDIVIDUAL CLUBS IN ENGLISH SOCCER

3.0. Abstract

Least squares is used to fit a model to the individual match results in English football and produce a home ground advantage effect for each team in addition to a team rating. We show that for a balanced competition this is equivalent to a simple calculator method using only data from the final ladder. The existence of a spurious home advantage is discussed. Home advantages for all teams in the English Football league from 1981-82 to 1990-91 are calculated, and some reasons for their differences investigated. A paired home advantage is defined and shown to be linearly related to the distance between club grounds.

Key **words:** football statistics, home ground advantage, soccer, performance measures, least squares.

3.1. Introduction

The existence of a home advantage (HA) in most sports is now well documented. Courneya & Carron (1992) give a summary of the work done on HAs. They made the point that future research should be directed to the causes of HA rather than document its existence. However this requires the calculation of the HAs of individual clubs, so differences can be related to the playing characteristics of the clubs. Pollard (1986b) quantified HA (in a competition where each team plays an equal number of matches at home and away) as the number of games won by teams playing at home expressed as a percentage of all games played, with 50% indicating no HA. Although this method is acceptable when averages over a whole competition are taken, it is obviously inadequate when the performance of individual clubs is studied. Here a team may win more (or less) than 50% at home because it is a relatively strong (or weak) team. Snyder & Purdy (1985) show the limitations of this approach, when in looking at a universities basketball competition they found that division 2 teams won only 40% of their home matches against division 1 teams. This implies that the quality of opposition effect overshadowed the HA effect. Because the quality of teams differ, we must allow for differences in ability and measure HA by comparing a team's home and away performance (See also Harville & Smith (1994).)

TABLE 3.1. End of season ladder for Division 1, 1986-87

TEAM	HW	HD	HL	Hf	Ha	HGD	AW	AD	AL	Af	Aa	AGD	GD	Pnts	h	u
Everton	16	4	1	49	11	38	10	4	7	27	20	7	45	86	0.84	0.92
Liverpool	15	3	3	43	16	27	8	5	8	29	26	3	30	77	0.49	0.76
Tottenham Hotsp.	14	3	4	40	14	26	7	5	9	28	29	-1	25	71	0.64	0.57
Arsenal	12	5	4	31	12	19	8	5	8	27	23	4	23	70	0.04	0.82
Norwich City	9	10	2	27	20	7	8	7	6	26	31	-5	2	68	-0.11	0.42
Wimbledon	11	5	5	32	22	10	8	4	9	25	28	-3	7	66	-0.06	0.51
Luton Town	14	5	2	29	13	16	4	7	10	18	32	-14	2	66	0.79	-0.03
Nottingham Forest	12	8	1	36	14	22	6	3	12	28	37	-9	13	65	0.84	0.20
Watford	12	5	4	38	20	18	6	4	11	29	34	-5	13	63	0.44	0.40
Coventry City	14	4	3	35	17	18	3	8	10	15	28	-13	5	63	0.84	0.01
Manchester United	13	3	5	38	18	20	1	11	9	14	27	-13	7	56	0.94	0.01
Southampton	11	5	5	44	24	20	3	5	13	25	44	-19	1	52	1.24	-0.28
Sheffield Wed	9	7	5	39	24	15	4	6	11	19	35	-16	-1	52	0.84	-0.12
Chelsea	8	6	7	30	30	0	5	7	9	23	34	-11	-11	52	-0.16	0.15
West Ham United	10	4	7	33	28	5	4	6	11	19	39	-20	-15	52	0.54	-0.29
Queens P. Rangers	9	7	5	31	27	4	4	4	13	17	37	-20	-16	50	0.49	-0.29
Newcastle United	10	4	7	33	29	4	2	7	12	14	36	-22	-18	47	0.59	-0.38
Oxford United	8	8	5	30	25	5	3	5	13	14	44	-30	-25	46	1.04	-0.77
Charlton Athletic	7	7	7	26	22	4	4	4	13	19	33	-14	-10	44	0.19	-0.00
Leicester City	9	7	5	39	24	15	2	2	17	15	52	-37	-22	42	1.89	-1.12
Manchester City	8	6	7	28	24	4	0	9	12	8	33	-25	-21	39	0.74	-0.53
Aston Villa	7	7	7	25	25	0	1	5	15	20	54	-34	-34	36	0.99	-0.95
Total						297						-297			14.14	0.00

H, home; **A,** away; W, win; D, Draw; L, Loss; f, goals for; a, goals against; GD, goal difference; h, Home advantage; u, team rating.

Table 3.1 shows a typical ladder (division 1, 1986) as published for English soccer, with the addition of two extra columns to be explained later. Sports followers have long recognised the importance of a HA, and soccer tables have traditionally separated a team's home and away performance. At the bottom of the table, Aston Villa, under the definition of percentage of games or points scored at home, could be construed as having no HA. Counting a draw as 0.5 of a win, they have won exactly 10.5 or 50% of games at home and scored exactly 50% of the goals in their home matches. However, this is largely due to their low team ability relative to their opponents. The away matches can be used to allow for team ability, and we see that Aston Villa have won only 3.5 of their away matches and scored 34 fewer goals than their opponents. A supporter might say that over the season they enjoyed a 10.5–3.5=7 game and a 0–(–34)=34 goal HA. Virtually all teams, irrespective of ability, have a HA when measured in this way. However, we shall show that this includes the HA of all the other teams, and so overestimates the true individual HA effect. It is not generally appreciated that each team with a 'real' home ground advantage automatically gives each other team in the competition a 'spurious' or apparent HA. This is best demonstrated by a simple example which is given in Appendix 3.1.

3.2. Modelling team ability and home advantage

To estimate HA correctly we need to model for the ability of team. We use a model similar to that used by Stefani (1983, 1987), Stefani & Clarke (1992), and Clarke (1993b) that has proved successful in predicting match results. The winning margin w_{ij} in a match between team i and team j played at the home ground of team i is modelled as

$$w_{ij} = u_i - u_j + h_i + e_{ij}\,, \tag{3.1}$$

where u_i is a measure of team i's ability, h_i is a measure of team i's HA and e_{ij} is a zero mean random error. We assume that the u_i and h_i are constant throughout the season.

The w_{ij} can be measured either in 'win margin' (1, 0 or -1 depending on whether the home team won, drew or lost) or goals margin. The second case is usually followed in prediction models, and that is the method we prefer here as it is more sensitive to HA. For example a team that wins 3-0 at home and wins 2-1 away shows no HA if win margins are used, but shows a 2 goal advantage if goal margins are used.

This model, with the additional constraint that the u_i (being relative) sum to zero, can be fitted to the individual match results with a standard regression package by using dummy variables for the us (1 if a team is home, -1 if the team is away and 0 for the other teams) and hs (1 for the home team and 0 for the others). In this case the REG procedure from SAS 6.08 gave the values for u and h shown in the last two columns of Table 3.1. The overall model was significant at the 0.01% level, with $R^2 = 0.19$. The low value of R^2 reflects the high variability in soccer, and for the other seasons analysed increased with the unevenness of the season. Each u had a standard error of 0.33 and the hs a standard error of 0.49. A Q-Q plot of the residuals indicated they were normally distributed, which was confirmed with a Shapiro-Wilk test of normality statistic of 0.99 (p=0.63). A plot of the residuals against predicted values showed no evidence of heteroskedasticity.

Alternatively, the Lagrange multiplier technique can be used to derive the values of u_i and h_i that minimise the sums of the squares of the errors. This is shown in Appendix 3.2. Surprisingly, the equations derived only use simple arithmetic on information contained in the end of year ladder. Thus, provided that the draw is balanced, instead of using complicated regression procedures on the individual match results, ability and HA effects are easily found by using only a calculator and data obtained from the final ladder.

The procedure is as follows. Given a season's results in an N team competition, where each team plays the other $N-1$ teams once at home and once away, we can obtain measures u_i and h_i that describe each team's level of performance on a neutral ground and their home ground advantage.

(a) $H = \Sigma h_i = \Sigma \text{HGD}_i / (N-1)$ is the total of all the individual teams' HAs, i.e. H is the total of the team's HGD column, divided by $N-1$. In the ladder of Table 3.1, H = 297121 = 14.14.

(b) For each team, the HA $h_i = (\text{HGD}_i - \text{AGD}_i - H) / (N-2)$, i.e. for each team, their HA is the difference in their home and away goal difference, less the total of all the teams' HAs, all divided by (N–2). For example for Everton h_1= (38–7–14.14)/20 = 16.86120 = 0.84, and for Aston Villa h_{22} = {0 – (-34) –14.14}/20 = 0.99.

(c) For each team, the ability measure $u_i = \{\text{HGD}_i - (N-1)\, h_i\}/N$. In Everton's case u_1 = (38–21x0.84)/22 = 0.92, and for Aston Villa u_{22} = (0–21x0.99)/22 = -0.95.

These equations could be explained quite simply to a layman by replacing each match result in the usual home and away grid with the expected or model result and using simple addition. This derivation is given in Appendix 3.3.

The source of the spurious HA is now clearly shown in (b) above. The difference in a team's home and away performance is given by $(N-2)h_i + \sum h_i$. The difference is made up of one component due to that individual team's HA, and a second due to the total of all the teams' HAs. Thus, although a team does better at home than away, this may be due to the collective advantage enjoyed by the other teams.

The final two columns of Table 3.1 show the results for 1986 division 1. Although all teams do better at home than away, the sum of the HAs of all teams is 14.14 goals. Teams with 14 goals or less difference in their home and away performance will consequently have a negative HA. For Norwich, Wimbledon and Chelsea their better home than away performances are spurious and due entirely to the HA of the other teams. The us have a range of about 2, so their difference has a range of about 4, whereas the hs have a range of about 1.4. This implies that in equation (3.1) ability is about three times more important than HA in determining goal difference. For the above ladder, the correlation between actual ladder position and the ladder position determined by $u_i + kh_i$ is best for k of about 0.5. As u affects a team's performance every match, and h only for half the matches, this is perhaps not unexpected.

3.3. Data and results

Data were collected for all English soccer matches from season 1981-82 to season 1990-91, comprising 920 teams and 20,306 matches. The Official Football Association Yearbook published by Penguin contains a summary of the previous year's match results and final ladders for each division. Individual results were entered and a computer program used to produce the end of year ladders, which were checked with those published. This often showed up about a 1% error rate in the actual match results (i.e. about three results per year per division were incorrectly reported). Results were checked with the newspapers if necessary until agreement with the ladder was obtained. All computing work was performed with SAS.

The home teams won 9894 (48.7%) and drew another 5415 (26.7%) of their matches. Of the total 54,378 goals the home team scored 32,556 or 59.9%, which is very close to but just under the percentage of wins 48.7 + 0.5x26.7 = 62.1%. This may suggest that HA factors are slightly better at producing wins than larger margins. The proportion of wins, draws and losses was remarkably consistent across divisions, with a chi square test for independence of results and division producing a p value of 0.949.

However, our main interest here is in calculating individual HAs. Using the methods shown above, the HAs were calculated for all teams playing from 1981-82 to 1990-91, and are given in Table 3.2. Table 3.2 is sorted in order of decreasing average HA, which is shown in the last column. By looking at the HAs of individual clubs we may discover the mechanism behind HA.

3.4. Discussion

The results show that a team's HA is quite variable from year to year and that in some years some teams have a negative HA. In fact 126 of 920 or about 14% are negative - in any one division in any year about 3 teams actually have a negative HA. Because of the inherent variation in soccer matches, an average over several years is necessary to obtain a reasonable measure of HA.

On average the home ground advantage is worth just over 0.5 of a goal, and that is amazingly constant over the divisions (0.521, 0.529, 0.529, 0.533 for divisions 1 to 4). The general linear models framework was used to perform various analysis-of-variance (ANOVA) tests which indicated that the year is significant, division is not, and differences between the clubs were only borderline significant. For example a test on all the data, 920 values, gave for year effect $p = 0.014$, division effect $p = 0.990$ and club effect $p = 0.085$. The residuals passed the usual tests for normality, and the R^2 value for this model was 14% so there is a large variation in HA. The following is clear from these results.

(a) There is no division effect. This is in contrast with the results of Pollard (1986b) and seems to negate crowd factor as a cause of HA.

(b) There is a highly significant year effect. The HA is about 10% higher than average in 1982, 1983 and 1985 seasons, and 10% lower in 1981, 1987 and 1989.

(c) There is some evidence for a significant club effect, but this is not conclusive. Certainly the club effect is weaker than the year effect.

TABLE 3.2. Home ground advantages for all teams in English soccer, 1981-82 to 1990-91, in order of decreasing average home advantage

Rank	Team	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	Aver
1	Plymo	0.36	0.72	1.38	0.50	1.03	0.84	1.22	1.08	0.38	1.59	0.911
2	Alder	0.43	0.76	0.79	0.93	1.52	0.86	1.14	0.75	0.58	0.73	0.849
3	Maids									1.31	0.32	0.816
4	Luton	0.31	0.20	-0.22	1.33	0.75	0.79	1.60	1.44	0.94	0.89	0.803
5	Exete	1.31	0.81	-0.35	0.70	-0.03	1.18	0.54	1.63	1.18	1.00	0.798
6	Leeds	0.99	-0.26	1.06	0.94	0.96	1.39	0.75	0.26	0.83	0.94	0.788
7	Leice	0.21	0.09	0.58	1.13	0.50	1.89	0.79	0.72	0.51	1.28	0.770
8	Evert	0.49	0.65	0.23	1.18	0.70	0.84	0.60	1.06	1.55	0.39	0.768
9	Carli	0.95	1.09	0.56	0.74	0.76	0.47	1.27	-0.55	1.08	1.27	0.765
10	Colch	0.52	1.35	1.38	0.25	0.84	1.09	-0.23	0.68	0.81		0.743
11	Millw	0.27	1.36	1.11	0.96	0.91	0.79	0.56	0.61	0.71	0.00	0.728
12	Bolto	0.91	0.74	1.11	1.19	0.43	0.33	1.13	0.93	0.20	0.27	0.724
13	Soton	1.24	0.75	0.73	-0.02	1.35	1.24	-0.13	0.33	0.49	1.22	0.720
14	Orien	0.56	0.86	0.83	0.05	0.29	0.77	1.04	1.59	0.06	1.14	0.719
15	Stock	0.89	1.12	0.29	1.65	0.34	0.68	0.13	0.32	0.86	0.87	0.714
16	Queen	1.41	0.84	0.43	1.38	1.25	0.49	0.71	0.11	-0.01	0.44	0.705
17	Hartl	0.43	1.26	1.29	0.47	1.20	-0.05	-0.19	0.68	1.58	0.32	0.700
18	Bourn	0.02	1.00	0.65	1.32	0.57	0.60	0.56	0.90	0.29	1.09	0.700
19	Stoke	0.44	1.00	1.08	0.38	0.76	0.79	0.89	0.81	0.42	0.37	0.694
20	York	1.21	1.62	0.52	0.69	1.39	1.06	-0.18	0.91	-0.14	-0.27	0.679
21	BrisR	0.49	1.13	0.92	0.82	0.98	0.42	1.14	-0.16	0.52	0.46	0.673
22	Oldha	-0.04	-0.06	1.21	0.79	0.51	0.34	0.56	0.81	1.38	1.23	0.673
23	ShefU	0.52	1.50	1.15	0.39	0.51	0.54	0.70	0.84	0.06	0.44	0.666
24	Scunt	0.30	0.22	0.88	1.02	0.47	1.09	0.36	0.13	0.36	1.77	0.659
25	Gilli	0.81	-0.00	0.88	0.96	0.93	1.28	0.95	0.30	0.22	0.23	0.657
26	Rochd	0.39	1.35	0.93	0.34	1.34	0.50	0.27	1.13	0.08	0.14	0.646
27	BrisC	0.22	0.49	1.20	0.69	1.12	0.74	0.73	0.11	0.06	1.09	0.645
28	Brigh	0.64	0.95	1.36	0.49	0.56	0.44	0.41	0.90	0.33	0.32	0.641
29	Newca	0.91	0.49	1.01	0.78	1.15	0.59	0.39	0.39	0.65	0.00	0.636
30	Shrew	1.11	-0.11	1.11	1.14	0.86	0.99	0.08	0.08	0.56	0.50	0.633
31	Charl	1.26	1.44	1.26	0.49	0.61	0.19	0.50	0.56	0.16	-0.13	0.633
32	Bpool	0.48	1.03	1.02	0.34	0.57	0.19	0.73	0.21	0.25	1.41	0.622
33	Prest	0.08	1.27	0.56	0.96	0.11	0.04	0.64	0.80	1.25	0.50	0.620
34	Burnl	0.13	1.29	0.97	0.59	0.02	0.13	0.54	1.18	0.04	1.23	0.612
35	Rothe	1.11	-0.16	0.61	0.55	1.62	0.47	0.50	-0.09	0.93	0.59	0.611
36	MancU	-0.11	1.15	0.43	1.23	-0.00	0.94	0.34	0.67	0.88	0.56	0.608
37	Halif	0.25	0.40	1.84	0.11	0.56	0.22	0.81	1.04	0.04	0.73	0.600
38	Oxfor	-0.19	0.04	0.61	1.79	1.00	1.04	0.39	0.31	0.29	0.46	0.574
39	Walsa	0.54	0.91	0.88	0.14	1.43	1.01	0.23	-0.19	0.34	0.23	0.551
40	Soend	0.08	0.36	1.06	0.20	0.70	0.41	1.14	0.89	0.81	-0.18	0.547
41	WestH	0.44	0.65	0.68	0.58	0.90	0.54	0.39	-0.17	1.01	0.41	0.544
42	Ipswi	0.69	0.05	0.63	0.58	0.15	0.49	0.98	0.67	0.92	0.23	0.540
43	Bburn	0.31	1.09	0.36	0.54	0.81	0.69	0.18	1.26	0.01	0.09	0.535
44	Chfld	0.13	0.36	0.29	0.56	0.48	0.38	0.55	0.98	0.90	0.68	0.531
45	Coven	0.59	0.90	0.43	0.88	0.15	0.84	-0.29	-0.06	0.49	1.33	0.527
46	Swans	1.04	0.35	0.96	-0.04	1.34	0.68	-0.14	0.48	0.15	0.41	0.524
47	Torqu	0.48	0.49	0.79	0.02	1.11	0.18	-0.01	0.91	0.45	0.77	0.519

TABLE 3.2. (Continued)

Rank	Team	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	Aver
48	Arsen	0.24	0.75	-0.07	1.03	0.70	0.04	0.55	-0.61	1.60	0.83	0.507
49	Wigan	0.57	-0.05	0.38	0.91	1.16	0.38	0.27	-0.16	0.88	0.59	0.494
50	Hull	0.25	0.58	0.52	0.46	0.86	0.39	0.65	0.67	-0.53	1.05	0.490
51	Carnbr	0.86	1.19	0.46	-0.72	0.65	0.91	0.31	0.82	0.40	0.00	0.488
52	PVale	-0.02	-0.24	1.02	0.70	0.79	-0.12	0.95	0.30	0.83	0.59	0.480
53	Wrexh	0.21	1.41	0.06	0.65	0.79	-0.05	0.81	0.23	0.22	0.46	0.479
54	Psmth	1.04	-0.09	0.26	-0.06	0.56	0.74	0.50	0.90	0.20	0.69	0.473
55	Heref	0.21	0.67	0.15	0.15	1.56	0.50	0.13	0.86	-0.32	0.77	0.469
56	Cryst	-0.14	1.24	-0.19	0.24	-0.19	0.99	0.89	0.67	1.27	-0.11	0.467
57	Wolve	1.04	0.79	0.78	0.19	-0.43	-0.23	0.27	1.07	0.74	0.41	0.463
58	Linco	0.54	1.04	0.52	0.59	0.25	0.72		0.41	-0.23	0.32	0.462
59	North	1.16	1.08	0.43	0.38	-0.25	0.54	0.27	0.48	-0.07	0.59	0.460
60	Hudde	0.49	1.09	-0.34	-0.01	0.86	0.54	0.51	0.71	0.20	0.55	0.460
61	Bradf	0.16	0.54	0.29	-0.09	1.16	0.29	0.65	0.36	0.79	0.46	0.460
62	Aston	-0.11	1.75	0.73	0.73	0.10	0.99	-0.92	0.61	0.21	0.50	0.460
63	Readi	0.63	0.72	1.38	-0.63	0.16	0.74	0.18	0.30	0.43	0.68	0.459
64	Watfo	0.71	1.20	0.38	0.63	0.85	0.44	-0.13	0.31	0.79	-0.59	0.458
65	Mansf	0.30	0.31	1.02	0.20	0.47	0.28	0.36	0.48	1.02	0.14	0.458
66	Scarb							0.86	-0.37	0.72	0.59	0.452
67	Grims	-0.09	1.24	0.76	0.59	0.66	0.09	-0.41	0.50	0.36	0.73	0.443
68	MancC	0.54	0.60	0.61	0.79	-0.00	0.74	0.56	0.13	0.55	-0.11	0.441
69	NottC	-0.11	1.15	0.03	0.19	0.30	1.10	0.64	0.30	0.61	0.09	0.430
70	Peter	1.02	0.44	1.47	0.11	0.97	-0.05	-0.28	0.13	-0.01	0.46	0.428
71	Derby	1.06	0.14	1.11	0.78	-0.16	0.49	0.18	0.11	0.21	0.28	0.421
72	Norwi	0.91	0.65	0.88	0.63	0.61	-0.11	0.18	-0.17	0.49	0.11	0.418
73	Swind	0.36	0.67	0.20	1.15	0.56	-0.22	0.51	0.54	0.42	-0.04	0.416
74	NottF	-0.21	0.15	0.88	0.63	-0.10	0.84	0.50	0.17	0.16	0.89	0.390
75	Tranm	-0.11	0.44	-0.07	1.29	-0.25	-0.09	1.22	0.45	0.75	0.27	0.389
76	Totte	0.14	1.90	0.18	-0.52	0.50	0.64	0.39	-0.33	0.05	0.78	0.372
77	WBA	0.04	0.55	0.63	0.83	0.60	-0.01	0.60	0.58	-0.49	0.37	0.371
78	Newpo	0.18	0.04	0.70	-0.04	0.12	0.47	1.13				0.370
79	Donca	0.77	0.50	0.20	0.37	-0.43	0.97	0.50	0.73	-0.19	0.18	0.359
80	Barns	0.41	0.19	0.11	0.99	-0.19	-0.16	0.41	0.49	0.42	0.91	0.359
81	Bury	0.98	0.17	-0.39	0.61	1.34	0.15	0.09	0.39	0.02	0.23	0.358
82	Liver	-0.31	0.55	1.18	-0.62	1.20	0.49	0.44	0.06	-0.23	0.56	0.332
83	Sunde	-0.11	0.60	0.78	-0.17	0.46	0.09	0.14	0.95	0.06	0.50	0.329
84	Chels	0.11	0.79	0.61	0.38	-0.15	-0.16	1.23	-0.42	-0.18	0.89	0.311
85	Middl	0.09	0.24	0.41	-0.11	0.11	0.28	0.79	0.39	0.38	0.37	0.296
86	ShefW	0.01	-0.26	0.51	0.48	0.25	0.84	-0.03	0.17	0.82	0.00	0.279
87	Fulha	0.36	0.14	0.06	0.19	0.26	-0.22	0.18	0.57	0.56	0.68	0.279
88	Brent	-0.32	0.72	0.65	0.91	-0.52	0.24	0.14	0.66	0.43	-0.27	0.264
89	Crewe	0.80	0.44	0.52	-0.03	0.47	-0.09	-0.14	0.18	0.11	0.37	0.262
90	Cardi	-0.19	0.86	0.46	-0.71	0.07	0.18	0.63	1.21	-0.07	0.14	0.257
91	Birmi	0.39	0.40	0.08	-0.31	0.35	0.59	0.22	0.36	0.75	-0.27	0.256
92	Darli	0.25	-0.47	0.43	0.38	0.66	0.69	0.68	-0.46		0.00	0.241
93	Chstr	0.18	-0.06	-0.12	0.43	0.38	-0.22	-0.09	1.07	0.70	0.09	0.236
94	Wimbl	0.72	0.12	0.33	0.79	0.51	-0.06	0.23	0.67	-1.01	-0.33	0.198
	Aver	0.45	0.66	0.64	0.52	0.59	0.51	0.47	0.50	0.45	0.50	0.528

3.4.1 Special Clubs

Pollard (1986b) singled out five clubs for special attention when looking at the effect of local conditions: Bristol Rovers and Halifax (small pitch), Manchester City and Carlisle (large pitch), Queen's Park Rangers (artificial turf). Pollard found that the points gained at home were not significantly different for these clubs. However, as we argued above, this would be affected greatly by the relative strengths of those clubs. Table 3.2 shows that most of them are in the top third - with ranks 9, 16, 21, 37 and 68 in a table with 94 values. A rank sum test gives R = 151 which has a p value of 0.076 which is some evidence that these teams have a higher-than-average HA.

What can we conclude about the 13 London clubs - Millwall (ranked 11), Leyton Orient (14), Queen's Park Rangers (16), Charlton (30), West Ham (41), Arsenal (48), Crystal Palace (56), Watford (64), Tottenham (76), Chelsea (84), Fulham (86), Brentford (88) and Wimbledon (94)? Because of their proximity we might expect them to have low HAs, and there are four in the bottom 11 rankings. Again, a rank sum test gives R = 708 with a p value of 0.161. Since Queen's Park Rangers has already been singled out as having special properties that may give it a large HA, it might be argued that we should exclude it in this analysis. Doing this gives R = 692 which has a p value of 0.072. So again there is some evidence that the London clubs have a lower than average HA.

Barnett & Hilditch (1993) looked specifically at the effect of artificial pitch on HA. Table 3.2 confirms their findings of an artificial pitch effect. Queen's Park Rangers from 1981-82 to 1987-88, Luton from 1985-86 to 1989-90, Oldham from 1986-87 to 1990-91 and Preston from 1986-87 to 1990-91 all had artificial pitches. The 22 seasons played on an artificial pitch had a mean HA of 0.889, compared with 0.519 for the other 898 seasons - significant at the 0.1% level. As this may be due to a year or team effect, including type of pitch along with year, division and club in an ANOVA test showed that the type of pitch was significant at the 1% level ($p = 0.0013$).

3.4.2 HA versus time in division

A possible reason often advanced for HA is the home team's familiarity with the 'quirks' of their home ground. Alternatively we could argue that the visiting teams are unfamiliar with the home team's facility. If this was so we would expect the HA to be greatest when a team is new to the division. To test this the current length of time continuously in the division for all teams was calculated. The results were the reverse of that expected, with a small non significant positive correlation between HA and years in division. Perhaps when teams are new to the competition the opposition put effort into counteracting their peculiarities, but relax this effort after they (mistakenly) believe that they are familiar with the opposition. To make sure this was not due to a year effect, averages by year and continuous time in divisions were looked at and the results confirmed the finding. For 0 or 1 years in the division, most of the averages were below the yearly average, whereas for 2 to 3 years in the division, most of the averages were above the yearly averages. Contrary to expectation, the teams that are new or have been in the division for only one year do not appear to have higher-than-average HAs.

3.5. Paired home advantage

The arguments advanced earlier for going from a competition level to a club level can be extended one stage further. Just as the competition level HA is an average of the HAs of all the clubs, so an individual club's HA is an average of its paired HA with all the other clubs it plays. For example, suppose that the HA is due entirely to distance travelled. A particular club would travel a short distance to some clubs (with no HA) and a long distance to others (with a consequently large HA). As its HA is an average of these it would have an average HA and the effect of distance would be lost. Thus the HA of one club is really the average of all its paired HAs with the HAs of the other clubs removed. It includes matches with nearby clubs, far clubs etc and so to some extent averages out the effects (of distance, crowd etc). Can we obtain a more refined measure by looking at the paired HA?

Stefani & Clarke (1992) state that the home ground advantage can be thought of as h_{ij}. For each pair of teams the difference in home and away matches - in our previous notation $w_{ij} + w_{ji}$ which is equivalent to $h_i + h_j$ - gives a measure of this for each year. Note that we need the actual match results for this - it cannot be calculated from the ladder. For our data, this gave 10153 match pairs, played between 2865 club pairs, with an average of 1.057 paired HA. This agrees with the previous estimate that a HA is worth about 0.5 of a goal, as the paired HA incorporates two individual HAs. However, the values are highly variable, ranging from -7 to +11.

To investigate whether distance had an effect, the grid coordinates of the home ground of each club on a map were estimated and the straight line distance between each pair calculated. The correlation between distance and HA was 0.07 - this is very small but because of the great number of observations is highly significant with $p = 0.0001$. The low correlation is due to the high variation in the individual data, making it difficult to explain high proportions of the variation. Several averaging methods were tried to reduce the variation, and all showed a clear relationship between paired HA and distance. For example, by averaging the paired HAs for each of the 2865 club pairs the correlation between average paired HA and distance became 0.11. By restricting the analysis to the 1303 pairs of teams which played each other for four seasons or more, the correlation increased to 0.14. The pairs of clubs were separated into groups in multiples of 50 miles apart and the average paired HA was calculated for each group. The results are shown in Figure 1 and clearly show increasing paired HA for increasing distance. To some extent this effect is reflected in Table 3.2. One referee pointed out that three clubs in the top ten, Plymouth, Exeter and Carlisle, are geographically isolated.

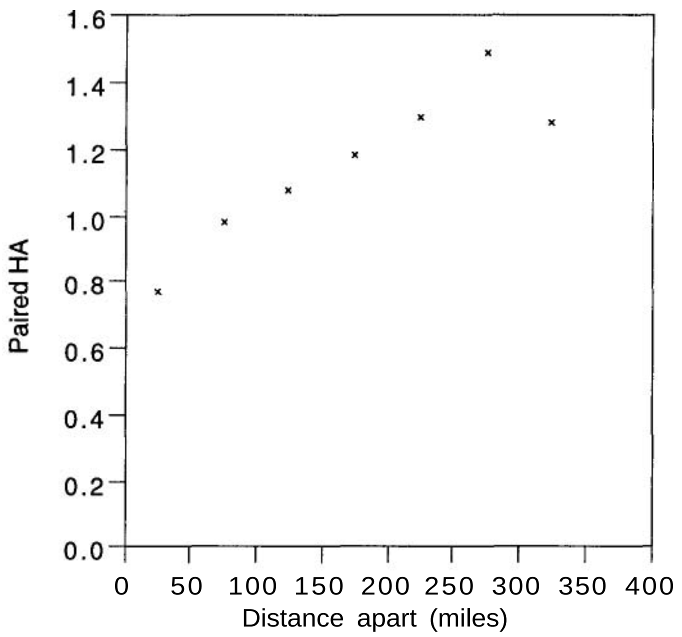


Figure 3.1. Average paired HA (goals) versus distance apart of clubs.

3.6. Win/loss home advantage

In the above analysis we have used goal difference to measure a team's performance. However, the analysis can be repeated using win (or point) margins as the measure of performance. Although goal difference should be a more sensitive measure than wins, it may be that HA works to produce wins rather than large margins. Replacing a team's score by 1 for a win, 0.5 for a draw and 0 for a loss produces win margins of 1, 0 and -1. The analysis can be repeated exactly, but the measures obtained would now be in terms of win margins rather than goal margins. Alternatively, using 3, 1 and 0 points produces point margins of 3, 0 and -3, but this is only an exact multiple of the win or lose case.

Using win margins produced similar results to the above, with a tendency to produce slightly more significant results. For example the overall average HA is 0.472, or nearly 0.5 of a win. The ranks of the five clubs singled out for special characteristics by Pollard (1986b) now go to 4, 9, 11, 41 and 80 with a rank sum 145 now significant with $p = 0.06$. The ranks of the 12 London clubs (without Queen's Park Rangers) are 28, 31, 34, 40, 47, 70, 72, 75, 81, 83, 91 and 93. This gives a rank total of 745 significant with $p = 0.02$. The fact that both these have moved in the direction indicating enhanced HA suggests that HA may have a greater effect on winning than on goal difference. Thus whatever it is that produces HA tends to operate more effectively in determining winners rather than just larger winning margins.

3.7. Conclusion

At a competition level, variations in percentage of home matches won may arise because of differences in team ability as well as variations in HA. To calculate HA at a club level, we need to take account of team ability, by looking at the difference in home and away performances. Least squares can be used on the match results to estimate team and HA effects. However, for a balanced competition such as English soccer, this is equivalent to simple calculation methods on the final ladder results.

Using ten years' data we have calculated HA in terms of goal and win difference for all 94 clubs in English soccer. These showed no division effects but significant year effects. There was some evidence that clubs with special facilities have significantly higher HA, and that London clubs have less-than-average HAs. There was no evidence that clubs new to a division have a higher HA. It also appeared that HA effects have more leverage on winning than on goal margins.

Paired HA is a more sensitive measure of HA, but individual match results are needed for its calculation. A definite linear relationship exists between a pair of clubs' paired HA and their distance apart.

Note: Some correspondence appeared following publication. See Bland & Bland (1996), Clarke & Norman (1996) and Longford (1997), Clarke & Norman (1997a).

3.8. Commentary. English soccer 1991-92 to 1995-96

The ease of applying these results via a spreadsheet is demonstrated by applying the methods to English soccer from 1991-2 to 1995-6. Final ladder results are more easily obtained than individual match results. In fact the final ladder results are archived on the internet for every year from 1887 onwards. These were copied and read into a Microsoft Excel spreadsheet. This spreadsheet was created using the same form as the archive, with the above formulas used to calculate two additional columns containing the u_i and hi. It is only necessary to cut and paste a year's results from the archive into the Excel file to produce the us and hs. This was done for each division and year for 1990-91 to 1995-6. The first year was used as a check against previous results. The columns for year, division, club, team rating and home advantage were then copied into a single spreadsheet and to SAS/JMP for further analysis. The calculated HAs are incorporated into Table 3.7 which extends Table 3.2 to include the years through to 1995-96. The table is sorted in alphabetical order.

In general the results as reported in this chapter for the years 81-82 to 90-91 are repeated for the years 1990-91 to 1995-6. The average HA was 0.43 goal per match. The yearly HAs ranged from -1.1 to 2.0 goals per match, with about 18% negative. However the average HA over 5 years ranged from -.11 to 1.1. Analysis again shows that HA is not dependant on division, nor year. However the team effect is significant ($p=.0386$). Of the 10 clubs with the lowest HA, five are London clubs. Clearly the mean HA of 0.29 for the 13 London clubs is significantly lower than the mean HA of 0.44 for the 81 non-London clubs. A surprising fact was the lack of consistency in the HAs from one year to the next. The correlations between HA from one year to the next are very small or even negative, while the correlation between the average HAs for the two periods was only 0.15. This suggests that teams do not enjoy a large HA over many years, and that opponents may quickly counteract perceived HAs. This may also be due to the promotion and relegation system in English soccer.

TABLE 3.3. Home ground advantages for all teams in English soccer, 1981-82 to 1995-96, in alphabetical order

Team	81-82	82-83	83-84	84-85	85-86	86-87	87-88	88-89	89-90	90-91	91-92	92-93	93-94	94-95	95-96	Aver
Aldershot	0.43	0.76	0.79	0.93	1.52	0.86	1.14	0.75	0.58	0.73						0.849
Arsenal	0.24	0.75	-0.07	1.03	0.70	0.04	0.55	-0.61	1.60	0.83	0.67	-0.14	-0.56	-0.07	0.11	0.338
Aston Villa	-0.11	1.75	0.73	0.73	0.10	0.99	-0.92	0.61	0.21	0.50	0.87	0.61	0.39	0.08	0.44	0.465
Barnet											1.01	1.30	0.64	0.92	0.57	0.888
Barnsley	0.41	0.19	0.11	0.99	-0.19	-0.16	0.41	0.49	0.42	0.91	0.18	0.56	-0.09	1.02	0.44	0.379
Birmingham C.	0.39	0.40	0.08	-0.31	0.35	0.59	0.22	0.36	0.75	-0.27	0.49	0.29	0.13	0.64	1.03	0.343
Blackburn R.	0.31	1.09	0.36	0.54	0.81	0.69	0.18	1.26	0.01	0.09	0.59	0.36	0.34	0.78	1.50	0.594
Blackpool	0.48	1.03	1.02	0.34	0.57	0.19	0.73	0.21	0.25	1.41	1.71	0.92	0.42	0.23	0.16	0.645
Bolton Wand.	0.91	0.74	1.11	1.19	0.43	0.33	1.13	0.93	0.20	0.27	0.04	0.78	0.31	1.16	-0.39	0.609
Bournemouth	0.02	1.00	0.65	1.32	0.57	0.60	0.56	0.90	0.29	1.09	0.63	0.15	-0.22	0.14	1.07	0.585
Bradford City	0.16	0.54	0.29	-0.09	1.16	0.29	0.65	0.36	0.79	0.46	-0.06	0.46	0.42	-0.36	0.84	0.394
Brentford	-0.32	0.72	0.65	0.91	-0.52	0.24	0.14	0.66	0.43	-0.27	0.63	0.15	-0.40	0.32	0.57	0.261
Brighton & HA	0.64	0.95	1.36	0.49	0.56	0.44	0.41	0.90	0.33	0.32	0.41	0.37	0.64	0.45	-0.03	0.549
Bristol City	0.22	0.49	1.20	0.69	1.12	0.74	0.73	0.11	0.06	1.09	0.82	0.65	0.41	0.21	0.25	0.586
Bristol Rovers	0.49	1.13	0.92	0.82	0.98	0.42	1.14	-0.16	0.52	0.46	0.96	-0.17	0.10	0.77	-0.30	0.539
Burnley	0.13	1.29	0.97	0.59	0.02	0.13	0.54	1.18	0.04	1.23	0.31	1.10	1.92	0.84	0.66	0.730
Bury	0.98	0.17	-0.39	0.61	1.34	0.15	0.09	0.39	0.02	0.23	0.31	0.90	0.80	0.32	-0.15	0.385
Cambridge U.	0.86	1.19	0.46	-0.72	0.65	0.91	0.31	0.82	0.40	0.00	0.09	0.15	0.05	0.82	0.39	0.425
Cardiff City	-0.19	0.86	0.46	-0.71	0.07	0.18	0.63	1.21	-0.07	0.14	0.46	0.30	0.64	0.32	0.80	0.340
Carlisle United	0.95	1.09	0.56	0.74	0.76	0.47	1.27	-0.55	1.08	1.27	0.51	0.50	0.10	-0.23	1.52	0.669
Charlton A.	1.26	1.44	1.26	0.49	0.61	0.19	0.50	0.56	0.16	-0.13	-0.54	0.15	0.86	0.52	-0.47	0.457
Chelsea	0.11	0.79	0.61	0.38	-0.15	-0.16	1.23	-0.42	-0.18	0.89	0.17	0.31	0.99	0.08	0.27	0.328
Chester City	0.18	-0.06	-0.12	0.43	0.38	-0.22	-0.09	1.07	0.70	0.09	0.04	0.33	0.20	0.00	0.80	0.249
Chesterfield	0.13	0.36	0.29	0.56	0.48	0.38	0.55	0.98	0.90	0.68	-0.09	0.20	0.30	-0.08	0.88	0.435
Colchester U.	0.52	1.35	1.38	0.25	0.84	1.09	-0.23	0.68	0.81			1.25	0.20	-0.13	0.48	0.653
Coventry City	0.59	0.90	0.43	0.88	0.15	0.84	-0.29	-0.06	0.49	1.33	0.32	-0.19	0.39	0.23	0.27	0.419
Crewe Alex.	0.80	0.44	0.52	-0.03	0.47	-0.09	-0.14	0.18	0.11	0.37	0.06	1.05	0.20	0.23	0.16	0.289
Crystal Palace	-0.14	1.24	-0.19	0.24	-0.19	0.99	0.89	0.67	1.27	-0.11	-0.13	0.31	0.13	-0.42	-0.15	0.294
Darlington	0.25	-0.47	0.43	0.38	0.66	0.69	0.68	-0.46		0.00	0.26	-0.95	0.35	0.37	-0.43	0.126
Derby County	1.06	0.14	1.11	0.78	-0.16	0.49	0.18	0.11	0.21	0.28	-0.27	-0.40	0.95	0.66	1.08	0.415
Doncaster R.	0.77	0.50	0.20	0.37	-0.43	0.97	0.50	0.73	-0.19	0.18	-0.64	-0.25	0.10	-0.38	0.62	0.203
Everton	0.49	0.65	0.23	1.18	0.70	0.84	0.60	1.06	1.55	0.39	0.42	-0.54	0.34	0.68	0.16	0.583

TABLE 3.3 (continued). Home ground advantages for all teams in English soccer, 1981-82 to 1995-96, in alphabetical order

Team	81-82	82-83	83-84	84-85	85-86	86-87	87-88	88-89	89-90	90-91	91-92	92-93	93-94	94-95	95-96	Aver
Exeter City	1.31	0.81	-0.35	0.70	-0.03	1.18	0.54	1.63	1.18	1.00	1.31	-0.22	1.01	0.17	0.16	0.693
Fulham	0.36	0.14	0.06	0.19	0.26	-0.22	0.18	0.57	0.56	0.68	0.44	-0.08	-0.18	0.97	1.03	0.331
Gillingham	0.81	-0.00	0.88	0.96	0.93	1.28	0.95	0.30	0.22	0.23	1.21	0.80	0.40	1.07	0.71	0.717
Grimsby Town	-0.09	1.24	0.76	0.59	0.66	0.09	-0.41	0.50	0.36	0.73	-0.04	0.15	0.13	0.71	0.44	0.388
Halifax Town	0.25	0.40	1.84	0.11	0.56	0.22	0.81	1.04	0.04	0.73	0.36	-0.75				0.468
Hartlepool U.	0.43	1.26	1.29	0.47	1.20	-0.05	-0.19	0.68	1.58	0.32	0.26	-0.08	0.51	0.97	1.03	0.645
Hereford U.	0.21	0.67	0.15	0.15	1.56	0.50	0.13	0.86	-0.32	0.77	0.86	0.65	0.70	0.72	0.39	0.533
Huddersfield	0.49	1.09	-0.34	-0.01	0.86	0.54	0.51	0.71	0.20	0.55	0.40	0.51	-0.27	0.41	1.21	0.457
Hull City	0.25	0.58	0.52	0.46	0.86	0.39	0.65	0.67	-0.53	1.05	-0.10	0.69	0.32	1.00	0.38	0.479
Ipswich Town	0.69	0.05	0.63	0.58	0.15	0.49	0.98	0.67	0.92	0.23	0.46	0.41	-0.26	1.38	0.53	0.527
Leeds United	0.99	-0.26	1.06	0.94	0.96	1.39	0.75	0.26	0.83	0.94	0.22	2.01	0.29	0.48	0.44	0.753
Leicester City	0.21	0.09	0.58	1.13	0.50	1.89	0.79	0.72	0.51	1.28	0.77	0.88	0.22	0.38	-0.38	0.638
Leyton Orient	0.56	0.86	0.83	0.05	0.29	0.77	1.04	1.59	0.06	1.14	0.63	1.37	1.23	0.91	1.07	0.827
Lincoln City	0.54	1.04	0.52	0.59	0.25	0.72		0.41	-0.23	0.32	-1.09	0.50	-0.10	0.82	0.85	0.367
Liverpool	-0.31	0.55	1.18	-0.62	1.20	0.49	0.44	0.06	-0.23	0.56	0.92	1.41	0.49	0.63	1.16	0.529
Luton Town	0.31	0.20	-0.22	1.33	0.75	0.79	1.60	1.44	0.94	0.89	1.97	0.10	0.81	0.02	0.35	0.752
Maidstone U.									1.31	0.32	0.26					0.630
Manchester C.	0.54	0.60	0.61	0.79	-0.00	0.74	0.56	0.13	0.55	-0.11	0.72	-0.29	0.44	0.98	1.11	0.491
Manchester U.	-0.11	1.15	0.43	1.23	-0.00	0.94	0.34	0.67	0.88	0.56	0.17	0.16	0.19	0.83	0.39	0.522
Mansfield T.	0.30	0.31	1.02	0.20	0.47	0.28	0.36	0.48	1.02	0.14	0.11	0.74	-0.10	0.12	-0.34	0.341
Middlesbrough	0.09	0.24	0.41	-0.11	0.11	0.28	0.79	0.39	0.38	0.37	0.96	1.11	0.81	0.21	0.33	0.425
Millwall	0.27	1.36	1.11	0.96	0.91	0.79	0.56	0.61	0.71	0.00	-0.13	1.20	0.77	0.71	0.08	0.661
Newcastle U.	0.91	0.49	1.01	0.78	1.15	0.59	0.39	0.39	0.65	0.00	1.09	0.92	1.34	1.13	1.11	0.797
Newpo	0.18	0.04	0.70	-0.04	0.12	0.47	1.13									0.371
Northampton	1.16	1.08	0.43	0.38	-0.25	0.54	0.27	0.48	-0.07	0.59	0.21	0.00	0.95	0.27	0.16	0.413
Norwich City	0.91	0.65	0.88	0.63	0.61	-0.11	0.18	-0.17	0.49	0.11	0.47	0.86	-0.81	0.98	-0.38	0.353
Nottingham F.	-0.21	0.15	0.88	0.63	-0.10	0.84	0.50	0.17	0.16	0.89	0.37	-0.29	-0.23	-0.12	1.05	0.313
Notts County	-0.11	1.15	0.03	0.19	0.30	1.10	0.64	0.30	0.61	0.09	0.17	1.24	1.18	0.21	0.29	0.493
Oldham Ath.	-0.04	-0.06	1.21	0.79	0.51	0.34	0.56	0.81	1.38	1.23	0.77	1.31	0.09	0.61	0.62	0.675
Oxford United	-0.19	0.04	0.61	1.79	1.00	1.04	0.39	0.31	0.29	0.46	0.68	0.33	0.41	0.05	1.25	0.564
Peterborough	1.02	0.44	1.47	0.11	0.97	-0.05	-0.28	0.13	-0.01	0.46	0.76	0.20	0.81	0.00	0.97	0.467
Plymouth Arg.	0.36	0.72	1.38	0.50	1.03	0.84	1.22	1.08	0.38	1.59	0.55	0.60	-0.13	0.05	0.62	0.719

TABLE 3.3 (continued). Home ground advantages for all teams in English soccer, 1981-82 to 1995-96, in alphabetical order

Team	81-82	82-83	83-84	84-85	85-86	86-87	87-88	88-89	89-90	90-91	91-92	92-93	93-94	94-95	95-96	Aver
Port Vale	-0.02	-0.24	1.02	0.70	0.79	-0.12	0.95	0.30	0.83	0.59	0.14	0.33	0.55	0.25	0.03	0.407
Portsmouth	1.04	-0.09	0.26	-0.06	0.56	0.74	0.50	0.90	0.20	0.69	1.55	1.47	0.36	0.16	0.17	0.563
Preston N.End	0.08	1.27	0.56	0.96	0.11	0.04	0.64	0.80	1.25	0.50	0.85	0.24	1.00	0.72	-0.25	0.585
QPR	1.41	0.84	0.43	1.38	1.25	0.49	0.71	0.11	-0.01	0.44	-0.08	0.01	-0.06	0.43	0.44	0.519
Reading	0.63	0.72	1.38	-0.63	0.16	0.74	0.18	0.30	0.43	0.68	0.13	0.96	0.01	-0.02	-0.15	0.368
Rochdale	0.39	1.35	0.93	0.34	1.34	0.50	0.27	1.13	0.08	0.14	0.51	0.50	0.65	0.92	-0.34	0.581
Rotherham U.	1.11	-0.16	0.61	0.55	1.62	0.47	0.50	-0.09	0.93	0.59	0.06	-0.26	0.46	0.68	0.84	0.527
Scarborough							0.86	-0.37	0.72	0.59	0.81	0.05	0.05	0.12	0.39	0.358
Scunthorpe U.	0.30	0.22	0.88	1.02	0.47	1.09	0.36	0.13	0.36	1.77	1.31	0.75	0.65	0.32	-0.15	0.632
Sheffield U.	0.52	1.50	1.15	0.39	0.51	0.54	0.70	0.84	0.06	0.44	0.07	0.81	0.69	0.39	-0.15	0.564
Sheffield Wed.	0.01	-0.26	0.51	0.48	0.25	0.84	-0.03	0.17	0.82	0.00	0.42	0.06	0.99	-0.07	0.11	0.287
Shrewsbury T.	1.11	-0.11	1.11	1.14	0.86	0.99	0.08	0.08	0.56	0.50	0.04	-0.05	-0.45	0.59	0.29	0.449
Southampton	1.24	0.75	0.73	-0.02	1.35	1.24	-0.13	0.33	0.49	1.22	-0.73	0.71	0.44	0.23	0.83	0.579
Southend U.	0.08	0.36	1.06	0.20	0.70	0.41	1.14	0.89	0.81	-0.18	0.55	0.92	0.18	1.02	0.76	0.593
Stockport C.	0.89	1.12	0.29	1.65	0.34	0.68	0.13	0.32	0.86	0.87	0.90	1.01	0.69	0.45	-0.25	0.663
Stoke City	0.44	1.00	1.08	0.38	0.76	0.79	0.89	0.81	0.42	0.37	0.44	0.24	1.00	0.48	0.67	0.651
Sunderland	-0.11	0.60	0.78	-0.17	0.46	0.09	0.14	0.95	0.06	0.50	0.91	0.65	0.77	-0.39	0.44	0.379
Swansea City	1.04	0.35	0.96	-0.04	1.34	0.68	-0.14	0.48	0.15	0.41	0.90	0.56	1.14	-0.05	0.93	0.581
Swindon Town	0.36	0.67	0.20	1.15	0.56	-0.22	0.51	0.54	0.42	-0.04	0.37	0.42	0.34	0.39	-0.30	0.358
Torquay United	0.48	0.49	0.79	0.02	1.11	0.18	-0.01	0.91	0.45	0.77	1.54	-0.10	-0.15	0.72	0.30	0.500
Tottenham H.	0.14	1.90	0.18	-0.52	0.50	0.64	0.39	-0.33	0.05	0.78	-0.38	1.26	-0.46	-0.17	-0.39	0.239
Tranmere Rov.	-0.11	0.44	-0.07	1.29	-0.25	-0.09	1.22	0.45	0.75	0.27	0.00	0.92	1.00	1.57	0.62	0.534
Walsall	0.54	0.91	0.88	0.14	1.43	1.01	0.23	-0.19	0.34	0.23	0.21	-0.05	0.10	0.22	0.43	0.429
Watford	0.71	1.20	0.38	0.63	0.85	0.44	-0.13	0.31	0.79	-0.59	-0.41	-0.17	0.45	0.61	0.62	0.379
West Brom. A.	0.04	0.55	0.63	0.83	0.60	-0.01	0.60	0.58	-0.49	0.37	0.58	1.01	0.50	0.52	0.44	0.450
West Ham U.	0.44	0.65	0.68	0.58	0.90	0.54	0.39	-0.17	1.01	0.41	0.47	0.65	-0.26	0.63	0.44	0.491
Wigan Athletic	0.57	-0.05	0.38	0.91	1.16	0.38	0.27	-0.16	0.88	0.59	0.81	0.06	0.60	-0.28	0.66	0.452
Wimbledon	0.72	0.12	0.33	0.79	0.51	-0.06	0.23	0.67	-1.01	-0.33	0.77	0.36	0.94	0.38	-0.34	0.272
Wolverham.W.	1.04	0.79	0.78	0.19	-0.43	-0.23	0.27	1.07	0.74	0.41	0.27	0.42	0.22	0.61	0.44	0.439
Wrexham	0.21	1.41	0.06	0.65	0.79	-0.05	0.81	0.23	0.22	0.46	1.06	0.65	1.10	0.55	0.70	0.590
Wycombe W.													0.25	0.50	0.20	0.317
York City	1.21	1.62	0.52	0.69	1.39	1.06	-0.18	0.91	-0.14	-0.27	0.61	0.85	0.23	0.32	0.07	0.593

Appendix 3.1. Spurious and real home ground advantage

Consider three teams, A, B and C. Suppose that A is better than B which is better than C, and there are no home ground advantages. Suppose that both home and away A beats B 2-1 and C 3-1, whereas B beats C 2-1. Final results would be as in Table 3.4 with the final ladder as in Table 3.5. Obviously each team has the same home performance as away both in terms of wins and goals.

TABLE 3.4. Final results

Home team	Away team		
	A	B	C
A	-	2-1	3-1
B	1-2	-	2-1
C	1-3	1-2	-

TABLE 3.5. Final ladder

Team	Home				Away				Home–Away	
	Wins	Draws	Losses	Goals	Wins	Draws	Losses	Goals	Wins	Goals
A	2	0	0	5-2	2	0	0	5-2	0	0
B	1	0	1	3-3	1	0	1	3-3	0	0
C	0	0	2	2-5	0	0	2	2-5	0	0

However, we now give C a 2 goal home ground advantage so that C will perform 2 goals better at home than anywhere else. Thus at home it will draw against A and beat B 3-2. The results and end-of-year ladder are now as in Tables 3.6 and 3.7 respectively. The final ladder shows that, even though only C has a HA, all teams had better results at home than away, both in terms of wins and goal difference.

A 'naive' analysis of goal difference would incorrectly conclude that each team had a home ground advantage - A and B performing better at home than away over the season by a total of 2 goals, whereas C performed better by a total of 4 goals.

TABLE 3.6. Final results when C has a 2-goal HA

Home team	Away team		
	A	B	C
A	-	2-1	3-1
B	1-2	-	2-1
C	3-3	3-2	-

TABLE 3.7. Final ladder when C has a 2-goal HA

Team	Home				Away				Home–Away	
	Wins	Draws	Losses	Goals	Wins	Draws	Losses	Goals	Wins	Goals
A	2	0	0	5-2	1	1	0	5-4	0.5	2
B	1	0	1	3-3	0	0	2	3-5	1	2
C	1	1	0	6-5	0	0	2	2-5	1.5	4

Appendix 3.2. Derivation of formula for calculation of home advantage and team performance by using least squares

Let w_{ij} be the winning margin for home team i against away team j, (negative if loss). For N teams, this gives an N x N matrix with no diagonals. Adding across a row gives the home goal difference HGD, whereas adding down **a** column gives the negative of away goal difference AGD,

i.e. for team I $\qquad \mathrm{HGD}_I = \sum_{j=1(j\neq I)}^{j=N} w_{Ij}$, $\qquad \mathrm{AGD}_I = -\sum_{i=1(i\neq I)}^{i=N} w_{iI}$

Thus since we are merely summing all the w_{ij} in a different order

$$\sum_{i=1}^{i=N} \mathrm{HGD}_i = -\sum_{i=1}^{i=N} \mathrm{AGD}_i$$

If u_i is a measure of team ability, rating or skill level etc. of team i and h_i is the home ground advantage of team i, and e_{ij} is random error, then as before in equation (3.1) we model the winning margin by

$$w_{ij} = u_i - u_j + h_i + e_{ij} \tag{3.1}$$

Since only differences of the u_i are used, they are relative, and we make the arbitrary restriction that $\sum_{i=1}^{i=N} u_i = 0$. So minimising the sums of squares of the errors subject to this condition, we have using the usual Lagrange multiplier expression.

$$\text{Minimise} \quad S = \sum_{i=1}^{i=N} \sum_{j=1(j\neq i)}^{j=N} (w_{ij} - u_i + u_j - h_i)^2 + \lambda \sum_{i=1}^{i=N} u_i$$

In the normal manner, partial differentiating with respect to u_I, I = 1 to N, h_I, I = 1 to N, and λ we get 2N + 1 equations.

$$\sum_{j=1(j\neq I)}^{j=N} 2(w_{Ij} - u_I + u_j - h_I)(-1) + \sum_{i=1(i\neq I)}^{i=N} 2(w_{iI} - u_i + u_I - h_i) + \lambda = 0, \ I = 1 \text{ to } N \tag{3.2}$$

$$\sum_{j=1(j\neq I)}^{j=N} 2(w_{Ij} - u_I + u_j - h_I)(-1) = 0, \quad I = 1 \text{ to } N \tag{3.3}$$

$$\sum_{i=1}^{i=N} u_i = 0 \tag{3.4}$$

Expanding *(3.3)* gives $\displaystyle\sum_{j=1(j\neq I)}^{j=N} w_{Ij} = (N-1)\, u_I + (N-1)\, h_I - \sum_{j=1(j\neq I)}^{j=N} u_j$

$$\text{i.e. } \mathrm{HGD}_I = N\, u_I + (N-1)\, h_I + \sum_{j=1}^{j=N} u_j$$

$$\mathrm{HGD}_I = N\, u_I + (N-1)\, h_I \tag{3.5}$$

So adding for I = 1 to *N*,

$$\text{i.e. } \sum_{I=1}^{I=N} \mathrm{HGD}_I = N \sum_{I=1}^{I=N} u_I + (N-1) \sum_{I=1}^{I=N} h_I$$

$$\mathbf{\mathit{HGD}} = (\mathrm{N}-1)\ H \tag{3.6}$$

where $\mathrm{H} = \sum_{i=1}^{i=N} h_i$ is the total of all the individual team's home ground advantages.

From (3.2), substituting *(3.3)* eliminates the first summation term, so

$$-\lambda/2 = \sum_{i=1(i\neq I)}^{i=N} (w_{iI} - u_i + u_I - h_i)$$

$$= \sum_{i=1(i\neq I)}^{i=N} w_{iI} - \sum_{i=1(i\neq I)}^{i=N} u_i - \sum_{i=1(i\neq I)}^{i=N} h_i + (N-1)\, u_I$$

$$= -\mathrm{AGD}_I + u_I - H + h_I + (N-1)\, u_I$$

$$-\lambda/2 = -\mathrm{AGD}_I - H + h_I + N\, u_I \tag{3.7}$$

$$\text{So } \sum_{I=1}^{I=N} -\lambda/2 = \sum_{I=1}^{I=N} -\mathrm{AGD}_I - NH + \sum_{I=1}^{I=N} h_I + N \sum_{I=1}^{I=N} u_I$$

$$-N\,\lambda/2 = \sum_{I=1}^{I=N} \mathrm{HGD}_I - (N-1)\, H + 0$$

$$= 0 \text{ from } \mathit{(3.6)}$$

So $\lambda = 0$ and *(3.7)* becomes

$$AGD_I = -H + h_I + N u_I \tag{3.8}$$

So subtracting *(3.8)* from *(3.5)*

$$HGD_I - AGD_I = N u_I + (N-1) h_I + H - h_I - N u_I$$

$$HGD_I - AGD_I = H + (N-2) h_I \tag{3.9}$$

Thus H is calculated from (3.6), h_i from *(3.9)* and u_i from *(3.5)*.

Appendix 3.3. Derivation of formula for calculation of home advantage and team performance from final ladder using simple explanation

As before we model the winning margin by

$$w_{ij} = u_i - u_j + h_i$$

The error term is neglected for simplicity. It could be included, and discarded later under the assumption that it sums to zero.

Adding across row i gives the home ground performance of team i as

$$\mathrm{HGD}_i = \sum_{j=1(j\neq i)}^{j=N} w_{ij} = \sum_{j=1(j\neq i)}^{j=N} (u_i - u_j + h_i),$$

$$= (N-1)u_i - \sum_{j=1(j\neq i)}^{j=N} u_j + (N-1)h_i$$

$$= Nu_i - \sum_{j=1}^{j=N} u_j + (N-1)\, h_i$$

Now as the u_i are relative, and it is only the difference that matters, we can require that they sum to zero.

$\mathrm{HGD}_i = Nu_i + (N-1)h_i$ as in *(3.5)* from Appendix *3.2.*

If we sum all the home performances for the whole competition we obtain

$$\sum_{i=1}^{i=N} \mathrm{HGD}_i = \sum_{i=1}^{i=N} \{Nu_i + (N-1)h_i\} = (N-1) \sum_{i=1}^{i=N} h_i$$

$$= (N-1)H$$ as in *(3.6)* from Appendix *2*

In a similar manner, a team's away performance is obtained by adding up the negatives of a column. For column j we have

$$\mathrm{AGD}_j = \sum_{i=1(i\neq j)}^{i=N} -w_{ij} = \sum_{i=1(i\neq j)}^{i=N} (-u_i + u_j - h_i)$$

$$= - \sum_{i=1(i\neq j)}^{i=N} u_i + (N-1)u_j - \sum_{i=1(i\neq j)}^{i=N} h_i$$

$$= -\sum_{i=1}^{i=N} u_i + Nu_j - H + h_j$$

$$= Nu_j - H + h_j \quad \text{as in } (3.8) \text{ from Appendix } 3.2.$$

The difference between home and away performance for any team now becomes

$$\mathrm{HGD}_i - \mathrm{AGD}_i = Nu_i + (N-1)\,h_i - Nu_i + H - h_i$$

$$= H + (N-2)h_i \quad \text{as in (3.9) from Appendix } 3.2.$$

COMPUTER FORECASTING OF AUSTRALIAN RULES FOOTBALL FOR A DAILY NEWSPAPER

4.0. Abstract

An exponential smoothing technique operating on the margins of victory was used to predict the results of Australian rules football matches for a Melbourne daily newspaper from 1981-86 and again for a competitor in 1991-92. An initial 'quick and dirty' program used only a factor for team ability and a common home ground advantage to predict winning margins. Probabilities of winning were accumulated to predict a final ladder, with a simulation to predict chances of teams finishing in any position. Changes to the competition forced a more complicated approach, and the current version uses several parameters which allow for ability, team/ground interaction, team interaction, and a tendency for team ability to regress towards the mean between seasons. A power method is used to place greater weight on the errors in closer matches, and errors across the win-lose boundary. While simple methods were used originally, the Hooke and Jeeves method was used in optimising the parameters of the current model. Both the original model and the improved version performed at the level of expert tipsters.

Key words: sports, forecasting

4.1. Introduction

The major winter sport of the southern states of Australia is Australian rules football, played between teams of 18 players on oval grounds (the same grounds used for cricket during the summer). A match is played for four quarters, each of 25 minutes duration plus about five minutes of extra time. Players can run with the rugby shaped ball, but it is moved forward more quickly by kicking or punching it to a team-mate, and with no off-side rule, the game is reasonably fast. The scoring region consists of four upright posts. Kicking the ball between the two centre posts scores a goal worth 6 points, while the region between either centre post and the corresponding outside post scores a 'behind' worth 1 point. Draws are rare; a typical score might be 18 goals 12 behinds, 120 points, to 12 goals 15 behinds, 87 points, for a winning margin of 33 points. In 1981 the major competition in Australia was organised by the Victorian Football League (VFL), in which 12 Melbourne based clubs played for 22 home and away rounds with a final series of six matches between the top five teams.

In common with most team sports, Australian rules football uses a ladder which accumulates points for winning matches to rank the individual teams throughout the year. Such methods, in which the total number of points never diminishes, have limited use for prediction, as no account is taken of the ability of the opponent, nor of how recently wins occurred. In adjustive methods, the level of performance above or below that predicted is used to adjust the current rating up or down. Harville (1980), Stefani (1977, 1980, 1987), and Stefani & Clarke (1992) give examples of adjustive least square ratings methods applied to soccer, American football and Australian rules football. In 1981, it was decided to predict VFL results using an adjustive scheme similar to the Elo system used by the World Chess Federation, where a simple exponential smoothing technique is used to adjust player ratings (Elo, 1978).

4.2. Initial program

About two months before the start of the 1981 football season, work began on developing a computer prediction model for Australian rules football. Because of the time constraints, a relatively simple method was used. There are many factors which football followers believe affect performance - team ability, current form, the opposition, team personnel, home ground advantage, weather etc. The initial program used only a rating for each team, and a common home ground advantage. Thus, if team i played at home to team j, the predicted winning margin P for the home team i was

$$P = u_i + h - u_j \tag{4.1}$$

where u_i is the rating of team i, and incorporates team ability and current form, while h is a home ground advantage common to all teams. A negative value of P implies a win of $|P|$ for the away team. While the published draw always specified a nominal home team (for the purposes, where necessary, of choice of rooms, colour of shorts etc) a few teams actually shared grounds, and one match each week was played on the league's own ground VFL park. For these matches there was no home ground advantage and h was set to zero. To update the ratings a simple exponential smoothing algorithm was used. If the actual margin of the match was w points the prediction is in error by

$$e = w - P \tag{4.2}$$

points and u_i is then increased by αe and u_j decreased by αe, where α is the smoothing constant. Thus, if the margin for the home team is greater than predicted, its rating would be increased, and that of its opponent decreased, and vice versa. In practice it was decided arbitrarily to limit this change to a maximum value - if the magnitude of the error was greater than 75 points it was reset to 75.

In many so called computer ratings, the computer may not be necessary to calculate the ratings once the form of the algorithm is decided. The above algorithm is so simple it can be performed on a hand calculator or even mentally. However the computer is necessary in finding the values of the parameters that optimise performance. In this case, we need some starting values for the ratings, and values for the smoothing constant and the common home ground advantage. With little time to prepare the program, some short cuts were necessary. As starting values for the ratings the premiership points (four times the number of wins) gained by each team the previous year were used. A short program was written in BASIC to run through the 1980 results and calculate the number of correct winning predictions, using values of the smoothing constant, α, of 0.0 to 0.5 in steps of 0.05, and home ground advantage, h, of 0 to 10 points in steps of 1. While the integral values of the objective function allowed for some judgement in the final selection, the optimal values were a home ground advantage of 7 points with an α value of 0.15.

With the parameter values decided, a cumulative relative frequency histogram of the absolute prediction errors was charted which allowed conversion of a predicted point margin into a probability of winning. For example, if 24% of predictions are in error by more than 40 points, then a team predicted to win by a 40 point margin has a 12% chance of winning by more than 80 points and a 12% chance of not winning. Thus, a predicted winning margin of 40 points translates to an 88% chance of winning. In the resultant computer program a five section straight line approximation joining the points (0, 0), (5, 0.58), (15, 0.68), (40, 0.88), (65, 0.95), (□, 0.95) was used to convert margins to winning probabilities. As the predicted probability of a team winning a match can also be interpreted as the expected number of matches that team will win on that day, a simple accumulation of actual wins in past matches and expected wins in future matches was used to produce a predicted final ladder. Because the VFL use the 'percentage' (100 times the ratio of the total points each team scores to the total points scored against it) to separate ties on the ladder, a separate smoothing of the points each team scored and had scored against it each week was used in conjunction with the margin prediction to produce a predicted score for each team in a match. These were not printed out but accumulated to estimate the percentage at the end of the season.

Until 1991, when a sixth team was introduced, the top five teams at the end of the season played off in a final series. Because of the structure of this series, there is a big advantage in finishing top at the end of the home and away matches. In turn, second and third have a big advantage over fourth and fifth. To estimate the chances of teams finishing the home and away series in any position a simulation was introduced. For any unplayed matches, a uniform random number was generated. This was used in the inverse of the margin distribution function to generate an actual winning margin for the match. As before, the wins could now be accumulated to obtain the final ladder. Thus, in the ladder prediction we accumulated the probability of winning each match, whereas in the simulation we replaced this with a 0 or 1 depending on a random number. This was repeated for 1000 years to estimate the probabilities of teams finishing in any ladder position. Once the final series began, a separate program used the ratings from the prediction program to calculate win probabilities for all possible matches in the final series, which it then used to evaluate the probabilities of a range of final series outcomes.

All the analysis was performed on a FACOM mainframe and the final program consisted of about 600 lines of BASIC. Upon completion, the results of the program were offered to a Melbourne daily newspaper under a consulting agreement, and so began a six-year association. The computer's predicted winners and margins were published each week, and the final ladder predictions a couple of times each season.

Surprisingly, in view of the quick and dirty development of the program, it performed quite well in its first year. Clarke (1981) showed that with 99 correct winners from 132 matches its 75% correct placed it equal 22nd out of 56 tipsters. It averaged 26 points in error, predicted 71% of matches within 36 points, and for each round after the 12th round it predicted at least 10 out of 12 teams to finish within one place of their actual finishing position.

Some of the experiences of this period are discussed in Clarke (1988c). Minor adjustments to the printout were necessary to allow for the readership's level of expertise. These ranged from referring to probabilities as percentage chances to actually suppressing information from the printout. The original printout showed for each team the expected result (win or loss) for each of the remaining matches, in addition to the expected number of wins for the remainder of the season. This often resulted in seemingly contradictory material. For example, a team's predicted results would be shown as WWWW if it was rated a 0.75 chance to win each of its remaining four matches, whereas the final ladder prediction would show it was expected to win only three (4x0.75) of its last four matches. The immediacy and nature of the forecasts also meant the predictions were often judged harshly. The margin prediction is technically a line which divides the possible margins into two equally likely regions. Thus, a margin prediction of 40 points implies the team is just as likely to win by more than 40 as it is to win by less than 40 or lose. The general public consider it as the actual margin the computer is predicting will occur, so when the result is a win by 80 points they consider the computer has performed badly. If the above team loses, the computer is considered to be completely wrong, whereas in fact it would have estimated the team's chance of losing as 12%. Unfortunately, in six years the predicted probabilities of winning were never published, whereas the margins, which are rarely correct, were always published. At the end of the season, the final judgement by the public of the computer's performance would be the number of correct winners for the season. However, predicting 132 matches with roughly a 75% success rate results in a high variability in the number of correct winners. Success by the public's usual measure owed as much to good fortune as to good forecasting. However, the program continued to perform so well that in spite of the simple nature of the model, it was five years before other factors forced a rethink of its development.

4.3. Second program

In 1986 the computer program was reorganised. The VFL over the previous few years had been modernising their draw to maximise crowds. Previously, all matches were played on a Saturday. With the introduction of a team from outside Victoria, the league began a move towards Sunday matches, Friday night matches, splitting rounds over long weekends etc. A round-by-round program was no longer appropriate as it often meant providing predictions half way through a round that were based on out of date information. In addition, the league was making greater use of its large capacity grounds by ground sharing schemes and even shifting popular matches after the draw was published. Not only was an individual home ground advantage required, but also a measure of how teams performed on grounds other than their home ground became more important.

It was decided to change the program from a round-by-round prediction to match-by-match. At the same time the prediction equation and updating equation were changed and re-optimised to take into account the above and other factors.

The prediction equation used, if team i played team j at ground k, became

$$P = u_i + h_{ik} + I_{ij} + u_j - h_{jk} \tag{4.3}$$

where h_{ik} is a ground measure of how team i performs at ground k, and I_{ij} is a measure of interaction between team i and team j. I_{ij} was introduced as it was a widely held view of supporters that some teams always performed well (or badly) against some other teams irrespective of their relative ladder positions.

The rating, ground and interaction measures were updated using the same method as before, but with different smoothing constants for each measure. However, a radically different measure of the 'error' was chosen.

In the previous formulation, a prediction of a 49 point win that resulted in an 81 point win was in error by 32 points. The same error resulted from a 4 point prediction and a 36 point win, or a 16 point predicted loss with a 16 point win. However, the significance of the error increases with each case. In the first example, a match predicted as one sided was just that, in the second a predicted close win became a comfortable one, and in the last a predicted loss was actually a win. A measure of the error that reflected the increasing seriousness was needed. It was decided to use a power function to reduce the relative errors of matches with large actual or predicted margins, and to increase the weighting across the 'win-loss' boundary of zero points. For example, the use of a square root power in the above would give errors of 9-7 = 2, 6-2 = 4, and 4-(-4) = 8 respectively. Thus, we have

$$e = \mathrm{Sgn}(w).|w|^x - \mathrm{Sgn}(P).|P|^x \tag{4.4}$$

where x is the chosen power.

One other factor needed to be taken into account. At the beginning of each year, starting values for the ratings were needed. This always caused some stress, as the chosen values virtually selected the margins in the opening round, and the process needed automating. The practice had arisen of simply using the ratings at the end of the previous year, but shrinking them relative to the mean to allow for a tendency to regress towards the mean. Thus, because of team changes, injuries and a host of other random effects we expect, on average, the best teams to get weaker, and the weak teams to get better. Since the ratings averaged about 70, at the beginning of each year the following equation would be applied.

$$\text{Rating at start of year} = 70+k(\text{rating at end of previous year} - 70) \qquad (4.5)$$

The shrinkage factor, *k*, was around 0.8, but as it was chosen on subjective grounds it varied slightly from year to year. An optimal value was needed.

With the form of the method settled, but with six unknown parameters (three smoothing constants, the power *x*, the start of year shrinkage factor, and one other not detailed here) there was six years of past data on which to optimise the parameters. By this time the program had been transferred to a PC and took about five minutes to run through six years' data for one set of parameters and evaluate the total sum of the absolute errors. With a possible grid of over 10^6 sets of parameter values, special hill climbing techniques were necessary to find the optimum values and the Hooke and Jeeves' method outlined in Walsh (1975) p 76 was used. Running overnight on the PC, this gave optimal values to any desired accuracy. The average absolute error in margin prediction was used as the objective function rather than the number of winners, as it was more sensitive to small changes in the parameter values. It was assumed that good predictions of winners would follow from accurate margin predictions. In addition, after trying several alternatives, the government had finally settled on a legal gambling system for football that involved selecting the correct winning margin band, so accurate margins had become relatively more important to the football public.

Several parameters came out very close to the values that had previously been chosen. The power parameter was 0.75, the main smoothing value was 0.2, with a much smaller value for the ground factor. The interaction parameter of zero suggested that supporters were misled in their belief in an interaction effect, and the end of season shrinkage came out near the 0.8 that had previously been used. The suggested values were implemented for the season 1986. It was considered not worthwhile to spend time in developing better probability and end-of-season ladder predictions as these were rarely published. The simulation was discarded as the results had never been published and it slowed the running on the PC considerably.

Although the program again tipped more winners than the paper's major football writer, at the beginning of 1987 the Sports Editor decided to dispense with the computer tip and concentrate on human tipsters. This suggests that the client probably judged the success of the project on different criteria from the public or the practitioner. The newspaper's interest in the publicity the computer tips created may overshadow the need for accurate forecasts. In fact the major football writer of the paper was known for his sometimes outlandish predictions, which created huge public interest. Although it is possible in an exponential smoothing forecast to select parameters that give conservative or controversial forecasts, this issue was never discussed with the client and the parameters were always chosen to optimise accuracy.

The program was maintained in the vain hope of renewal in 1988 and 1989. Then, at the beginning of 1991, one week before the start of the season, a request to supply tips to Melbourne's other daily paper was made. By this time, the VFL in going national, had become the Australian Football League (AFL) and the competition had increased to 15 teams including four from interstate. The program was dusted off, minor alterations made to allow for 15 teams, 1990 data entered and run, and the 1991 draw entered. Although there was no time to perform any re-optimising, the league had introduced a pre-season knockout night series which could be used to test the program and allow the ratings to adjust to a suitable level. The program predicted 12 out of 14 of these correctly, so the season was approached with some confidence.

In fact the computer had an excellent year, beating all the paper's ten human tipsters with an average 70.3% correct. For the first time a comparison of the program's margin tipping accuracy with humans was possible, as the now rival paper carried margin tips for some nine celebrities and 12 experts. The average margin of error for the celebrities was 37.3 points, and for the experts 36.7 points, and only one of the celebrities and one of the experts had a lower average margin of error than the computer's 35.4 points. Clarke (1992a) has a detailed comparison of the computer's and the human tipsters' performances and also explores reasons why the computer performed better than humans. Interestingly, tipsters performed worst for the team they knew most about - virtually all tipsters selected the teams they supported more often than the team won. In tipping margins, most tipsters avoided margins close to zero, producing a distinctly bi-modal distribution of forecasts, in contrast to the normal distribution of actual results.

Although the computer performed better than human tipsters in 1991, one might ask why the new 'improved version' performed worse in terms of percentage of correct winners and average margin of error in 1991 than the original version in 1981? The answer lies in the changing face of league football - one facet due to off-the-ground action by administrators, another due to changing tactics on the ground. To increase crowds, the AFL have been attempting to make the competition more even. In particular, salary caps and drafting of players were introduced to try and reduce the gap between the weak and strong clubs. Their success can be judged by the relative performance of the bottom sides in 1981 and 1991. In 1981, the bottom four sides, in the main, only won matches against each other and, in fact the bottom two sides won one and two matches out of 22. This made it easy to select matches in which these teams participated, and in fact the computer selected the bottom four teams correctly, 21, 20, 19 and 19 times out of 22, or an average of 90%. Because the top teams consistently beat the lower teams a team winning 13 matches out of 22 only finished seventh out of 12 teams. By contrast, in 1991, the bottom two sides had three and four wins respectively, including victories over the top and third team. The computer only managed to get the bottom four teams correct 19, 17, 12, and 15 times, an average of 72%. The evenness of the competition was illustrated by the fact that thirteen wins out of 22 matches was now enough to finish fifth out of 15 teams. Because of the evenness of the competition, picking winners was much more difficult in 1991 than a decade earlier.

Paradoxically, this evenness over the season between teams was accompanied by a one-sidedness on the field in individual matches. Over the decade, the way football was played had steadily changed. Play had become much faster, and teams had become more attacking and less defensive. This had resulted in larger scores and larger margins. In 1981, the median winning margin was 31 points with an upper quartile of 52. By 1991 these had risen to 36 and 58. At the upper end, in 1981 there were 11 matches over the 75 point margin, while in 1991 there were 23 in this category (of these, 18 involved an interstate team, an effect not even present in 1981). Margins were thus becoming more difficult to predict accurately. The Australian rules football tipster is faced with a situation similar to Olympic athletes - as distances thrown or jumped increase, they need to find better methods and improve performance just to keep the same relative position. Here, despite the new, more difficult circumstances, the revisions to the forecasting method have allowed the computer to perform better relative to human tipsters than before.

However, even if the new formulation had showed no improvement over the old for questions they both answer, it is still worthwhile as it allows the computer to answer a new class of questions. Home ground advantage is often discussed among football followers. To maximise crowds, the league sometimes shifts popular matches to large capacity grounds. This often causes uproar from the fans, due to a perceived loss of a home ground advantage. Stefani & Clarke (1991) give details of individual home ground advantage over a period of a decade. The methods used there compare a team's home performance with their away performance against the same team. In the AFL, where home and away matches are not balanced, this results in excluding many matches from the analysis. There is also no attempt to measure a team's performance on grounds other than its home ground. However, as a by-product of the computer tips, the new formulation gives the h_{ik}, a measure of a team's performance on all grounds. These team/ground effects are of interest to supporters, and can be used to indicate to administrators possible advantages and disadvantages in shifting matches, or in scheduling finals matches on certain grounds.

4.4. Possible applications

One disappointment in the project has been that the final ladder predictions have received little publication, and the results of the computer simulation and finals program were never reported. If the computer can match or out-perform humans in the relatively straight-forward task of selecting winners, it should perform even better when complications such as differing match schedules come into play. However, there is little chance of testing this hypothesis as the publication of even the human predictions of these events is rare. It seems that such predictions when made are only to promote discussion or controversy, rather than any real attempt to forecast the outcomes. However, with the introduction of betting on ladder positions such a model could be useful in assisting punters or bookmakers. The drawing power of games depends on the closeness of the ladder position of the teams. Probability estimates of the likely ladder positions of teams on the day of the matches could be used as input into computer or human estimates of crowds to assist in the forward planning of match requirements.

Another interesting possibility is the use of the model to obtain more powerful tests of statistical hypothesis. For example, in 1991 there was an odd number of teams, which required a bye to be introduced for the first time. It was noticed that teams often lost the week after the bye - and the journalists quickly dubbed it the killer bye. In fact serious consideration was given to introducing a 'bye effect' into the prediction equation. However the teams that lost may have been playing better teams the week after the bye. For example, Russel (1980) discusses the preparation of a draw to minimise carry-over effects, which arose from a football draw in which one team played another team's previous opponent 18 out of 21 weeks. However, we could test for a bye carry-over effect by comparing how teams performed following the bye with the computer's prediction. Because the prediction takes into account team strength, ground advantage etc, a more sensitive test should result. Similar methods could be used to quantify the effect of key players, weather, night performance etc. Even if a formal test could not be derived, a simple non parametric argument could convince a supporter or administrator of the existence or absence of such effects.

4.5. Conclusions

In terms of a consultancy project the study has proved quite successful. Because the computer is predicting events only a few hours away, its performance is often judged harshly by supporters. An objective analysis has shown the computer's performance in predicting the winner and margins is at least as good as the human expert. However this forecasting project is rather unusual in that the success of the project is probably judged by the client by the publicity generated rather than the accuracy of the forecasts.

Australian rules football shares with other football codes the high degree of passion and subjectivity supporters bring to the game. Most commentators have previous club affiliations, and it is difficult to obtain objective opinions on football matters. Over several years, the relatively simple computer algorithm described has provided winners and margins with at least the accuracy of human experts. In 1991, in addition to their major writer, the client newspaper had seven extra human tipsters and one computer tip. However, all the experts share much the same information. Morrison & Schmittlein (1991), using the notion of equivalent number of independent experts developed by Clemen & Winkler (1985) show that ten experts whose forecasts show a correlation of 0.6 are equivalent to only 1.56 independent forecasts. The computer uses only the previous match results. It does not read the papers all the tipsters read, does not hear the rumours all the experts hear, does not peruse the team selections as all the experts do. As such, it is likely to be more independent than the experts, and the single computer tip may provide more extra information to followers than the many additional human experts. Computer forecasts of sporting events provide an interesting, objective and useful alternative to the human expert.

4.6. Commentary. Current state of play

Since publication of the above paper, the computer predictions have gone from strength to strength. Publication continued in The Age up to 1995. Midway through that season Channel 7 in Adelaide began broadcasting the tips each week on their current affairs program Today/Tonight, This has continued through 1997. In addition the Adelaide advertiser published the tips in 1996 as a celebrity tip in conjunction with their tipping competition.

However the major change has occurred in Melbourne. In 1996, because of the success in The Age the previous year, the tips were again invited back to the Sun. Instead of the bare tips being published, they were accompanied by an explanatory write up highlighting particular aspects. Also, for the first time the estimates of HA and probability of winning were published each week, as well as regular final ladder and simulation predictions. The computer's predictions are used to contribute to the debate on other football topics, such as movement of grounds and the make up of finals. This demonstrates an increasing awareness in the general public of the importance of such effects, and a desire for more quantitative information. In 1997 the tips moved to The Australian Financial Review (a national paper), and the internet. The computer has proved successful, both in terms of correct predictions and publicity generated. It has made the transition from an interesting oddity to being seen as providing objective and accurate 'value added' information.

COMPUTER AND HUMAN TIPPING OF AFL FOOTBALL - A COMPARISON OF 1991 RESULTS

5.0. Abstract

For over a decade the author has been involved in computer tipping of VFL and now AFL football. Evidence suggests that the computer, although ignoring much information available to human tipsters, is at least as accurate. This paper explores the difficulty of predicting, analyses the accuracy of the computer in 1991, compares the relative accuracy of human and computer tipping in 1991, and investigates some reasons for limiting human performance.

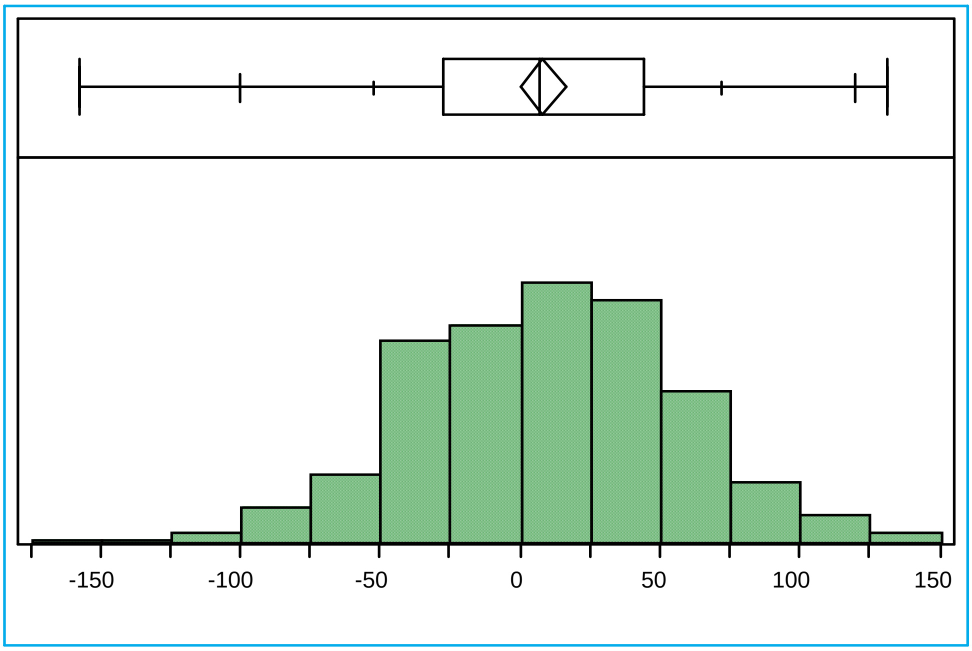
5.1. Introduction

In 1981 The Sun News Pictorial began publishing the results of a computer tipping program written by the author. This continued until 1986, when The Sun decided to concentrate on human tipsters. Some details of this period are contained in Clarke (1981, 1988c). In 1991 The Age published the now updated computer program tips for winners and margins along with the predictions of winners by several experts. The Sun meanwhile published both the predicted winners and margins for 12 experts and 12 celebrities. This allows an opportunity to compare the accuracy of the computer with those of so called experts, and the general public.

Details of computer methods for tipping football are contained in Clarke (1988a), Harville (1980), Stefani (1977, 1980, 1987), Stefani & Clarke (1992). The program discussed here uses an exponential smoothing algorithm, to produce team ratings and team/ground interaction factors for each team. Of relevance to the present paper is that the algorithm uses only the names of the teams playing, the ground the match is played on, and the previous final results of the matches. It ignores all other data, many of which the average and expert follower believe is important. The computer knows nothing of such things as team personnel (absence of key players), weather, time of day (e.g. night matches), previous team played (e.g. bye), time since last match, etc. One would therefore expect the humans to out-perform the computer.

5.2. Distribution of margins

Before looking at how the computer has performed, it is worth looking at how difficult the task has become. Figure 5.1 shows the home ground margins for home and away matches in 1991. The distribution of scores is reasonably symmetric. The mean home ground advantage for the (nominal) home teams is 8.3 points. Note the large spread of scores - standard deviation of over 50 points. Stefani & Clarke (1991) show that prediction of winners in football has become more difficult in the latter half of the eighties. In terms of margins this is even more apparent. A comparison of 1980 and 1991 absolute margins is shown in Figure 5.2. Clearly the proportion of large winning margins has increased. Most percentiles have increased by 10 to 20%, with both the mean and median margins increasing by over seven points.



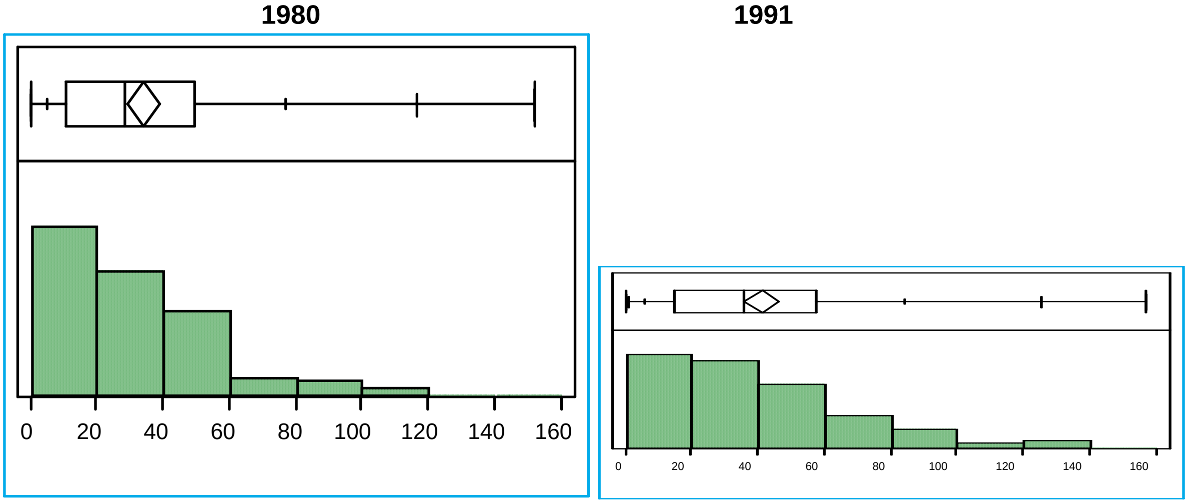
Quantiles

maximum	100.0%	131.00
	90.0%	72.40
quartile	75.0%	44.00
median	50.0%	7.00
quartile	25.0%	-27.50
	10.0%	-52.00
minimum	0.0%	-157.00

Moments

Mean	8.3697
Std Dev	51.6123

Figure 5.1. Distribution of home team winning margins in 1991



Quantiles

	1980			1991	
maximum	100.0%	152.00	maximum	100.0%	157.00
	97.5%	116.70		97.5%	125.55
	90.0%	77.00		90.0%	84.80
quartile	75.0%	49.75	quartile	75.0%	58.00
median	50.0%	29.00	median	50.0%	36.00
quartile	25.0%	11.00	quartile	25.0%	15.00
	10.0%	5.00		10.0%	6.00
minimum	0.0%	0.00	minimum	0.0%	0.00

Moments

	1980		1991
Mean	34.2500	Mean	41.2667
Std Dev	29.4474	Std Dev	32.0021
N	132.0000	N	165.0000

Figure 5.2. Comparison of absolute margins in 1980 and 1991

Selecting matches with the greatest margins gives a possible reason for the change. The matches with the greatest winning margins (over 75 points) are shown in Table 5.1. Eighteen out of 21 of these matches involve an interstate team - an effect entirely absent when the author started tipping. (In addition, the round 21 match was actually played in Tasmania).

TABLE 5.1. Matches resulting in a margin greater than 75 points

Round	Home team	Away team	Result
1	Adel	Haw	86
1	WC	Melb	79
2	Haw	Syd	91
2	Fitz	Melb	-131
4	Bris	Geel	-102
6	Fitz	Haw	-157
7	Haw	WC	-82
7	St.K	Adel	131
8	Fitz	Syd	-77
9	WC	Fitz	99
11	Geel	Adel	84
13	WC	Foot	118
13	Haw	Bris	87
14	Coll	Syd	99
15	Coll	Adel	123
15	Syd	Melb	-83
17	WC	Coll	81
19	Geel	Bris	101
20	Bris	Coll	-101
21	Haw	Fitz	126
23	Carl	Haw	-96
23	St.K	Bris	120
24	Ess	Haw	-80

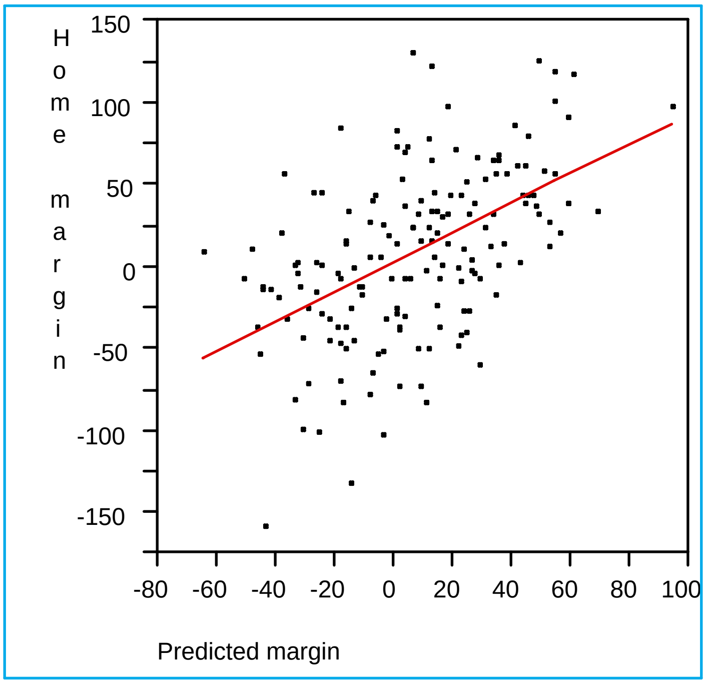
5.3. Prediction accuracy

5.3.1. Winners

In 1991 the computer correctly selected 116 winners out of 165 home and away matches, and five out of seven finals. At just over 70% correct this is slightly better than the decade average for a computer tip reported in Stefani & Clarke (1991).

5.3.2. Margins

Figure 5.3 shows the relationship between the predicted and the actual margin. The fit accounts for about 25% of the variation. Given that the prediction takes account of team ability, current form and ground advantage there is still a large degree of unexplained or random variation. Computer predictions, because they are predicting the expected score, will never have the variation shown by the actual values. Figure 5.4 demonstrates this, but also gives an idea of the spread of results for predictions in given ranges.



Summary of Linear Fit

Rsquare	.2589047
Root Mean Square Error	44.59554
Mean of Response	8.357575

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	2.1448566	3.56804	0.60	0.5486
Predicted margin	.89372157	.118433	7.55	0.0000

Figure 5.3. Actual margin versus predicted margin

We now look at the distribution of errors, defined as the difference between forecast and actual home ground margin. Figure 5.5 shows the distribution of errors. Note that the mean error is still slightly negative although not significantly so, and the median error is -5.00. This implies that the HA is possibly not large enough - the computer may still be adjusting to interstate teams and their large HA. The table shows the median absolute error is 30, with a mean of 36. Thus half the time the computer is less than five goals out.

Predicted home margin **Actual Home margin**

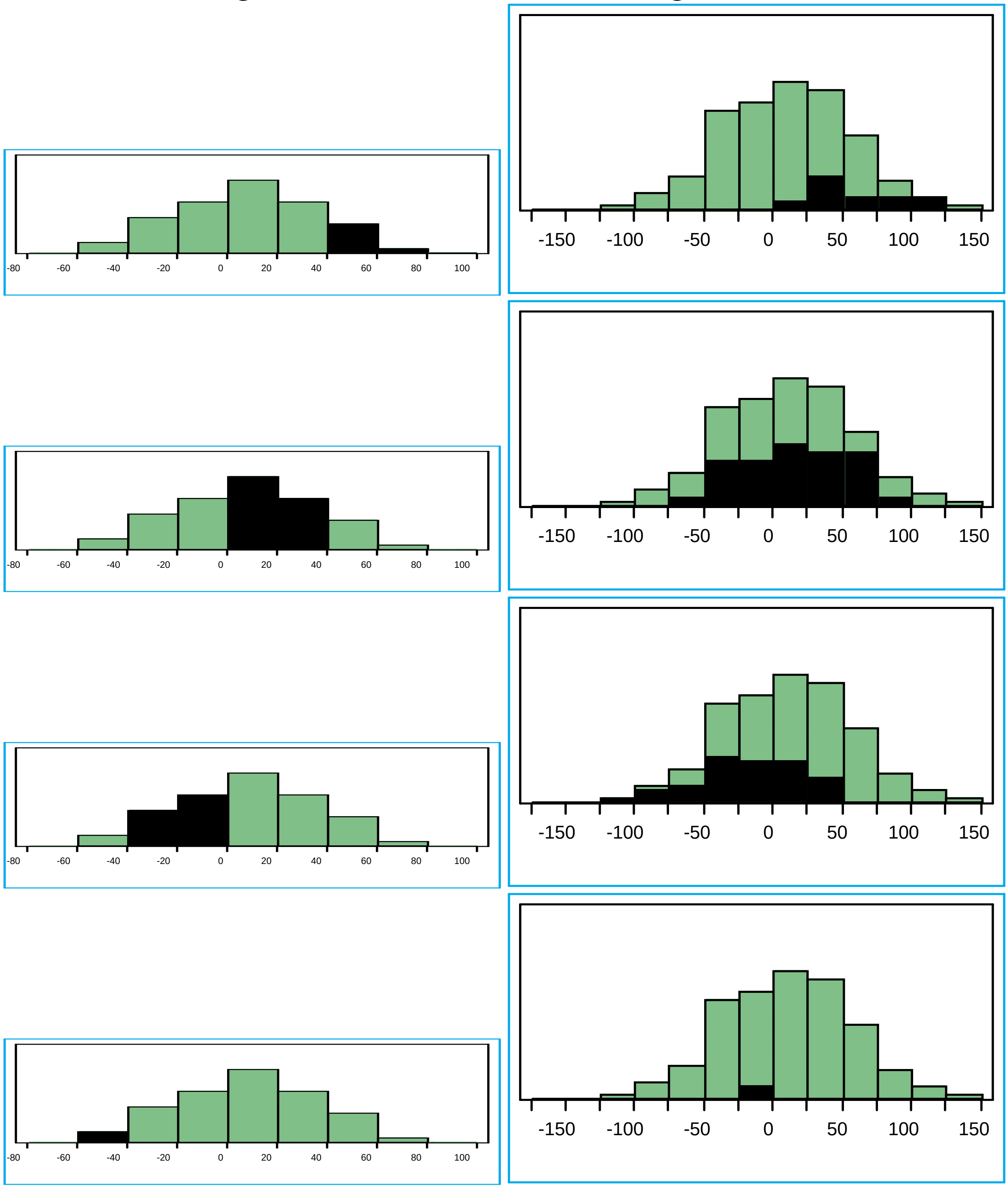
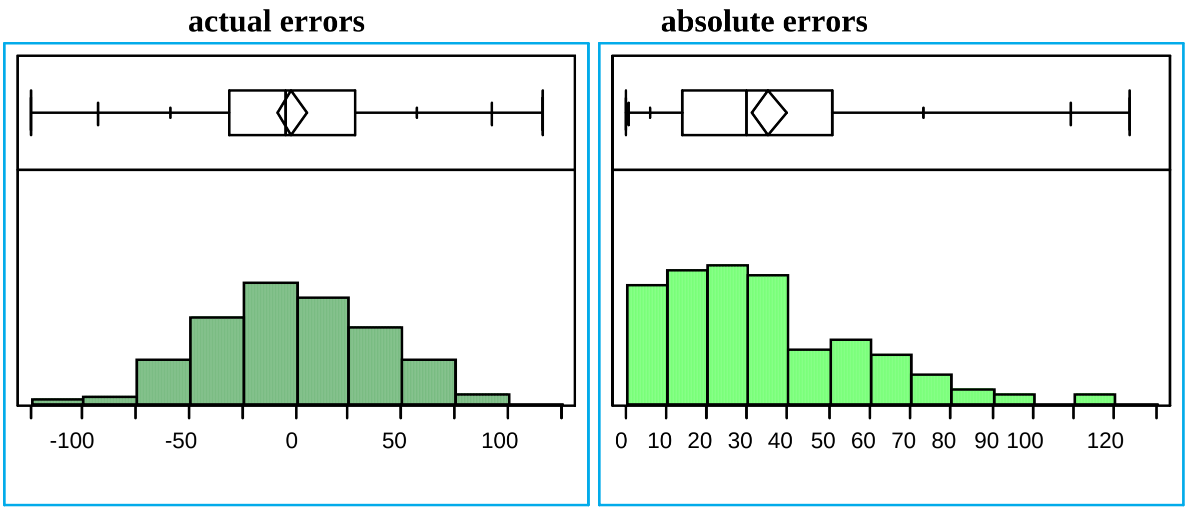


Figure 5.4. Distribution of actual margins for ranges of predicted margins



Quantiles

maximum	100.0%	117.00	maximum	100.0%	124.00
	90.0%	57.40		90.0%	73.40
quartile	75.0%	29.00	quartile	75.0%	51.00
median	50.0%	-5.00	median	50.0%	30.00
quartile	25.0%	-30.50	quartile	25.0%	14.00
	10.0%	-59.00		10.0%	6.00
minimum	0.0%	-124.00	minimum	0.0%	0.00

Moments

Mean	-1.4061	Mean	35.4424
Std Dev	44.5691	Std Dev	26.9177
N	165.0000	N	165.0000

Figure 5.5. Distribution of errors

5.3.3. Final ladder predictions

Although not usually published, the computer also predicts in each round the final ladder at the end of the home and away season. Given the intricacies of the draw, this is one area where the computer should have advantages over human tipsters. Unfortunately, expert predictions of final ladder position are usually only published at the beginning of the season. Figure 5.6 shows the final ladder predictions before each of the 24 rounds. The teams are in order of actual finishing position. The computer clearly has more trouble with the middle of the ladder rather than the very top and bottom. Defining a prediction to be close if within one of the true final position, the final row shows the steady improvement through the season. After 4 rounds over half the teams are predicted closely, and by round 17 about 12 out of 15 teams are closely predicted.

Because ladder position can alter drastically due to just one game, it is also worth looking at predicted final premiership points. Again, if we look at a close prediction as within four premiership points (one game), the final row shows that from round 16 onwards the computer has closely predicted the final ladder position of almost all the teams.

```
           |
           |
     15 +     s s s s f f f f f f f f f f f f f f f f f f f f B
 P          |
 o   14 +     R R f f B B B B B B B B B B B B B B B B B B B B f
 s          |
 i   13 +     f f R B R R s R R R R s s R s R R s s R R R R R R
 t          |
 i   12 +     B B B R s s R s s s s R R s R s s R R s s s s s s
 o          |
 n   11 +     G G F F F F F F F F F C C F F A F F c c F c A c c
            |
     10 +     F F G G S N N A A A C F F C C F A c F F A A c F F
            |
      9 +     S M S S N A S G G C G A A c A M c A A A c F F A A
            |
      8 +     A S A M A S A N c c A c c A M C M M C C C M C N N
            |
      7 +     W c E A c c c c N G c G N M c c C C M M N C N C C
            |
      6 +     M E M N G G G S C N N N G N N N N E N E E E M M E
            |
      5 +     c W c c M M C H S S S M M G G E E N E S M S E E M
            |
      4 +     E A N E C H M C M E M S H H H G G S S N S N S S S
            |
      3 +     N N W W E C E M H M E H S E E H H H H H H G G G G
            |
      2 +     C H C C W E H E E H H E E S S S S G G G G H H H H
            |
      1 +     H C H H H W W W W W W W W W W W W W W W W W W W W
            |
            ---+---------+---------+---------+---------+---------+--
 ROUND        0         5         10        15        20        25

number >1        1   1
position away  9 1 9 1 6 7 6 7 3 6 5 5 5 8 8 6 4 2 3 3 2 2 2 0 0

number >1      1 1   1 1   1
game away      0 1 9 1 0 9 0 9 5 6 3 6 5 5 7 3 4 4 1 2 1 1 0 0 0
```

Chart symbol is the first letter of the team name

Figure 5.6. Predicted final position by round

5.4. Comparison with human tipsters

Table 5.2 shows the number of correct winners and percentage correct for all the tipsters in The Age and The Sun. In some cases (such as leader of the Opposition) selections from different people have been combined. Draws are counted as half correct. For the home and away matches the computer correctly selected 116 winners out of 165 matches, a success rate of 70.3%. Of The Age tipsters, the nearest to this was Ron Carter with 111 or 67.3%. Only two of The Sun experts, and one of the celebrities beat the computer, with another two celebrities choosing the same number of winners. In interpreting a table such as this, it should be borne in mind that in selecting 165 matches, each with a probability of success of 0.7, the number of correct choices will have a standard deviation of about 6. As the computer gives its own estimate p_k of the probability of success for the prediction for match k, the mean and variance of the number correct over the season is $\sum p_k$ = 121.7 and $\sum p_k(1-p_k)$. = 29.35, giving a standard deviation of 5.4. Thus by the computer's own estimates it had an unlucky year. (In fact the high value of $\sum p_k$ is probably an indication that the probability estimates need updating. With the general increase in margins as discussed earlier, a predicted win of 20 points (say) implies a lesser chance of winning than it did 10 years ago. Thus the computer is probably over estimating the chance of selected teams winning)[1]. I suspect that differences between commentators in number of winners less than about five are probably insignificant. Nevertheless, the general public don't see it this way, and it is better to be on top of the table than on the bottom.

Table 5.2 also shows the total and average absolute errors of the margin predictions for The Sun tipsters. Only one expert and one celebrity performed better than the computer. (Although perhaps the computer is more intelligent than we give it credit for, and thought it politic to come in just behind the Prime Minister).

---

[1] Commentary: The algorithm has now been updated. Figure 5.5 shows the prediction errors are approximately normal. The computer keeps a record of the standard deviation of the prediction errors, which it then uses to estimate the chance of an incorrect result prediction. Dowe et al (1996) give details of a probabilistic tipping competition, where the probability of winning is selected, and a Gaussian competition, where the mean and variance of the signed margin is selected. The evaluation system they use could provide an alternative measure of the comparative accuracy of the computer's estimates.

TABLE 5.2. Accuracy of The Age and The Sun tipsters

Tipster	Number tipped	Number correct	Percentage correct	Total deviation	Average deviation
Computer	165	116	70.3	5848	35.4
Age experts					
Ron Carter	165	111	67.3		
Greg Baum	110	74	67.3		
Nick Johnson	76	51	67.1		
Gary Linnel	74	49.5	66.9		
Martin Blake	153	102	66.7		
Steve Linnel	102	67.5	66.2		
Len Johnson	156	103	66.0		
Penny Crisp	95	62.5	65.8		
Patrick Smithers	55	36	65.5		
Peter Schwab	7	3.5	50.0		
Sun Experts					
Geoff Poulter	158	115	72.8 *	5476	34.7 *
Ron Reed	158	109	69.0	5702	36.1
Ron Barassi	165	117	70.9 *	5898	35.8
Bruce Matthews	158	109	69.0	5611	35.5
Niall/Pierce	165	113	68.5	6333	38.4
Don Scott	165	111	67.3	6040	36.6
Tony De Bolfo	165	110	66.7	6038	36.6
Daryl Timms	165	109	66.1	6135	37.2
Crackers Keenan	165	107	64.9	5941	36.0
Michael Stevens	165	107	64.9	6170	37.4
Lou Richards	165	103	62.4	6514	39.5
Eva/Atkins/West.	158	101	61.2	5750	36.4
Sun Celebrities					
Joan Kirner	165	118	71.5 *	5909	35.8
Bob Hawke	165	116	70.3	5839	35.4 *
Wynne/Meldrum	165	116	70.3	5943	36.0
David Johnston	165	113	68.5	6111	37.0
John Hewson	165	112	67.9	6019	36.5
Daryl Somers	165	111	67.3	6001	36.4
Mary Delahunty	165	110	66.7	6223	37.3
Steve Vizard	165	104	63.0	6455	39.1
Brown/Kennett	165	98	59.4	6863	41.6

* Better performance than the computer

5.4.1. Reasons for computer supremacy

Figure 5.4 shows that the distribution of the computer margin prediction is roughly the same shape as that of the actual margins, with the same mean but a lesser variance. This is not true of many human tipsters, who often have a distinctly bi-modal distribution of predicted margins. There appears to be an aversion to predicting close margins. In addition, some tipsters tend to choose multiples of 10 or 6 points for the margins. One reason the computer may perform better than experts is that it has no loyalties to particular teams. While no data is available on the teams followed by many of the experts, there is evidence to suggest that tipsters are certainly influenced (to their detriment) by the teams they follow. Figure 5.7 shows a graph of the number of times Lou Richards selected each team and the number of wins for each team. Clearly Lou favours Collingwood, the team he barracks for. This graph is typical of all the celebrities. With the exception of Bob Hawke, all celebrities selected the team they followed more often than they won, the excess ranging from 5 to 9 wins.

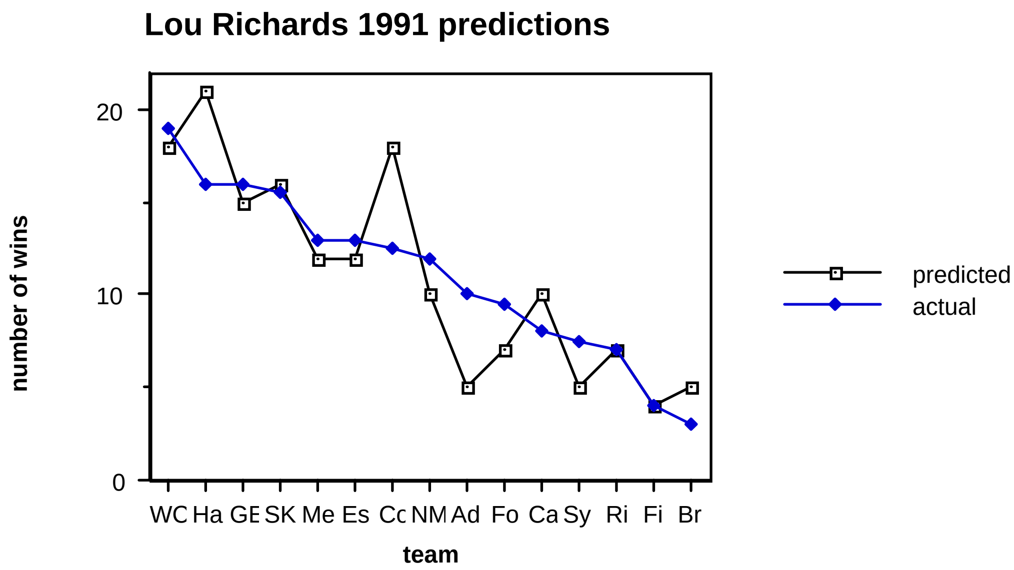


Figure 5.7. Lou Richards' predicted and actual number of wins for each team

It is well known that supporters look for any reason to convince themselves that their team will win next week. Nevertheless it is interesting that football followers predict most poorly the performance of the team they know most about. One reason humans may choose poorly is that they know too much information, and they overrate the importance of much of it. The return of a player from absence due to injury, good training form, a perceived after effect of a bye, etc might also be given too much weight by experts. However all the experts share much the same information. Morrison & Schmittlein (1991), show that 10 experts whose forecasts show a correlation of 0.6 are equivalent to only 1.56 independent forecasts. It would be interesting to look at the correlations between the margin tips of experts, to see if the tips of those with shared information (such as expert tipsters from The Sun), are more closely correlated within groups than between groups.

5.5. Conclusion

An analysis has shown the computer's performance in predicting the winner and margins in 1991 was better than the average expert or football follower. The computer uses only the previous match results and is not influenced by publicity surrounding particular events, nor club loyalties. As such it is likely to be more independent than the experts, and the single computer tip may provide more extra information to followers than the many additional human experts. Computer forecasts of sporting events provide an interesting, objective and useful alternative to the human expert.

Acknowledgments: Some of the data used in this report was collected and computerised by my students during an undergraduate project 'Football Tipping'. My thanks to Cameron Howell, Brad Patterson, Gabriele Sorrentino, Andrew Moar, David Thomas and Graeme Wilson.

5.6. Commentary

While no formal studies have been undertaken since 1991, the computer has generally remained in the upper half of the expert tipsters range. In 1995 the computer was second with 127 winners out of all the expert tipsters in The Sun and The Age, and then selected eight out of nine finals correctly. In 1996 it was again in the top few tipsters with 126 winners, and selected all nine finals correctly. 1997 proved a difficult year for all tipsters, with the computers 108 winners beating about a quarter of the expert tipsters.

PREDICTIONS AND HOME ADVANTAGE FOR AUSTRALIAN RULES FOOTBALL

6.0. Abstract

In a previous paper, it was demonstrated that distinctly different prediction methods when applied to 2435 American college and professional football games resulted in essentially the same fraction of correct selections of the winning team and essentially the same average absolute error for predicting the margin of victory. These results are now extended to 1446 Australian rules football games. Two distinctly different prediction methods are applied. A least-squares method provides a set of ratings. The predicted margin of victory in the next contest is less than the rating difference, corrected for home ground advantage, while a 0.75 power method shrinks the ratings compared with those found by the least-squares technique and then performs predictions based on the rating difference and home-ground advantage. Both methods operate upon past margins of victory corrected for home advantage to obtain the ratings. It is shown that both methods perform similarly, based on the fraction of correct selections of the winning team and the average absolute error for predicting the margin of victory. That is, differing predictors using the same information tend to converge to a limiting level of accuracy. The least squares approach also provides estimates of the accuracy of each prediction. The home advantage is evaluated for all teams collectively and also for individual teams. The data permit comparisons with other sports in other countries. The home team appears to have an advantage (the visiting team has a disadvantage) due to three factors: the visiting team suffers from travel fatigue; crowd intimidation by the home team fans; lack of familiarity with the playing conditions.

6.1. Introduction

A variety of schemes exist to rank athletes or teams so as to predict the outcome of a subsequent competition and for seeding competitors in a tournament. These schemes are either accumulative or adaptive. An accumulative scheme results in the accumulation of points and rankings based on those points. The point total never diminishes and may be subject to some limiting process. Most soccer tables of standings are ranked with two or three points accruing for a win and one point for each draw. The World Cup of Skiing uses an accumulative system with provisions for limiting the total number of points.

An adaptive scheme causes ratings to rise or fall as performance is above or below some predicted level. For example, the World Chess Federation uses the Elo system in which the rating difference between each competitor and the average opponent provides an estimate for the number of victories for each competitor. The competitor's rating changes as a function of the actual number of victories compared with that target.

Adaptive schemes are also used to rate teams in a variety of sports and then to predict the outcome of the next competition. A few schemes use relatively large quantities of offensive and defensive statistics, although most schemes operate only on the margin of victory adjusted for home advantage (HA). The latter approach is especially efficient where a large number of teams are in competition or where a large number of games are played. Stefani (1977, 1980, 1987) applied a least-squares (LS) rating scheme to American college football, American professional football and American college basketball. The method selected the correct winning team 70% of the time when applied to 10000 games over a 10 year period and tends to provide ratings whose differences consistently exceed the actual margin of victory of the next set of opponents. He shrank the rating differences to provide unbiased predictions.

In addition to the LS approach, Stefani (1987) applied three other schemes: weighted LS with a shrinking factor, James-Stein (1981) (which automatically shrinks predictions compared with LS) and another method (Harville's (1980)) that preshrinks ratings. The four methods were applied to a common set of 2435 American college and professional football games with virtually the same accuracy for selection of the correct winning team and nearly the same average absolute error for selection of the margin of victory. The conclusion was that the information content of the margin of victory adjusted for HA appeared to limit the accuracy of all four estimators with respect to the prediction of American college and professional football games.

In order to extend the conclusion, a survey of other sports was performed and it was determined that a database was available for at least 10 seasons of Australian rules football, including 1446 games. It is therefore the intention of this paper to apply two different schemes to Australian rules football to determine whether the schemes also perform similarly while operating only on the margin of victory corrected for HA. One scheme (Stefani's) is the LS method that shrinks predictions compared with the rating difference, while the second method by Clarke (1981, 1988c) uses the 0.75 power of error to pre-shrink ratings compared with the LS so that each prediction depends on the actual rating difference. A by-product of both schemes is HA data that permit comparisons of Australian rules football with other sports.

6.2. Australian rules football

Australian rules football is a high scoring continuous-action game. There are 18 players on each side. The dimensions of the playing surface vary from ground to ground, but the shape is generally oval and longer than for soccer or rugby. There are no offside rules; hence, each player can quickly advance the ball down the field by carrying it himself or by kicking or punching the ball forward to a team mate. If the ball is kicked between the centre goalposts a six point 'goal' is scored and action returns to the centre of the ground. If the ball is kicked between a centre goalpost and one of the two outer goalposts, a one point 'behind' is scored and action resumes from the goal area. A game consists of four 25 minute quarters, during which over 200 points are commonly scored.

Major professional activity in Australia began in 1896 with the formation of the Victorian Football League (VFL), consisting originally of teams in the greater Melbourne area. Because one team moved to Sydney in 1982, and Brisbane and West Coast joined the VFL in 1987, the VFL has been renamed the Australian Football League.

The season is divided into a 22 game home-away schedule followed by a five team, six game ladder play-off, culminating in the grand final game. There are currently 14 teams in the league. The 22 game home-away schedule does not result in an equal number of home away pairs because of the number of teams, ground sharing and the use of a neutral league ground.

6.3. Modelling game results

In the following, the index i represents some reference team, m represents the week of the season and $j(m)$ represents the index of the opponent for team i during week m. The independent variables are i and m since **j** is dependent on the schedule and is completely determined by i and m. Primary interest focuses on data for the winning margin, the number of points scored by a team minus those scored by the opponent and on the HA. **A** model for the winning margin is

$$w_{ij(m)} = h_{ij(m)} + u_i^m - u_{j(m)}^m + e_{ij(m)} \tag{6.1}$$

where $w_{ij(m)}$ represents the winning margin for team i against opponent **j** in week m,

$h_{ij(m)}$ is the HA for team i, u_i^m is the rating for team i using data including week m, $u_{j(m)}^m$ is the similar rating for the opponent and $e_{ij(m)}$ is a zero-mean random error.

6.4. Home advantage

A HA may be found so as to minimise the sum of the squared errors from equation (6.1) such that

$$J = \sum_{i=1}^{N} \sum_{m=1}^{K} e_{ij(m)}^2 \tag{6.2}$$

where there are N teams and K weeks have been completed.

There are three possible generations of HAs: $h_{ij(m)}$ can be interpreted as a single h for all teams, as a distinct h_i for each team or as a distinct value for each combination of teams. Only the first two interpretations are used here. If a single value is to be found, then $h_{ij(m)}$ is interpreted as $+h$ if i plays at home, as $-h$ if i plays away or as 0 if i and j play on a neutral ground or on a ground that both teams use as a home ground.

6.5. Single h for all teams

It follows that the LS value of a single h which minimises equation (6.2) for the M games played at home is

$$h = \frac{1}{M}\left[\sum_{\substack{i=1 \\ i \text{ at home}}}^{N} \sum_{m=1}^{K} w_{ij(m)} + \sum_{\substack{i=1 \\ i \text{ at home}}}^{N} \sum_{m=1}^{K} (u_{j(m)}^m - u_i^m) \right] \tag{6.3}$$

In order to minimise equation (6.2) it is also necessary to select the team ratings, so that the calculation of team ratings and HA is coupled. The value of h can be uncoupled from the calculation of the team ratings by adopting a suboptimal approach in which the right-hand double summation in equation (6.3) is assumed to be zero. During a given week, some home teams contribute positive, negative and zero values to that summation. Over subsequent weeks, these teams shift location. Over subsequent seasons, or perhaps during the same season, most teams will play at home and away against each opponent. Over a large enough collection of games, it is reasonable to assume that the right-hand double summation in equation (6.3) is nearly zero. In that case, the calculation of h can proceed independently from any rating generator, so that h depends only on the average margin of victory by the home team.

Table 6.1 contains HA data for seven types of competition, including Australian rules football. The data include the number of games, the fraction of games won by the home team, the fraction of draws, the home-win-minus-away-winfraction, the HA h expressed in score units (goals, points, etc.) per game, the total score per game and the total-score-to-home-advantage ratio R. The seven types of competition in Table 6.1 are organised by decreasing values of R. The home-win-minus-away-winfraction generally declines as R increases - an intuitive tendency.

TABLE 6.1. Home advantage

Sport	Games	Home win	Draw	HW-AW[a]	Home advantage (h)	Total score (T)	R (T/h)
Soccer: 3 European cups	1079	0.603	0.198	0.404	0.97	2.76	3
Soccer: 6 nations[b]	6601	0.485	0.281	0.251	0.45	2.72	6
Hockey (USA)	2840	0.505	0.166	0.176	0.68	6.76	10
College football (USA)	1669	0.574	0.017	0.165	3.71	43.0	12
Profess. football (USA)	671	0.574	0.003	0.151	3.27	40.7	12
Australian rules football	1109	0.580	0.007	0.167	9.8	206.5	21
Baseball (USA)	2106	0.538	0.000	0.076	0.26	8.9	34

[a] HW, home win; AW, away win.

[b] England, Germany, Italy, Norway, Spain and Switzerland.

The data in Table 6.1 are an updated and expanded version of data collected in Stefani (1983, 1987). The three Europe Cups (Champions Cup, Cup Winners Cup and EUFA Cup) data for international club competition were collected from newspaper results during six seasons (1981-82 and 1985-86 through 1989-90). The six nation soccer data for club competition within each nation was collected from soccer tables of standings for England, Germany, Italy, Norway, Spain and Switzerland during two seasons (1980-81 and 1981-82). Those six nations provide a reasonable cross-section of club competition within a given nation. Hockey data are from National Hockey League (USA/Canada) results for four seasons (1975-76 through 1978-79) and were provided by Cleroux (Univ. of Montreal). American college and professional football data are for four seasons (1979-80 through 1982-83) taken from Stefani (1987). Australian rules football data are from Clarke's database of ten seasons (1980 through 1989). A total of 1446 games were played (1386 during the home-away schedules and 60 during the play-offs) of which 1109 featured a home ground. Finally, the baseball data are for the 1982 American League and National League (USA) season from the 1983 league yearbooks.

Authors such as Pollard (1986b) suggest a number of causes of HA (visiting team disadvantage). These causes may be placed into three groups: physiological factors such as the travel fatigue of the visiting team; psychological factors such as crowd intimidation by the home team fans; tactical factors such as lack of familiarity of the visiting team with the playing conditions. Each competition in Table 6.1 has a specific absolute and relative mix of the three factors. Home advantages can be seen to rise or fall as the absolute amount of the factors change or as one or more factors are absent.

For example, competition between European club soccer teams from different nations for the three European cups appears to have the largest HA. One goal in three is attributable to the HA. The home team wins 40.4% more games than it loses. These cup matches are generally scheduled at mid-week, with league competitions also being played during the weekends. Travel is therefore exceptionally fatiguing. It is an understatement that interest is high and that crowd intimidation is a psychological likelihood given the tragic loss of life that has accompanied European cup competition. Therefore, the greatest absolute incidence of fatigue and crowd intimidation correlates with the largest HA, while a lower amount of those factors for competition within each of the six European nations generates a lower HA (travel is of shorter duration, and followers are of the same nation, hence reducing the absolute amount of crowd intimidation in general, while recognising a few rather intense local rivalries). One goal in six is attributable to the HA and the home team wins 25.1% more games than it loses.

The lowest HA in Table 6.1 occurs for professional baseball games played in the USA. Each team plays 81 games at home and 81 games away. Travel expenses are reduced by playing three or four consecutive games at each location and travelling for extended periods. Travel fatigue may differentially affect the visiting team only for the first game of a series when the home team has not also returned from a trip. The majority of the HA is attributable to crowd intimidation and lack of familiarity with the ground due to various playing surface compositions, the location of walls and wind conditions, although lack of familiarity would diminish in importance with successive games. The HA amounts to one run out of each 34 scored. The home team wins 7.6% more than it loses.

Professional hockey in the USA (including several Canadian teams), American college football, American professional football and Australian rules football comprise the remainder of Table 6.1. The home-win-minus-away-win fraction is nearly the same for each competition, although differences exist in the ratio R. Perhaps the relatively compact nature of a hockey arena generates more intimidation and, therefore, a greater amount of HA as measured by R. Conversely, most Australian rules football teams compete in the greater Melbourne area, reducing travel fatigue and balancing the mix of team supporters. It follows that most of the HA is attributable to lack of familiarity with the playing conditions, since there are a variety of grounds from oval to round-shaped and a variety of ground sizes. The HA in Australian rules football creates a margin of 9.8 points per game for the home team which wins 16.7% more games than it loses. The HA amounts to one point in 21, since 206.5 points per game are scored.

The data in Table 6.1 support the hypothesis that there is a HA due to travel fatigue, crowd intimidation and lack of familiarity with the playing conditions which affect the visiting team negatively. As the absolute amount of each rises, so also does the HA. The absence of one or more factors tends to diminish the HA. No effort is made here to rank the relative importance of the three factors.

6.6. Distinct h_i for each team

If a HA is to be found for each team, then $h_{ij(m)}$ in equation (6.1) is interpreted as h_i when i is at home, as $-h_j$ when j is at home and as 0 when the teams play on a neutral ground or on a ground that both teams normally consider to be a home ground. To facilitate an LS value which minimises equation (6.2) independently of any rating scheme, the results are considered only when two teams have a home away pair during the home-away season. For simplicity, the value of m is not shown and the ratings are assumed to be the same after each match. The two results may be added so that

$$w_{ij} = h_i + u_i - u_j + e_{ij} \tag{6.4a}$$

$$w_{ji} = h_j + u_j - u_i + e_{ji} \tag{6.4b}$$

$$d_{ij} = h_i + h_j + e1_{ij} \tag{6.4c}$$

where $d_{ij} = w_{ij} + w_{ji}$, i.e. is the sum of the home team win margins which is the home–away differential for each team, and $e1_{ij}$ is the sum of the two errors. For example, if team i wins by five points at home and then team j wins the return match by seven points at team j's home ground, then d_{ij} is 12 from the perspective of either team, i.e. each team was 12 points more successful at home. The information contained in equation (6.4c) can be collected over an entire season and then an algorithm can be used to estimate the HA of each team.

Using this method, a HA was found for each of the teams currently competing in the Australian Football League. Table 6.2 contains the average of the HAs from the end of each season weighted by the number of home games that could be paired with a return match. The notes section identifies teams in the greater Melbourne area and two groups of teams currently sharing the same ground. The South Melbourne team moved to Sydney at the beginning of 1982. For simplicity, the two years of South Melbourne results are combined and listed with Sydney's eight years of results. West Coast and Brisbane joined the league in 1987; hence, each played during only three of the 10 years.

TABLE 6.2. Home advantage by team for the 1980s

Rank	Team	Home advantage	Home games	Notes[a]
1	West Coast	36.8	25	
2	Footscray	19.9	73	M
3	Carlton	16.3	58	M, S1
4	Essendon	13.7	72	M
5	Sydney	13.1	97	
6	St Kilda	12.5	74	M
7	Fitzroy	10.8	64	M, S1
8	Melbourne	9.9	63	M, S2
9	Brisbane	8.9	23	
10	Geelong	5.5	77	
11	Hawthorn	3.7	63	M, S1
12	Collingwood	0.3	62	M
13	North Melb.	-1.0	66	M,S2
14	Richmond	-1.6	63	M, S2
	All teams	9.8	880	

[a] M, Greater Melbourne; S, shared home field.

About 80% (880) of the 1109 home games could be paired. The West Coast Eagles exhibited the largest HA. Since that team plays in Perth, about 3000 km (1900 miles) by air from Melbourne, travel fatigue and crowd intimidation would affect the visiting team more than at matches in Melbourne, and the relatively large HA is logical. Conversely, Melbourne, North Melbourne and Richmond share the Melbourne Cricket Ground (MCG). The relatively minimal **HAs** may be explained by the minimal visiting-team travel (except for Brisbane, Sydney and West Coast), by the fact that the MCG is too spacious to generate as much crowd intimidation as would be possible on a smaller field and by the fact that a large enough number of games are played at the MCG so that the field configuration is well known.

6.7. Rating systems

Home advantage can be removed from the winning margin modelled in equation (6.1) by defining an adjusted winning margin $wa_{ij(m)}$ such that

$$wa_{ij(m)} = w_{ij(m)} - h_{ij(m)} \tag{6.5}$$

Two rating systems are considered, one based on LS and one based on exponential smoothing using the 0.75 power of error. The LS rating for team **i** which minimises equation (6.2) by smoothing over all K weeks of data is

$$u_i^K = \frac{1}{n(i)} \sum_{m=1}^{K} \left[wa_{ij(m)} + u_{j(m)}^m \right] \tag{6.6}$$

where team **i** has played $n(i)$ games during the K weeks of the season over which the ratings are computed. The recursive equivalent for equation (6.6) is

$$u_i^K = \frac{1}{n(i)} \left\{ [n(i) - 1] u_i^{K-1} + wa_{ij(K)} + u_{j(K)}^K \right\} \tag{6.7}$$

A similar equation can be written for team **j** by reversing the indices **i** and **j.** There are two unknowns in the two equations: the ratings u_i^K and $u_{j(K)}^K$. These equations can be solved simultaneously to yield the LS recursive algorithm for u_i^K.

$$u_i^K = u_i^{K-1} + \frac{n(j)-1}{n(i)n(j)-1} \left\{ wa_{ij(K)} - \left[u_i^{K-1} - u_{j(K)}^{K-1} \right] \right\} \tag{6.8}$$

The adaptive term in braces in equation (6.8) may be limited to two standard deviations to reduce large rating changes resulting from anomalous results.

Unfortunately, the differences between the LS ratings tend to exceed future margins of victory, so that a model for the margin of victory in the next game (to be played during week $K+1$) is

$$w_{ij(K+1)} = h_{ij(K+1)} + L\left[u_i^K - u_{j(K+1)}^K\right] + e_{ij(K+1)} \tag{6.9}$$

where L is a shrinking factor between zero and one which reduces the predicted winning margin compared with the rating difference when the rating difference is consistently more than the subsequent winning margin. The sum of squared errors in equation (6.9) can be minimised by selecting L as

$$L = \frac{\text{sample covariance}\left[wa_{ij(K+1)}, (u_i^K - u_{j(K+1)}^K)\right]}{\text{sample variance}\left[u_i^K - u_{j(K+1)}^K\right]} \tag{6.10}$$

then, using previously calculated L, the predicted winning margin for the next game becomes

$$\hat{w}_{ij(K+1)} = h_{ij(K+1)} + L\left[u_i^K - u_{j(K+1)}^K\right] \tag{6.11}$$

In summary, each LS rating is computed by equation (6.8) using previously calculated HAs and then equation (6.11) is used for prediction using previously calculated L.

A substantially different method suggested by Clarke exponentially smooths past results to obtain ratings which are closer together than LS ratings. Clarke used a learning set of previous Australian rules football games. By trial-and-error he applied a 0.75 power to the adaptive term in equation (6.8) and he found that the result was an effective pre-shrinking of ratings whose differences are then used for prediction so that L is unity in equation (6.11). Again by trial-and-error, Clarke determined that the result should be multiplied by 0.2 to smooth it exponentially. The rating for team i becomes

$$u_i^K = u_i^{K-1} + 0.2\left\{\left|wa_{ij(K)}\right|^{0.75}\text{sign}(.) - \left|u_i^{K-1} - u_{j(K)}^{K-1}\right|\text{sign}(.)\right\} \tag{6.12}$$

where $|.|$ denotes an absolute value operator which avoids taking the root of a negative number and sign(.) restores the sign.

The LS and 0.75 power schemes were both applied to predict the 1446 Australian rules football games played during the entire decade of the 1980s. The results are summarised in Table 6.3. The table shows ***h*,** L, the fraction of games correctly predicted (a draw is considered half correct and half incorrect) and the average absolute error between the true and predicted winning margin.

Both methods used a common home advantage h for all teams. Table 6.3 shows the value of h used for prediction and the actual value calculated at the end of each season. The value for h at the end of the 1979 season was used for the 1980 predictions; then the average value of h from the end of the 1979 and 1980 seasons was used for the 1981 predictions. A five-year moving average was used when more than five previous values of h were available. The value of L used for prediction by the LS method was calculated similarly to the averaging of past values of *h*. **A** value of L was calculated for the 0.75 power method but predictions assumed L was unity.

The fraction of games correctly predicted by the LS method was higher by 0.5% (seven games out of 1446), while the 0.75 power method had a lower average absolute error by 0.7 points per game (32.2 compared with 32.9). Neither difference was statistically significant. The LS method required an average L of 0.66, while the 0.75 power method exhibited an L of 0.94, indicating that the preshrunk rating difference was an unbiased predictor of the next game.

A comparison of different prediction schemes was also made in Stefani (1987). The LS method predicted the correct winning team in 69.8% of 2435 American college and professional football games with an average absolute error of 12.16 points per game. Other predictors such as James-Stein were also used, resulting in nearly the same results as for the LS method. For example, the James-Stein predictor selected the correct winning team in 69.8% of the games with an average absolute error of 12.17 points per game.

In summary, the 1446 Australian rules football predictions verify the conclusion that the information content of the margin of victory and HA limits the accuracy of predictions, so that additional accuracy most likely requires the use of additional data. That is, dissimilar rating methods using the same data tend to converge to a level of accuracy limited by the data and not necessarily by the structure of the algorithms.

Compared with human predictors in both Australia and the USA, the computer predictions perform at the better-than-average to expert level.

TABLE 6.3. Least squares and 0.75 power predictions for Australian rules football, 1980-1989

Year	Games	h used	h computed	L used	L computed	Proportion correct	Average absolute error
Least squares							
1980	138	8	1.5	0.60	0.78	0.700	29.8
1981	138	5	10.1	0.69	0.79	0.732	27.5
1982	138	7	12.6	0.72	0.73	0.700	32.9
1983	138	8	8.9	0.73	0.62	0.667	34.1
1984	138	8	9.4	0.70	0.48	0.667	32.4
1985	138	9	5.7	0.68	0.58	0.678	34.7
1986	138	9	11.7	0.63	0.81	0.681	35.6
1987	160	9	14.2	0.64	0.62	0.675	36.5
1988	160	9	10.0	0.62	0.54	0.641	32.9
1989	160	10	13.7	0.60	0.67	0.678	31.8
	1446	8.2	9.8	0.66	0.66	0.681	32.9
0.75 power predictions							
1980	138	8	1.5	1.00	0.90	0.692	29.3
1981	138	5	10.1	1.00	1.06	0.754	27.0
1982	138	7	12.6	1.00	0.97	0.670	32.6
1983	138	8	8.9	1.00	0.78	0.667	35.1
1984	138	8	9.4	1.00	0.78	0.644	31.3
1985	138	9	5.7	1.00	1.12	0.634	33.2
1986	138	9	11.7	1.00	1.00	0.659	34.9
1987	160	9	14.2	1.00	0.96	0.731	35.1
1988	160	9	10.0	1.00	0.80	0.659	32.2
1989	160	10	13.7	1.00	1.04	0.647	30.8
	1446	8.2	9.8	1.00	0.94	0.676	32.2

6.8. Estimating the accuracy of the predictions

The popular print media and many gambling establishments often estimate that Team A is a 3:1 favourite to defeat Team B; that is, Team A is 75% likely to defeat Team B. A gambling establishment provides an estimate to divide the money bet in such a way that the gambling establishment should show a profit whether or not Team A wins. In order to use the predictions of equation (6.11) to estimate the probability that a team should win, it is necessary to know the probability density for equation (6.11). In Stefani (1980,1987) a Gaussian assumption was shown to provide accurate estimates for American college football. The Gaussian assumption also provides accurate estimates for Australian rules football.

Under the assumption that each rating is an unbiased estimate of the true rating, the probability that team **i** will actually win the next game becomes the probability that the estimated margin of victory of equation (6.11) is greater or equal to zero, since equation (6.11) would then be an unbiased estimate of the true margin of victory. An estimate of the standard deviation of equation (6.11) facilitates estimation of that probability. Since the variance of the rating for team **i** can be estimated by

$$S_i^2 = \text{sample variance}\left[w a_{ij(K+1)} + u_{j(K)}^K\right] \tag{6.13}$$

then the standard deviation of a prediction using equation (6.11) can be estimated by

$$S_{\hat{w}} = L\left[S_i^2 + S_{j(k+1)}^2\right]^{0.5} \tag{6.14}$$

Assuming that the distribution of equation (6.11) is Gaussian, the probability that team **i** will win is 0.5 plus the additional integrated area due to the ratio $\hat{w}/S_{\hat{w}}$.

The predicted accuracies and actual accuracies for the 1446 predictions are gathered into five ranges of predicted accuracy in Table 6.4. The actual accuracies agree closely with the predicted accuracies in each of the five ranges as well as overall. The overall accuracy was predicted to be 0.670, while the actual overall accuracy was 0.681. It is noteworthy that 51 games were predicted to have more than a 0.9 probability of being correct. The actual result was that 49 were predicted correctly, one game was drawn and one game was incorrectly predicted.

TABLE 6.4. Predicted and actual accuracies for Australian Rules football, 1980-1989

Predicted accuracy	Games	Actual correct	Actual wrong	Actual accuracy
0.50 - 0.59	474	256.5	217.5	0.541
0.60 - 0.69	428	283.5	144.5	0.662
0.70 - 0.79	314	235.5	78.5	0.750
0.80 - 0.89	179	159.5	19.5	0.891
0.90 - 0.99	51	49.5	1.5	0.971
0.670	1446	984.5	461.5	0.681

6.9. Conclusions

It appears that the accuracy of a prediction depends primarily upon the information content of the data used to construct the ratings and much less on the algorithm used to compute the ratings, assuming that each algorithm is properly applied. That is, differing predictors using the same data tend to converge to a limiting level of accuracy.

Home advantage appears to depend on the three negative influences upon the visiting team: travel fatigue (a physiological effect); intimidation by home team fans (a psychological effect); lack of familiarity with the playing conditions (a tactical effect). The absolute influence of these three factors varies from sport to sport.

6.10. Commentary. Comparison of Clarke's Models 1 and 2

The exponential smoothing algorithm used above is a slightly simplified version of the one described in Chapter III and used regularly for computer tipping. For example it uses a common HA. How do program 1 and 2 as described in Chapter III compare? While complete records of all tips actually published have not been kept, it is possible to regenerate the tips by using the data base of matches described in Chapter II. In this case we are not using the results of pre-season matches as was usually the case for the actual tip, and the final matches have not been included. However the results would not differ markedly from those published. Table 6.5 gives the proportion correct each year, for both models described above. Note there is some variation from year to year. Some years are more predictable than others, but the average correct for both models is about 68%. Note the average errors for Program 1 are generally positive, which indicates the HA is not large enough. Program 2, with its automatic adjustment of ground effects performs better in this regard. Program 2 is also slightly more accurate in margin prediction, with an average absolute error nearly 1 point better than Program 1. This is perhaps not surprising as it was initially optimised on absolute margin of error. However generally there is not a lot of difference between the two methods. These results are in accord with a comparison between the methods of Stefani and Clarke.

TABLE 6.5. Proportion correct and average absolute error for Clarke's two prediction programs

Year	No. of Games	Program 1			Program 2		
		Proportion Correct	Average Error	Average absolute error	Proportion Correct	Average Error	Average absolute error
80	138	0.772	-3.9	29.6	0.699	-3.0	29.1
81	138	0.754	3.6	27.5	0.754	3.0	27.5
82	138	0.692	4.1	33.4	0.678	3.1	32.7
83	138	0.645	2.7	36.3	0.659	1.0	35.6
84	138	0.652	1.9	32.6	0.652	0.0	31.1
85	138	0.656	-1.4	35.5	0.692	-0.7	33.4
86	138	0.659	6.6	36.4	0.659	6.6	35.4
87	160	0.719	9.8	36.1	0.694	7.8	35.5
88	160	0.653	3.7	34.1	0.653	2.1	33.1
89	160	0.647	6.3	32.0	0.647	3.6	31.0
90	161	0.643	4.1	35.2	0.668	1.1	35.0
91	172	0.715	3.2	36.2	0.709	-0.3	35.4
92	172	0.599	2.2	35.9	0.640	-0.7	34.7
93	157	0.685	2.1	34.0	0.678	0.2	32.5
94	174	0.670	5.6	33.9	0.652	3.8	33.3
95	185	0.684	0.0	34.3	0.705	-1.6	33.2
80-95	2467	0.677	3.2	34.0	0.677	1.6	33.1

CHAPTER VII

CALCULATING PREMIERSHIP ODDS BY COMPUTER: AN ANALYSIS OF THE AFL FINAL EIGHT PLAY-OFF SYSTEM

7.0. Abstract

The Australian Football League's final eight play-off system is explained and the premiership chances of all teams are evaluated and compared with previous final systems. The importance of matches, and the playing order necessary to avoid dead finals is discussed. A computer tipping program published weekly in a daily newspaper produces the probability of any team beating any other team on any ground. This was extended to calculate the chances of every possible finishing order in the finals. The suitability and fairness of the play-offs is evaluated under various criteria. The use of such a program to assist in framing premiership and quinella odds is discussed, and inconsistencies between bookmakers' odds on individual matches and winning the premiership highlighted. Home ground advantage in finals is assessed. It is shown that the knockout nature of the finals magnifies any home advantage possessed by a team.

Key ***words:*** sport, football, finals, play-off, probabilities, odds

7.1. Introduction

Australian rules football is the major winter sport of the southern states of Australia. From a base of 12 Melbourne clubs, by 1994 the Australian Football League (AFL) had expanded to 15 teams including teams from Sydney, Brisbane, Adelaide and Perth. The season consists of a 22 game home and away schedule followed by a final series between the top teams culminating in a grand final to determine the premier team. For many years the author has operated a computer program to predict the results of individual games Clarke (1988c, 1991a, 1993b). The predictions are published each week in the daily press and Clarke (1992a) shows that they compare favourably with human tipsters. In addition to selecting the winner of any match, the computer program estimates the chances of either team winning. Near the end of the 1992 home and away series a program was written that would produce, at the beginning and during the final series, each team's chance of winning the grand final. In 1994 this needed updating for the new final eight system introduced by the AFL.

Since 1931 the football finals were played under the 'Page final four' system. Unlike the normal knockout system of three matches (two semi-finals and a final) used in most competitions with four finalists, the Page final four gives an advantage to the top teams by introducing a bye and a double chance. The top two teams play, with the winner gaining direct entry into the grand final by virtue of a bye the following week; the loser is not eliminated but has a second chance by playing the winner of the lower two teams for the right to play in the grand final. This became known as 'the double chance'. Thus the top two teams are advantaged by either gaining the bye or receiving the double chance. Over the years, several variations on this theme were used as the number of teams in the finals was increased. In 1972 the league introduced a final five played under the 'McIntyre Final Five' system (see Schwertman & Howard, (1989, 1990)) and in 1991 the AFL introduced a new finals system played between the top six teams. After some criticism, they adjusted the system again for 1992 with the 'McIntyre Final Six' system. As Clarke (1993b) points out, this introduced a controversial aspect into the finals as the path a team takes, and hence its chances of winning the premiership, is determined by the results of matches in which it does not participate. This aspect was entrenched when, for 1994, the AFL introduced the final eight, with a series of nine finals matches over four weeks, called the McIntyre Final Eight system (MF8).

Monahan & Berger (1977), in discussing the fairness of play-off structures in hockey suggest three criteria for measuring their suitability: maximise the probability that the highest ranked team wins, maximise the expected number of points of the premier, and maximise the chance the best two teams meet in the final. In their conclusions they point out that in some of their proposed structures, a lower-ranked team has a higher chance of reaching the semi-final and final than a higher-ranked team, and this is unacceptable to players who are rewarded according to final position. In such competitions (the AFL prize money and the following year's draft are affected by final position), the play-off system should not be judged solely on how fairly it determines the premier, but how fairly it determines all positions. This suggests several other criteria - the probability of a team finishing in any position or higher should be greater than for any lower-ranked team; the expected final position should be in order of original ranking; the probability of a team finishing above a team of lower rank should be greater than 0.5 and should increase as the difference in ranks increases; the probability of any two teams appearing in the grand final should monotonically decrease as the ranks of the two teams increases. To these we might add one of consistency between years - teams of the same ranking that perform in a similar manner in different years should finish in similar positions.

7.2. The McIntyre Final Eight system

The MF8 consists of nine matches: four qualifying finals the first week, two semi-finals the second week, two preliminary finals the third week and the grand final the fourth week. For the first three weeks, two teams are eliminated each week. The original ladder ranking before the finals determines the draw for the first week of the finals, and also determines the relative position of the winners and losers each week.

During week 1, four Qualifying finals are played: 1v8, 2v7, 3v6 and 4v5. This produces four winners who go to the top of the ladder and four losers who go to the bottom. Within these two groups they preserve their original ranking. Thus after week 1, we have a new ranking of Winner 1, Winner 2, Winner 3, Winner 4, Loser 1, Loser 2, Loser 3, Loser 4, although in the following we will usually use the terminology of current ranking one to eight. An example will illustrate. In 1994 the qualifying finals were 1v8 West Coast v Collingwood, 2v7 Carlton v Melbourne, 3v6 North Melbourne v Hawthorn and 4v5 Geelong v Footscray.

MF8 week 1 matches from 1994 and resulting week 2 draw

Original ladder ranking	Team	Week 1 result	Current ranking	Team	Week 2 match
1	West Coast	win	1- Winner 1	West Coast	Bye
2	Carlton	loss	2- Winner 2	North Melb	Bye
3	North Melb	win	3- Winner 3	Geelong	Semi 2
4	Geelong	win	4- Winner 4	Melbourne	Semi 1
5	Footscray	loss	5- Loser 1	Carlton	Semi 2
6	Hawthorn	loss	6- Loser 2	Footscray	Semi 1
7	Melbourne	win	7- Loser 3	Hawthorn	Eliminated
8	Collingwood	loss	8- Loser 4	Collingwood	Eliminated

The bottom two teams are eliminated, and from week 2 onwards the system is a knockout tournament with the current top two teams gaining a bye in week 2. Under the old final four, in week 2 one team has a bye straight through to the final in week 3, while two other teams play to see which of them continues. This system is the same but in two halves (teams currently ranked 1, 4 & 6 in one half and teams 2, 3 & 5 in the other). Teams 1 and 2 get a bye straight through to the two preliminary finals in week 3, while 4 & 6 and 3 & 5 play in the two semi finals to determine who joins them. In week 4 the two winners of the preliminary finals play in the grand final.

A feature of the MF8 is the degree to which a team's progress depends on results in other matches. One year team 3 could win the first week and not gain the bye, while another year team 6 could win and gain the bye. Similarly team 3 could lose and be eliminated one year whereas another year team 6 could lose but not be eliminated. Under the assumption that all teams are equal, the chances of teams 1 to 8 gaining the bye if they win are respectively 100%, 100%, 75%, 50%, 50%, 25%, 0% and 0%. The same list reversed gives the chances of the teams being eliminated if they lose. This is the first finals system where a team's elimination has depended on other match results.

Note that the positions after the first week are symmetrical. If after the first week a team is now ranked position N, their opponent from the first week will now be in position $9N$. Put another way, a team's opponent will be as far off the bottom as the team is from the top. It is easily proved using this symmetry that there can be no repeat finals matches until the grand final since respective opponents from week I go into opposite halves of the draw.

The symmetry also means that the first round opponents of the teams eliminated gain the bye. This implies these two matches were very important to the participants - the winner gained the bye, the loser was eliminated. However it is not known beforehand which of the matches are elimination matches as they are determined by the results in other matches. Thus in the most perverse cases, if both 1 and 2 lose, the 3v6 and 4v5 matches become the elimination matches with 3 and 4 eliminated if they lose and 5 and 6 gaining the bye. On the other hand if 1 and 2 both win, these matches are virtually irrelevant, as the winner cannot gain the bye nor the loser be eliminated. The result merely determines in which semi-finals the four teams will play.

7.3. Premiership chances - comparison with previous systems

For the case when all teams are considered equal, the chances of winning the premiership can be calculated easily by first principles. For teams which make the grand final the chances of winning the premiership are 50.0%, from the preliminary finals 25.0%, and from the semi-finals 12.5%.

The chances at the beginning of the finals can now be calculated as weighted averages of these probabilities. For example teams 7 & 8 have a 50% chance of making the semi-finals, hence a 6.25% chance of the flag. Teams 1 & 2 have a 50% chance of making the preliminary final directly and a 50% chance of making the semi-final, to give a 18.725% chance of being premier. The others can be calculated in a similar way and are given in Table 7.1 along with probabilities for all previous final systems.

Note the importance of gaining the bye. **A** team doubles their chance of winning by getting direct access to the preliminary final. For teams 3 to 6 this depends as much on the outcomes of other matches as on their own. The saying 'you make your own luck' cannot be said to apply to the MF8. Other teams make it for you. Consider team 3. Before the finals they have a 15.6% chance of winning the flag. Suppose they win their qualifying final. If 1 and 2 both win they now have a 12.5% chance; less than before. On the other hand if 1 or 2 lose, the chances of team 3 increase to 25%. Similarly teams 4 and 5 do not increase their chances by winning the qualifying final if they do not make the preliminary final direct. That a team's chances could alter so dramatically from year to year, depending on the result of a third party, may be considered by some to be a flaw in the system.

TABLE 7.1. Premiership chances for MF8 and previous final systems

Team	Final 4	Final 5	Final 6 (1)	Final 6 (2)	MF8
1	37.50	37.50	25.00	25.00	18.750
2	37.50	25.00	25.00	25.00	18.750
3	12.50	25.00	18.75	18.75	15.625
4	12.50	6.25	18.75	12.50	12.500
5		6.25	6.25	12.50	12.500
6			6.25	6.25	9.375
7					6.250
8					6.250

With each new system, the chances of teams 1 and 2 have been steadily eroded until they are now exactly half of that under the final four. Team 3's chances doubled with the introduction of the final five, but have since been eroded although they are still greater under the final eight than under the final four. Thus even though the number of finalists has doubled, team 3's chances have increased. Team 4's chances have enjoyed a roller coaster ride, but have settled on exactly the same probability as for the final four. Team 6's chances have increased as have teams 7 and 8. In economic terms, we have seen a great redistribution of probability from the rich top order to the poor lower order, with the middle class largely unaffected. Teams 7 and 8 now have as much chance of winning as 4 and 5 had under the final five system. However one should note that this is only 1 in 16, actually less than they had at the start of the season (1 in 15 if there are 15 equal teams). Clearly the AFL have not been interested in maximising the chance of the highest-ranked team winning, but they have produced a system in which a team's chances increase steadily with their ranking.

One consequence of the diminution of the top team's chances is that the league should consider recognising the team who finishes at the top of the ladder before the finals, perhaps with a minor premiership cup, since their chances of turning that position into a premiership is now much smaller than under earlier final systems.

7.4. Importance of matches

Football supporters know the grand final is the most important match of the year. It would be desirable if finals matches built up in importance, but how can we quantify this notion of importance? Morris (1977) defines the importance of a point in tennis as the difference in the probability of winning the match if a player wins the point and the probability of winning the match if a player loses the point. Thus the grand final is the most important match at 100%, with the preliminary final at 50%. The calculations are shown below.

Grand final	100%-0%	=	100%
Preliminary final	50% -0%	=	50%
Semi-final	25% -0%	=	25%
Qualifying final between 1&8 (or 2&7)			
For team 1,2	25% - 12.5%	=	12.5%
For team 8,7	12.5% - 0%	=	12.5%
Qualifying final between 4&5 (or 3&6)			
Depends if 1 and 2 lose	25%- 0%	=	25%
or if 1 and 2 win	12.5%-12.5%	=	0%
or if 1 and 2 win and team 1 plays interstate			negative?

In general the matches are in order of increasing importance. However as we have said, the qualifying finals also have importance to other teams. The definition of importance needs extending to take this into account, perhaps to the total expected absolute change in probability of all the teams. When this is done the importance of the qualifying finals increases. Note that a win by the lower-ranked teams in the matches 1v8, 2v7 and 3v6 is good for the winner of the other qualifying finals and bad for the loser - so it makes those matches much more important. For this reason, the importance of the qualifying finals 3v6 and 4v5 is more difficult to calculate, as it depends on the results of other finals.

One special case is worth discussing. In order to maximise crowd attendance and television coverage, the finals are played at different times over a weekend. Thus it is possible the league could schedule the match between 4 & 5 (or 3 & 6) after the other qualifying matches. If the other qualifying matches have both gone to the higher-ranking team, then these matches would be of zero importance, since the winner cannot make the bye and the loser cannot be eliminated. The only factor hinging on the match is which half of the draw the teams go into. We could even have the situation where one or even both teams are trying to lose, to avoid the half of the draw containing specific teams. For example, two interstate teams may already be in the half of the draw into which the winner will go. It is highly likely matches against these interstate clubs would be scheduled on their home grounds, a large disadvantage to any Melbourne team drawn to play them. A 4v5 qualifying final between two Melbourne clubs could see the winning team having to play two finals interstate to make the grand final - with the loser having the easier draw of two games in Melbourne. To avoid this situation the qualifying finals need to be played in a certain order. The match 4v5 must be played in the first two finals, and the match 3v6 in the first three finals. The problem is equivalent to ordering the digits 1, 2, 3, and 4 so that no digit is preceded by two or more lower numbers. The possible orders are shown in Table 7.2. In 1994, the AFL chose the second last shown. Note that this does not stop a final being 'dead' in retrospect. In 1994 the 4v5 clash between Geelong and Footscray, won by the final kick of the match, was in fact in this category. Neither team was eliminated, so in fact the result did not matter.

TABLE 7.2. Possible playing order of matches to avoid 'dead' finals

4v5	3v6	1v8	2v7
4v5	1v8	3v6	2v7
4v5	3v6	2v7	1v8
4v5	2v7	3v6	1v8
3v6	4v5	1v8	2v7
1v8	4v5	3v6	2v7
3v6	4v5	2v7	1v8
2v7	4v5	3v6	1v8

7.5. Development of program using a word processor

Schwertman & Howard (1989, 1990) look at a probability model for the AFL Finals series as it was played up to 1990 - a series of six games between the top five teams. They list the four paths that result in the fourth team winning the grand final, and the 16 paths that result in the second team winning. For the top team they say "Direct computation of the probability that team A wins the grand final is quite involved, with many different paths" and suggest indirect methods. The MF8 is a series of nine matches with the extra complication that the position of a winning team now depends not just on their match but on the results of other matches. Clarke (1993a) gives a method for determining all possible finishing orders for the final six, and that is extended here.

For the MF8, we wish to calculate not only the chance of each team winning the grand final, but also some other probabilities of interest such as the chance of pairs of teams making the grand final and the chance of each team finishing in any position. All the probabilities would follow from the chance of all possible finishing orders. So the problem was: "Given the original order before the finals, what is the probability of any final finishing order?" Although in studying the fairness of the finals system it is of interest to assume all teams are of equal ability, for the computer tip we had different probabilities for any team beating any other. Furthermore, because the computer tipping program takes grounds into account, the probabilities changed from week to week depending on the grounds at which the matches were scheduled.

Suppose we designate each team by their finishing position at the end of the home and away matches. Using the actual results from 1994, we have a sequence of matches as follows, where positions and teams are: I West Coast, 2 Carlton, 3 North Melbourne, 4 Geelong, 5 Footscray, 6 Hawthorn, 7 Melbourne, 8 Collingwood.

Week 1	Week 2	Week 3	Week 4
1-8 (1 wins)	4-2 (4 wins)	3-4 (4 wins)	1-4 (1 wins)
2-7 (7 wins)	5-7 (7 wins)	1-7 (1 wins)	
3-6 (3 wins)			
4-5 (4 Wins)			

The final finishing order produced by this particular sequence of results and its associated probability is:

$$14372568 \quad p1(1,8)p1(7,2)p1(3,6)p1(4,5)p2(4,2)p2(7,5)p3(4,3)p3(1,7)p4(1,4) \tag{7.1}$$

where $pk(i,j)$ is the probability of team i beating team **j** in week k. There is no obvious pattern between the final order and the probabilities that produce that order. It would be tedious in the extreme to work out all possible $2^9 = 512$ sequences and their probabilities. The fact that positions of teams depend not just on the results of their matches but on the results of others, further complicates matters. However a word processor comes to our aid. **A** modern word processor allows the copying and movement of columns of text as well as rows of contiguous text. Using this facility, the 512 lines of code that formed the crux of the program were written in about 1 hour.

The method involves keeping on the left side the 'current' order and on the right side the probabilities of that order arising. Each match result has a certain probability and produces an associated change in the order. Before the final series the order is 1,2,3,4,5,6,7,8 so we have:

Order	Probability
1,2,3,4,5,6,7,8	$p1()$

Consider the match between 4 and 5 This can have two results, so we copy the whole row. Now if 4 beats 5 the order stays the same, so we leave the first row alone, but if 5 beats 4, 5 moves to 4th and 4 moves to 5th, so we do this to the second row. This gives us:

1,2,3,4,5,6,7,8	$p1()$
1,2,3,5,4,6,7,8	$p1()$

We now want on the right side the probabilities of these results - i.e. $p1(4,5)$ in the first row and $p1(5,4)$ in the second. But these are just the numbers in the middle two columns of the left side. Thus we can use the column copy and insert facility on the word processor to copy them across. This gives:

1,2,3,4,5,6,7,8	$p1(4,5)$
1,2,3,5,4,6,7,8	$p1(5,4)$

We have added the *pk*()s for ease of reading, but in practice, these were all inserted at the end of the process. We repeat the procedure for the remaining matches. Each match iteration results in a doubling of the number of rows with a duplication of the whole table, the movement of columns on the left side, and the addition of another set of pks by copying parts of columns from the left to the right of the table. With column copy and insert and global replace it was about an hour's work to produce 512 lines similar to (7.1) and convert to code. In this case SAS was used, but other packages such as Excel could be utilised. **A** SAS data set with two variables, order and probability was created, and the above lines of code produced 512 observations. Programming and SAS procedures could then be used to calculate any required probabilities.

In the author's original problem, the normal weekly computer tipping program was used to provide probabilities for each team beating any other, and the above program was used both before and during the final series to predict estimated probabilities of teams winning the flag or finishing the year in different positions. These predictions were included with the usual ones of winners and margins. The program thus served its original purpose.

However the program can also be used to investigate the degree to which the MF8 satisfies the criteria for fairness discussed earlier.

7.6. How fair is the MF8?

Schwertman & Howard (1990) suggest several suitable models to investigate the fairness of finals systems. Here we assume that all teams are equal. The qualifying finals in 1994 showed this is not an unreasonable model, with team 8 losing to team 1 by two points, team 6 and team **3** drawn at full time, team 4 beating team 5 with the last kick of the match and team 7 beating team 2 comfortably. Tables 7.3, 7.4, 7.5 and 7.6 give some output from the program that demonstrates the MF8 performs well on the fairness criteria. The chance of winning and the expected final position (EFP) are in order of original ranking. In fact the chance of finishing in position **j** or higher is in monotonic order of original ranking for every **j**. The chance of team *i* finishing above team *j* is generally in increasing order of **j** for every i, with a couple of exceptions for teams widely separated in the rankings.

TABLE 7.3. Percentage chance of teams finishing in any position, and expected final position (EFP) - Equal probability model

Team	Final position								EFP
	1	2	3	4	5	6	7	8	
1	18.75	18.75	37.50	0.00	25.00	0.00	0.00	0.00	2.94
2	18.75	18.75	23.44	14.06	18.75	6.25	0.00	0.00	3.14
3	15.63	15.63	18.75	12.50	18.75	6.25	12.50	0.00	3.72
4	12.50	12.50	9.38	15.63	18.75	6.25	18.75	6.25	4.41
5	12.50	12.50	6.25	18.75	6.25	18.75	18.75	6.25	4.56
6	9.38	9.38	3.13	15.63	6.25	18.75	25.00	12.50	5.19
7	6.25	6.25	1.56	10.94	6.25	18.75	25.00	25.00	5.86
8	6.25	6.25	0.00	12.50	0.00	25.00	0.00	50.00	6.19

TABLE 7.4. Percentage chance of teams finishing in any position or higher - Equal probability model

Team	Final position							
	1	2	3	4	5	6	7	8
1	18.75	37.50	75.00	75.00	100.00	100.00	100.00	100.00
2	18.75	37.50	60.94	75.00	93.75	100.00	100.00	100.00
3	15.63	31.26	50.00	62.50	81.25	87.50	100.00	100.00
4	12.50	25.00	34.38	50.00	68.75	75.00	93.75	100.00
5	12.50	25.00	31.25	50.00	56.25	75.00	93.75	100.00
6	9.38	18.75	21.88	37.50	43.75	62.50	87.50	100.00
7	6.25	12.50	14.06	25.00	31.25	50.00	75.00	100.00
8	6.25	12.50	12.50	25.00	25.00	50.00	50.00	100.00

TABLE 7.5. Percentage chance of team i (row) finishing above team **j** (column) - Equal probability model

Team i	Team j							
	1	2	3	4	5	6	7	8
1	00.00	60.16	65.23	70.51	70.51	75.78	81.25	82.81
2	39.84	00.00	62.89	70.51	70.51	78.13	82.81	81.25
3	34.77	37.11	00.00	66.80	66.80	70.31	76.56	75.78
4	29.49	29.49	33.20	00.00	57.81	65.23	72.07	72.07
5	29.49	29.49	33.20	42.19	00.00	65.23	72.07	72.07
6	24.22	21.88	29.69	34.77	34.77	00.00	67.58	68.36
7	18.75	17.19	23.44	27.93	27.93	32.42	00.00	66.41
8	17.19	18.75	24.22	27.93	27.93	31.64	33.59	00.00

The chance of grand finals between the teams in various positions is shown in Table 7.6, and is roughly in accord with the sum of the teams' rankings. **A** grand final between 1 & 2 is the most likely result, although it only has about a 1 in 8 chance of occurring, compared with 1 in 2 under the old final four system. The chance of a grand final between 2 & 3 is relatively low because of the high probability they will end up in the same half of the draw. Note also that two grand finals are impossible - between 2 & 8 and 1 & 7. However this did not stop the National Sportsbook from offering odds of 100-1 on a Collingwood-Carlton grand final, and 50-1 on a West Coast-Melbourne grand final in the week preceding the qualifying finals in 1994.

TABLE 7.6. Percentage chance of pairs of teams playing in grand final - Equal probability model

Team	2	3	4	5	6	7	8
1	14.1	8.6	5.1	5.1	1.6		3.1
2		3.9	5.1	5.1	6.3	3.1	
3			5.5	5.5	3.1	3.1	1.6
4				3.1	2.3	2.0	2.0
5					2.3	2.0	2.0
6						0.8	2.3
7							1.6

7.7. Comparison of bookmakers' and computer's odds

It is interesting to see the extent to which bookmakers' odds reflect the intricacies of the finals draw. The head to head odds for the first match along with the premiership team and quinella (the two teams that play in the grand final) odds offered by the National Sportsbook as published in The Melbourne Herald Sun, September 10, 1994 are given in Tables 7.7-7.9. Odds of a/b are converted to percentage chances as $100*b/(b+a)$. As these usually sum to more than 100 due to the bookmaker's percentage, adjusted chances which are proportional but sum to 100 are shown. For comparison the head to head chances as estimated by the computer tipping program and the consequent premiership and quinella chances are also given.

It is clear the Sportsbook odds often do not reflect the intricacies of the MF8. Although Sportsbook give North Melbourne a greater chance of winning the first match (63% as against 56% by the computer), they are given less chance (16% as against 19%) of winning the premiership. This apparently underestimates their chance of gaining the double chance. Geelong is treated in the same way. Although Melbourne is given less chance of winning the first match than Hawthorn, they have the same chance of winning the flag, completely discounting Hawthorn's possible bye or double chance.

TABLE 7.7. Comparative chances of winning first match

	Computer	Sportsbook		
Team	chances	Odds	Chances	Adjusted
WC	80.00	1/5	83.33	76.16
Carl	55.70	4/9	69.23	63.27
NthM	55.50	4/9	69.23	63.27
Geel	56.50	1/2	66.67	60.93
Foot	43.50	11/8	42.10	38.48
Haw	44.50	11/8	42.10	38.48
Melb	44.30	614	40.00	36.56
Coll	20.00	3/1	25.00	22.85

TABLE 7.8. Comparative premiership chances

Team	Computer chances	Sportsbook Odds	Sportsbook Chances	Sportsbook Adjusted
WC	29.80	7-4	30.76	25.53
Carl	24.51	9-4	36.36	30.18
NthM	19.41	4-1	20.00	16.60
Geel	11.52	7-1	12.50	10.38
Foot	7.15	14-1	6.67	5.54
Haw	4.06	16-1	5.88	4.88
Melb	2.61	16-1	5.88	4.88
Coll	0.94	40-1	2.43	2.02
Total	100.00		120.18	100.00

TABLE 7.9. Computer chances and Sportsbook odds on quinellas

Team	WC	Carl	NthM	Geel	Foot	Haw	Melb	Coll
WC		29.02	19.38	9.59	6.57	1.92		1.22
Carl	7-4		2.59	3.80	2.64	3.37	1.72	
NthM	7-2	10-1		4.24	3.12	2.22	1.75	0.34
Geel	10-1	8-1	8-1		2.07	0.95	1.04	0.30
Foot	16-1	16-1	25-1	50-1		0.64	0.64	0.19
Haw	20-1	20-1	33-1	40-1	100-1		0.32	0.22
Melb	50-1	50-1	50-1	66-1	200-1	200-1		0.14
Coll	80-1	100-1	100-1	125-1	200-1	200-1	250-1	0.14

The quinellas also show inconsistencies. An obvious case is the Carlton-Collingwood and West Coast-Melbourne quinellas. These are both impossible, yet are not only given odds but are shown at shorter odds than many possible quinellas. The Carlton-North quinella is also misquoted. Tables 7.6 and 7.9 show this to be quite unlikely, due to the high probability of team **3** and 4 ending up in the same half. A more detailed analysis shows that team 2 and **3** will only end up in opposite halves if in week 1, team 1 wins and 2 and 3 both lose, or team 1 loses and 2 and **3** both win. Using the adjusted Sportsbook odds on this occurring from Table 7.7 we obtain a probability after week 1 of only 0.2 that a Carlton-North grand final will still be possible. Either two or four matches will still have to fall the correct way for the grand final to eventuate. The odds of

10-1 are thus extremely poor and overestimate the chance of this particular grand final. In a similar way, the extra difficulty of obtaining a West Coast-Hawthorn quinella over a West Coast-Carlton quinella that is shown in both Table 7.6 and Table 7.9 is not reflected in the Sportsbook odds.

7.8. Home advantage

There are two large capacity Melbourne grounds on which finals have been traditionally played, the MCG and Waverley Park. However, finals in the first three weeks involving interstate teams are likely to be played at their home venue. Since the MCG and Waverley are shared as home grounds by several Melbourne teams, there is a considerable home advantage (HA) element in the finals. Stefani & Clarke (1991, 1992) have calculated home ground advantages in home and away matches in AFL football. The computer program developed here can also be used to quantify HA in the finals. By altering the venues for matches in a particular week, the effect on a team's premiership chances can be evaluated. Table 7.10 shows the final program's estimated chances of winning the premiership using the weekly computer tipping program's probabilities of winning each match under various venue assumptions. The first column shows the program's calculated premiership chances if all matches are played at the MCG, the second with matches involving West Coast moved to West Coast the first three weeks, and the third with all matches in the first three weeks played at Waverley. In general this shows a multiplier effect if teams play several matches at home. Thus a team that wins 10% more matches at home may win 30% more premierships if all matches are played at home.

The effects are quite dramatic. West Coast's chances almost double by having the preliminary matches at home, although their average chance of winning at the West coast is 'only' 47% higher than at the MCG. By moving all preliminary matches to Waverley, Hawthorn's chance increases by 31% although their average chance of winning at Waverley is only 11% higher than the MCG. Similarly Melbourne's chance would reduce to 67% although their average chance of winning at Waverley is 89% of their chance at the MCG. While one could argue about the magnitude of individual home effect, the general point is that, because of the knockout nature of the finals series, individual home ground effects are magnified when considering the chances of winning the premiership.

TABLE 7.10. Premiership percentage chances if matches in the first three weeks at different grounds

Team	All matches at the MCG	West Coast at West Coast, others at MCG	All matches at Waverley Park
WC	16.83	29.74	17.50
Carl	24.71	24.43	27.79
NthM	22.15	20.03	20.23
Geel	14.39	11.37	10.58
Foot	9.03	7.09	10.96
Haw	5.77	3.79	7.55
Melb	5.25	2.61	3.53
Coll	1.87	0.94	1.87

7.9. Conclusion

In football, subjective judgements are often used to rate team chances of winning the premiership. These often tend to reflect the relative strengths of the teams, and ignore the current ladder position. With the complicated structure now in place for the finals series, a mathematical analysis using a simple model can shed light on the chance of teams winning or finishing in any position, given their original or current ranking in the final series. One aspect of mathematics is recognition of patterns. In this case, there was no obvious pattern between the final order and the probabilities that produced that order. However, there was a pattern in the way these orders and probabilities were built up when individual matches were considered. The functions of a word processor could be used to exploit this pattern to write the required equations and subsequent computer code. The code generated was flexible enough to handle many different models.

It is important to investigate how a sport's draw operates, rather than complaining when a specific unforseen case arises. Teams cannot be blamed when they 'exploit' weaknesses in the rules of a competition that organisers have allowed to creep in. There are many criteria that a final series should satisfy. For many competitions the relative chances of teams finishing in any position, not just first, should be considered. The MF8 passes most of the tests given here. The higher a team finishes at the end of the home and away matches, the greater their chances of being the premier team, the greater their chances of finishing in any position or higher, the higher their final expected position, and the greater their chances of finishing above lower teams. However two major flaws exist. The first is that a team's chances depend so much on matches in which it does not participate, which results in a lack of consistency from year to year. There seems little, other than a new system, that the AFL could do to redress this. The other major weakness that may need to be addressed is the possible lack of importance of the qualifying finals between 4 & 5 and 3 & 6. It seems a great pity that a final would degenerate into a 'dead' match, or worse still a farce where both teams were trying to lose. The AFL must always schedule these matches early so the chance of the winner making the preliminary final direct or to avoid elimination exists to give the teams incentive. Of course matches will still often be dead in retrospect, as occurred in 1994. Because of the greatly reduced chances of a team winning the flag from top position, the league should consider recognising the leader after the home and away matches by a minor premier cup.

It is also clear that a computer program, such as detailed here, could be useful in assisting with framing the odds for premierships and quinellas under a system as complicated as the MF8.

7.10. Commentary. Increasing influence of HA in finals

Since this paper was written the home ground advantage in finals has continued to have a strong influence. As interstate teams grow in number and strength, more finals include a home ground factor. In 1996, due to the presence of Essendon and North Melbourne (both if which have the MCG as home ground), and three interstate sides, each of the nine finals included exactly one team playing on their home ground. All nine finals were won by the team with the HA.

7.11. Commentary. Quantifying the effect of AFL decisions on the home and away draw

All sports are affected by the overall rules of the competition. The previous paper quantified the effects of the various final systems. The models discussed in the thesis can also be used to quantify AFL decisions that affect the home and away draw. The League along with individual clubs makes many decisions affecting the running of the competition. These are often based on financial aspects such as to maximise crowds or television exposure, but they also affect teams' chances of success in the competition. They range from relatively minor changes such as moving the venue of a single match or moving the home ground of a club for an entire season, through to decisions having major ramifications such as organising an unbalanced draw. What effect do these have on a team's chances? In the past these have not been quantified. The remainder of this chapter shows how some of these aspects can be investigated.

7.11.1. Change of venue

By using the estimated ground effects developed by the computer tipping program, the effects of changing venues on the chance of teams winning can be evaluated. The AFL often move particular matches. This may be to allow for anticipated large crowds, or the poor state of the surface of a particular ground or for other reasons. Sometimes it is done with the approval of the affected clubs, but often against their wishes. Clubs will often cite the loss of their HA as a reason against the move, but never is this quantified. It now can be.

For example, suppose it is mooted that the 1996 round 14 Footscray-Melbourne match be moved from the MCG to Footscray. In fact several of Melbourne's home matches have been moved, to leave the MCG free for matches expected to draw large crowds. As at round 11, 1996, the computer rated Melbourne at 52.2 and Footscray 56.2. However the ground effect for Melbourne at the MCG was 1.8, and for Footscray -6.8. Thus the HA to Melbourne at the MCG was 1.8 + 6.8. At the Whitten oval it was 9.9 + 0.4 to Footscray. Thus, using equation 4.3, at the MCG the expected result is a 3.6 point win to Melbourne, whereas at Footscray it is 14.3 win to Footscray. In terms of percentage chances, this changes Melbourne from a 53% chance to a 38% chance. Thus the change of venue resulted in a decrease of 15% in Melbourne's chances of winning. If this was repeated over 11 home matches it would be almost two extra losses by the club.

A similar analysis can be performed for a team that changes its home ground permanently. In 1993 St Kilda moved from Moorabbin where they enjoyed one of the highest HAs of 12.5 (compared with an average of -0.5 for the other teams) to Waverley Park where they had a negative ground effect (-4.8 at the end of 1995, compared with the average for the other teams of 0). By the methods of the previous paragraph the change cost them on average over 17 points each home match, or a decrease in percentage chance of winning of about 15%. This results in an expected decrease in the number of wins in their 11 home matches of 1.7. While this may not seem a lot, in 1993, although finishing 12th, they were only two wins behind Adelaide who finished 5th. In 1995, an extra 1.5 wins would have taken them from 14th to at least 9th on the final ladder. It is clear the move to

Waverley has been costly to St. Kilda in terms of its on-field success.

7.11.2 Fairness of draw - average strength of opponents

A major drawback of the League competition is that the draw is not balanced. It is unbalanced with respect to strength of opposition (each team plays a different set of opponents twice) and with respect to grounds (teams play a different number of matches on their home grounds). While the general public recognise this is inequitable, again it has never been quantified in a proper manner. At the very most a football writer may tabulate the number of times each team plays a weak team, or a finalist from the previous season, but never is it done at the end of a season when the true strengths of the teams is better able to be estimated. This unfairness will not necessarily even out over the years. For example at one time the draw was made on the basis that the top teams in one year played each other twice the following year. Thus there was an ongoing bias in the draw.

It is relatively simple in principle to quantify the unfairness of the draw after the season. An analysis such as performed in section 2.14 gives team rating and HAs. Unlike other measures such as final ladder positions or percentage, these are independent of the toughness of the draw. Summing the ratings of the opponents of each team gives a measure of the difficulty of the draw for that team. This is equivalent to the approach of (Leake, 1976) who suggested the average rating of opponents as a measure of schedule difficulty. The home ground advantage of opponents could also be included as this contributes to the draw difficulty. However there are problems with this approach. Since the good teams do not play themselves, they will appear to have an easier draw than the others. Thus even in a balanced competition this method would give a measure of unbalance. For this reason we need to subtract the average strength of the opponents. Thus we are measuring the excess strength of the actual opponents over the average strength of all possible opposition. This is equivalent to adding a proportion of a team's own rating to account for the above bias.

If the measure of team ability in an N team competition is u_i, i = 1 to N, where $\Sigma u_i = 0$, then opponent j will exceed the average strength of all possible opponents of team i by

$$u_j - \frac{\sum_{j \ne i} u_j}{N-1} = u_j - \frac{-u_i}{N-1} = u_j + \frac{u_i}{N-1}$$

Summing this for all opponents is a measure of the total strength of opposition to team i. While we could use the us derived earlier, or better still $u_i + 0.5$ hi, there are advantages in using a measure that the general football follower would understand. For this reason, percentage, which Figure 2.4 showed was highly correlated with $u_i + 0.5$ hi, may be a good choice.

A well understood measure of a teams ability is final ladder position. This incorporates both team ability and some measure of HA. Unfortunately it also includes a component due to the factor we are measuring - draw difficulty, but we bear with this in the interests of having a simple measure. While it would be more accurate to use (say) the regression estimates of a team's ability rather than ladder ranking, the latter has the advantage of being understood by the average administrator and supporter. Table 7.11 was obtained using the ladder ranking at the end of the year. Because a low number indicates a high ranking and strong opposition, a negative total indicates the draw was more difficult than average, a positive number easier than average. Note that during the years 80 to 86 all teams had a balanced draw. In other years, the difference between highest and lowest is generally about 35. This is clearly a significant amount, particularly for two teams in a similar position on the ladder, where the difference cannot be attributed to the different rankings of the two teams. For example in 1988, Geelong, one position on the ladder ahead of Richmond, had a more difficult draw by 36 ranking points. That is the equivalent of playing the top three teams instead of the bottom three teams. In the same year West Coast finished one spot above Melbourne with the same number of wins. However Melbourne's draw was 29 ranking points harder than West Coast. A similar draw could have given Melbourne three extra wins and put them second on the ladder. (They actually did win their way through to the grand final). In 1995, the two teams with the hardest draw, Melbourne and Collingwood both missed the final eight by one game, even though they had better percentages than Footscray who finished seventh and Brisbane who finished eighth. Again the difference in their draw difficulty could easily account for the difference. It is clear that the degree of imbalance that exists in the draw is enough to have a significant effect on the final ladder outcomes. Individual clubs should also look at their draw difficulty in assessing the measure of success they have achieved through the year.

Clearly the draw difficulty does not even out through the years. Richmond and Sydney appear to have had a long run of good draws, while Carlton has had a long run of more difficult draws. Many AFL clubs have criticised the level of financial support given to Sydney. They have also, it appears, received support from the schedule.

TABLE 7.11. Measure of draw difficulty for AFL teams, 1980-1995

Year	Number of Teams	Team															
		Adel	Car	Coll	Ess	Fitz	Foot	Geel	Haw	Melb	NthM	Rich	Syd	StK	Bris	WC	Frem
80	12		0	0	0	0	0	0	0	0	0	0	0	0			
81	12		0	0	0	0	0	0	0	0	0	0	0	0			
82	12		0	0	0	0	0	0	0	0	0	0	0	0			
83	12		0	0	0	0	0	0	0	0	0	0	0	0			
84	12		0	0	0	0	0	0	0	0	0	0	0	0			
85	12		0	0	0	0	0	0	0	0	0	0	0	0			
86	12		0	0	0	0	0	0	0	0	0	0	0	0			
87	14		-10	0	-2	-12	-9	-11	0	16	5	3	-1	14	13	-6	
88	14		12	6	2	-12	12	-20	13	-19	-12	16	10	-9	-9	10	
89	14		5	3	17	12	-7	13	14	-12	-16	-7	12	-14	-14	-6	
90	14		-9	-7	-4	2	0	3	-4	5	1	3	4	-9	8	7	
91	15	-9	-5	0	6	-3	5	1	9	4	1	5	-3	9	-6	-14	
92	15	-3	-4	14	-14	4	8	6	-5	-2	1	6	6	-9	-10	2	
93	15	4	0	-12	-4	7	-9	-14	2	9	20	4	-12	3	2	0	
94	15	-7	-3	-6	-1	10	-11	-6	-6	-2	-9	3	12	16	18	-8	
95	16	-1	-7	-17	9	-4	6	19	-6	-16	-1	15	9	4	-5	3	-8
Total		-18	-21	-19	9	5	-4	-8	16	-17	-9	47	38	4	-1	-12	-8

7.11.3 Fairness of home and away draw - as assessed by the computer prediction

In this section we demonstrate how the computer prediction program can be used to obtain another estimate of the effect of the draw on the success of clubs. This gives a measure of the effect in ladder positions. Particular goals for clubs would be to make the finals, make the grand final and win the premiership. Any sporting competition is designed to produce a winner, and the rules should ensure the expected final positions reflect the abilities of the participants. The ladder prediction model described in section 4.2.1 provides a perfect tool to investigate this. Given the ratings of each team and the draw it provides the expected finishing position of all teams. This can be compared with the ratings. This is demonstrated with a detailed look at 1995.

The computer prediction program was used to predict the 1995 results using the initial ratings derived from the previous year. The ratings of each team at the end of each round were recorded and plotted. This showed that the rating of a team is certainly not constant over a year, that teams have periods of good and poor form. (To ensure this was not just an apparent affect because the initial rating was in error, the average rating was calculated for each team and the program run again with these ratings as the initial ratings. Most teams still showed the same general pattern - as the rating is in effect a smoothed average, the initial ratings only affect the first few week's ratings). Figures 7.1a and 7.1b show the week by week ratings of a couple of teams along with their average rating. Clearly Collingwood has played the last half of the of the season much better than the first half, while Brisbane has made a tremendous improvement from about Round 15 onwards. Many of these graphs are interesting in their own right, and could be used to investigate the effects of changes, such as injury to star players, changes of coach etc. For example, the Brisbane coach announced his decision to retire at the end of the year about Round 15. In the round 15 match, Brisbane overcame a 45 point deficit at three quarter time to beat Hawthorn by 7 points. The victory was put down to the visiting team wilting in the heat, but the effects were obviously seen for the remainder of the year. Brisbane's rating shows a steady rise from that point and they made the finals for the first time. Hawthorn's ratings decreased just as steadily, and they finished second last, missing the finals for the first time in since 1981. While such a spectacular change in fortune can be picked up by other means, graphs such as these can be used to study a team's form. Their advantage is they have allowed for ability of opposition and HA. Currently winning and losing streaks are often used, but these are affected to a great degree by quality of opposition.

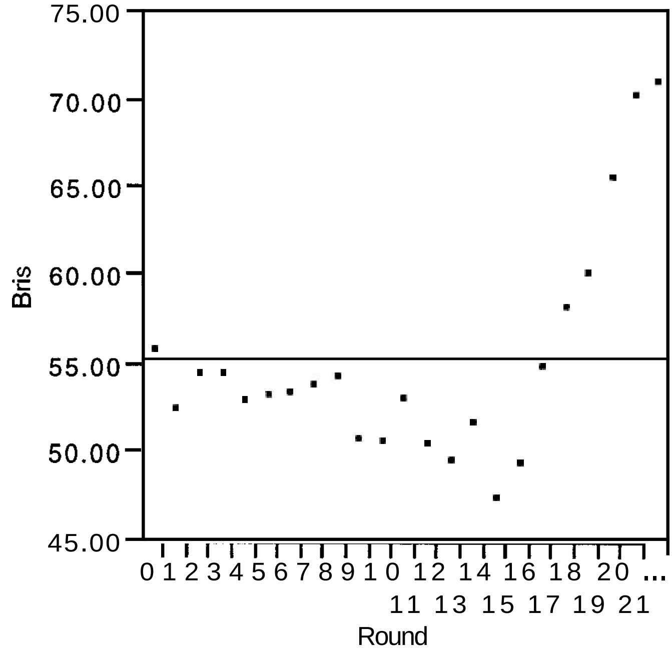


Figure 7.1a. Weekly ratings of Brisbane during 1995

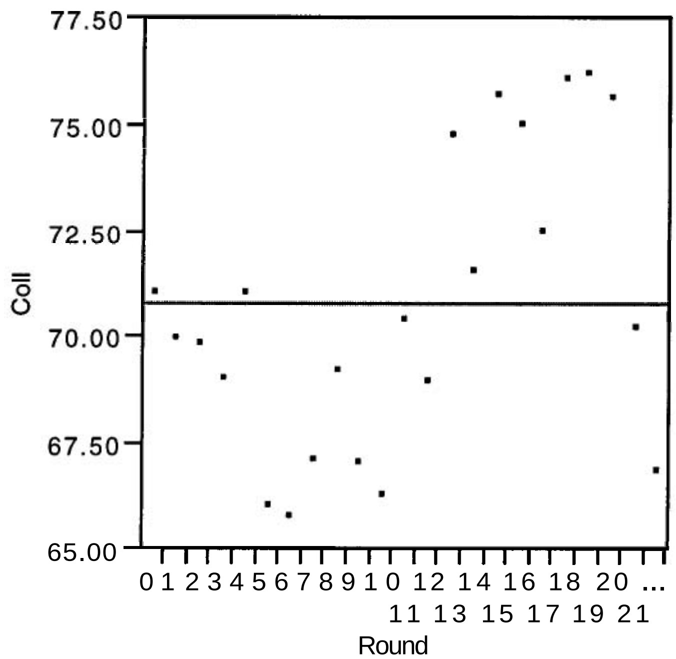


Figure 7.1b. Weekly ratings of Collingwood during 1995

We cannot expect a draw to be balanced for this change of form for different teams. Always some teams will be lucky and play opposition when they are down on form or missing star players. However it should be balanced for average ability. The average ratings for the year were used as initial ratings for the computer predictions. (That actually increased the predicted number of winners for the home and away matches for the year from 127 to 134, and reduced the margin of error from 32.6 points to 31.2 points. This gives a measure of the value of the unknown information, average form level, at the start of the year.) The final ladder predictions assume current form and HAs for the remainder of the year, so the computer predictions before round 1 of the final ladder can be used to estimate the expected finishing positions of each team. In a perfectly balanced competition balanced for quality of opposition and HA, the expected final ladder would be roughly in the same order as $u + 0.5h$. Variations from this reflect unfairness in the draw. The expected ladder produced by the program, along with the team ratings and HAs are given in Table 7.12.

Note the predicted order is in good agreement with the fair order. However teams that have done better than fair are Geelong, Essendon, Richmond and Hawthorn. Those that have done worse are West Coast, North Melbourne, Melbourne and Footscray.

Some of these results are consistent with the previous results on draw difficulty. The two teams with the easiest draws both do better than expected and Melbourne with a hard draw does worse. Again, there are probably competing effects here. Some can be attributed to their poor ground effect at the MCG. The major difference is that the previous was based on actual ladder position whereas this method used average rating for the year.

TABLE *7.12*. Expected final ladder for *1995*, with team ratings and HAs shown

Predicted order	Team	Premier. points	Percentage	u	h	$u + 0.5h$	Fair order
1	Car	60	130.5	92.6	14.3	99.8	1
2	Geel	60	125.6	86.9	10.7	92.3	3
3	WC	56	122.5	82.3	22.5	93.6	2
4	Ess	56	118.2	84.8	3.5	86.6	5
5	Nth	52	115.7	89.0	-0.6	88.7	4
6	Rich	48	106.7	76.4	2.6	77.7	7
7	Melb	48	102.9	77.3	3.1	78.9	6
8	Coll	40	96.6	71.1	1.2	71.7	8
9	Haw	40	95.8	65.8	5.3	68.5	10
10	Foot	40	95.4	66.5	8.9	71.0	9
11	Adel	40	93.7	58.8	16.7	67.2	11
12	Syd	40	93.2	63.2	3.9	65.2	12
13	Bris	36	89.3	55.9	18.5	65.2	13
14	Frem	36	88.5	54.0	14.2	61.1	14
15	StK	32	83.8	57.9	-4.8	55.5	15
16	Fitz	20	64.6	36.3	-11.5	30.6	16

7.11.4. Using simulation to measure the efficiency of the draw

Of course the actual ladder positions will be due to some extent on random variation. We might wish to investigate the extent to which the final ladder position is affected by random variation. A season of football has a large random element, and most supporters recognise that luck plays some part in the success of their club. Also club success is not a linear function of ladder position. For example, obviously two seconds would not be equivalent to a first and third. For both these reasons it is appropriate to look at the probabilities of teams achieving certain goals. In racquet sports, for instance, this has resulted in the concept of efficiency of scoring systems, where the length of matches is traded off against the probability of the better player winning.

While it is outside the scope of this thesis to investigate alternatives, we do want to give an idea of the effects of random variation on the final ladder. The simulation discussed in section *4.2.1* will give us an indication of its extent. Table *7.13* is the result of simulating the 1995 season 1000 times using the average rating for the teams as initial ratings. The table shows the number of seasons the team finished in the given position. While the most likely position was generally close to the ranking based on $u + 0.5h$ given in Table 7.12, the probability of this was often quite low. The table shows the huge variation possible in a season of football and demonstrates the dangers in putting too much emphasis on the final ladder position as a measure of the team's performance. It is possible for almost any team to finish anywhere from last to first due to the random effects. The range within which a team was an 80% chance to fall within was about four positions for the very best and worst teams, up to about 10 positions for some of the middle teams. This dependence on chance can be demonstrated by looking at individual matches. In round 9, Adelaide beat Hawthorn 9.06 to 7.16 by two points. Had just one of Hawthorn's 16 behinds been a goal, Hawthorn's final ladder position would have been three places higher and Adelaide 4 places lower. In contrast, the u_i and h_i for those teams as developed in section 2.14 would have hardly altered at all. For this reason, measures as suggested in this thesis are a more accurate reflection of a team's performance through the season.

TABLE 7.13. Chances in 1000 of ending in any position after home and away matches

Team	Position															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Adelaide	2	10	13	31	50	54	68	83	107	93	112	90	110	101	63	13
Carlton	302	218	136	113	83	63	35	26	11	6	3	4	0	0	0	0
Collingwood	10	15	29	36	55	60	87	73	109	87	105	93	90	95	48	8
Essendon	107	145	163	151	123	81	76	51	40	11	18	13	10	9	2	0
Fitzroy	0	0	0	0	0	0	0	1	2	1	3	9	14	44	145	781
Footscray	3	5	21	39	36	66	93	86	102	99	113	107	95	71	53	11
Geelong	236	185	170	121	93	59	52	31	22	11	7	5	6	2	0	0
Hawthorn	2	13	27	44	50	70	78	110	78	107	91	103	76	76	61	14
Melbourne	18	38	53	68	82	106	98	98	95	91	66	59	59	49	18	2
Nth Melbourne	103	114	124	114	125	104	88	66	41	43	33	21	9	9	5	1
Richmond	26	53	73	90	117	114	95	110	79	91	37	46	30	24	14	1
Sydney	4	6	15	22	36	59	59	85	99	87	114	110	104	104	76	20
St. Kilda	1	2	4	7	12	18	21	30	63	62	88	102	116	152	241	81
Brisbane	1	4	10	16	24	36	48	56	68	100	102	102	147	126	137	23
West Coast	184	189	156	139	91	77	63	32	25	18	14	5	5	1	1	0
Fremantle	1	3	6	9	23	33	39	62	59	93	94	131	129	137	136	45

CHAPTER VIII

WHEN TO RUSH A BEHIND IN AUSTRALIAN RULES FOOTBALL: A DYNAMIC PROGRAMMING APPROACH

8.0. Abstract

In Australian rules football, points are scored when the ball passes over the goal line. Six points are awarded for a 'goal', when the ball passes between the two centre posts and one point for a 'behind', when the ball passes between a centre post and an adjacent outer post. After a behind, the defending team has a free kick from the goal line. It may be worthwhile, particularly in the closing stages of a game, for a defending team voluntarily to concede a behind, by themselves passing the ball between the two outer posts, either to avert the possibility of an imminent goal or to increase the probability of scoring a goal themselves. A dynamic programming model is used to analyse this situation.

Key words: sports, dynamic programming, Australian rules football

8.1. Introduction

When a game between two teams nears the end of the scheduled time, it may be advantageous for a team, particularly if it is losing, to change its strategy. A well known example is that of 'pulling the goalie' in ice hockey, treated by Washburn (1991) in a paper which demonstrates how a dynamic programming formulation can take account of all possible policies and not restrict consideration to a subset, as previous papers had done. End game tactics in other sports have been analysed, for example by Clarke & Norman (1998b) on cricket and by Kohler (1982) on darts.

In this paper we consider the end game in Australian rules football, the major winter sport of the southern states of Australia, played between teams of 18 players on oval grounds (the same grounds used for cricket during the summer). A match is played in four quarters, each of 20 minutes plus about 10 minutes of extra time. Players can run with the rugby-shaped ball, but it is moved forward more quickly by kicking or punching it to a team-mate, and with no off-side rule, the game is fast. The scoring region consists of four upright posts. Kicking the ball between the two centre posts scores a goal worth six points, while the region between either centre post and the corresponding outside post scores a 'behind' worth one point. A 'rushed' behind is scored by a defender kicking, punching or carrying the ball over his own goal line, conceding a point in the hope of a possible territorial advantage. Draws are rare: a typical score might be anything between 50 and 150 points.

A particular feature of the closing stages of a match played under Australian rules is the decision by a defender whether or not to rush a behind. A defender might prefer to give away one point for a rushed behind to eliminate the possibility of the attacking team scoring a goal (six points), particularly when the scores are fairly close and a goal scored by the attacking team would win the game for them. After a rushed behind, the defending team also gains the advantage of possession with a kick from the goal line. However, the wisdom of rushing a behind might depend on the time left until the end of the match. What might be sensible with an hour left might not be wise in the closing minutes.

8.2. Description of model

We propose a simple model which is intended to capture the essential features of the game, particularly in regard to the decision whether or not to rush a behind. Such a model may be useful in generating and testing ideas before practical trials. To fix ideas, suppose we (our team) play up the ground, which we divide into seven parts, numbered as shown in Figure 8.1.

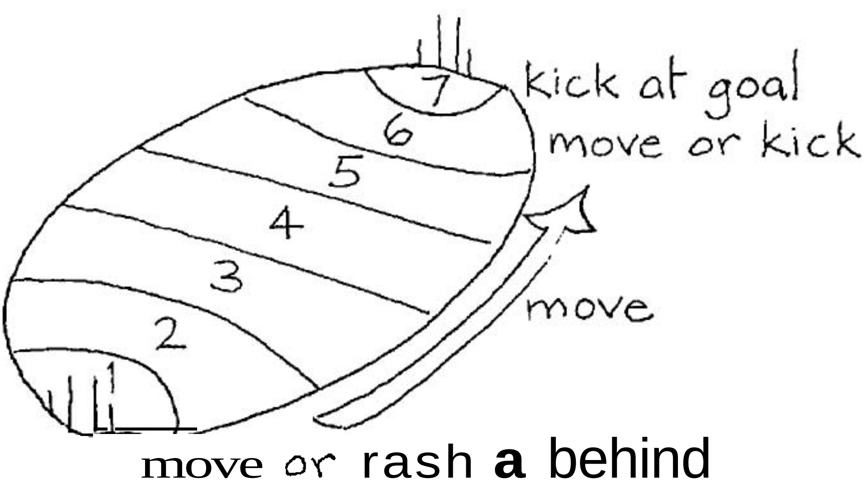


Figure 8.1. Tactical choices in a model of Australian rules football

We try to move the ball into areas 6 and 7 in which we can attempt to kick a goal. k_i^6 and k_i^1 are the probabilities of scoring 6 (for a goal) or 1 (for a behind) from area i (i = 6 or 7), ($k_i^6 + k_i^1 = k_i$). If a goal is scored (+6), the ball is bounced on the centre spot in area 4. If a behind is scored (+1) or a kick at goal misses completely (no score) or the defending side rush a behind (-1) the defending side kicks from the goal area and has possession in area **3** with probability π.

When a team moves the ball, generally forward, it may maintain or gain ground (one area at a time) and keep or lose possession with the following probabilities:

p_1	keep ball, gain ground
p_2	keep ball, maintain ground
p_3	lose ball, maintain ground
p_4	lose ball, gain ground.

$p_1 + p_2 + p_3 + p_4 = 1$, and normally $p_1 < p_2$ and $p_3 < p_4$.

We take it that each of these transitions takes place between stages - decision epochs occurring at constant time intervals throughout the duration of the match.

To facilitate analysis, we suppose that a team will always move the ball in areas 2, 3, 4 and 5, and will always kick for goal when in area 7. In area 6 the team may either move the ball or kick for goal. In area 1 the team may either move the ball or rush a behind.

Let $f_n(s, \text{p}, 1)$ be the probability of our winning the game, with n stages remaining, when our team leads by s points and we have the ball in area p, under an optimal policy. Optimal here means maximising the probability of winning the game. Similarly, let $f_n(s, \text{p}, 2)$ be the probability of our winning the game, with n stages remaining, when our team leads by s and the opposing team has the ball in area p. Suppose also that the probabilities {p}, {k} and π are the same for both teams.

Then $f_0(s, .\,, 1) = 1$ if $s > 0$
$\qquad = 0$ if $s \leq 0$

$$\begin{aligned} f_n(s, p, 1) = {} & p_1 f_{n-1}(s, p+1, 1) \\ & + p_2 f_{n-1}(s, p, 1) \\ & + p_3 f_{n-1}(s, \text{p}, 2) \\ & + p_4 f_{n-1}(s, p+1, 2) \quad \text{for } p = 2, 3, 4, 5 \end{aligned}$$

$$f_n(s, 7, 1) = k_7^6\,\{0.5f_{n-1}(s+6, 4, 1) + 0.5f_{n-1}(s+6, 4, 2)\}$$
$$+ k_7^1\,\{(1-\pi)f_{n-1}(s+1, 5, 1) + \pi f_{n-1}(s+1, 5, 2)\}$$
$$+ (1 - k_7)\{(1-\pi)f_{n-1}(s, 5, 1) + \pi f_{n-1}(s, 5, 2)\}$$

$$f_n(s, 6, 1) = \max\{\text{ move: } p_1f_{n-1}(s, 7, 1) + \ldots + p_4f_{n-1}(s, 7, 2)$$
$$\text{kick:}k_6^6\,\{0.5f_{n-1}(s+6, 4, 1)+\ldots+\pi f_{n-1}(s, 5, 2)\}\}$$

$$f_n(s, 1, 1) = \max\{\text{ move:}p_1f_{n-1}(s, 2, 1)+ \ldots + p_4f_{n-1}(s, 2, 2)$$
$$\text{rush: } \pi f_{n-1}(s-1, 3, 1) +(1-\pi)f_{n-1}(s-1, 3, 2)\ \}$$

$f_n(s, p, 2) = 1 - f_n(-s+1, 6\text{-p}, 1)$ by symmetry, for

The probability that we win when they have the ball in area p and we lead by s
= the probability that they win when we have the ball in area 6-p and they lead by s
= 1-probability that we tie or win when we have the ball in area 6-p and they lead by s
= 1 - probability that we win when we have the ball in area 6-p and they lead by s-1

8.3. Initial results

A short computer program was written in BASIC to evaluate $f_n(s, \text{p}, 1)$ and $f_n(s, \text{p}, 2)$ for $-25 \le s \le 25$ and p = 1 to 7 for successive values of n. The following parameter values were assumed:

$p_1 = 0.3$	$k_6^6 = 0.2$	$\pi = 0.7$
$p_2 = 0.4$	$k_6^1 = 0.3$	
$p_3 = 0.1$	$k_7^6 = 0.5$	
$p_4 = 0.2$	$k_7^1 = 0.2$	

Although calculations were carried out for $n = 1$ to 50, the main purpose of the program run was to confirm that appropriate behaviour occurred at small values of n (near the end of the match). Table 8.1 shows when to rush a behind, indicated by a tick, depending on the score difference and the number of stages left. A tick indicates that in area 1, a defending team will increase its chances of winning by rushing a behind rather than moving the ball. For example, with ten stages left, a team which is two points ahead should rush a behind, a team which is one point ahead should not. Not surprisingly, when a team is up to four points behind, it can be worthwhile for it to rush a behind if there are enough stages left for it to have a chance of moving the ball the length of the ground and scoring a goal. At five points behind, this tactic is not worthwhile, as the point given up through the rushed behind makes it impossible to do better than tie.

When a team is between two and six points in front, it can be worthwhile for it to rush a behind because even if the opposing team gain the ball from the goal line kick-off, they are further away from the goal line and less likely to score a goal. When scores are level, it is sometimes worthwhile to rush a behind, despite the one point penalty; this is a consequence of the objective, to maximise the probability of winning.

Instead of maximising the probability of winning, a team may wish to maximise the probability of not losing. It would be surprising for a player to rush a behind in the closing stages of a match if the scores were level. No further calculations are needed for the probability of our not losing a match when we have the ball in area p and we lead by s = the probability that we win the match when we have the ball in area p and we lead by $s + 1$.

8.4. Checking the model

Any operational research model should be checked to ensure that it conforms in a real situation. In this case we need to check the realism of our model of the Australian rules game. The division of the ground into seven areas enables us to mimic the centre spot bounce (area 4) and the possession possibilities following a kick from the goal area (areas **3** and 5). Areas 1 and 7 are defined by the areas in which a defending team may rush a behind and an attacking team must try for a goal. Areas 2 and 6 represent the area in which a team has a choice whether to kick for a goal or to move further towards its target goal. The transitions are limited, but they conform to real life: teams rarely move the ball backwards.

It is more difficult to validate the probabilities we have assumed, although they are derived from a long term study of the game. Obviously, the probabilities $\{p\}, \pi$ and $\{k\}$ will vary from team to team, from match to match and even during a match. We believe that the chosen values are fit for an exploratory study. A check on their appropriateness, albeit limited, may be made by comparing the results derived from the model applied to the entire game with the kind of results that occur in practice.

TABLE 8.1. When to rush a behind (1 point penalty)

Stages	Points difference																																																	
	Behind																										Ahead																								
left	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1																																																			
2																												√	√	√	√	√																			
3																												√	√	√	√	√																			
4																												√	√	√	√	√																			
5																						√	√	√	√	√		√	√	√	√	√		√	√	√	√	√													
6																						√	√	√	√	√		√	√	√	√	√		√	√	√	√	√													
7																						√	√	√	√	√		√	√	√	√	√		√	√	√	√	√													
8																√	√	√	√	√		√	√	√	√	√		√	√	√	√	√		√	√	√	√	√			√	√	√	√							
9																√	√	√	√	√		√	√	√	√	√		√	√	√	√	√			√	√	√	√			√	√	√	√							
10																√	√	√	√			√	√	√	√	√		√	√	√	√	√			√	√	√	√			√	√	√	√							
20				√	√	√	√			√	√	√	√			√	√	√	√			√	√	√	√				√	√	√	√			√	√	√	√			√	√	√	√			√	√	√	√	
30	√	√		√	√	√	√			√	√	√	√			√	√	√	√			√	√	√	√				√	√	√	√			√	√	√	√			√	√	√	√			√	√	√	√	
40	√	√		√	√	√				√	√	√	√			√	√	√	√			√	√	√	√				√	√	√	√			√	√	√	√			√	√	√	√			√	√	√	√	√
50	√	√		√	√	√				√	√	√				√	√	√				√	√	√	√				√	√	√	√			√	√	√	√			√	√	√	√	√		√	√	√	√	√

TABLE 8.2. Transition matrix

S	T	Y	1		2		3		4		5		6		7		Expected single stage gain
			Us	Them	Us	Them	Us	Them	Us	Them	Us	Them	Us	Them	Us	Them	
1	Us	M	p_2	p_3	p_1	p_4											
	Us	R					π	$1-\pi$									-1
	Them						$\pi(1-k_7^6)$	$(1-\pi)(1-k_7^6)$	$.5k_7^6$	$.5k_7^6$							$-6k_7^6-1k_7^1$
2	Us				p_2	p_3	p_1	p_4									
	Them	M	p_4	p_1	p_3	p_2											
	Them	K					$\pi(1-k_6^6)$	$(1-\pi)(1-k_6^6)$	$.5k_6^6$	$.5k_6^6$							$-6k_6^6-1k_6^1$
3	Us						p_2	p_3	p_1	p_4							
	Them				p_4	p_1	p_3	p_2									
4	Us								p_2	p_3	p_1	p_4					
	Th						p_4	p_1	p_3	p_2							
5	Us										p_2	p_3	p_1	p_4			
	Them								p_4	p_1	p_3	p_2					
6	Us	M											p_2	p_3	p_1	p_4	
	Us	K							$.5k_6^6$	$.5k_6^6$	$(1-\pi)(1-k_6^6)$	$\pi(1-k_6^6)$					$6k_6^6+1k_6^1$
	Them										p_4	p_1	p_3	p_2			
7	Us								$.5k_7^6$	$.5k_7^6$	$(1-\pi)(1-k_7^6)$	$\pi(1-k_7^6)$					$6k_7^6+1k_7^1$
	Them	M											p_4	p_1	p_3	p_2	
	Them	R									$1-\pi$	π					+1

S = State; T = Team; Y = Strategy; K = Kick; M = Move; R = Rush

A transition matrix for the four possible policies (move or kick for goal in area 6, move or rush a behind in area 1) is given in Table 8.2. Each entry in the table is the probability of moving from the state identified by the row name, under the decision specified (if any), to the state identified by the column heading.

The steady state probabilities for possession by our team are shown in Table 8.3. Each entry in the table is the probability that our team, following the policy identified by the column heading, will be, at any stage, in the area specified by the row name. Thus if we always move in area 1 (and do not rush a behind) and move in area 6 (and do not kick for goal), we shall be in area 1, with possession, 2.3% of the time. Each column sums to 0.5, as each team has possession half the time.

TABLE 8.3. Steady state probabilities under four stationary policies

Area	Move/Move	Move/Kick	Rush/Kick	Rush/Move
1	0.023	0.000	0.000	0.013
2	0.057	0.043	0.043	0.048
3	0.100	0.134	0.134	0.113
4	0.124	0.151	0.151	0.132
5	0.104	0.129	0.129	0.112
6	0.069	0.043	0.043	0.064
7	0.023	0.000	0.000	0.019

In a match lasting, say, 100 minutes, it would be quite typical for each team to run with the ball about half the time and to make about 200 kicks and 100 hand passes. (Statistics on the number of kicks and hand passes are published as a matter of routine in the sports pages of Australian newspapers). Thus we might envisage stages in the model to occur at 5 second intervals, with 1200 stages in the match.

We now compute the expected scores under each of the four policies. In the move/move policy, we are in area 7 for 0.023x1200 = 28 stages and in half of these ($k_7^6 = 0.5$) we score a goal. We have a probability of 0.2 of scoring a behind. Thus the expected number of goals we score is 14 and the expected number of behinds is 6, making an expected score of 90. In the move/kick and rush/kick strategies, we are in area 6 for .043x1200 = 52 stages and score an expected 10 goals and 15 behinds, making a total score of 75 points. In the rush/move strategy, we are in area 7 for 0.019x1200 = 23 stages and score an expected 11 goals and 5 behinds. Our opponents are in area 7 for $0.013 \times 1200 = 16$ stages and thus concede an expected 16 behinds, so that our total score is 87 points.

These scores are typical of those occurring in actual play and confirm that a 5-second interval between stages is reasonable. They also help to confirm the reasonableness of the initial probability estimates.

8.5. A suggested change in scoring

The results of Table 8.1 suggest that if the objective is simply to maximise the probability of winning, it may be worthwhile for teams to rush a behind for many score differences and at many stages of the game. However, if such a strategy were generally adopted, it would change the nature of the game. The initial results shown in Table 8.1 are thus relevant to the view held by some commentators that "teams seem to have adopted the attitude that it is preferable to give up one point to eliminate the possibility of conceding a goal (six points) ... reinforced ... by the belief that one point is a small price to pay for possession". Mike Sheahan of the Herald Sun suggests a three point penalty for a rushed behind (Sheahan, 1996). This suggestion was adopted in the computer program with the results shown in Table 8.4.

For small values of n (near the end of the game) the effect of increasing the penalty from one point to three is to reduce, in a simple way, the blocks of score differences where it is worthwhile to rush a behind: if it is worthwhile to rush a behind if the score difference is between L and U and incur a one point penalty, then it is worthwhile to rush a behind if the score difference is between L+2 and U and incur a three point penalty. For $n > 7$ this is not always exactly the case, largely because of the longer term effects of the difference in penalty points incurred. Generally, the three point penalty has a more lasting deterrent effect.

8.6. Possible extensions

When the attacking team scores a behind, the ball is kicked into play from behind the goal line by the defending team. In the model this possession by the defending team has not been considered a possible state: only possessions on the field of play have been counted.

Sometimes it is not clear who has possession. After a kick, a player may fumble and drop the ball, after which both teams may struggle for possession. 'Possession in dispute' could be a possible extension of the state description.

As well as running with the ball, a player may play the ball by kicking or handballing (punching) it. A possible extension to the model is to allow a team to move the ball either within its present area (running or handballing) or to the next forward area (kicking). In trying to move the ball further forward there would be a greater risk of losing possession.

An obvious extension would be allow for differences in team capabilities. It would be surprising if the two teams were equally competent in all areas and hence had the same values of $\{p),\{k)$ and π. There is clearly scope for a sensitivity analysis to investigate the dependence of tactical choices on these probability values. Finally, it would be simple to incorporate differing probabilities of winning possession from the centre circle bounce.

8.7. Conclusions

A computer program has been used to determine optimal end-game strategies for the decision whether or not to rush a behind, and to consider the possible effect of a change in the associated penalty. The initial results support the view that it is often to a team's advantage to concede a point through a rushed behind in order to obtain the probable advantage of possession. The main objection to such a policy is that it runs counter to the spirit of the game. Australian rules football is an attacking game and a rushed behind is a kind of 'own goal' with perhaps an inappropriate penalty. Just as in soccer, a deliberate hand ball to stop a certain goal is now penalised more severely than by a simple penalty kick, maybe in Australian rules a different kind of penalty is needed than a simple one-point penalty.

Further Reading: Readers interested in the subject matter of this paper will find material of interest in Clarke (1993b) and Norman (1995).

TABLE 8.4. When to rush a behind (3 point penalty)

Stages	Points difference																																																		
	Behind																									Ahead																									
Left	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1																																																			
2																														√	√	√																			
3																														√	√	√																			
4																														√	√	√																			
5																								√	√	√				√	√	√				√	√	√													
6																								√	√	√				√	√	√				√	√	√													
7																								√	√	√				√	√	√				√	√	√													
8																		√	√	√				√	√	√				√	√	√					√	√					√	√							
9																		√	√	√				√	√	√				√	√	√					√	√					√	√							
10																		√	√					√	√	√				√	√	√					√	√					√	√							
20						√	√					√	√					√	√					√	√						√	√					√	√					√	√					√	√	
30	√	√				√	√					√	√					√	√					√	√						√	√					√	√					√	√					√	√	
40	√	√				√						√						√							√						√	√					√	√					√	√					√	√	
50	√	√				√						√						√													√	√					√	√						√						√	

TEST STATISTICS

9.1. Introduction

With origins that can be traced back to the 13th century, the first set of rules for cricket were written in 1744. One hundred years later, on September 24-25 1844, Canada played the USA at St George's Cricket Club Ground, Manhattan, New York. The game had spread from England throughout the British Empire and beyond. Cricket is now administered by the International Cricket Council, and matches between countries of a suitable standard are called Test Matches and are usually scheduled as 30 hours of play over five days. About 250 Tests were played between 1877 and 1935, but with the expansion in the number of Test playing countries a similar number were played in the 1980s. Currently, nine Test match playing countries (England, Australia, West Indies, India, Pakistan, Sri Lanka, New Zealand, South Africa, and Zimbabwe) play irregular series against each other consisting of between one and six Tests. While the major interest for statisticians is Test cricket, many official matches scheduled to be played over at least three days are deemed First-Class and these constitute the majority of cricket records. The domestic competitions of countries with Test status, of which the English County Cricket Championship is the best known, are in this category.

Cricket is played between two teams of 11 players on large oval shaped grounds of various sizes. The main action takes place in the centre of the ground on a grass pitch 22 yards long and about six feet wide. A wicket consisting of three stumps forms a target for the bowlers at each end of the pitch. Unlike pitchers in baseball, bowlers are not allowed to throw but use a stiff arm action to deliver the ball on the run, and usually bounce the ball off the pitch before it reaches the batsman. Batsmen play in pairs, one at each end, and score a run each time they run the length of the pitch, thus changing ends. A line on the pitch about three feet in front of each wicket is known as the crease, and a batsman is dismissed if he fails to ground his bat or part of his body over the crease before the fielding side hit the corresponding wicket with the ball. There is no foul area, and the batsmen do not have to run when they hit the ball. A long hit may give batsmen time for up to four runs, while hitting the ball to the boundary automatically scores four, and over the boundary on the full scores six. The many means of dismissal include being bowled (the ball hitting the wicket), Leg Before Wicket or LBW (the ball hitting the batsman's legs when it would otherwise have hit the wicket), caught (a fielder catching the ball off the bat), and run out (failing to make an attempted run). Bowlers bowl balls in sets of six called overs, with alternate overs bowled by different bowlers from opposite ends. A team's innings ends when 10 wickets have fallen, (leaving one batsmen not dismissed or not out) or the captain 'declares' the innings closed (leaving two batsmen not out).

In the traditional form of the game, each team has two innings, and a match is played over a fixed maximum period of two to five days. In most domestic round robin competitions first innings points would be given for the team leading after each team has batted once, but outright victory is the major goal. This is achieved by obtaining a higher total score than the opponent obtains in two completed innings. In Test cricket, since only outright victory counts as a win, the losing team generally must be dismissed twice to get a result. This means games lasting five days often end in a draw (i.e., are unfinished). The famous 'Timeless Test' (England v South Africa, Durban, 1939) scheduled to be played to the finish, was abandoned and declared a draw on the tenth day of play because the England team had to board its ship home. Cricket uses a tie to distinguish the extremely rare event when two teams finish on the same score after all four innings have been completed. To overcome the high frequency of draws, and generally improve spectator interest, a one-day form of the game was introduced in the 1960s. Each team has only one innings in which they each face a maximum number of balls (usually 300) over a total of about six hours playing time. While less tactically subtle, failing severe interruption by the weather a decision is always achieved.

The difficulty of scoring runs depends very much on the quality of the pitch, which varies greatly from match to match and generally deteriorates during a match. Thus, while 300 may be a respectable score in a Test innings, the highest team score in a Test match was 952 for 8 wickets by Sri Lanka against India in 1997, while the lowest was 26 by New Zealand against England in 1955. A typical scorecard for one innings as usually published in the daily press or cricketing almanacs contains a list for each batsman in batting order with their total score, the method of dismissal and the bowler responsible. Note the bowler is still credited with taking the wicket even if the batsman is caught. For each bowler, the number of overs bowled, the number of maidens (overs which were not scored off), the total runs scored off their bowling and the number of wickets taken is given. Also included is the team score when each batsman was dismissed. This allows the calculation of the partnership - the total score made while each pair of batsmen were batting. The batting order is usually determined by ability, with the first six being recognised batsmen and the last four players selected for their bowling. Milestones for batsmen are the multiples of 50, in particular a century. For bowlers, five wickets in an innings or 10 in a match are more difficult achievements. While the traditional scoresheet used by officials has more information, such as a batsman's individual scoring shots, it is a non trivial task to reconstruct a ball-by-ball account of the match from the score sheets. Since official score sheets are usually only available from the particular association running the match, and these may not be archived, almost all analysis is done from published scorecards. (Officials actually burned the official score book used for the historic 1960 first ever tied Test). Some televised matches use computer systems that keep ball-by-ball data, including where on the ground the ball was hit, but such data is not freely available.

Table 9.1. shows a typical Test career record, that of English all rounder Ian Botham.

TABLE 9.1. Test career record of Ian Botham

	M	I	NO	Runs	HS	Ave	100	50	Ct	St
Batting	102	161	6	5200	208	33.54	14	22	120	-
	Balls	M	R	W	Ave	Best	5	10	SR	Econ
Bowling	21815	788	10878	383	28.40	8-34	27	4	56.9	2.99

There is a lack of symmetry in the scorecard information which is reflected in the career records. For a bowler, we know how many balls he bowled (Balls), runs allowed (R) and wickets taken (W). For a batsman, we are told runs scored (Runs) and whether he lost his wicket or not, but are not told how many balls he faced. So for bowlers, we can calculate a 'strike rate' equal to the number of balls bowled per wicket taken (SR = Balls/W), but cannot calculate the run rate (Runs/Balls) for batsmen. Only in recent years has balls faced by the batsman been collected, but it is rarely published in the newspapers. Even if balls faced by each batsman is published, it is impossible to reconstruct balls faced by a particular partnership. In the same way, the number of maiden overs gives some distributional information about bowlers which is not available for batsmen. Recent scorecards sometimes show the number of fours and sixes hit by batsmen, but a bowlers figures are never broken up in this way.

Like all sports, changes to rules require some care when statistics from different eras are compared. While an over is now standardised as six balls, Australia used eight ball overs for sixty years prior to 1978. The rules for LBW have undergone several modifications, and the definition of a no ball and wide may alter from series to series, or from Test to one-day cricket. The treatment of sundries (no balls, wides and leg byes and byes) has also changed over the years. For example, no balls are now debited to a bowler's figures, but this was not always the case. In some competitions, a no ball has been credited as two runs, whereas traditionally it has been worth one run. It has recently been suggested to discontinue counting a wide as a ball faced by batsmen. This seems reasonable, since a wide by definition is a ball impossible to hit.

While earlier publications (Haygarth, 1862-1885, 1925) detail the history of cricket for each season from 1746, *Wisden Cricketers Almanack*, (Engel, 1997) has been the traditional Bible of cricket statistics since its first publication in 1864. Each annual edition contains statistics from the previous English domestic season, plus full scorecards of every Test match played around the world during the year, and coverage of cricket in 40 different countries. Over 40,000 first-class matches have been played, and the maintenance and publication of their records is a major aim of the Association of Cricket Statisticians and Historians (ACS). The association publishes a quarterly journal, *The Cricket Statistician*, containing non-technical articles on cricket statistics. All ACS scorecards are gradually being transferred to CricInfo, a fan-based organisation which aims to provide cricket scores and records via the World Wide Web. Further details on both these organisations and several similar ones can be obtained from the web. While such organisations provide cricket statistics, the capabilities of their memberships to perform analyses of a technical nature are limited. Thus, while a huge number of statistics and records are collected and published, there is little attempt at any serious statistical analysis.

Cricket has the distinction of being the first sport used for the illustration of statistics. In *Primer in Statistics*, Elderton and Elderton (1909), used individual scores of batsmen to illustrate frequency distributions and elementary statistics. Elderton (1927) used scores of batsmen to illustrate the exponential distribution, and Wood (1941) investigated the idea of consistency. These efforts resulted in Wood (1945) and Elderton (1945) reading separate papers at the same meeting of the Royal Statistical Society. These papers have some claim as the first full quantitative papers applying statistics to sport. The papers are accompanied by 17 pages of discussion, demonstrating the great interest generally created by papers in sport. Yet in spite of this interest, the topics raised were ignored in the professional statistical literature for over thirty years. In contrast to baseball, few papers in the professional literature analyse cricket, and two rarely examine the same topic. This allows us the luxury of looking in some detail at virtually all the published material using statistics in cricket.

9.2. Distribution of scores

The distribution of batting scores was first discussed by Elderton and Elderton (1909). Graphs of a batsman's scores were used to illustrate skewed distributions. While no formula was given a theoretical graph that could only be that of the negative exponential distribution is shown. Elderton (1927) formalises this by using the same scores to illustrate the Pearson Type X or negative exponential distribution. If a batsman scores only singles and his probability of dismissal is constant his scores should follow a geometric distribution, the discrete equivalent to the negative exponential. Elderton (1945) obtains a reasonable fit of the geometric distribution to the individual scores over three years of four early cricketers. Wood (1945) takes this further and compares the scores and several statistics of several groups and many individual batsmen with that expected using the geometric distribution. Although no significance tests are used, the fit is generally fair. However, the number of ducks and scores less than five is less than predicted while the number of centuries is greater than expected suggesting that the probability of dismissal is not constant. Cricket folklore says batsmen are more prone to dismissal early in their innings, perhaps get nervous or careful when their score reaches ninety, and tire or hit out later in their innings. In fact there are several reasons why scores should not be exactly geometric. The score does not increment by one, but advances by jumps of usually 1, 2, 3, 4, or 6. As well as changing throughout an innings, one would certainly expect the probability of dismissal to change from innings to innings, depending on the quality of the pitch, or the opposition. Wood raises the possibility that the discrepancies are the result of combining two or more geometric distributions and investigates this by looking at batsmen's scores over several years both individually and in combination. He finds that while combining several geometric progressions will understate the expected number of zeros and centuries, the effect is small and does not explain the discrepancy with the observed scores.

Other authors argue that because the chance of dismissal varies throughout an innings, and from one innings to another, the negative binomial distribution may describe the number of runs scored in completed innings. Using data from Wood for three batsmen, and data for three contemporary batsmen, Reep et al. (1971) found only one batsman's scores in each set were approximately negative binomial. Pollard (1977) applies a chi-square goodness of fit test to Elderton's data and finds the geometric distribution performs slightly better than the negative binomial. However, both fitted distributions have higher variances than the empirical data. Pollard et al. (1977) claim the excess of high scores will not be present if partnerships are investigated, as several partners may be dismissed while a player is making a high score. They obtain a good fit of the negative binomial to all partnerships for a team in the English County championships. On the other hand, Croucher (1979) found a negative binomial failed to fit the total number of runs for each completed Australian partnership in 82 England Australia Tests.

In the papers of Elderton and Wood, we first come across a continuing problem in cricket - how to handle the not out scores (i.e., scores by batsmen who are not dismissed). While Elderton and Elderton (1909) added the next score to a previous not out score, these early papers generally treated not outs as completed innings. Kimber and Hansford (1993) looked at the empirical discrete hazard and smoothed empirical hazards of the Test, first class and one-day international scores of several batsmen, and compared them with the constant hazard expected for the geometric distribution. For most batsmen, the region from 0 to 5 runs is higher than expected. However, they found no compelling evidence that the hazard is not otherwise constant and concluded that the tail of the score distribution for a batsman is roughly geometric.

Many arguments against the geometric distribution for scores do not apply to the number of balls a batsman faces. This count certainly increments by one. A batsman can alter the degree of risk he takes, in order to play each ball with the same chance of dismissal. Early in the innings he is just content to survive, whereas later when he is settled he will take risks in order to score. Similarly, on a bad pitch or against good opposition he may play more carefully, and adjust his scoring rate to keep roughly the same risk of dismissal. For these reasons, Clarke (1991) suggested the distribution of the number of balls faced by a batsman in an innings may be geometric. Unfortunately, in the past the number of balls faced has not been regularly published or even kept by scorers. However, using ball by ball data for Australian batsmen from all matches in the 1989-90 one day series involving Australia, Pakistan and Sri Lanka, he failed to obtain a better fit for balls faced than scores. All 60 innings of the top Australian batsmen were combined, and the distribution of both the number of balls faced and scores were compared to the geometric distribution using the Kolmogorov Test. Both scores and balls faced had p values of about 0.2. Strangely, the results for scores tended to be the reverse of those found in Test cricket; the number of very low values and very high values was less than expected for both scores and balls. The limited nature of one-day cricket obviously produces a distribution of scores different from the traditional game.

Because teams generally play under similar conditions throughout a match, a strong positive correlation between performances in a match may be expected. However, various investigations have failed to produce much evidence of this correlation. Elderton (1945) tabulated the scores of two pairs of county championship opening batsmen, using intervals of 10 runs, and by visual inspection found no correlation between their scores. Croucher (1982b) calculated the correlation for each team between twenty five pairs of completed innings in Tests between Australia and England. He found zero correlation for Australia, and a non-significant *negative* correlation for England. Testing a belief held by fans that long partnerships are generally followed by short ones Croucher (1979) found some evidence to the contrary, that partnerships following century partnerships tend to be *longer* than usual.

The distribution of scores is important as it affects the appropriateness of other statistics. Kimber and Hansford (1993) question the use of the batting average. This is defined as the total number of runs scored divided by the number of times the batsman has been dismissed. Thus, runs scored in innings in which the batter has not been dismissed are included in the numerator, but the denominator does not count those innings. Because of the handling of not outs, many cricket fans have a certain unease with the current statistic as it appears to give inflated averages - a player may have an average larger than his greatest score. Since more than 10% of scores are not outs, it is an important statistical issue. Kimber and Hansford demonstrated that the batting average has the desirable property of consistency irrespective of the censoring mechanism for not out scores only if the distribution of scores is geometric. Since this is not the case exactly, they claimed the batting average does not estimate the population mean. Applying the methods of survival analysis, they defined a non-parametric alternative which effectively distributes the not out scores using the empirical distribution of any higher scores. A parametric adjustment is needed when the batsman's highest score is not out, as there is no empirical information on his chances of dismissal above that score. The method has the disadvantage of needing re-calculation from scratch with each new score. They calculated the new average for several batsmen and also suggested other statistics such as selected centiles. They finally recommended a slight alteration to the career statistics of batsmen, by including with the number of 50s and 100s the percentage of 50s and 100s inflated for not outs.

Elderton and Elderton (1909) used the standard deviation of scores as a measure of consistency, with zero implying perfect consistency. By contrast, because the mean and standard deviation of the geometric distribution are (roughly) equal for a large mean, Wood suggested using the coefficient of variation (CV) (multiplied by 100) as a measure of consistency, with the closer to 100, the more consistent a batsman. In his analysis, Wood found CVs as low as 96 and as high as 139, but mainly in the range 100-109. Pollard (1977) claimed a high CV indicates a batsman has problems early, but scores runs more easily later in the innings. Using a model where the number of balls faced was geometric, while the score off each ball had any unspecified alternative distribution, Clarke (1991, 1994) showed that perfectly consistent batsmen will have CVs greater than 100, and that perfectly consistent batsmen with different scoring distributions off each ball will have different CVs. Thus it is not possible to have a single measure (CV closeness to 100 as proposed by Wood) which indicates perfect consistency for all batsmen. Still, the regular publication of the standard deviation or CV of scores would assist fans in judging the consistency of batsmen. Analogous questions on the distribution and consistency of runs and wickets for bowlers have not been analysed in the literature.

9.3. Rating players

The major system of rating players is the Deloitte ratings which were created by Deloitte, Haskins and Sells in 1987. After several mergers, they are now called the Coopers & Lybrand ratings. The system rates the current Test form in both batting and bowling for all international players with ratings updated after each Test match. The algorithm supposedly takes into account the playing conditions, the strength of opposition, and the results of the match. However, these are estimated objectively from the details in a typical summary score sheet. The ratings have been backdated to the late seventies.

Because of its proprietary nature, it is difficult to obtain details of the algorithm, and even more difficult to get details of the statistical work behind its derivation. Berkmann (1990) contains a few pages of description of the algorithm, along with ratings for all players from the 1980s. A rating is worked out for each performance (defined for the purposes of calculating the ratings as an innings for a batsman, and 30 runs conceded for a bowler), combined to give a player rating for each Test, which is then smoothed with past ratings to produce a new rating. The statistics used to calculate a batsman's rating are runs scored, whether not out or out, the team score, wickets fallen, and the match result. The opposition bowling strength is also used but this is estimated from the statistics. For bowlers, the statistics used are runs conceded, wickets taken (including which batsmen), balls bowled per wicket (strike rate), runs conceded per ball, team scores and match result. The ratings take a player's basic statistic (e.g., runs scored or wickets taken) and multiplies it by a series of factors depending on whether the performance is above or below average with respect to playing conditions. For example, a batsman's runs are increased if the pitch is 'difficult', as measured by the ratio of the number of runs per wicket in the match to the average of 31 found in previous Test matches.

A major point of contention is that the algorithm rewards performances of players on the winning side in a match more than in drawn or lost matches. Thus, above average players on the winning side are given a bonus, while below average players on the losing side are penalised. Ted Dexter, one of the initiators of the scheme, said “the pleasant surprises include ... the considerable accuracy of the ratings when used to compare the relative strengths of the Test playing countries” (Berkmann, 1990, page iii). This is hardly surprising for an algorithm that gives a greater rating to a performance on a winning team. Also the rating is not symmetrical with respect to batsmen and bowlers. A bowler’s strike rate is included in the calculation, but not a batsman's scoring rate. Similarly, the overall opposition bowling strength is taken into account for batsmen, but opposition batting strength is not used for bowlers. Nevertheless, the ratings seem to have gained some acceptance by cricket followers. However, the need for some proper analysis of the scheme is evident. For example, a simulation could be used to investigate the effect playing on good and poor teams has on ratings of players of similar calibre. The effect of batting against good and poor bowling sides should also be analysed. Surprisingly, alternatives based on models of how Test cricket is played have not been proposed.

The Deloittes ratings are only applied to Test cricket. Johnston (1992) and Johnston et al. (1992, 1993) looked at a method of rating players in one-day cricket in which traditional statistics are not as relevant. A quick score of 30 may be more valuable than a slow century; three wickets in the last over is of no more value than a maiden. By comparing the actual number of runs each ball with the optimal given by a dynamic programming formulation, both a batsman's and a bowler's contribution could be measured in a radically different way to normal methods. The context in which events occurred becomes important. Johnston et al. (1993) gave ratings for each player in a one-day series. One difficulty confronting the acceptance of the system is its need for ball-by-ball data. However, with the growing use of computer-assisted scoring, such methods become more viable.

9.4. Tactics

Some quantitative work has been done in the area of tactics. The limited nature of the innings in one-day cricket creates a trade-off between run rate and loss of wickets. Traditional tactics suggest an innings starts cautiously, with teams scoring slowly and preserving wickets. The run rate steadily increases, until near the end of the innings there is often a frenzy of runs scored and wickets fallen. Clarke (1988) analysed this by setting up a dynamic programming model with the stage the number of balls to go in the innings. The model assumed all batsmen were of the same ability, and the probability of dismissal on each ball depended only on the run rate. For the first innings, the states were the number of wickets to have fallen, and the objective function was the number of runs scored in the remainder of the innings. In the second innings, the state was the number of wickets and the number of runs to go, and the objective function was the probability of exceeding the opponent's score.

Under the assumed relationship between run rate and dismissal probabilities (necessitated by a lack of data), Clarke’s results suggested that teams should score more quickly early in the innings - in fact, at any stage they should score at a slightly greater rate than the expected rate for the remainder of the innings. Such tactics have become more accepted recently, and in particular used by the current World champions, Sri Lanka. Johnston (1992) later showed the recommendations were valid under a range of relationships between run rate and dismissal probabilities. Clarke and Norman (1995, 1997a, 1997b) have extended the models to allow for batsmen of unequal ability. Near the end of an innings in Test or one-day cricket, two batsmen of widely different ability will often refuse a possible run early in the over to protect the weaker batsman from the strike. They investigated the point in the over and in the innings when this tactic is optimal with respect to different objective functions.

The way batsmen are dismissed has received some attention. For example Croucher (1982a) analysed dismissals in the 96 Australia - England Test matches between 1946-80. He investigated the various types of dismissal with respect to batting position and location (England or Australia). About 20% of batsmen 1 to 7 are bowled, but this increases steadily to nearly 38% for number 11. The LBW dismissal rate is about 14% for batsmen 1 to 8 but much less for batsmen 9 to 11. In contrast the percentage of batsmen caught is reasonably constant. A thorough investigation of modes of dismissal would be interesting and could find application to vary tactics in different countries.

9.5. Umpiring decisions

An area that has received some attention is the contentious one of umpire bias. Only recently has the traditional practice of the home country providing both umpires for Test matches been modified. Because the rule is quite complicated, LBW is the most subjective type of dismissal and often creates controversy. For the period 1877 to 1980, Sumner and Mobley (1981) found significantly fewer LBW dismissals against home teams than visiting teams in India, Pakistan, and South Africa, but not in Australia. Croucher (1982a) found that, while on average about 12% of batsmen are dismissed LBW, this percentage varied from 10.6% in Australia to 14.0% in England. This could be due to the different behaviour of pitches in England and Australia. However when subdivided by team, LBW rates for England were fairly constant at 11.6% and 12.0% in Australia and England, but for Australia varied from 9.5% in Australia to 15.4% in England. This could be due to umpire bias, or Australian batsmen not adjusting to conditions. Breaking batsmen into two categories, 1-5 and 6-11, the frequency of LBW decisions showed a dependence of location and category for Australian batsmen but not for English. One interpretation of this is that umpires give decisions against top order batsmen but square the account against lower order batsmen.

Crowe and Middeldorp (1996) used logistic regression to compare LBW rates in Test matches played in Australia for visitors and Australians for the period 1977-1994. The odds of an LBW dismissal are defined as the ratio of LBW dismissals to all dismissals by other means. A logistic model was used to fit the logarithm of the odds for a series of matches to a linear expression using indicator variables for the various countries. Separate models were fitted using only the top six batsmen and all batsmen. An initial model found a significant difference for LBW rates for visiting teams and Australia. Since there was no evidence of the odds for countries changing over time, nor of Australia's odds changing depending on opponent, the overall odds for each country were calculated and compared with the common odds for Australia. For both the top six batsmen and all batsmen, only three of the seven countries that had visited Australia during the period (England, Sri Lanka and South Africa) had significantly proportionally more LBW dismissals than Australia. For example the odds ratio for England was 1.7, with a 95% confidence interval of 1.2 to 2.5. Interestingly India, whose captain had complained of umpire bias, had an odds ratio of only 1.059, barely greater than the expected value of 1 if there was no bias. Of course, as Crowe and Middeldorp point out, other factors such as a difference in playing style, more experience on home wickets or different tactics could account for the results.

9.6. Rain interruption in one-day cricket

Because of the limited number of balls in one-day cricket and the requirement that a result has to be achieved if at all possible, allowance has to be made for rain interruption. For example, if the innings of the team batting second is shortened because of rain, the target score for victory has to be reduced to compensate for the reduced number of overs. Various formulae to adjust the target score have been tried with varying degrees of success. These rules appear to be developed ad hoc by administrators and rarely are based on a proper quantitative study. They are used until a particular situation arises which makes a mockery of the rule, which is then replaced. For example, in one World Cup semi-final, South Africa, who had a reasonable chance of achieving their target when rain interrupted, were then required to face *one* ball and score 22 runs to win when play resumed.

Using a dynamic programming formulation to investigate tactics as described earlier, Clarke (1988) produced a table giving the number of runs a team could expect to make in the remainder of the innings, depending on the number of wickets fallen and the number of balls to go. For the second innings, the analysis gave the probability of scoring the required number of runs, again depending on the number of wickets fallen and the number of balls to go. He suggested the tables could be used to evaluate the effectiveness of rain interruption rules, so that teams would have the same chance of winning after the interruption that they had prior to the rain. He showed that the rule current at the time, where the second team had to score at the same rate for the shortened number of overs, was highly advantageous to the team batting second.

Duckworth and Lewis (1996) and Duckworth (1997) produced a method using tables similar to Clarke's first innings table, but which were derived from past statistics. They treated balls to go and wickets remaining as run-scoring resources, and fitted an exponential decay model for the average number of runs scored to past data. The resulting function used to give a measure of the proportion of these combined resources available at any stage of the innings. Targets are adjusted depending on the proportion lost due to rain. For example, suppose a team chasing 250 has lost two wickets after ten overs when a rain interruption reduces their available overs from 50 to 40. The table shows at the time of the interruption they had 77.6% of resources left, but on resumption had 68.2% left. They lost 9.4% so their target is reduced by 9.4% of 250 or 24 runs. The method can be applied to interruptions at any stage of a match. The system allows the calculation at any stage of the second innings of the score the batting side needs to win the match if abandoned at that point. This allows a team and the spectators to gauge the match run by run and wicket by wicket. This system was first used in international competitions in Zimbabwe in 1996, and is being used for domestic competitions in England in 1997.

A by-product of the system could be a better method of declaring the winning margin in one-day matches. Margins in one-day matches are still given using traditional measures of runs if won by the team batting first or wickets if won by the team batting second. These can be quite misleading. For example, a team batting second that scores the winning run on the last ball of the innings may be credited with a 6 wicket victory. This sounds comfortable, when in fact the team had used virtually all their resources. The method could also be investigated to provide alternative tie-breaking procedures in one-day round-robin tournaments.

9.7. Sundries

In 1981, in response to a request for data for a simulation study of batting order in one-day cricket, a student of the author received a firm refusal from a high ranking Australian cricket administrator which included " Any analysis that you suggest would be wholly hypothetical, and of no value...the analysis would only be greeted by scorn by those with a proper understanding of cricket. The inherent essence of cricket is its unpredictability; and an attempt to reverse this .. is something I would not personally encourage." With the increasing use of science in sport, one hopes this view is not widely held today. Cricket administrators now clearly seek assistance from academics to solve management problems that are not peculiar to cricket (Johnston, 1992; Willis and Armstrong, 1993; Willis, 1994; Wright, 1991, 1992).

The history of rain interruption rules suggest they are less reluctant to seek their help with on field and other problems. Few first class competitions have the luxury of allowing 5 days play as in test matches. Consequently, in a high proportion of these matches, neither team achieves an outright victory. In domestic round robin tournaments, the relative allocation of points for first innings and outright victory varies, and administrators have experimented with bonus points for fast batting or penetrative bowling. Rarely are these experiments based on, nor their effectiveness judged by, statistical studies. Bosi (1976) investigated the effect of the introduction of bonus points in county cricket. He claimed that a significant change in the correlation in ladder position using the traditional and new methods shows the rule alteration affected the way cricket was played.

However, statisticians should play a major role in developing rules, not just evaluating their effect after the fact. The problem of allocating points for unfinished matches should be investigated, with possibly some of the methods used for one day matches applied. There is currently discussion in cricket about the publication of a world ranking of Test teams. Wisden publish their own table (Engel, 1977, p19), based on each country scoring 2 points for winning a series and 1 for drawing, but the system does not take into account margin of victory nor home advantage. With countries playing intermittent series of different lengths, statisticians should investigate and recommend suitable ranking systems before cricket administrators decide on something inappropriate. Similarly, player evaluations should not be considered resolved, and rivals to the Cooper and Lybrand rating should be developed.

Cricket is highly variable. Clarke (1994) showed that with roughly geometric partnership score distributions, purely random variations give rise to team scores ranging from 100 to 500 in Test cricket. Johnston (1992) simulated one-day cricket using optimal batting rate policies and obtained scores ranging from 75 to 322. A game with so much variation provides ample scope for statisticians to assist participants, administrators and supporters to separate real effects from random noise.

However, topics that have proved fruitful areas of research in other sports have been largely ignored in cricket. While the difficulty of winning a Test series on foreign soil is recognised, home ground advantage has not been thoroughly studied. Pollard (1986) quoted home advantage in cricket county championships to be 56.1%, excluding drawn games. Because in domestic competitions some teams play on pitches that regularly produce results, and outright wins are rewarded more than first innings wins, Clarke (1986) showed that some teams not only win a greater proportion of the points awarded on their home ground, but compete for a greater number of points than other teams. Such a system would never be tolerated in other sports.

Surprisingly few alternative statistics have been suggested. Cricket fans seem to be satisfied with dividing the traditional ones into categories. So, a batsman is still judged by his average, although this may also be given against a particular country or at a particular ground. With the introduction of one-day cricket, strike rate, economy rate, and run rate have also become popular and the articles mentioned here have made several suggestions for alternative statistics. With these statistics currently being measured over the life of a player, the use of moving averages could be used to measure current form. There may be advantages in combining the statistics in various ways. For example Kimber (1993) compared bowling statistics using scatterplots, while

Ganesalingam et al. (1994) applied multivariate analysis techniques to classify players as batsmen, all-rounders, or bowlers. Such studies could suggest further statistics or indices to be used in evaluation of players and selection of teams.

9.8. Conclusion

In spite of a huge collection of statistics on cricket dating back over two hundred years, little attempt at serious analysis has appeared in the professional literature. Of all the sports in this text, cricket has the distinction of having statistics that stretch back the longest, the first use of sport in a statistics text to illustrate statistical principles, the first full quantitative paper, and yet probably the fewest serious papers analysing the statistics in the professional literature. It is surprising that more statistical analysis has not been undertaken. Many questions in cricket could be investigated using relatively elementary statistical techniques. Is one team better than another? Is one batsman better than another? Does the rate of dismissal vary?

A large number of papers have been written about baseball - best batting order, value of player, measurement of hot streaks etc. Similar research could be done in cricket. One of the neglected areas calling for study is bowling. When thanking Elderton, (Wood, 1945) said "At last a great statistician has discovered what is, I believe, the richest field of statistical material left untilled. I have scratched over its surface, but other statisticians will find in it materials for all sorts of statistical experiments, particularly in the bowling analyses". This statement is still valid today.

ANOTHER LOOK AT THE 1985/86 SHEFFIELD SHIELD COMPETITION CRICKET RESULTS

10.1. Introduction

This article aims to show how a simple analysis of competition results can yield valuable insights into team performance and competition rules. While the analysis has been done on the Sheffield Shield, similar analysis could be performed on any other cricket or sporting competition. While this simple analysis uses only the outcome of the matches (first innings win, outright or draw), similar calculations using wickets, runs or runs per wicket could be performed.

In the 1985/86 Sheffield Shield competition, the team leading on the first innings gained four points with eight points for an outright win. The results of the 1985/86 competition are shown below in Table 10.1.

TABLE 10.1. Results of the 1985/86 Sheffield Shield competition

Team	Outright wins	First innings wins	Total points
NSW	4	6	56
Qld	4	5.5	54
Vic	2	8	48
WA	2	6.5	42
SA	2	3	25
Tas	0	1	4

The two top teams, NSW and Qld, played in the final at Sydney which was drawn but NSW won the Shield because of its top position. While NSW performed best under the rules of the competition, this may not necessarily indicate that NSW was the best team under all scoring systems and ground allocations. To determine how well each state performed, a more critical analysis of the results is appropriate. Here we analyse the year's results in various ways and suggest alternative methods of allocating points.

10.2. First innings win - outright win weighting

In the 1985/86 season a first innings victory gained four points, and outright victory gained another eight points. In the past, other weightings have been used to give greater rewards to an outright victory such as four for a first-innings and 16 for an outright. In fact this weighting can be changed drastically without altering the above - order. Keeping first innings points at four, outright points can come down as low as five and go up as high as you like and the above order will be preserved. But the scoring system has again been changed for the 1986/87 season. (Note: For 1987/98, a team earns four points for an outright win and two for leading on the first innings. However, any team that is beaten outright after leading on the first innings, loses its first-innings points. In addition teams who do not bowl the required number of overs may be penalised points.)

10.3. Home ground advantage

A greater insight into the year's results can be achieved by looking at where each team gained its victories. Table 10.2 shows the points won/lost in each match.

The home team is shown at the top of the table and the away team down the side. The entry shows the number of points won by the home team and the number of points won by the away team. Thus 8-4 in the SA column/Tas row tells us SA (home team) gained eight points for the outright, although Tasmania won first innings points.

Table 10.2 allows us to easily see a team's home and away performance, or compare two teams' performances against other teams. The table shows that:

- None of the top four teams lost outright at home.
- When the top four teams played each other only three times out of 12 did the away side gain first innings points.
- Teams are far more likely to win at home (146 points) than when they are away (86 points).
- With 52 points being won in Sydney, the NSW ground is far more likely to produce an outright result than is the MCG at which only 20 points were won.

From these findings, we see that a side is given a huge advantage when the final is played at its home ground. Even if the final was decided on first innings points, the home side would have a 75 per cent chance of winning based on 1985/86 results. By insisting that the visiting (second placed) side win outright to win, the Shield organisers are virtually giving the Shield to the top side.

TABLE 10.2. Points won/lost in each match.

Away Team	Home Team						Total points away
	NSW	Qld	Vic	WA	SA	Tas	
	For-Ag	For-Ag	For-Ag	For-Ag	For-Ag	For-Ag	Ag-For
NSW	-	0-4	4-0	12-0	40	0-4	20-8
Qld	4-0	-	4-0	4-0	0-12	0-12	12-24
Vic	8-4	4-0	-	0-4	4-8	0-12	16-28
WA	12-0	2-2	4-0	-	0-4	0-4	18-10
SA	12-0	12-0	4-0	4-0	-	0-12	32-12
Tas	12-0	12-0	4-0	12-0	8-4	-	48-4
Total points at home	48-4	30-6	200	32-4	16-28	0-44	146-86

Table 10.3 illustrates the point by showing the total number of points awarded in matches each team played at home and away and the percentage of those points each team gained. It shows that:

- The four top teams gained the majority of their home game points - over 80 per cent for each of the top four teams and over 90 per cent for Victoria and NSW.
- The greatest number of points came from games played in NSW where 52 points were scored from a possible 60 points available.
- The least number of points came from games played in Victoria, where only 20 points were scored and all by the home side.
- Games played in NSW produced almost 50 per cent more home points than those played in Queensland and at least 20 per cent more home points than games played in any other State.
- In matches played away from home those involving Victoria and SA produced the most points and those involving NSW the least points.
- Both Victoria and Queensland gained more than 60 per cent of the points scored in their away matches whereas NSW gained only 30 per cent.

The above analysis shows quite clearly that there is a large home ground advantage in the Sheffield Shield competition. This advantage is in two parts. The first allows a team to perform better on its home ground than away - such an advantage is inherent in the game and while it appears to be greater for some teams than others, that is not a problem. It is up to each team to learn to successfully exploit the idiosyncrasies of their home ground. The second is more insidious - for successful exploitation of this home ground advantage, some teams are rewarded more than others. This type of home ground advantage is inherent in the scoring system and administrators should work towards its removal.

It is clear that NSW finished on top of the table because they exploited a home ground where outright victories were the norm. At the other grounds, perhaps because of flat pitches or loss of time due to rain, outrights were far less common. These teams were put at a severe disadvantage because of the scoring system used. It should be noted that if groundsmen wish to assist their teams, they should prepare pitches which will produce outrights. That way, their team will be playing for zero or 12 rather than zero or four points. Similarly, captains would be better off agreeing to declare their first innings closed at 0-0 and so make a first innings win an outright win. (This is not as outlandish as it seems. A similar happening occurred in a Victorian Cricket Association match when the teams declared their first innings closed at 2-69 and 3-69, allowing time for one side to gain an outright.)

TABLE 10.3. Comparison of points won at home and away

	Points won at home matches			Points won at away matches		
Team	Home team	Both teams combined	Home team x100 /combined	Away team	Both teams combined	Away team x100 /combined
NSW	48	52	92%	8	28	29%
Qld	30	36	83%	24	36	67%
Vic	20	20	100%	28	44	64%
WA	32	36	89%	10	28	36%
SA	16	44	36%	12	44	27%
Tas	0	44	0%	4	52	8%

10.4. Overcoming home ground advantage

It is obviously the organiser's responsibility to choose scoring methods which reduce the present home ground advantage. Two methods are suggested.

1. Sharing points for undecided outrights. This method treats undecided outrights in the same manner as undecided first innings points. The Queensland/WA match resulted in the four first innings points being split 2-2 as the two first innings were not completed. If outrights were treated the same way, the eight outright points would be split between the two teams in the event of the match being drawn. This would mean that every match resulted in the awarding of twelve points, and in 1986 would have resulted in the following final Shield table.

TABLE 10.4. Shield table resulting if points for undecided outrights shared.

Team	Points
Queensland	78
NSW	76
Victoria	76
WA	70
SA	44
Tasmania	16

2. Using point difference. In deciding the point allocation for games administrators influence the way the game is played. Rewards are given for outcomes that correspond to some ideal behaviour winning on the first innings or outright. For example, in the past, to encourage attractive play, bonus batting points have been given for scoring quickly. Now, while winning matches is behaviour that should be rewarded, it also is true that not losing matches should also be encouraged. A team is not penalised for losing matches outright. In Test cricket, the quality of not losing matches is important - teams sometimes fight for several days just to avoid defeat. In the present Shield competition, such a fight would not be rewarded. A team receives exactly the same number of points if it loses or draws. A simple way to incorporate rewards for not losing (or penalties for losing) is to award points as present but determine the table on points difference: points won less points lost.

Such a table is easily made up by referring back to Table 10.2. For example, NSW total points gained is 48+8 = 56, total points lost 4 + 20 = 24. Doing this for each team gives the Table 10.5

TABLE 10.5. Final Shield table using point difference

Team	For	Against	Difference
Queensland	54	18	36
NSW	56	24	32
Victoria	48	16	32
WA	42	22	20
SA	28	60	-32
Tasmania	4	92	-88

This gives a final order of Queensland on top with NSW and Victoria equal second. Note that NSW has dropped in the order because it lost more matches than Victoria and Queensland. Now while negative numbers are difficult to work with, it can be shown that this system is exactly the same as allocating outright points equally between two drawn teams. This system produces the same order as the method outlined in the previous section. In fact, the points under the previous system are always 60 + half the above differences.

The previous method of allocating outright points equally between drawn teams not only goes some way to reducing home ground advantage, but also rewards teams for not losing matches.

10.5. Alternative method

While it is not suggested as a method for determining a Shield winner, this alternative method does shed some light on the year's results, and is useful in looking at how well a particular State performed.

Consider each pair of matches that two teams play against each other, say SA and NSW. In one match SA beat NSW on the first innings, in the second NSW beat SA on first innings and outright. Clearly, NSW has performed better than SA. We would say that pair of matches shows NSW should finish higher than SA. On the other hand, in the two Queensland Victoria matches, both teams had one first innings victory each - that pair tells us nothing about the relative merits of Queensland and Victoria. Thus each pair of matches either gives us an ordering of the two teams or is inconclusive. The only combination which produces an arguable result is where one side has won both first innings points, but lost one on outright, as was the case with NSW and Victoria. Under last season's scoring system that would count as inconclusive (eight points each), although most observers would probably say that NSW had the better of Victoria. For any other combination, it is beyond dispute which team had the better of the other, independently of whatever decisions are made regarding relative merits of first innings and outright points. Working through the 15 pairs of matches we obtain the following:

NSW - Qld	NSW above Qld 8-0
NSW - SA	NSW above SA 16-0
NSW - Tas	NSW above Tas 12-4
NSW - Vic	arguable 8-8 - see above
NSW - WA	inconclusive 12-12
Qld - SA	Qld above SA 24-0
Qld - Tas	Qld above Tas 24-0
Qld - Vic	inconclusive 4-4
Qld - WA	Qld below WA 2-6
SA - Tas	SA above Tas 20-4
SA - Vic	SA below Vic 4-12
SA - WA	SA below WA 0-8
Tas - Vic	Tas below Vic 0-16
Tas - WA	Tas below WA 0-8
Vic - WA	Vic above WA 8-0

We are now looking for an order which preserves as many of these relationships as possible. Clearly, Tasmania is on the bottom, as it is below all other teams. Next is SA as it is above only Tasmania. Continuing in this way, we obtain the following order, which surprisingly is consistent with all the above results.

NSW, Vic, WA, Qld, SA, Tas

Surprisingly, Queensland drops to fourth on the table. Its effort in defeating both Tasmania and SA by outrights twice, which counts so much in last season's system, here only confirms that it is better than both SA and Tasmania. However, Queensland's performance against the other top teams is poor - 48 of its 54 points come from the SA and Tasmania matches. This method tends to judge top teams more on how they perform against each other rather than how they perform against weaker teams. Would we judge the relative merits of West Indies and England on how well they beat the United States, or how they perform against each other?

10.6. Conclusion

NSW is to be congratulated on winning the 1985/86 Sheffield Shield. Given a set of conditions for the running of any competition, good teams will play to maximise their score. However, it is clear that there are many methods of evaluating a team's performance. Team managers and supporters should not uncritically accept the final finishing order as some absolute measure. Looking at results in various ways can point coaches and captains to areas requiring improvement - e.g. NSW performances away, Victoria's need to get outrights on its home ground and Queensland's performance against top sides. Of course, such observations can be tempered with a more detailed knowledge of individual matches, such as interruptions by rain, and absence of Test players.

Administrators also can learn from analysis and work towards devising points systems that minimise unfair advantages to some teams and reward preferred outcomes. In the Shield competition, home ground advantage and lack of penalties for losing matches should be areas of concern.

The above analysis has been performed on the 1985/86 results only. Previous years results also could be analysed to see if the above results are normal or a one-off aberration.

While the above analysis has been done on the Sheffield Shield, a similar analysis could be performed on District cricket or any other cricket or sporting competition and provide useful insights to players and administrators.

Acknowledgments: My thanks go to Peter Spence, development manager for The Australian Cricket Board, for his comments on an earlier draft of this article.

10.7. Commentary

While this article was written some time ago, its conclusions still remain valid. Hopkins (1997b) used the points raised as the basis for a newspaper article. A summary of the last six years shield data reveals the home team won 60.2% of the points. Table 10.6 shows the breakup between points won at home and points won away. Clearly most teams enjoyed a large home advantage. For example SA won 61% of the points at their home venue against only 21% at the away venues. However they have also over the six years had nearly 30 more points awarded at their home ground than Queensland - virtually a one point per game unfair advantage. The away results show this cannot be attributed to an attacking style of play. The methods of Chapters II and III could be applied to this data, although more complicated models may need to be used to allow for the variation in the number of points allocated per match.

TABLE 10.6. Comparison of points won at home and away for seasons 91/92 to 96/97

	Points won at home matches			Points won at away matches		
Team	Home team	Both teams combined	Home team x100 /combined	Away team	Both teams combined	Away team x100 /combined
NSW	96.0	156.3	61.4%	71.8	167.8	42.8%
Qld	94.0	135.5	68.8%	60.0	154.0	39.0%
Vic	100.0	164.1	60.9%	30.0	146.3	20.5%
WA	61.0	130.8	46.6%	62.0	133.9	46.3%
SA	95.9	149.9	64.0%	45.0	131.0	34.4%
Tas	88.3	150.3	58.7%	84.0	155.0	54.2%
	535.2	888.0	60.2%			

DYNAMIC PROGRAMMING IN ONE-DAY CRICKET - OPTIMAL SCORING RATES

11.0. Abstract

Using a dynamic programming formulation, an analysis is presented of the innings of the team which bats first (here referred to as the first innings) and the innings of the team which bats second (here referred to as the second innings). This allows a calculation, at any stage of the innings, of the optimal scoring rate, along with an estimate of the total number of runs to be scored (in the first innings) or the chance of winning (in the second innings). The analysis is used to shed some light on possible batting tactics (in terms of the best run rate at any stage of the innings), to quantify the effects of selecting extra batsmen in a side, and to suggest a method for the development of alternative measures of player performance. Results suggest that scoring rates should be more uniform than at present, and that the team batting second has an advantage. Possible extensions to the model are discussed.

Key words: cricket, dynamic programming, sport

11.1. Introduction

The application of OR techniques in general, and dynamic programming (DP) in particular, to problems in sport is growing. Sphicas & Ladany (1976) and Hayes & Norman (1984) are just two examples of using a DP analysis to assist participants in the development of tactics. However, cricket seems to have escaped this notice, the few papers on cricket generally being descriptive in nature. Pollard (1977) gives a summary of the statistical work on cricket up to 1977. More recently, Croucher (1982b) compares dismissals of Australian and English batsmen, while Clarke (1986a) uses a simple statistical analysis of cricket data to investigate the effects of home-ground advantage in cricket. However, lacking in the literature is the application of OR techniques to assist the cricketer with tactics. This seems strange given the role Britain and the Commonwealth have played in the origins and continued practice of both OR and cricket.

The ball-by-ball nature of cricket makes it particularly suitable for a DP analysis. This is especially true for the one-day game. In this paper we present a DP formulation of both the first and the second innings of one-day cricket. This allows a calculation, at any stage of the innings, of the optimal scoring rate, along with an estimate of the total number of runs to be scored (in the case of the first innings) or the chance of winning (in the second innings). The results are used to suggest optimal batting tactics in terms of the best run rate at any stage of the innings, to quantify the effects of selecting extra batsmen in a side, and could be used to develop alternative measures of player performance. It is also shown that the side batting second has an advantage.

11.2. The problem

In one-day or limited-over cricket, each team has to score as many runs as possible off a limited number of overs, the team scoring the most runs winning the match. The innings finishes when the batting side loses 10 wickets or when the bowling side completes its allotted overs. In international matches in Australia, each innings is restricted to 50 overs of six balls each, with no bowler allowed to bowl more than 10 overs. In practice, the second innings also finishes if the batting side passes the other team's score. Thus, while the object of the team batting first is to score as many runs as possible, the object of the team batting second is to score at least as many runs as the first team scored. There are also restrictions on field placings, and various rules to cope with shortened matches owing to rain interruption.

At each stage of the innings, a batsman has to decide how fast to score. An increase in the rate of scoring entails taking greater risks, with a consequent increase in the chance of losing wickets. Loss of wickets increases the chance of the innings finishing prematurely, and so decreases the total score. The basic problem confronting a batting side is how to trade off an increased scoring rate with the possible loss of wickets. As in most sports, there are widely held beliefs on the correct strategy. In one-day cricket, a common strategy is to bat slowly during the early part of the innings, keeping wickets in hand. This usually allows a steady increase in the scoring rate, and often results in a last-minute orgy of runs and wickets during the final part of the innings. Analogous tactics by 12-hour runners might find them walking for the first 11 hours to conserve energy for a final sprint.

11.3. Run rate

For each ball, ignoring such things as no balls, runouts and overthrows, a batsman can either lose his wicket without scoring or keep his wicket while scoring 0 to 6 runs.

Let p_d be the probability of dismissal.
Let p_x be the probability of scoring x runs, x = 0-6, where

$$p_d + \sum_{0 \le x \le 6} p_x = 1 \tag{11.1}$$

These ps depend on the skill and batting style of the batsman, the state of the ball, the bowler, the run rate, etc. Apart from the run rate, these factors will be ignored for the time being. The run rate/ball r is the expected number of runs scored off each ball, and is given by

$$r = \sum_{0 \le x \le 6} (x.p_x) = 1 \tag{11.2}$$

To increase the run rate, a batsman will attempt to alter the distribution of the number of runs per ball, e.g. increase p_4 while reducing p_0. This would normally also increase p_d, the probability of dismissal. As run rate is usually discussed in terms of run rate per over, we shall use $R = 6r$.

11.4. First-innings formulation

For the first innings, let the stage n be the number of balls to go and the state i be the wickets in hand (i.e. the number of batsmen still to be dismissed). Let $f_n(i)$ be the maximum expected score under an optimal policy in the remaining n balls, with i wickets in hand. Note the important principle: no matter what the actual score, batsmen should be maximising the expected score in the remaining part of the innings. Each ball, a batsman either goes out and the team has one less ball and one less wicket in hand, or scores x runs and does not go out, so the team has one less ball to go and the same number of wickets in hand. Since he should choose R to maximise the expected score in the remaining n balls, we have

$$\begin{aligned} f_n(i) &= \underset{R}{Max}\left\{ p_d f_{n-1}(i-1) + \sum_{0 \le x \le 6} p_x (x + f_{n-1}(i)) \right\} \\ &= \underset{R}{Max}\left\{ p_d f_{n-1}(i-1) + R/6 + (1-p_d) f_{n-1}(i) \right\} \end{aligned} \tag{11.3}$$

using equations (11.1) and (11.2).

Since the innings finishes when there are either no more balls to be bowled or no wickets in hand, we get the boundary conditions

$f_0(i)=0$ for i = 0-l0
$f_n(0)=0$ for n = 0-300.

11.4.1. Evaluation of dismissal probabilities

For the first-innings formulation, we need only determine p_d, the probability of dismissal, as a function of the run rate R. It is generally accepted that p_d is an increasing function of R. Thus if a batsman attempts to score at two runs per over, he might have a 1% chance of dismissal each ball, whereas if he scores at 12 runs per over, he might have a 50% chance of dismissal. These probabilities might be estimated after a match by analysing the data, or before a match by expert opinion. In this case it is usually easier to determine $1/p_d$, the average number of balls faced before dismissal. Thus a 1% chance of dismissal implies facing on average 100 balls, whereas a 50% chance of dismissal implies on average facing only two balls. By looking at the expected score of a batsman or partnership, we can place further restrictions on p_d. By scoring at an expected rate of R per over for an average of $1/p_d$ balls, the average score before dismissal is $R/(6p_d)$. It is generally accepted that this is also a decreasing function of R for $R > 0$. In the above examples, average scores would be 33.3 and 4 for scoring rates of 2 per over and 12 per over. Thus any estimates of p_d derived from either expert opinion or statistical analysis should be adjusted to conform with this property.

For example, in the second final of the Benson & Hedges World Series Cricket played at Sydney on 11th February 1987, the six recognised English batsmen scored a total of 153 runs in 232 balls. Thus, at an average rate of 4 runs/over, a wicket was lost for each 39 balls. The seven recognised Australian batsmen averaged 3.4 runs/over for a wicket every 48 balls. In this manner, and using the property described above, we might guesstimate the figures shown in Table 11.1 for a reasonably difficult pitch.

11.4.2. Computation and results

Equations (11.3) can be solved by computer, calculating first $f_1(i)$ for all i = 1, 2,... 10, then $f_2(i)$, etc. A short program in interpreted BASIC takes about 45 minutes to run on an IBM XT. The output from the program gives the optimal action (the recommended run rate) and the value (total expected score in the remainder of the innings) for each stage and state (i.e. for each number of balls to go and wickets in hand). A selection is shown in Table 11.2.

TABLE 11.1. Dismissal probabilities

Scoring rate in runs/over (R)	Average number of balls faced ($1/p_d$)	Probability of dismissal each ball (p_d)	Average score before dismissal ($R/6p_a$)
0	300	0.003	0
1	250	0.004	42
2	100	0.010	33
3	60	0.017	30
4	40	0.025	27
5	28	0.038	23
6	20	0.050	20
7	15	0.067	18
8	10	0.100	13
9	7	0.143	11
10	4	0.250	7
11	3	0.333	6
12	2	0.500	4

11.4.3. Discussion

Since a team of 11 players must have at least five players who can bowl, plus a wicket-keeper, usually five players are selected solely for their batting ability. However, depending on the batting expertise of the bowlers and wicket-keeper, a team usually has at least six and sometimes eight or nine good batsmen. We assume here that if a team bats to number 7 (say), the first seven partnerships are the only ones that contribute to the score. In this case Table 11.2 shows the expected score to be 174 and the optimal scoring rate at the start of the match to be greater than the average scoring rate of 174/50 = 3.5. This is true no matter how many batsmen are in the team. Thus if it is assumed the team bat to number 10, the expected score is now 222, although the initial scoring rate is 5 per over. This holds at virtually all stages of the innings, and still holds if different pitch characteristics are tried (i.e. if the relationship between p_d and R is altered). This suggests that teams should try to score slightly faster than they expect their average rate for the rest of the innings to be, and if wickets are lost, slow up, rather than the current practice of scoring slower than average and speeding up if wickets are not lost. Thus the generally accepted view of scoring slowly at the beginning of the innings is not optimal under this model.

TABLE 11.2. Optimal run rate and expected score in remainder of innings under optimal policy

Optimal run rate								
Overs to go	Wickets in hand							
	2	4	5	6	7	8	9	10
1	9	12	12	12	12	12	12	12
5	6	8	9	9	9	9	11	11
10	4	7	7	7	8	8	9	9
20	3	5	5	6	6	7	7	7
25	3	4	5	5	5	6	7	7
30	2	4	4	5	5	5	6	6
40	2	3	4	4	4	5	5	5
50	1	3	3	3	4	4	4	5

Expected score in remaining balls								
Overs to go	Wickets in hand							
	2	4	5	6	7	8	9	10
1	9	12	12	12	12	12	12	12
5	27	39	42	45	47	49	51	53
10	38	59	67	73	77	82	85	88
20	49	83	96	107	117	126	134	141
25	53	91	106	119	131	142	152	160
30	56	97	114	129	142	155	166	176
40	60	106	126	144	160	175	189	202
50	63	113	135	155	174	191	207	222

Can we develop a simple rule for batsmen to follow? With 300 balls to go and seven recognised batsmen left, a team can afford to lose a wicket every 43 balls. Table 11.1 gives 4 as the nearest run rate with this dismissal probability, and Table 11.2 confirms this as the optimal run rate. Similarly, with 25 overs to go and five batsmen left, the batting team should aim to lose a wicket every five overs or 30 balls. Table 11.1 shows this is a run rate of about 5, which is again confirmed by Table 11.2 as being the optimal rate. This appears to hold also for tables derived using different dismissal probabilities, and could form the basis of a reasonable heuristic. The optimal run rate at any stage is the one that on average results in a wicket in the next *x* balls, where *x* is the ratio of the number of balls to go and the wickets in hand. Thus, for example, teams that bat to number 10 should begin the innings at a scoring rate that would, on average, lose a wicket in five overs.

Table 11.2 also allows the advantages of an extra batsman to be evaluated. By batting to number 8 rather than 7, a team could expect to increase its score from 174 to 191. If the advantage of an extra batsman or long batting tail is to be realised, an increased scoring rate is necessary right from the beginning of an innings. Table 11.2 highlights the folly of preserving wickets for a last-minute orgy of runs. The advantage of wickets in hand is minimal as the innings reaches the end. For example, five overs to go and six wickets in hand rather than four only results in an increase in expected runs from 39 to 45.

These comments hold not just for 'difficult' pitches. Under this model, the penalties for slow early batting in terms of foregone runs can be large for very good pitches. For example, in a one-day match against India on 7th September 1986, Australia scored 250 for 3, including a world-record opening partnership of 212 from 260 balls. However, an analysis similar to the above shows a score of over 350 should have been achieved. (In the second innings, India reached 251 for 3 off only 44 overs.)

Table 11.2 can also be used to compare the relative merits of alternative scores. For example, is it better to be 1 for 50 or 3 for 80 after 25 overs? Assuming a team bats to number 7, 1 for 50 should realise another 119 runs for a total of 169, whereas 3 for 80 should realise another 91 runs for a total of 171, marginally better. This allows the contribution of a batsman or partnership to a team to be assessed. Thus an opening partnership of 50 in 25 overs has actually decreased the expected score from 174 to 169. However, should the next partnership score 26 in the next five overs, they have increased the potential score from 169 to 76 + 96 = 172. Similar arguments could be applied to bowling performances. This method could be developed to produce measures of performance that reflect a player's contribution to team performance better than the currently used averages and run rate.

11.5. Second-innings formulation

When the second team bats, they know the total scored by the first team. If the first team scores 174 (say), the second team wins if it scores at least 175. For the second innings we wish to maximise the probability of achieving a certain score, and so need to introduce s, the number of runs to go, into the state.

Each ball, a batsman either goes out with probability p_d and the team still has s runs to score with one less wicket in hand and one less ball, or he scores x runs with probability p_x and so the team has s - x runs to score with one less ball to go and the same number of wickets in hand.

Thus if $P(s, i)$ is the probability, under an optimal policy, of scoring at least another s runs with i wickets in hand and n balls to go,

$$P_n(s,i) = \underset{R}{Max}\left\{ p_d P_{n-1}(s,i-1) + \sum_{0 \le x \le 6} p_x P_{n-1}(s-x,i) \right\} \qquad (11.4)$$

Since the second team wins when it has no more runs to score, but loses if it still has runs to score when there are no more balls to go or it has no more wickets in hand, we have the boundary conditions

$P_n(0, i) = 1$ for $n = 0 - 300$, i = 0 - 10
$P_0(s, i) = 0$ for $s = 1$ to s_{max}, i = 0 - 10
$P_n(s, 0) = 0$ for $n = 1 - 300$, $s = 1$ to s_{max}

11.5.1. Computation and results

For the first-innings problem, we saw that the actual distribution of the number of runs scored per ball is not required, only the average. For the second innings this is not so. However, apart from the last couple of overs, it is the average run rate rather than how these runs are scored that is important. Hence in this analysis, to reduce the computational load, only two values of x (0 and a) are allowed for each run rate. For each ball a batsman is either dismissed, or is not dismissed and scores 0 or a runs. Thus only p_d, p_0, and p_a are non-zero. Once p_d is given, provided a sensible value of a is chosen, p_0 and p_a can be determined by the equations (11.1) and (11.2), which give $p_a = R/6a$ and $p_0 = 1 - p_a - p_d$.

The second-innings equations can then be solved. A maximum value of s must be chosen (say 300), and is evaluated for s=1 and each value of i, then for s=2...300, then the values of $P_2(s, i)$ can be calculated, etc. The program had to be streamlined and compiled before it would run in about 10 hr on an IBM XT. The output now runs to thousands of pages, as for each ball (1-300) and each wicket in hand (1-10) we have for each number of runs to go (0-300) the chance of winning and the optimal run rate. For illustration, selected output in Table 11.3 shows the probability of winning at the start of the innings, i.e. with 300 balls to go.

TABLE 11.3. Probability of scoring a further s runs with 300 balls to go

	Wickets in hand							
s	2	4	5	6	7	8	9	10
100	0.200	0.638	0.805	0.907	0.961	0.985	0.995	0.998
125	0.094	0.432	0.622	0.776	0.881	0.942	0.975	0.990
150	0.042	0.262	0.430	0.600	0.743	0.848	0.917	0.958
175	0.016	0.137	0.257	0.403	0.555	0.691	0.799	0.878
200	0.006	0.065	0.137	0.241	0.367	0.502	0.629	0.739
220	0.002	0.031	0.073	0.140	0.234	0.348	0.471	0.591
225	0.002	0.025	0.061	0.120	0.206	0.312	0.431	0.551
250	0.000	0.009	0.024	0.053	0.101	0.170	0.258	0.359
275	0.000	0.003	0.008	0.020	0.043	0.079	0.132	0.200
300	0.000	0.001	0.003	0.007	0.017	0.034	0.061	0.100

11.5.2. Discussion

The comments made on the first innings in general also apply to the second, but the evaluation of team position, or contribution of batsmen or partnerships, would now be reflected in a change in the probability of winning rather than runs scored.

Note that the second team has an inherent advantage. If the first team score their expected maximum of 174, at the beginning of their innings the second team has a 0.555 chance of scoring 175 if they also bat to number 7. This is assuming both teams have full knowledge of the state of the wicket, i.e. the value of the p_ds. In practice, the first team would begin their innings with much less knowledge of the state of the wicket than the second. This would increase the advantage to the second team.

Output from the analysis could also be used to evaluate the effect of rain-interruption rules. Suppose, after the first innings is completed, rain causes a delay necessitating a reduction in the second innings to 20 overs. What is a fair target for the second team? A commonly used rule is that they should score at the same rate as the first team. If the first scored 174 (at a rate of 174/50 = 3.48 runs/over), this would give a target of 3.48*20 = 69.6 or 70 runs for the second team. However, a similar output to Table 11.3 with 120 balls to go shows the second team would have a probability of 0.974 of reaching that target, clearly an unacceptable advantage.

11.6. Extensions

The model we have chosen is a simple one. Cricket followers could easily suggest areas where it does not conform to reality. For example, many teams may bat during the first innings with a second-innings strategy in that they wish to maximise their chance of reaching some preconceived total. Other variables could be taken into account provided the effects could be quantified. For example, by altering the dismissal and scoring probabilities with respect to n, we can take account of different bowlers and ball deterioration. This involves no change in the formulation, and is easily incorporated in the program. In a similar way, by varying the probabilities with i, we could take account of different batsmen (or more correctly partnerships) without altering the basic formulation.

To account properly for different batsmen, we could introduce both i and j into the state, being the strike and non-strike batsmen. For the first five balls of an over, (i, j) would change to (j, i) when an odd number of runs were scored, remain (i, j) when an even number of runs were scored, and become $(\max(i,j) + 1, j)$ when a batsman was dismissed without scoring or crossing, etc. For the final ball of the over, the transition (i, j) to (j, i) due to the change of bowling ends would be superimposed on any other changes. The extra complication of such a model might be justified when evaluating tactics at the end of an innings, particularly with one free-scoring batsman and one poor batsman. For example, should the better batsman take a single?

Other extensions, such as allowing the dismissal probabilities to vary with time at the crease, involve including this factor as a state variable, and result in large increases in computing time.

11.7. Conclusion

Previously, cricket has escaped the attention of OR analysts. The model presented here shows that currently accepted tactics in one-day cricket may be incorrect. Batting sides should score more quickly in the early part of their innings. There is also evidence that teams should choose to bat second when they win the toss.

The model could also be used by selectors to quantify the effects of including extra batsmen in a team, used by coaches, captains and commentators (or bookmakers) to provide better measures of how teams are performing during the match, assist in evaluating different rules for deciding winners in rain-interrupted matches, and develop measures of player performance that better reflect the demands of one-day cricket.

More complicated models could be developed, allowing for different player characteristics. Such models could be used to investigate optimal tactics near the end of an innings, the effects of different batting orders, etc.

There is plenty of scope for operational research on applications in cricket. The major problem likely to be encountered is data collection. Official score sheets of matches contain little information of a ball-by-ball nature, and the information recorded even varies from scorer to scorer. However, the development of computer scoring methods should solve this. One proposal, CRICKETSTAT, Croucher (1987) developed in Australia, records 11 pieces of information for each ball. When such systems are common, operational researchers will have few excuses for not using their skills to assist cricketers.

11.8. Commentary. Testing of heuristic

This chapter suggested a simple heuristic that could be used by batsmen to approximate the optimal run rate. The suggested heuristic selects the run rate that will result in all wickets being lost as near as possible to the end of the 50th over, i.e. a wicket loss every (balls to go / wickets in hand) balls. For instance if there are 120 balls (20 overs) remaining in the innings and six wickets in hand then the heuristic suggests losing a wicket every 120/6=20 balls. Table 11.1 shows that the run rate that would result in the loss of a wicket in as close to 20 balls as possible is six runs per over. Therefore if there were 120 balls remaining in the innings and six wickets in hand the heuristic suggests a run rate of six runs per over, the same rate as given by Table 11.2.

In Johnston et al. (1992) we investigated the sensitivity of the heuristic to changes in the model.

The heuristic's validity was first tested by comparing the run rates suggested by the heuristic with those calculated by the DP formulation. Table 11.1 was extended to include run rates in increments of 0.5 per over, and the algorithm used to produce the optimal run rate and expected score. Table 11.4 shows the difference between the run rates recommended by the heuristic and the DP solution. The two run rates never vary by more than 0.5 runs per over.

Another test of the heuristic's ability is to compare the expected innings scores of teams scoring at heuristic run rates with the expected innings scores of teams scoring at DP run rates. In this case, the maximum difference (in runs) in the expected scores between the heuristic and DP was 0.3 runs. The closeness of the expected innings scores under the DP formulation and heuristic confirms the conclusion that the heuristic is a valid method of selecting near optimal run rates.

Although all the comparisons between the heuristic and DP formulation shown here have been made under the one relationship between run rate and probability of dismissal (that given in Table 11.1), several other relationships have been tested with almost identical results. The maximum difference between the totals under the two policies was less than four runs.

TABLE 11.4. Graphical comparison of DP and heuristic run rates

Overs to go	Wickets in hand 1	2	3	4	5	6	7	8	9	10
1										
5		+								
10			+	+						
15	-				+					
20					+					
25										
30		-								+
35	-			-	-					+
40		-								
45			-	-						
50										

: a blank indicates agreement

+ : the heuristic is 0.5 runs/over above the optimal

- : the heuristic is 0.5 runs/over below the optimal

DYNAMIC PROGRAMMING IN CRICKET - PROTECTING THE WEAKER BATSMAN

12.0. Abstract

A simple dynamic programming model of cricket is presented. The state is the facing batsmen and the number of runs on offer. The decision is whether to run or not, with the objective to maximise the chance the better batsman is on strike at the start of the next over. The model is solved analytically to find the optimal policy and the value of the objective function. The simple initial model is extended to a more realistic one requiring no further calculations and a numerical example is given. An alternative optimality criterion is investigated and we demonstrate that trying to put the better batsman on strike at the start of the over does not necessarily maximise the expected duration of the partnership. This alternative objective function is investigated numerically, and it is shown that the better batsman should generally run if possible off the second last or last ball of the over.

Key words: sports, cricket, dynamic programming, Markov processes

12.1. Introduction

There are several examples where a dynamic programming (DP) formulation has the potential to assist the sports person with decision making. Norman (1995) in giving one example of an application of DP in each sport lists 10 papers. There have been few applications of DP in cricket, which is surprising as the ball by ball nature of the game should lend itself to this structure. Clarke (1988b) uses a DP formulation to advise on optimal run rates in both the first and second innings, and Johnston (1992), Johnston et al. (1992, 1993) use the first innings formulation to provide measures of a batsman's performance. There may, of course, be several possible optimality criteria depending on the different stages of a cricket match. In the first innings batsmen are generally trying to maximise the expected number of runs, but they may prefer to maximise the probability of achieving a certain number of runs. The team batting last is usually trying to maximise the probability of achieving the opponent's score. However, in test cricket, teams may often be just trying to avoid a loss. In this case they may wish to avoid finishing in a certain state (team dismissed) with runs being immaterial. In other situations teams may wish to bat for as long as possible. For example, in the fourth test between Australia and the West Indies played on the 29th April to 3rd May 1995, in Australia's first innings on the third day, Steve Waugh was the last batsman dismissed for 200. His partners, after the last specialist batsman was dismissed, made 6, 8, 23, 0 and 3 not out. While runs were still important, some commentators made the point that it was also important to occupy the crease for as long as possible, to give the pitch time to break up and so assist the Australian spinner.

The above situation, where a top order batsman is paired with a batsman of lesser ability, often arises towards the end of an innings. In cricket, the facing batsman changes whenever an odd number of runs is scored, and also at the completion of each six ball over. The bowling team wish to bowl to the weaker batsman and will often set deep fields to concede a single to the good batsman early in the over. The good batsman in turn will sometimes decline to take the single, in the hope of protecting the weaker batsman for a few balls and taking a run nearer the end of the over. The desired result is to take a single off the last ball, so the better batsman is again on strike at the beginning of the next over. In most such situations runs are still important, but in other cases runs are immaterial except in that they allow the batsmen to change ends. Describing the last session of play in the famous drawn test between Australia and the West Indies in 1961, Lunn (1993) says

> But with half an hour to go Mackay was no longer the only person who thought he could do it. Taking just a single off the last ball of almost every over (eight ball overs in those days) he faced almost every ball instead of Kline. No attempt to score other than this. ... The last ball of the second last over: Mackay scores a single to face the last over against the fastest man in the world, giant Wes Hall.

In this case Mackay would clearly bat out the last over without taking a run. But in the second last over, if he wants to maximise the chance that he will be on strike for the final Wes Hall over, should he wait until the last ball before taking a single? We look here at a simple DP model to analyse this end play strategy - i.e. maximise the probability that the weaker batsman finishes the over on strike, so that after the change of ends he will be protected from the strike for the beginning of the new over. We leave until later discussion on whether this is a sensible objective function.

12.2. The model

Suppose you have two batsmen whom we will refer to as **G** and **B** for **G**ood and **B**ad, although they can be any two batsmen of different abilities such as **G**reg Chappel and **B**ruce Reid. We assume each batsman can score zero, one (more correctly hit a stroke for which he has the opportunity to run a single,) or be dismissed. (While this initial model is obviously not realistic it is presented for simplicity. We will see that it can be extended with no further calculations to something more realistic). For the good batsman these occur with probability p_0, p_1 and p_d, and for the bad batsman q_0, q_1 and q_d . Let $p = 1\text{-}p_d = p_0 + p_1$, $q = 1\text{-}\ q_d = q_0 + q_1$ and $p>q$ since the poor batsman has a greater chance of dismissal. There are nine wickets down, so once a batsman is dismissed the innings is **E**nded (E).

Since the batsmen have to decide whether to take the runs on offer or not, consider the situation *after* a ball is bowled but before a run is taken. The stage n is the number of balls still to be bowled in the over. There are five possible states S_n ε {G0, G1, B0, B1, E } representing the facing batsman and the runs on offer. Note that because the epoch in which the state is defined and decision is made is after the ball is bowled, stage 0 refers to the last ball of the over and stage 5 to the first ball in a six ball over. The only decisions occur at G1 and B1 and are whether the batsmen take the run on offer YES (Y) or not NO (N).

The transition diagram is shown in Figure 12.1.

The transition probabilities are easily calculated provided we ignore complicating factors such as runouts. Thus for example at G1, if the good batsman refuses the run, he will be on strike for the next ball and so after it is bowled there will be a probability p_0 of no run on offer (so with probability p_0 the next state is G0). Similarly, with probability p_1 the next state is G1 and with probability p_d the next state is E. Exactly the same transition probabilities arise if the state is B1 and the batsmen take the run. Thus in the transition diagram all lines entering G0 have probability p_0, all entering G1 have probability p_1, all entering B0 are q_0, and all entering B1 are q_1. Those lines entering E from G0, G1(N) and B1(Y) have probability p_d, those from B0, B1(N) and G1(Y) have probability q_d.

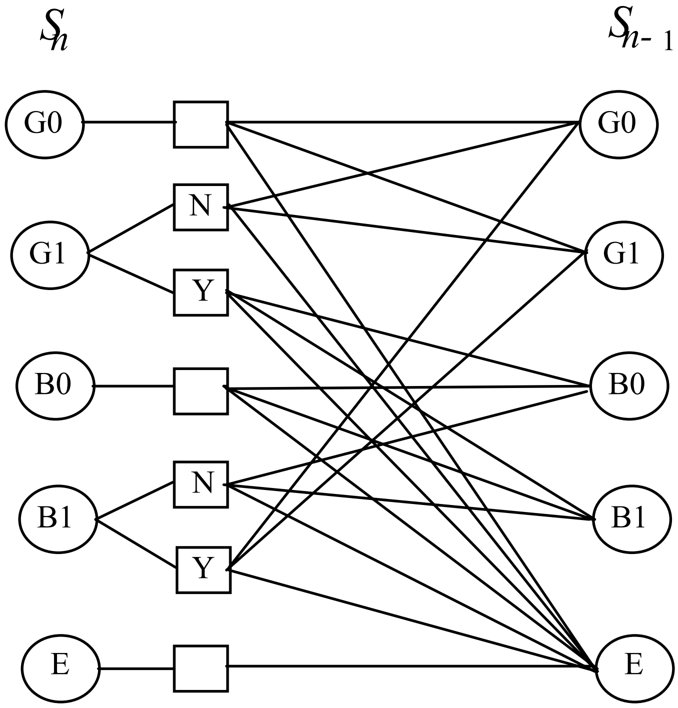


Figure 12.1. Transition diagram

We wish to maximise the chance that at the end of the over the bad batsman will be on strike. Note there is no need to consider in any special way the change of ends at the end of the over. The batsmen wish to maximise for each over the chance that the bad batsman finishes the over on strike. Whether they succeed or not, they have exactly the same problem in the following over.

Let $f_n(S_n)$ be the probability under an optimal policy of ending the over with the bad batsman on strike when in state S_n with n balls to go.

Initial conditions: For the last ball of the over, the bad batsman will not run, and the good batsman will definitely run if possible, so any of the states B0, B1, and G1 will put the bad batsman on strike at the end of the over. Thus

$$f_0(\mathrm{G0}) = f_0(\mathrm{E}) = 0$$
$$f_0(\mathrm{B0}) = f_0(\mathrm{B1}) = f_0(\mathrm{G1}) = 1 \qquad (12.1)$$

Functional Equations: In general

$$f_n(S_n) = \underset{\substack{\text{admissible}\\ \text{decisions}}}{Max} \sum_{S_{n-1}\,\varepsilon\{G0,G1,B0,B1,E\}} \text{prob}(S_n \to S_{n-1}) f_{n-1}(S_{n-1}) \text{ for } n = 1, 2, \ldots$$

where admissible decisions are defined in Figure 12.1.

In the particular cases this becomes

State E: $\quad f_n(\text{E}) = f_{n-1}(\text{E})$ □ $f_n(\text{E}) = f_0(\text{E}) = 0 \qquad (12.2)$

State G0: $\quad f_n(\text{G0}) = p_0 f_{n-1}(\text{G0}) + p_1 f_{n-1}(\text{G1}) + p_d f_{n-1}(\text{E})$

$$= p_0 f_{n-1}(\text{G0}) + p_1 f_{n-1}(\text{G1}) \qquad (12.3)$$

State B0: $\quad f_n(\text{B0}) = q_0 f_{n-1}(\text{B0}) + q_1 f_{n-1}(\text{B1}) \qquad (12.4)$

State G1: $\quad f_n(\text{G1}) = \text{MAX}\begin{cases}\text{YES}: q_0 f_{n-1}(\text{B0}) + q_1 f_{n-1}(\text{B1})\\ \text{NO} \ : p_0 f_{n-1}(\text{G0}) + p_1 f_{n-1}(\text{G1})\end{cases}$

$$= \text{MAX}\begin{cases}\text{YES}: f_n(\text{B0})\\ \text{NO} \ : f_n(\text{G0})\end{cases} \text{ from (12.3), (12.4)} \qquad (12.5)$$

State B1: $\quad f_n(\text{B1}) = \text{MAX}\begin{cases}\text{YES}: p_0 f_{n-1}(\text{G0}) + p_1 f_{n-1}(\text{G1})\\ \text{NO} \ : q_0 f_{n-1}(\text{B0}) + q_1 p_{n-1}(\text{B1})\end{cases}$

$$= \text{MAX}\begin{cases}\text{YES}: f_n(\text{G0})\\ \text{NO} \ : f_n(\text{B0})\end{cases} \qquad (12.6)$$

So $\quad f_n(\text{G1}) = f_n(\text{B1}) \quad$ for $n = 1, 2\ldots$ from (12.5) & (12.6),
and if optimal decision is YES at G1 then it is NO at B1 and vice versa. (12.7)

This is a sensible result. Since runs are not important, clearly if it is optimal for the good batsman to change ends at some stage, it is optimal for the bad batsman not to change ends at the same stage. From now we will refer to the optimal policy as simply YES or NO, meaning the decision at G1 only, as B1 is implied to be the opposite.

12.3. Model solution

This model is completely solvable analytically. The appendix shows that the values of the objective functions can be calculated in a closed form for all stages and states.

Case 1: If $q \le p_1$ the optimal decision is NO for all $n > 0$. The better batsman should obtain and remain on strike for the whole over and only run off the last ball. In this case $f_n(\mathrm{G1}) = f_n(\mathrm{G0}) = f_n(\mathrm{B1}) = p^{n-1} p_1$

$$f_n(\mathrm{B0}) = q_0^{n-1} q + \frac{q_1 p_1}{q_0 - p}\left(q_0^{n-1} - p^{n-1}\right).$$

Case 2: If $q > p_1$ there is a stage $N \ge 1$, below which the decision at G1 is YES, and for $n \ge N$ the optimal decision is NO. Thus there will always be a number of balls to go in the over (which may be greater than 6 or 8), before which the optimal strategy is to protect the weaker batsman, and after which the optimal strategy is to put the weaker batsman on strike. N is given by the smallest integer satisfying the inequality

$$n > \frac{\ln\left(\dfrac{p-q}{p_1}\right)}{\ln\left(\dfrac{p_0}{q}\right)}.$$

For $n = 1,2,3,\ldots N\text{-}1$, $\quad f_n(\mathrm{G0}) = \dfrac{p_1}{p_0 - q}\left(p_0^{\,n} - q^n\right)$

$$f_n(\mathrm{G1}) = f_n(\mathrm{B0}) = f_n(\mathrm{B1}) = q^n.$$

For $n = N, N+1, \ldots$ $\quad f_n(\mathrm{G0}) = f_n(\mathrm{G1}) = f_n(\mathrm{B1}) = \dfrac{p^{n-N} p_1}{p_0 - q}\left(p_0^{\,N} - q^N\right)$

$$f_n(\mathrm{B0}) = q_0^{n-N} q^N + \frac{q_1}{q_0 - p}\left(q_0^{n-N} - p^{n-N}\right) f_N(\mathrm{G0}).$$

12.4. Prior probabilities

It could be argued that probabilities before a given ball is bowled would be more useful. These are easily obtained by weighting the above state probabilities with the chances of them arising. Thus if we let $F_n(\mathrm{G})$ and $F_n(\mathrm{B})$ be the probabilities of finishing in the required state with G and B facing *before* the *n*th last ball of the over is bowled we have

$$F_n(\mathrm{B}) = q_0 f_{n\text{-}1}(\mathrm{B0}) + q_1 f_{n\text{-}1}(\mathrm{B1}) = f_n(\mathrm{B0}) \text{ from equation (12.4)}$$

Similarly $$F_n(\mathrm{G}) = f_n(\mathrm{G0}).$$

Thus although in the above formulation the stage n only goes to 5, it is useful to calculate $f_6(\mathrm{G0})$ and $f_6(\mathrm{B0})$ as they give the probabilities before the beginning of the over.

12.5. Extension

With no further calculations this model can be extended by considering the state G0 to be the good batsman facing and a score of zero or a boundary (four or six) is made, so no decision on running can be made; G1 that there are 1, 2, 3 runs or a 4 all run on offer so that a decision on running is possible; and the decision to be made is NO to 'run' an even number (quotes because it includes not taking any runs when 1 is on offer) or YES run an odd number of runs. These decisions could be rephrased as NO don't change ends and YES change ends. Thus p_0 becomes the probability of a score of 0, 4 (boundary) or 6, p_1 becomes the probability of a score of 1, 2, 3 or 4 all run, and p_d remains the probability of dismissal. Similar states and probabilities apply for the bad batsman. The state E can be thought of as the partnership ending rather than the innings ending. The above model and results are then directly applicable.

12.6. Discussion

The probabilities p_0 etc should be estimable from scoresheets when batsmen are taking all available runs. The optimal strategy can then be calculated. This is done in the example following. However some of the ramifications of the above seem quite interesting and can be used to make general statements irrespective of the scoring profiles of the batsmen.

For example, equation 12.7 says that if at any stage, one batsman takes an odd number of runs, at the same stage the other batsman should take an even number of runs. Thus the policy of taking all the runs is *never* optimal. (Except of course when $p = q$).

If $q > p_1$ then the model says that there is always some value N balls to go (admittedly N may be greater than 5) where the optimal strategy changes from NO in one ball to YES for the next. Now if N is greater than 5 this implies the weak batsman will never be protected from the strike (in fact he should be given it). At any stage in the over the good batsman will put the weak on strike if possible, and for the rest of the over the weaker batsman will stay on strike. Note this only occurs when p_1 is so low that it is highly unlikely that the good batsman will ever be able to put the bad batsman back on strike. This seems unlikely to occur in practice. (It also suggests that we should perhaps look at other objective functions, such as minimise the proportion of times we are in a given state, or maximise the chance of not finishing in State E).

However if N is any number less than 6, the results confirm what is done in practice. It implies that at some stage in the over the optimal decision changes from NO (protect the weak batsman) to YES (put the weak batsman on strike). This is the common strategy used but has an interesting implication when the bad batsman is on strike at the beginning of the over. It implies the bad batsman if on strike should change ends if possible, only to again change ends the next ball if possible. Table 12.1 gives the optimal policy for the good and bad batsman when N=3. If the bad batsman is on strike, he should change ends if possible with 3 balls to go. In this case the good batsman will be on strike and should now also immediately change ends if possible with two balls to go. This would probably be criticised by commentators.

TABLE 12.1. Optimal strategies for the case N=3.

Ball of over	Balls to go n	Good Batsman	Bad batsman
1	5	No	Yes
2	4	No	Yes
3	3	No	Yes
4	2	Yes	No
5	1	Yes	No
6	0	Yes	No

12.6.1. Example

In the World cup match between SA and Australia on 26th Feb 1992, for two batsmen we had from the official score sheets:

Steve Waugh - 1311121114121111112, 27 runs from 51 balls

Bruce Reid - 1211 - 5 runs from 10 balls.

Although Bruce Reid was not out, we might give him an honorary dismissal and from these figures estimate the probabilities as follows.

Steve Waugh: $p_d = 1/51 = 0.02$,
p_1 = probability of 1, 2 or 3 = 18/51 = 0.35,
p_0 = probability of 0, 4 or 6 = 32/51 = 0.63

Bruce Reid: $q_d = 1/10 = 0.1$,
$q_1 = 4/10 = 0.4$,
$q_0 = 5/10 = 0.5$

Now $q = 0.9 > p_1 = 0.35$, $N >$ ln(0.08/0.35)/ln(0.63/0.9)= 4.14, so $N = 5$. Thus if on strike at the beginning of the over, Steve Waugh should only take an even number of runs for the first ball, but an odd number thereafter. If Reid is on strike he should take an odd number of runs for the first ball, but an even number thereafter. If Waugh is on strike at the start of the over then the chance of Reid finishing the over on strike is given by

$$F_6(\text{G}) = f_6(\text{G0}) = pf_5(\text{G0}) = p\frac{p_1}{p_0 - q}(p_0^5 - q^5) = 0.624$$

If Reid starts the over on strike his chance of finishing on strike is

$$F_6(\text{B}) = f_6(\text{B0}) = q_0 q^5 + q_1 f_5(\text{G0}) = 0.550$$

The optimal policy to put Reid on strike very early is here caused by the high probability of Waugh scoring zero. If we alter p_0 to 0.35 and p_1 to 0.63, we get $N = 3$, and the optimal strategy would revert to that given in Table 12.1. The new values of Reid finishing on strike become 0.740 and 0.626. Note how the probability of success alters depending on the good batsman's distribution of probability. Two batsmen could have the same average and run rate, but the one that has a higher probability of scoring runs rather than boundaries is the more flexible and may be a better batsman with lower order players. Maybe more extensive statistics on cricketers should be published. This has also been suggested in the context of measuring consistency by Clarke (1991b; 1994c).

12.7. An alternative criterion of optimality

By providing a counter example we show that attempting to put the better batsman on strike at the start of the over does not necessarily maximise the expected duration of the partnership.

Consider a Markov chain where the states are B, G and E being the facing batsman when a ball is bowled and the partnership Ended. If the policy is NO (for the good batsman, implying YES for the bad batsman), then we have using the probabilities as before of a good and bad batsman scoring runs, the following transition matrix.

$$P_{\mathrm{N}} = \begin{array}{c} B \\ G \\ E \end{array} \begin{pmatrix} q_0 & q_1 & q_d \\ 0 & p & p_d \\ 0 & 0 & 1 \end{pmatrix}$$

For example, for the good batsman, the probability of not being dismissed is $p_0 + p_1 = p$, which as he does not run is the probability of the good batsman being on strike next ball.

If the policy is YES we get the following transition matrix.

$$P_{\mathrm{Y}} = \begin{array}{c} B \\ G \\ E \end{array} \begin{pmatrix} q & 0 & q_d \\ p_1 & p_0 & p_d \\ 0 & 0 & 1 \end{pmatrix}$$

Then using the example given above with Waugh and Reid, we have

$$P_{\mathrm{N}} = \begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0 & 0.98 & 0.02 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad P_{\mathrm{Y}} = \begin{pmatrix} 0.9 & 0 & 0.1 \\ 0.35 & 0.63 & 0.02 \\ 0 & 0 & 1 \end{pmatrix}$$

This gives, for the optimal policy of NO for the first ball and YES for the last 5 balls a

transition matrix for the over of $P_N^1 P_Y^5 = \begin{pmatrix} 0.550 & 0.040 & 0.410 \\ 0.624 & 0.097 & 0.279 \\ 0 & 0 & 1 \end{pmatrix}$

Note the answers 0.550 and 0.624 agree with that given above for the bad batsman ending the over on strike.

Alternatively, if we look at a strategy of the good batsman running if possible only off the last ball, i.e. NO from the first 5 balls and YES from the last ball, we get a transition matrix for the over of $P_N^5 P_Y^1 = \begin{pmatrix} 0.283 & 0.458 & 0.259 \\ 0.316 & 0.569 & 0.114 \\ 0 & 0 & 1 \end{pmatrix}$

This certainly gives much lower probabilities of the bad batsman ending the over on strike. It also (not surprisingly) gives lower probabilities of the partnership ending by the end of the over. Our hope was these would be cancelled out by the higher chance of the partnership ending earlier because the bad batsman has a higher probability of facing up to the next over.

However, by swapping the first and second columns to allow for the change of ends at the completion of the over, we now have a transition matrix where the stage is an over and the states are the batsmen on strike at the beginning of the over. This gives for the optimal policy

$\begin{pmatrix} 0.040 & 0.550 & 0.410 \\ 0.097 & 0.624 & 0.279 \\ 0 & 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0.458 & 0.283 & 0.259 \\ 0.569 & 0.316 & 0.114 \\ 0 & 0 & 1 \end{pmatrix}$ for the other policy.

The expected number of times in each state can be calculated by applying normal methods for absorbing Markov chains. The fundamental matrix gives for the two cases $\begin{pmatrix} 1.222 & 1.788 \\ 0.316 & 3.123 \end{pmatrix}$ and $\begin{pmatrix} 3.264 & 1.349 \\ 2.719 & 2.587 \end{pmatrix}$. Again, the smaller numbers in the first column show the optimal strategy clearly produces a much lower expected number of times the bad batsman is on strike at the beginning of the new over. However, the expected number of completed overs the partnership will last is given by summing the rows to give

$\begin{pmatrix} 3.011 \\ 3.439 \end{pmatrix}$ and $\begin{pmatrix} 4.613 \\ 5.306 \end{pmatrix}$. Thus the second case gives longer expected partnership times, whoever is facing the first over, than the previous optimal policy. These calculations can be repeated for each of the seven possible policies and are shown in Table 12.2. This shows that to optimise the length of the partnership the good batsman should run on the fifth ball, although there is little difference in the two options of running on the fifth or sixth ball.

TABLE 12.2. Expected number of completed overs partnership lasts if good batsman runs on ball n

Facing batsman at beginning of over	n						
	1	2	3	4	5	6	Never
Bad	2.573	3.011	3.545	4.145	4.629	4.613	3.892
Good	2.960	3.439	4.050	4.753	5.325	5.306	4.447

The best policy under the criteria of maximising the expected number of completed overs was investigated numerically in this manner for a range of values. Using SAS/IML the optimal ball on which the good batsman should run was calculated for each value of p_d, q_d = 0.01 to 0.10 in steps of 0.01, $p_d < q_d$; p_0, q_0 = 0.1 to 0.9 in steps of 0.1. In 65% of these cases the optimal strategy was to run off the last ball, 32% the second last ball with the remaining 3% of cases giving the 4th ball. Thus the simple strategy of the good batsman getting off the strike if possible on the second last or last ball of the over, and the bad batsman doing the same off the first four balls, generally optimises the expected number of completed overs the partnership will last.

12.8. Conclusion

A simple DP model was set up to solve a specific and very limited problem. The model can be solved completely analytically, and the solution used to suggest a suitable strategy for every over except the last. In practice, the model is probably deficient in a couple of respects. The fielding side often set widespread fields to the good batsman early in the over to encourage a single, and bring the field in later in the over to prevent the batsmen taking a single. In this case the values of p_0 etc would depend on the stage n as well as the state. The model would still possibly be soluble analytically, and certainly numerically. However the model can also suggest reasons why some cricketers' scoring profile could make them more suited than others to certain situations such as playing with tail-enders. Commentators often comment on the ability (or lack of it) of a player at rotating the strike, but commonly kept statistics do not measure this.

It is clear that minimising the exposure of the weaker batsman to the first ball of the new over is not necessarily the appropriate objective function. It does not necessarily minimise his exposure to the strike, the chance of the partnership ending within a certain number of balls, or maximise the number of balls until the partnership is broken. The optimal strategy to maximise the number of completed overs for the partnership can be found numerically, and usually requires the better batsman running an odd number of runs if possible on the second last or last ball of the over.

The ball by ball nature of cricket makes it particularly suitable for a DP approach. Clarke & Norman (1997c) have constructed several other models using alternative objective functions such as maximising the expected number of runs in the remainder of the innings, which take into account the number of runs, run rate, the number of wickets down and the change of ends between overs. Such models could be used to assist with end play strategies in both the first and second innings.

Note: An earlier version of this paper was presented at the 13th ASOR conference in Canberra, 1995.

Appendix 12.1.

The functional equations (12.1) to (12.6) can be solved to obtain analytical solutions.

Theorem 1. If there is some stage n for which the optimal decision is NO, then it is also NO for the previous stage n + 1.

Proof: If decision is NO at n then from (12.5)

$$f_n(\mathrm{G0}) > f_n(\mathrm{B0}) \tag{12.8}$$

and $$f_n(\mathrm{G1}) = f_n(\mathrm{G0}) = f_n(\mathrm{B1}) \tag{12.9}$$

Then $f_{n+1}(\mathrm{G0}) - f_{n+1}(\mathrm{B0})$

$$\begin{aligned}
&= p_0 f_n(\mathrm{G0}) + p_1 f_n(\mathrm{G1}) - q_0 f_n(\mathrm{B0}) - q_1 f_n(\mathrm{B1}) && \text{from (12.3) and (12.4)}\\
&\geq p_0 f_n(\mathrm{G0}) + p_1 f_n(\mathrm{G0}) - q_0 f_n(\mathrm{G0}) - q_1 f_n(\mathrm{G0}) && \text{from (12.8) and (12.9)}\\
&= (p_0 + p_1 - q_0 - q_1)f_n(\mathrm{G0})\\
&= (p - q) f_n(\mathrm{G0})\\
&> 0 \quad \text{since} \quad p > q
\end{aligned}$$

ie. $f_{n+1}(\mathrm{G0}) > f_{n+1}(\mathrm{B0})$ and by (12.5) the optimal decision at n+1 is NO. ■

Consider the stage n= 1

$$\begin{aligned}
f_1(\mathrm{B0}) &= q_0 f_0(\mathrm{B0}) + q_1 f_0(\mathrm{B1}) \quad \text{from (12.4)}\\
&= q_0 + q_1 \quad \text{from (12.1)}\\
&= q && (12.10)\\
f_1(\mathrm{G0}) &= p_0 f_0(\mathrm{G0}) + p_1 f_0(\mathrm{G1}) \quad \text{from (12.3)}\\
&= p_1 \quad \text{from (1)} && (12.11)
\end{aligned}$$

So using (12.5), if $q > p_1$ then decision is YES and $f_1(\mathrm{G1}) = q$

and if $q \leq p_1$ then decision is NO and $f_1(\mathrm{G1}) = p_1$. (12.12)

We thus have two cases.

Case 1: $q \le p_1$

From (12.12) and Theorem 1, decision is always NO and so for $n \ge 1$

$$\begin{aligned}
f_n(\mathrm{G1}) &= f_n(\mathrm{G0}) \quad \text{from (12.5)} \\
&= p_0 f_{n\text{-}1}(\mathrm{G0}) + p_1 f_{n\text{-}1}(\mathrm{G1}) \quad \text{from (12.3)} \\
&= (p_0 + p_1) f_{n\text{-}1}(\mathrm{G0}) \quad \text{from above} \\
&= p f_{n\text{-}1}(\mathrm{G0}) \\
&= p^{n\text{-}1} f_1(\mathrm{G0}) \\
&= p^{n\text{-}1} p_1 \quad \text{from (12.11)} \\
f_n(\mathrm{B1}) &= f_n(\mathrm{G0}) = p^{n\text{-}1} p_1 \quad \text{from (12.6)} \\
f_n(\mathrm{B0}) &= q_0 f_{n\text{-}1}(\mathrm{B0}) + q_1 f_{n\text{-}1}(\mathrm{B1}) \quad \text{from (12.4)} \\
&= q_0 f_{n\text{-}1}(\mathrm{B0}) + q_1 p^{n\text{-}2} p_1 \quad \text{from above} \\
&= q_0^{n\text{-}1} q + \frac{q_1 p_1}{q_0 - p}(q_0^{n-1} - p^{n-1})
\end{aligned}$$

Case 2: $q > p_1$

From (12.10) decision at n=1 is YES and $f_1(\mathrm{G1}) = q$. Let $n = N$ be the first time the decision is NO, then for n = 1,2,... N–1 the decision is YES and

$$f_n(\mathrm{G1}) = f_n(\mathrm{B0}) = f_n(\mathrm{B1}) \text{ for } n = 1,2,...N\text{-}1 \quad \text{from (12.5)} \tag{12.13}$$

So for n = 1 to N

$$\begin{aligned}
f_n(\mathrm{B0}) &= q_0 f_{n-1}(\mathrm{B0}) + q_1 f_{n\text{-}1}(\mathrm{B1}) \quad \text{from (12.4)} \\
&= q_0 f_{n-1}(\mathrm{B0}) + q_1 f_{n\text{-}1}(\mathrm{B0}) \\
&= q f_{n-1}(\mathrm{B0}) \\
&= q^2 f_{n-2}(\mathrm{B0}) .... = q^n f_0(\mathrm{B0}) \\
&= q^n
\end{aligned} \tag{12.14}$$

So by (12.13), for n = 1, 2.....N-1,

$$f_n(\mathrm{G1}) = f_n(\mathrm{B0}) = f_n(\mathrm{B1}) = q^n \tag{12.15}$$

Now $\quad f_1(\mathrm{G0}) = p_1$ by (12.11)

So $\quad f_n(\mathrm{G0}) = p_0 f_{n\text{-}1}(\mathrm{G0}) + p_1 f_{n\text{-}1}(\mathrm{G1})$

$$\begin{aligned}
&= p_0 f_{n\text{-}1}(\mathrm{G0}) + p_1 q^{n\text{-}1} \quad \text{by (12.15)} \\
&= \frac{p_1}{p_0 - q}(p_0^n - q^n) \quad \text{for } n = 1, 2, 3...., N \text{ by induction}
\end{aligned} \tag{12.16}$$

Now, since N is smallest n for which decision at G1 is NO, we have from (12.5)

$$
\begin{aligned}
f_N(\mathrm{G0}) &> f_N(\mathrm{B0}) \\
\frac{p_1}{p_0-q}\,[p_0^N - q^N] &> q^N \quad \text{from (12.14), (12.16)} \\
p_1\, p_0^{\,N} - p_1\, q^N &> p_0\, q^N - q^{N+1} \quad \text{where sign reverses if } p_0 < q \\
\Rightarrow p_1 p_0^N &> (p_0 + p_1)\, q^N - q^{N+1} \\
&= (p - q)\, q^N \\
\Rightarrow \left(\frac{p_0}{q}\right)^N &> \frac{p-q}{p_1} \\
N &> \frac{\ln\frac{p-q}{p_1}}{\ln\frac{p_0}{q}} \quad \text{where sign reverses back if } p_0 < q \\
&= \frac{\ln\frac{q_d - p_d}{p_1}}{\ln\frac{p_0}{1-q_d}} \qquad (12.17)
\end{aligned}
$$

Now for $n = N, N + 1,...$ (i.e. $n \geq N$), decision is NO

So $\quad f_n(\mathrm{G0}) = f_n(\mathrm{B1}) = f_n(\mathrm{G1})$ by (12.5), (12.6) $\qquad$ (12.18)

$$
\begin{aligned}
\text{But} \quad f_n(\mathrm{G0}) &= p_0 f_{n-1}(\mathrm{G0}) + p_1 f_{n-1}(\mathrm{G1}) \text{ by (12.3)} \\
&= p_0 f_{n-1}(\mathrm{G0}) + p_1 f_{n-1}(\mathrm{G0}) \text{ by (12.18) if } n-1 \geq N \\
&= (p_0 + p_1)\, f_{n-1}(\mathrm{G0}) \text{ if } n \geq N + 1 \\
&= p^{n-N} f_N(\mathrm{G0}) \qquad (12.19)
\end{aligned}
$$

OR $\quad f_{m+N}(\mathrm{G0}) = p^m f_N(\mathrm{G0})$

so $\quad f_n(\mathrm{G0}) = f_n(\mathrm{B1}) = f_n(\mathrm{G1}) = p^{n-N} \dfrac{p_1}{p_0-q}\,(p_0^N - q^N)$ for $n > N$ (12.20)

OR $\quad f_{m+N}(\mathrm{G0}) = f_{m+N}(\mathrm{B1}) = f_{m+N}(\mathrm{G1}) = p^m \dfrac{p_1}{p_0\, q}\,(p_0^N - q^N)$ for $m > 0$ (12.21)

$$
\begin{aligned}
f_n(\mathrm{B0}) &= q_0 f_{n\text{-}1}(\mathrm{B0}) + q_1 f_{n\text{-}1}(\mathrm{B1}) \text{ from (12.4)} \\
&= q_0 f_{n\text{-}1}(\mathrm{B0}) + q_1 p^{n\text{-}N\text{-}1} f_N(\mathrm{G0}) \\
&= q_0^{\,n-N} q^N + \frac{q_1}{q_0 - p}\,(q_0^{\,n-N} - p^{n-N})\; f_N(\mathrm{G0}) \\
&\text{for } n = N+1, N+2, ... \qquad (12.22)
\end{aligned}
$$

OR $\quad f_{m+N}(\mathrm{B0}) = q_0^{\,m}\, q^N + \dfrac{q_1}{q_0 - p}\,(q_0^{\,m} - p^m)\; f_N(\mathrm{G0})$ for $m > 0$.

So we have an analytic solution for $f_n(\mathrm{G0}), f_n(\mathrm{G1}), f_n(\mathrm{B0})$ and $f_n(\mathrm{B1})$ for all n.

TO RUN OR NOT?
SOME DYNAMIC PROGRAMMING MODELS IN CRICKET

13.0. Abstract

In cricket, particularly near the end of an innings, batsmen of different abilities need to manage the rate at which they score runs. Either batsman can choose to bat aggressively or defensively, which alters their chance of scoring runs and being dismissed. Since they change ends when they score a run and at the end of an over, by scoring an odd or even number of runs the two batsmen also determine which of them will face the next ball. It may be worthwhile to refuse a run to keep the slower or lower scoring batsman from the strike. Some dynamic programming models are developed which could be used to maximise the total number of runs scored.

Key words: sports, cricket, dynamic programming, Markov processes

13.1. Introduction

A cricket team usually consists of six specialist batsmen, a wicket keeper and four specialist bowlers. Two players from the batting side are at the wicket together, one facing the bowler and the other at the non strikers end. As the batting ability of the wicket keeper and the bowlers may vary from good to terrible, near the end of an innings a good batsman often finds himself with a poor batsman as a partner. The facing batsman changes whenever an odd number of runs is scored, and also at the completion of each six ball over. The batsmen are then often faced with the possibility of sacrificing a possible run to avoid the weaker batsman being placed on strike. Batsmen also need to decide whether they should attempt to score quickly, which will tend to produce more fours that do not require a change of ends but does involve greater risk of dismissal, or play more slowly and cautiously.

Because of the ball by ball nature of cricket, dynamic programming (DP) is the natural choice to analyse cricket. Surprisingly there are few papers using this technique to analyse cricket tactics. In cricket, in the first innings a team will generally want to maximise the number of runs scored, while the team batting last will want to maximise the chance of passing the first team's score. However depending on the state of the game, other optimality criteria are possible. For example a team trying to obtain a draw may wish to maximise the time an innings takes. Clarke & Norman (1998a) look at possible strategies to maximise innings length with two batsmen of different abilities at the crease. Here we look at some possible models to assist in maximising the number of runs scored with batsmen of different abilities.

Clarke (1988b) used a DP formulation to investigate the optimal run rates in one day cricket. Scoring rates from 1 to 12 per over were allowed. To solve the DP formulations, the probability of dismissal and the scoring profile for the various run rates were required. However such data is very difficult to obtain. For example, while it is generally recognised that on a particular pitch to score more quickly involves taking a greater risk of dismissal, when analysed over many matches the data suggests the reverse. This is due to confounding variables. On good pitches and against weak attacks, batsman score quickly and at low risk, while on difficult pitches and good bowling the scoring rate drops while the rate of dismissal goes up. However by using a relationship between dismissal rates and run rate that seemed reasonable and at least satisfied certain logical criteria (such as monotonic increasing) this paper derived some useful criteria. Johnston (1992) later showed some of these criteria were valid under a range of possible relationships. It is in this vein we present the following models. While data to implement them in detail may be difficult or impossible to obtain, it is hoped that by using such models some general conclusions valid under a wide range of scoring profiles might be found.

Here we relax the restriction that all batsmen are of equal ability. This implies batsmen have to consider the effects of different run rates on the chances of the weaker or slower scoring batsman being put on strike. It also raises the possibility of batsmen refusing possible runs.

13.2. Model 1

We begin by looking at a model where batsmen have the opportunity to take runs and need to decide whether to take the maximum possible, or one less to put a certain batsman on strike. For the moment we ignore the change of ends at the completion of an over.

A team consists of 11 batsmen designated by $i = 1 – 11$. Let $j = 0$ to 6 be an index that defines the scoring possibilities of a batsman, namely 0, 1, 2, 3, 4 (all run), 4 (boundary), 6. Then batsman i can score these runs with probability p_{i0}, p_{i1}.......p_{i6} or be dismissed with probability p_{id}. If a batsman is dismissed, assume the next batsman comes in and takes strike at the same end. (In fact this is not always the case).

Define j^*, j_{odd}, j_{even} as below and also shown in Table 13.1.

j^* is the number of runs available $j^* \quad = j$ for $j \neq 5$,
$= j - 1$ for $j = 5$.

j_{odd} is the maximum number of runs taken if batsmen always change ends if possible.
Then $\quad j_{odd} = j^* - 1$ if $j = 2, 4$,
$= j^*$ otherwise.

j_{even} is the maximum number of runs taken if batsmen never change ends.
Then $j_{even} = j^* - 1$ if $j = 1, 3$,
$= j^*$ otherwise.

TABLE 13.1. Number of runs scored off single ball.

Index j	No of Runs on offer j^*	j_{odd}	j_{even}
0	0	0	0
1	1	1	0
2	2	1	2
3	3	3	2
4	4 (Run)	3	4
5	4 (Boundary)	4	4
6	6	6	6

We can now set up a DP model. Define the state to be (i, k, j) where i is the batsman on strike, k is the batsman at the other end and j is the index of the number of runs on offer *after* the ball is bowled and there are n more balls to be bowled. If the number of either batsman is 12 the innings is completed. Thus $i = 1$ to 12, $k = 1$ to 12 ($i \,\square\, k$) , $j = 0$ to 6.

Note there is no need to consider dismissal as a special state. Let $l = \max(i,k) + 1$ be the next batsmen in. Then if (say) state is $(i,k,2)$ and the batsmen run a two, then before the next ball is bowled i will be on strike with k at the other end. So *after* the next ball is bowled, there will be j runs on offer (state (i,k,j)) with probability p_{ij}, or i will have been dismissed with probability p_{id}, so l will be facing and there will be no runs on offer (state $(l,k,0)$).

Define the value function $f_n(i,k,j)$ = total expected number of runs in the remainder of the innings after the ball is bowled and with n more balls to go.

The decisions are YES, Change ends, take j_{odd} number of runs or
NO, Do not change ends, take j_{even} number of runs.

Note that a decision is only necessary when $j_{\text{even}} \neq j_{\text{odd}}$; i.e. for j = 1, 2, 3, 4.

The transition probabilities are easily evaluated. If the current state is (i,k,j) and decision is No, batsmen will not change ends, so i will still be on strike for the next ball, and will score j runs with probability p_{ij} or be dismissed with probability p_{id}. Thus state $(i,k,j) \varnothing (i,k,j)$ with probability p_{ij}, and state $(i,k,j) \varnothing (l,k,0)$) with probability p_{id}. If current state is (i,k,j) and decision is Yes, batsmen will change ends, so k will be on strike next ball, and will score j runs with probability p_{kj} or be dismissed with probability p_{kd}. Thus state $(i,k,j) \varnothing (k,i,j)$ with probability p_{kj} and state $(i,k,j) \varnothing (l,i,0)$ with probability p_{kd}.

The functional equations then become

$$f_n(i,k,j) = \text{MAX} \begin{cases} \text{YES}: & j_{\text{odd}} + \sum_j p_{kj} f_{n-1}(k,i,j) + p_{kd} f_{n-1}(l,i,0) \\ \text{NO}: & j_{\text{even}} + \sum_j p_{ij} f_{n-1}(i,k,j) + p_{id} f_{n-1}(l,k,0) \end{cases} \tag{13.1}$$

$$= \text{MAX} \begin{cases} \text{YES}: & j_{\text{odd}} + f_n(k,i,0) \\ \text{NO}: & j_{\text{even}} + f_n(i,k,0) \end{cases} \tag{13.2}$$

Note that for j = 0, 5 and 6 the option of changing ends is not available so the No choice is always taken. A more detailed derivation of these equations is given in Appendix 13.1. Equation (13.2) makes sense. If batsmen decide to change ends, they get j_{odd} runs, and are then in exactly the same position as if they had been at opposite ends and had no runs available. Similarly if they decide not to change ends.

For the last ball the batsmen will take all the runs, and when the 12th batsman is needed the innings is ended. This results in initial conditions

$$f_0(i,k,j) \quad = j^*,\ \ i = 1\text{–}11,\ k = 1\text{–}11,\ j = 0\text{–}6 \qquad (13.3)$$

$$f_n(12,k,j) \quad = f_n(i,12,j) = 0 \ \text{ for all } i = 1\text{–}11,\ k = 1\text{–}11,\ j = 0\text{–}6 \qquad (13.4)$$

In keeping with the example in Hastings (1973, p 100) we could define the state to be (i,k) and $F_n(i,k)$ = expected number of runs *before* a ball is bowled, with n balls to go. Then

$$F_n(i,k) \quad = \sum_{j=0}^{6} p_{ij}\, f_{n-1}\,(i,k,j) + p_{id}\, f_{n\text{-}1}(l,k,0) \qquad (13.5)$$

Hastings calls (i,k,j) and (i,k) primary and secondary states.
From (13.1) and (13.5) it follows that

$$F_n(i,k) \quad = f_n(i,k,0) \qquad (13.6)$$

13.3. Simplifying results

Several simplifying results are given in Appendix 13.2. We show that the optimal decision for any odd j is the same. Similarly for even j. Thus we need only discuss the two cases when j is odd (1,3) or j is even (2,4). For each facing batsman the policy can be stated as an ordered pair (•,•) taking the values (Yes, Yes), (Yes, No), (No, Yes) or (No, No), where the first string represents the decision when j is 1 or 3, and the second when j is 2 or 4. We then show that (No, Yes) is never optimal. This can also be argued from common sense grounds. If batsmen refuse a run to keep batsman k from the strike, they would certainly not at the same stage refuse a run to put batsman k on strike.

So we have only three possible policies for each facing batsman, (No, No), (Yes, Yes) and (Yes, No). These three policies could be stated as

(a) keep the strike batsman i on strike
(b) put the opposite batsman k on strike; and
(c) take all the runs on offer.

We next show that an optimal decision of Yes for (i,k) when j is odd □ an optimal decision of No for (k,i) when j is even (and vice versa.) This simply shows that if (a) above is optimal for one batsmen, (b) is optimal for the other. If (c) is optimal for one batsman it is also optimal for the other. Thus we can make table 13.2.showing the relationship between optimal policy at stage n for state (i,k) and state (k,i).

TABLE 13.2. Relationship between policies

Optimal decision for State (i,k)	Optimal decision for State (k,i)	Combined Policy
(Yes, Yes)	(No, No)	keep k on strike
(Yes, No)	(Yes, No)	take all runs
(No, No)	(Yes, Yes)	keep i on strike

Note that the first and third are mirror images of each other; they are both putting a particular batsman on strike. The middle one is taking all possible runs. Note also for two even batsmen, the only possible optimal policy is (Yes, No) 'take all the runs' as otherwise the symmetry would result in a contradiction.

Thus, the problem is much simplified. For a given pair of batsmen instead of a Yes/No decision for each of two possibilities for the facing batsman at each of four possible runs = 4^4 =16 policies, there are only three possible policies (take all runs, keep i on strike, keep k on strike). If we group batsmen as 'recognised batsmen' or 'duffers' we only need to determine the optimal policy when a recognised batsman and a duffer are together.

13.4. Some analytic results

It is possible to generate some analytic results. Details are in Appendix 13.3. A summary follows.

With one ball to go optimal policies are

(Yes, Yes) Keep k on strike if $\mu_k > \mu_i + 1$

(No, No) Keep i on strike if $\mu_k < \mu_i - 1$

(Yes, No) Take all runs if $\mu_i - 1 < \mu_k < \mu_i + 1$

where $\mu_i = \sum_j j^* p_{ij}$ is the mean run rate of batsman i.

These policies make sense and could be argued from first principles. Thus if a non strike batsman has a run rate greater than one more than the batsman on strike, it is worth sacrificing a run to put him on strike for the last ball. If the run rates are within one of each other, all the runs would be taken.

The question to be answered, is how far back into the over do these optimal polices remain. While the optimal policies can not be found analytically for all values of n, we can find recurrence formula for the value function under particular polices. Thus if for $\mu_i-1 < \mu_k < \mu_i+1$ and the batsmen take all the runs for all n, we get

$$F_n(i,k) = \mu_i + p_{iu} F_{n-1}(k,i) + p_{ie} F_{n-1}(i,k) + p_{id} F_{n-1}(l,k)$$

where $p_{iu} = p_{i1} + p_{i3}$ = probability batsman i will score an uneven number of runs, and $p_{ie} = p_{i0} + p_{i2} + p_{i4} + p_{i5} + p_{i6}$ = probability batsman will score an even number of runs.

Similarly, for $\mu_i > \mu_k + 1$ if the batsmen attempt to keep k from the strike for all n,

$$F_n(i,k) = \mu_{ie} + (1-p_{id}) F_{n-1}(i,k) + p_{id} F_{n-1}(l,k)$$

$$F_n(k,i) = \mu_{ku} + p_{kr} F_{n-1}(i,k) + p_{kn} F_{n-1}(k,i) + p_{kd} F_{n-1}(l,i)$$

where $\mu_{ie} = \sum p_{ij}\, j_{\text{even}}$ = mean scoring rate if batsman i always takes j_{even} runs and

$\mu_{iu} = \sum p_{ij}\, j_{odd}$ = mean scoring rate if batsman i always takes j_{odd} runs.

$p_{ir} = p_{i1} + p_{i2} + p_{i3} + p_{i4}$ = probability batsmen will be able to run something

$p_{in} = p_{i0} + p_{i5} + p_{i6}$ = probability batsman will not be able to run.

These expressions can be solved recursively. Thus we could generate the expected number of runs if batsmen consistently follow certain strategies.

However, for the last wicket, when $l = 12$, some progress can be made analytically. In this case $F_{n-1}(l,i)$, $F_{n-1}(l,k)$ are zero, the recursion equations can be solved, and although complicated, closed expressions can be found for $F_n(i,k)$.

Note also that several performance measures such as p_{ir} naturally arise. These are currently not kept on cricketers. However it is sometimes said that a particular cricketer is adept at batting with tail enders, or is good at rotating the strike. If this is due to his scoring profile, such ability may be indicated by a high p_{ir} , which is the probability the batsman will have a choice of changing or keeping the strike.

13.5. Model 2

The extension to allow for change of ends at the end of each 6 ball over follows simply with a slightly more complex notation.

Let $f_{m,n}$ (i,k,j) be the expected number of runs with m overs to go, n more balls to go in the over (after the ball is bowled), j runs on offer, m=0,1,... n=0-5.

Let $F_{m,n}$ (i,k) be the expected number of runs with m overs and n balls in the over to go (before the ball is bowled), m=0,1,... n=1–6.

For n=1 to 5: Similar to the previous section we have

$$f_{m,n}(i,k,j) = \text{MAX} \begin{cases} \text{YES}: & j_{\text{odd}} + \sum_j p_{kj}\, f_{m,n-1}(k,i,j) + p_{kd}\, f_{m,n-1}(L,i,0) \\ \text{NO}: & j_{\text{even}} + \sum_j p_{ij}\, f_{m,n-1}(i,k,j) + p_{id}\, f_{m,n-1}(L,k,0) \end{cases} \tag{13.7}$$

For j = 0, 5 & 6 there is no choice, only the No option is available.

Of special importance is j = 0, which gives

$$f_{m,n}(i,k,0) = \sum_j p_{ij}\, f_{m,n-1}(i,k,j) + p_{id}\, f_{m,n-1}(L,k,0) \tag{13.8}$$

so

$$f_{m,n}(i,k,j) = \text{MAX} \begin{cases} YES: & j_{\text{odd}} + f_{m,n}(k,i,0) \\ NO: & j_{\text{even}} + f_{m,n}(i,k,0) \end{cases} \tag{13.9}$$

For n=0: Last ball of the over so change ends if runs are even, overs go down by one

$$f_{m,0}(i,k,j) = \text{MAX} \begin{cases} YES: & j_{\text{odd}} + \sum p_{ij}\, f_{m-1,5}(i,k,j) + p_{id}\, f_{m-1,5}(L,k,0) \\ NO: & j_{\text{even}} + \sum p_{kj}\, f_{m-1,5}(k,i,j) + p_{kd}\, f_{m-1,5}(L,i,0) \end{cases} \tag{13.10}$$

and again: $$f_{m,0}(i,k,0) = \sum p_{kj}\, f_{m-1,5}(k,i,j) + p_{kd}\, f_{m-1,5}(L,i,0) \tag{13.11}$$

so $$f_{m,0}(i,k,j) = \text{MAX} \begin{cases} YES: & j_{\text{odd}} + f_{m,0}(k,i,0) \\ NO: & j_{\text{even}} + f_{m,0}(i,k,0) \end{cases} \tag{13.12}$$

The relationship between $F_{m,n}\,(i,k)$ and $f_{m,n}\,(i,k,0)$ is

$$\text{For } n\text{=1 to 5:} \qquad F_{m,n}\,(i,k) \;=\; \sum_j p_{ij}\, f_{m,n-1}\,(i,k,j) \;+\; p_{id}\, f_{m,n-1}\,(\mathrm{L},k,0) \qquad (13.13)$$

$$= f_{m,n}\,(i,k,0)$$

$$\text{For } n\text{=6:} \qquad F_{m,6}\,(i,k) \;=\; \sum_j p_{ij}\, f_{m,5}\,(i,k,j) \;+\; p_{id}\, f_{m,5}\,(\mathrm{L},k,0) \qquad (13.14)$$

$$= f_{m+1,0}\,(k,i,0)$$

$$\text{or} \qquad f_{m,0}\,(i,k,0) \;=\; F_{m-1,6}\,(k,i)$$

So that at the end of the over *only* the $f_{m,n}(i,k,0)$ matrix becomes the transpose of $F_{m,n}(i,k)$ matrix. This also makes sense as between the stage $f_{m,0}$ (after the last ball of the m^{th} over is bowled) and $F_{m-1,6}$ (before the first ball of the m-1^{th} over is bowled) there is a change of ends due to the end of over.

Most of the general results of the previous model still hold. Thus there are still basically only two decisions for each pair of batsmen for each ball.

13.5.1. Computer implementation

A computer program has been written to implement Equations 13.7 and 13.10 of Model 2. The various probabilities required could be estimated from standard score sheets, as these always show the number of runs for each scoring shot and usually show the total number of balls faced. They may not differentiate between a four all run and a boundary. We only take a very simple case for illustration here. We assume there are two types of batsmen, batsmen 1 to 7 are recognised and 8 to 11 are duffers with scoring probabilities as shown in Table 13.3. Then the basic program will produce the expected number of runs in the remainder of the innings before the ball is bowled and the optimal strategy after the ball is bowled for as far back in the innings as required. For example, from ball by ball data developed by Johnston (1992) from the 1989 Australia, Pakistan and Sri Lanka one day series in Australia we have the runs scored by batsmen at each position. For the Australian batsmen batting at 1 to 7, of 2819 balls faced 113 fours were scored. This gives the probability of a four at just on 4%. For 8-11 batsmen 50 balls faced produced one four to give a 2% chance. Now in this series the Australians were well on top, so the weaker batsmen did not bat very often. Hence some adjustment of their probabilities was necessary to produce sensible figures. For example the probability of scoring a 0 was increased to at least the level of the recognised batsmen, as this seems a reasonable assumption. This produced the probabilities in Table 13.3.

TABLE 13.3. Scoring profile of batsmen

Batsmen	p_d	p_0	p_1	p_2	p_3	p_4	p_5	p_6	μ
Recognised	0.02	0.50	0.32	0.10	0.02	0.00	0.04	0.00	0.74
Duffer	0.06	0.52	0.32	0.08	0.00	0.00	0.02	0.00	0.56

Note this gives an average run rate for recognised batsmen of 0.74, so if they bat through a one day innings of 300 balls they should score about 224. In fact, the printout for 300 balls to go with 1 and 2 batting gives an expected score of 207.3. This is less than 224 due to the possibility that the slower scoring of the batsmen 8-11 will be needed, or worse still the team dismissed before 50 overs. By repeating with one fewer or one more extra batsmen, the effect on expected score of selection decisions to choose extra batsmen could be evaluated. The model could be used by media to give estimates of expected score and effects of dismissals. For example, with 299 balls to go and batsman 2&3 at the crease the expected score reduces to 195.5, so a first ball duck costs the team about 12 runs. Johnston et al. (1993) use this approach for the basis of a player performance measure in one day cricket.

We give in Table 13.4 a sample output with 10 overs, 0 balls to go. Thus the decisions are for the last ball of the eleventh last over, and the expected runs for the last 10 overs.

The strategies give us the runs that should be taken off the last ball of the 11th last over. The batsmen should generally take all the runs, except if a recognised batsman is batting with the number 11 batsman. In this case they should put the number 11 bat on strike, so the recognised batsman will be on strike at the beginning of the next over. The transpose of the numbers give us the expected score in the last 10 overs. Thus at the beginning of the 10th last over, if say 4 and 5 are in they can expect to score another 44.2 runs. On the other hand if 4 and 11 are in they can expect 15.5 runs if 4 is facing and 14.4 if 11 is facing. The table can be used to estimate the value of wickets in hand. For example it is worth roughly 18 runs (41.7- 23.9) to have 6 & 7 batting rather than 8 and 9 with 10 overs to go.

TABLE 13.4. Expected score and optimal strategy with 10 overs to go

Facing batsman	Non strike batsman										
	1	2	3	4	5	6	7	8	9	10	11
1		44.4 YN	44.4 YN	44.4 YN	44.2 YN	43.4 YN	41.1 YN	34.6 YN	31.1 YN	24.7 YN	14.4 YY
2	44.4 YN		44.4 YN	44.4 YN	44.2 YN	43.4 YN	41.1 YN	34.6 YN	31.1 YN	24.7 YN	14.4 YY
3	44.4 YN	44.4 YN		44.4 YN	44.2 YN	43.4 YN	41.1 YN	34.6 YN	31.1 YN	24.7 YN	14.4 YY
4	44.4 YN	44.4 YN	44.4 YN		44.2 YN	43.4 YN	41.1 YN	34.6 YN	31.1 YN	24.7 YN	14.4 YY
5	44.2 YN	44.2 YN	44.2 YN	44.2 YN		43.4 YN	41.1 YN	34.6 YN	31.1 YN	24.7 YN	14.4 YY
6	43.4 YN	43.4 YN	43.4 YN	43.4 YN	43.4 YN		41.1 YN	34.6 YN	31.1 YN	24.7 YN	14.4 YY
7	41.1 YN	41.1 YN	41.1 YN	41.1 YN	41.1 YN	41.1 YN		34.6 YN	31.1 YN	24.7 YN	14.4 YY
8	34.8 YN	34.8 YN	34.8 YN	34.8 YN	34.8 YN	34.8 YN	34.8 YN		23.9 YN	17.3 YN	9.1 YN
9	31.5 YN	31.5 YN	31.5 YN	31.5 YN	31.5 YN	31.5 YN	31.5 YN	23.9 YN		17.3 YN	9.1 YN
10	25.4 YN	25.4 YN	25.4 YN	25.4 YN	25.4 YN	25.4 YN	25.4 YN	17.3 YN	17.3 YN		9.1 YN
11	15.5 NN	15.5 NN	15.5 NN	15.5 NN	15.5 NN	15.5 NN	15.5 NN	9.1 YN	9.1 YN	9.1 YN	

13.6. Model 3

In addition to deciding whether to take all possible runs, a batsman may have choices in the sort of stroke he plays. For example, when a recognised batsman is paired with a duffer, fielding sides sometimes have deep set fields to encourage the good batsman to take a single. In this case a batsman may decide to gently stroke the ball to a deep set fielder for a certain single and sometimes a two, with almost no risk of being dismissed. On the other hand he may decide to belt the ball in the hope of beating the fielder for a boundary. This shot carries a greater risk of dismissal, but also a greater chance that no run will be scored. A batsman may also just wish to avoid dismissal and block the ball, a shot that rarely scores a run. Superimposed on these shots are the decisions whether to run all the available runs or not.

Thus we let the state (i,k) be the two batsman, i facing the bowler, and define as before $F_{m,n}$ (i,k) the expected number of runs under optimal policy with m overs and n balls in the over to go m = 1,2,... n=1–6,. There are now five decisions, to Block, Stroke or Hit, and for each of the latter two to take all the runs or not. The model is similar to the previous except that we have different p_{ij} for each type of shot.

It is difficult to progress very far analytically with this model, but simple to solve numerically via a basic program. This allows the number of recognised batsmen to be altered. It is difficult to obtain estimates of the p_{ij} as it is not known from score sheets what type of stroke a batsman is trying to play. One could perhaps split up an innings into sections. For example, in one day matches where teams do not lose a lot of wickets, it could be assumed that batsmen were trying to thrash the ball in the last few overs. Here we take probabilities similar to that we had before for the recognised batsman as equivalent to stroking, and adjust the others up and down as necessary for blocking and hitting. This gives the probabilities in Table 13.5.

TABLE 13.5. Scoring profile for batsmen when blocking, stroking and hitting

		Probabilities								
Batsmen	Shot	p_d	p_0	p_1	p_2	p_3	p_4	p_5	p_6	μ
Recognised	Block	0.01	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Stroke	0.02	0.50	0.34	0.14	0.00	0.00	0.00	0.00	0.62
	Hit	0.10	0.30	0.30	0.20	0.00	0.00	0.10	0.00	1.10
Duffer	Block	0.06	0.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Stroke	0.10	0.40	0.40	0.10	0.00	0.00	0.00	0.00	0.60
	Hit	0.30	0.30	0.20	0.10	0.00	0.00	0.10	0.00	0.80

The program was run with 7 recognised batsmen for 300 balls. This gave an expected score at the beginning of the match as 181. Some sample output for 19 balls to go is shown in Tables 6 and 7. The expected score shows that if they still have two recognised batsmen at the crease they can expect to score about a run a ball. However this drops to under 8 if number 11 is batting. In the above strategy output, the first letter indicates the type of shot, and the others whether to take all the runs or not as previously used. Thus in this case they should always take all the runs. When recognised batsmen are batting with 10 and 11 they should stroke, otherwise hit. Presumably this is because of the greater chance of getting a single if they stroke, and so protecting 10 and 11 from the strike at the beginning of the next over. Duffers should stroke when playing with duffers. When playing with a recognised batsman, number 8 should hit, 9 & 10 stroke, and 11 block.

While the strategy here is a little complicated, it is possible that by studying the output for a range of values some general principals may be enunciated. For example, for the above scoring profile, for all stages the *only* time batsmen should refuse a run is to protect batsman 11 from the strike.

This output would apply for one day cricket when the number of balls remaining is known; in test cricket the length of the match is unknown. However 300 balls to go would represent a good approximation for test cricket, at least certainly in the cases when a duffer is at the crease. In this case the optimal strategy is for both batsman to stroke the ball, and to take all the runs except when 11 is at the crease. In this case they should refuse runs as necessary to keep the recognised batsman on strike for the first 4 balls of an over, take all the runs on the 5th and refuse runs to put the bad bat on strike for the last ball of the over.

13.7. Conclusion

DP models are useful in analysing optimal strategies in cricket. The common practice of refusing runs to protect weaker batsmen from the strike has been shown to be sensible under certain conditions. Clearly more work needs to be done before these models can be generally applied. The optimal strategies for a range of batsmen scoring profiles need to be determined to see if any general recommendations can be made. The models might also be extended to a second innings formulation. However this involves the addition of the number of runs to go to the state variable, and usually increases the computational requirements to an unacceptable level.

TABLE 13.6. Expected score with 19 balls to go.

Facing batsman	Non strike batsman										
	1	2	3	4	5	6	7	8	9	10	11
1		20.9	20.9	20.8	20.5	19.8	18.2	15.1	14.0	11.9	8.0
2	20.9		20.9	20.8	20.5	19.8	18.2	15.1	14.0	11.9	8.0
3	20.9	20.9		20.8	20.5	19.8	18.2	15.1	14.0	11.9	8.0
4	20.8	20.8	20.8		20.5	19.8	18.2	15.1	14.0	11.9	8.0
5	20.5	20.5	20.5	20.5		19.8	18.2	15.1	14.0	11.9	8.0
6	19.8	19.8	19.8	19.8	19.8		18.2	15.1	14.0	11.9	8.0
7	18.2	18.2	18.2	18.2	18.2	18.2		15.1	14.0	11.9	8.0
8	15.0	15.0	15.0	15.0	15.0	15.0	15.0		11.0	8.9	5.2
9	13.9	13.9	13.9	13.9	13.9	13.9	13.9	11.0		8.9	5.2
10	11.7	11.7	11.7	11.7	11.7	11.7	11.7	8.9	8.9		5.2
11	7.5	7.5	7.5	7.5	7.5	7.5	7.5	5.2	5.2	5.2	

TABLE 13.7. Optimal strategy with 19 balls to go

Facing batsman	Non strike batsman										
	1	2	3	4	5	6	7	8	9	10	11
1		HYN	HYN	HYN	HYN	HYN	HYN	HYN	HYN	SYN	SYN
2	HYN		HYN	HYN	HYN	HYN	HYN	HYN	HYN	SYN	SYN
3	HYN	HYN		HYN	HYN	HYN	HYN	HYN	HYN	SYN	SYN
4	HYN	HYN	HYN		HYN	HYN	HYN	HYN	HYN	SYN	SYN
5	HYN	HYN	HYN	HYN		HYN	HYN	HYN	HYN	SYN	SYN
6	HYN	HYN	HYN	HYN	HYN		HYN	HYN	HYN	SYN	SYN
7	HYN	HYN	HYN	HYN	HYN	HYN		HYN	HYN	SYN	SYN
8	HYN	HYN	HYN	HYN	HYN	HYN	HYN		SYN	SYN	SYN
9	SYN	SYN	SYN	SYN	SYN	SYN	SYN	SYN		SYN	SYN
10	SYN	SYN	SYN	SYN	SYN	SYN	SYN	SYN	SYN		SYN
11	BYN	BYN	BYN	BYN	BYN	BYN	BYN	SYN	SYN	SYN	

Appendix 13.1. Derivation of functional equations for Model 1

For $j = 0, 5, 6$:

No decision to be made, no change of ends, so *i* is still on strike next ball

$$f_n(i,k,j) = j^* + \sum_J p_{ij} f_{n-1}(i,k,j) + p_{id} f_{n-1}(l,k,0) \qquad (13.15)$$

$$\text{So } f_n(i,k,0) = \sum_J p_{ij} f_{n-1}(i,k,j) + p_{id} f_{n-1}(l,k,0) = F_n(i,k) \qquad (13.16)$$

For $j = 1, 3$:

$$f_n(i,k,j) = \text{MAX} \begin{cases} \text{Yes}: \ j^* + \sum_j p_{kj} f_{n-1}(k,i,j) + p_{kd} f_{n-1}(l,i,0) \\ \text{No}: \ j^* - 1 + \sum_j p_{ij} f_{n-1}(i,k,j) + p_{id} f_{n-1}(l,k,0) \end{cases} \qquad (13.17)$$

$$= \text{MAX} \begin{cases} \text{Yes}: \ j^* \quad + f_n(k,i,0) \\ \text{No}: \ j^* - 1 + f_n(i,k,0) \end{cases} \qquad (13.18)$$

$$= \text{MAX} \begin{cases} \text{Yes}: \ j^* \quad + F_n(k,i) \\ \text{No}: \ j^* - 1 + F_n(i,k) \end{cases} \qquad (13.19)$$

For $j = 2, 4$:

$$f_n(i,k,j) = \text{MAX} \begin{cases} \text{Yes}: \ j^* - 1 + \sum_j p_{kj} f_{n-1}(k,i,j) + p_{kd} f_{n-1}(l,i,0) \\ \text{No}: \ j^* \quad + \sum_j p_{ij} f_{n-1}(i,k,j) + p_{id} f_{n-1}(l,k,0) \end{cases} \qquad (13.20)$$

$$= \text{MAX} \begin{cases} \text{Yes}: \ j^* - 1 + f_n(k,i,0) \\ \text{No}: \ j^* \quad + f_n(i,k,0) \end{cases} \qquad (13.21)$$

$$= \text{MAX} \begin{cases} \text{Yes}: \ j^* - 1 + F_n(k,i) \\ \text{No}: \ j^* \quad + F_n(i,k) \end{cases} \qquad (13.22)$$

Appendix 13.2. Proof of theorems for Model 1

Theorem 1: Optimal decisions for all odd (even) values are the same. i.e. the optimal policy for some odd j is the same as for any other odd j. Similarly for even j.

This follows trivially from (13.2)

e.g. If the decision is Yes for some odd j

Then $j_{odd} + f_n(k,i,0) > j_{even} + f_n(i,k,0)$

□ $j^* + f_n(k,i,0) > j^* - 1 + f_n(i,k,0)$

□ $f_n(k,i,0) > -1 + f_n(i,k,0)$

□ $j'^* + f_n(k,i,0) > j'^* - 1 + f_n(i,k,0)$

□ $j'_{odd} + f_n(k,i,0) > j'_{even} + f_n(i,k,0)$ if j' is odd

so Yes is optimal for j' ■

Thus we need only discuss the two cases when j is odd (1,3) or j is even (2,4).

Thus for each facing batsman the policy can be stated as an ordered pair (·,·) taking the values (Yes, Yes), (Yes, No), (No, Yes) or (No, No), where the first string represents the decision when j is (1 or 3), and second when j is (2 or 4).

Theorem 2 : A policy of (No, Yes) is never optimal.

If decision at odd j is No, then for odd j we have from (13.2)

$j_{even} + f_n(i,k,0) > j_{odd} + f_n(k,i,0)$

□ $j^* - 1 + f_n(i,k,0) > j^* + f_n(k,i,0)$

□ $f_n(i,k,0) > 1 + f_n(k,i,0)$

So for even j :

$$
\begin{aligned}
j_{even} + f_n(i,k,0) &= j^* + f_n(i,k,0) \\
&> j^* + 1 + f_n(k,i,0) \text{ from above} \\
&> j^* - 1 + f_n(k,i,0) \\
&= j_{odd} + f_n(k,i,0)
\end{aligned}
$$

So optimal decision for even is No from (13.2). ■

Theorem 3: An optimal decision of Yes for (i,k,j) when j is odd □ an optimal decision of No for (k,i,j) when j is even (and vice versa.)

eg. If optimal decision is Yes for odd j

Then $j_{odd} + f_n(k,i,0) \quad > j_{even} + f_n(i,k,0)$ from (2)

i.e. $j^* + f_n(k,i,0) \quad > j^* - 1 + f_n(i,k,0)$

□ $f_n(k,i,0) \quad > -1 + f_n(i,k,0)$

So for even j:

$j^* + f_n(k,i,0) \quad > j^* - 1 + f_n(i,k,0)$

i.e. $j_{even} + f_n(k,i,0) \quad > j_{odd} + f_n(i,k,0)$

□ optimal decision is No for state (k,i,j) when j is even.

Similarly, for a decision of No for odd j, > becomes < in above

□ decision is Yes for state (k,i,j) when j is even. ■

Appendix 13.3. Some Analytic results for Model 1

First Stage: n=1

With one ball to go:

From (13.1) $\quad f_1(i,k,0) \quad = 0 + \sum_j p_{ij}\, f_0(i,k,j) + p_{id}\, f_0\,(l,k,0)$

$$= \sum_j p_{ij}\, j^* + p_{id}0 \qquad \text{from (13.3)}$$

$$= \mu_i$$

This is the mean run rate for batsman *i.*

(Note that equation 13.2 □ $f_1(i,k,5) = 4 + \mu_i$ and $f_1(i,k,6) = 6 + \mu_i$)

For other *j*, from (13.2) $f_1(i,k,j) = \text{MAX} \begin{cases} \text{YES } j_{\text{odd}} + \mu_k \\ \text{NO } j_{\text{even}} + \mu_i \end{cases}$

If j *is* even**:** Decision is Yes if $\quad j_{\text{odd}} + \mu_k > j_{\text{even}} + \mu_i$

i.e. $\quad j^* - 1 + \mu_k > j^* + \mu_i$

i.e. $\quad \mu_k > \mu_i + 1$

and No if $\quad \mu_k < \mu_i + 1$

If j is odd: Decision is Yes if $j^* + \mu_k > j^* - 1 + \mu_i$

i.e. $\quad \mu_k > \mu_i - 1$

and No if $\quad \mu_k < \mu_i - 1$

So for $n = 1$ optimal policies are (Yes, Yes) if $\quad \mu_k > \mu_i + 1$

(No, No) if $\quad \mu_k < \mu_i - 1$

(Yes, No) if $\quad \mu_i - 1 < \mu_k < \mu_i + 1$

The question to be answered is how far back into the over do these optimal policies remain. Our feeling is that you could not extend this to all values purely based on μ_k and μ_i, but that it would depend on p_0, p_1 ...etc, or more particularly p_u and p_e. Is there a formula that gives N = least number of balls to go when this policy is no longer optimal?

Special Case (Yes, No) decision - take all runs on offer

Consider case when decision is (Yes, No) for $n = 1$

i.e. $\mu_i - 1 < \mu_k < \mu_i + 1$

This is the case when batsmen take *all* the runs and is the same for the batsman at the other end.

$$f_n(i,k,j) = j^* + f_{n-1}(k,i,0) \text{ if } j \text{ is odd} \qquad \text{from (13.2)}$$
$$= j^* + f_n(i,k,0) \qquad \text{if } j \text{ even}$$

$$f_n(i,k,0) = 0 + \sum_j p_{ij}\, f_{n-1}(i,k,j) + p_{id}\, f_{n-1}(l,k,0) \qquad \text{from (13.1)}$$

$$= \sum_{\mathrm{u}} p_{ij}\, \{j^* + f_{n-1}(k,i,0)\} + \sum_{\mathrm{e}} p_{ij}\, \{j^* + f_{n-1}(i,k,0)\} + p_{id}\, f_{n-1}(l,k,0),$$

where $\sum_{\mathrm{u}}$ means sum over odd $j = 1,3$ and $\sum_{\mathrm{e}}$ means sum over even $j = 0, 2, 4, 5, 6$

$$= \mu_i + p_{iu}\, f_{n-1}(k,i,0) + p_{ie}\, f_{n-1}(i,k,0) + p_{id} f_{n-1}(l,k,0)$$

where $p_{iu} = p_{i1} + p_{i3}$ = probability batsmen will score an odd number of runs, and

$p_{ie} = p_{i0} + p_{i2} + p_{i4} + p_{i5} + p_{i6}$

= probability batsman will score an even number of runs.

$(p_{id} + p_{iu} + p_{ie} = 1)$

This expression can be solved recursively.

Special Case (No, No) decision - Keep batsman k off strike

Consider case when decision is (No, No) when state is (i,k). Then decision is (Yes,Yes) when state is (k,i). The batsmen are trying to keep batsman k from the strike. This is the case for $n = 1$ if $\mu_k < \mu_i - 1$. i.e. $\mu_i > \mu_k + 1$

Assume this is true for $n = 1,2,...N$.

Then from (13.2) $f_n(i,k,j) = j_{\text{even}} + f_n(i,k,0)$ for $n = 1,2,...N$

From (13.1) $f_n(i,k,0)$

$$\begin{aligned} &= \sum_j p_{ij} f_{n-1}(i,k,j) + p_{id} f_{n-1}(l,k,0) \\ &= \sum p_{ij} \{j_{\text{even}} + f_{n-1}(i,k,0)\} + p_{id} f_{n-1}(l,k,0), \\ &= \mu_{ie} + (1\text{-}p_{id}) f_{n-1}(i,k,0) + p_{id} f_{n-1}(l,k,0) \end{aligned}$$

where $\mu_{ie} = \sum p_{ij}\, j_{\text{even}}$ = mean score if batsmen always stay at same end.

Similarly $f_n(k,i,j)$

$$\begin{aligned} &= j_{\text{odd}} + f_n(i,k,0) \quad \text{if } j = 1 \text{ to } 4 \\ &= j_{\text{odd}} + f_n(k,i,0) \quad \text{if } j = 0,5,6 \end{aligned}$$

From (13.1) $f_n(k,i,0) = \sum_r p_{kj} \{ j_{\text{odd}} + f_{n-1}(i,k,0)\} + \sum_n p_{kj} \{j_{\text{odd}} + f_{n-1}(k,i,0)\} + p_{kd}\,(l,i,0)$

where $\sum_r$ means sum over the j where it is possible to take runs j = 1,2,3,4, and $\sum_n$ means sum over the j where it is not possible to take runs j =0,5,6.

$$\begin{aligned} &= \sum p_{kj}\, j_{\text{odd}} + \sum_r p_{kj}\ f_{n-1}(i,k,0) + \sum_n p_{kj} f_{n-1}(k,i\ ,0) + p_{kd} f_{n-1}(l,i,0), \\ &= \mu_{ku} + p_{kr} f_{n-1}(i,k,0) + p_{kn} f_{n-1}(k,i,0) + p_{kd} f_{n-1}(l,i,0) \end{aligned}$$

where $\mu_{iu} = \sum p_{ij}\, j_{\text{odd}}$ = mean score if batsmen always change ends if possible

$p_{ir} = p_{i1} + p_{i2} + p_{i3} + p_{i4}$ = probability batsmen will be able to run something

$p_{in} = p_{i0} + p_{i5} + p_{i6}$ = probability batsman will not be able to run.

$(p_{id} + p_{ir} + p_{in} = 1\ )$

This expression can be solved recursively.

ASSESSING PLAYER PERFORMANCE IN ONE DAY CRICKET USING DYNAMIC PROGRAMMING

14.0. Abstract

A dynamic programming formulation is used to develop a method of calculating the contribution, in runs, made by each player to the team's score in a game of one-day cricket. The advantages of such measures over the currently used statistics are outlined as well as their possible use to choose 'man of the match' award winners, to rank the best batsmen and bowlers and to aid selectors of one-day sides. Possible extensions to this work are also discussed.

Key words: dynamic programming, sport, cricket, computer ratings

14.1. Introduction

There has always been great interest in the ranking of individual and team achievements; from the richest ten people or the top ten companies in the world, to the listing of the poorest and the least taxed countries in the world. Comparisons of this type have carried over to the sporting arena. For instance, in tennis, players' ranking will help determine their seeding for tournaments and thus affect their chance of success, and in golf a high world ranking allows automatic entry to many of the top tournaments. Being on top of the ranking list for any of the major sports also gives an individual immediate access to remuneration from advertisement, product endorsement and appearance money.
Although Deloittes (1988) has developed a rating system for test cricketers, in one-day cricket there is no such system. This paper describes the use of dynamic programming to develop a new measure of player performance. The model developed calculates the contribution to the team in runs by each player in a one-day match.

These ratings could be used to determine the 'best' players for the interest of the general public, to determine the player to receive the player of the match award, as well as to assist team selectors.

14.2. The problem

Limited-over or one-day cricket is played by two teams of eleven players. Each team has 50 six-ball overs from which to score as many runs as possible. The innings of each team is terminated either when the 50 overs is completed or when 10 wickets have been lost, with the team that scores the most runs being the winner. Batsmen must therefore score as many runs as possible within the constraints of the number of overs to be bowled and the number of wickets to be lost. There is a constant trade-off between fast scoring rates and the risk of losing wickets.

The most common statistic kept on one-day batsmen is their *average*, or runs scored per dismissal, which is the same statistic used for test cricketers where the time factor is of less importance. Another statistic kept on one-day batsmen is the *strike rate* or runs scored per 100 balls faced. While these measures, as tabled in Dundas (1991), give an indication of how many runs were scored and how quickly, the stage of the innings at which runs were scored is not considered. The statistics kept on bowlers, *bowling average* or runs conceded per wickets taken and *strike rate* or balls bowled per wickets taken, as tabled in Dundas (1991), suffer the same problem. None of these statistics take account of the constraints that are present in the game of one-day cricket.

For example, four runs off one over by a batsman may be an excellent result in the first over of the innings, but a poor result from the last over. Similarly a wicket by a bowler off the last ball of the innings is no more value than a maiden ball, whereas a wicket off the first ball of the innings significantly decreases the expected score of the batting side. A measure of player performance is needed that takes account of the number of runs scored, the speed of scoring and the stage of the innings that the scoring took place.

14.3. Model formulation

Clarke (1988b) developed the dynamic programming formulation for the first innings in one-day cricket which is being used in this paper to calculate the performance measures. This formulation calculates the optimal run rate (the rate that will lead to the largest expected score in the remainder of the innings) and the expected score in the remainder of the innings at each stage and state of an innings. The functional equation is shown as equation 14.1.

$$f_n(i) = \underset{R}{Max}\left\{p_d\, f_{n-1}(i-1) + R/_6 + (1-p_d)\, f_{n-1}(i)\right\} \qquad (14.1)$$

where i = wickets in hand, i = 0 to 10.
n = balls to go, n = 0 to 300.
$f_n(i)$ = expected score with n balls to go and i wickets in hand.
p_d = probability of dismissal per ball.
R = runs per over.

Using a standard relationship between R and p_d, each $f_n(i)$ can be calculated. The actual number of runs scored at each stage and state of the innings can then be compared to the expected scores and a measure of how many extra runs each player has contributed can be calculated. For the batsman facing when there are n balls to go and i wickets in hand, the expected score in the remainder of the innings is $f_n(i)$. After that ball, the expected score is the score off that ball plus the expected score in the remaining n-1 balls. This is given by $f_{n-1}(i)$ if no wicket fell, or $f_{n-1}(i-1)$ if a wicket fell. Thus the contribution of the batsman to the score is given by:

No Wicket Loss: score off that ball + $f_{n-1}(i)$ - $f_n(i)$

Wicket Loss: score off that ball + $f_{n-1}(i-1)$ - $f_n(i)$

For the bowler of that ball, the contribution to his team's performance would be the negative of the above.

For instance if the expected score in the remainder of the innings with 200 balls to go and 5 wickets in hand is $f_{200}(5)$ =120.00 and the expected score in the remainder of the innings with 199 balls to go and 5 wickets in hand is $f_{199}(5)$ =119.20, then the batsman on strike when 200 balls remain in the innings must score 0.8 runs for the side to have the same expected innings score. If the batsman scores 2 runs in this situation then he has advanced his team's expected innings score by 1.2 runs. In this case the batsman's performance measure would increase by 1.2 and the bowler's measure would decrease by 1.2. All performance measures start at zero at the beginning of the innings and increase or decrease as each ball is bowled.

Because the measures for bowlers and batsmen are essentially the same (extra runs over what is expected) they can be added and so the performance measures for each player in a match is the sum of his batting and bowling performance measures in each innings.

Although equation 14.1 was developed for the first innings of a one-day cricket match, it has been used for both the first and second innings in this paper. This is due to the complication involved in using a model for the second innings where teams are not attempting to maximise the score, as in the first innings, but are attempting to score more runs than the team that batted first. The dynamic programming model therefore has another state variable which is the score that must be made in the remainder of the innings to win the match. In addition, the second innings formulation involves calculating the effect each player had on their sides probability of winning (instead of the effect on the expected score). Therefore each of the measures obtained from the two innings would not be in the same units and could not be directly compared. This diminishes the relevance of the performance measures generated.

14.4. Data collection

The calculation of the values of $f_n(i)$ using equation 14.1 requires that the relationship between R and p_d be determined. Although a season's data was used in an attempt to determine the actual relationship that existed between R and p_d, it was not possible to isolate the effect R had on p_d from all of the other factors, and therefore the relationship used had to be determined in some other way.

It was decided to develop a relationship between R and p_d which was 'fair'. In other words a relationship which when used to calculate performance measures did not favour batsmen over bowlers or vice-versa. Each player could then be measured against a known standard. This is common to the current practice where no account is taken of the standard of the opposition, the pitch or the ground dimensions. Several relationships were developed and used to calculate the performance measures on actual matches. The relationship that performed best in producing measures that agreed with a subjective appraisal of performance was chosen and is shown in Table 14.1. In addition to the probability of dismissal p_d for each run rate the average number of balls faced before dismissal $1/p_d$ and the expected partnership size $(1/p_d)$ $(R/6)$ are given for that run rate. With these values, if batsmen choose the optimal run rate at each stage of the innings the expected score would be 215.

TABLE 14.1. Relationship between R and p_d used to solve equation 14.1

Run Rate	p_d	Average number of balls faced before dismissal	Expected partnership size
1.0	0.005	200.0	33.3
1.5	0.008	127.8	31.9
2.0	0.011	91.7	30.6
2.5	0.014	70.0	29.2
3.0	0.018	55.6	27.8
3.5	0.022	45.2	26.4
4.0	0.027	37.5	25.0
4.5	0.032	31.5	23.6
5.0	0.038	26.7	22.2
5.5	0.044	22.7	20.8
6.0	0.051	19.4	19.4
6.5	0.060	16.7	18.1
7.0	0.070	14.3	16.7

14.5. Results

In order to calculate performance measures for a match the following data must be available for each ball bowled: batsman, bowler and result (e.g. batsman number 3, bowler number 2, no dismissal and 4 runs scored). Unfortunately official score sheets do not keep such data and therefore a computer program was written enabling the information contained on scoresheets to be transformed to the required form and saved to a 'match file' for later use. Another computer program was written to perform the dynamic programming calculations and calculate the $f_n(i)$ values.

The output from these two programs (the match file and the $f_n(i)$ values) are then used as input to a third program which performed all of the performance measure calculations as detailed above. All programs were written in the programming language of Turbo Pascal 5.0 on an IBM personal computer. The calculations involved in the programs are simple ones and since the values of $f_n(i)$ used are the same for every innings and need only be calculated once, the program could be run in real time (i.e. as the match is being played). Therefore the performance measure of each player could be shown each ball, as they increase or decrease, and the expected score of the team could be displayed as the innings progresses. The 1989/90 Benson & Hedges series of one-day matches played in Australia between Australia, Sri Lanka and Pakistan has been used to illustrate the use of the performance measures.

Table 14.2 shows the performance measures for all of the players in the match played in Melbourne on 26/12/89 between Australia and Sri Lanka. Australia won this match by scoring 5 for 228 to Sri Lanka's all out for 198.

There are several cases where players have performed at a similar standard according to the usual measures, but received quite different performance measures. For instance O'Donnell (57 runs off 60 balls not out, performance measure 10.73) has performed better in terms of the current statistics then Ranatunga (55 runs off 70 balls, performance measure 14.32). The reason for Ranatunga's performance measure being greater than O'Donnell's is that Ranatunga's innings took place between overs 15 and 41 while O'Donnell's innings took place between overs 34 and 49. At the end of the innings, when O'Donnell was batting, Australia still had 5 wickets in hand. The wickets in hand constraint thus becomes less relevant (as the innings ends after 50 overs regardless of the number of wickets lost), and therefore the model expects O'Donnell to score quickly. On the other hand Sri Lanka lost wickets regularly, so Ranatunga could not afford to score as quickly.

The bowlers J.Ratnayeke (1/47 from 9 overs) and R.Ratnayake (1/43 from 9.5 overs) have recorded very similar figures however their performance measures are quite different (-8.28 and 1.65 respectively). This is due to the stage of the innings at which the bowlers bowled their overs. J.Ratnayeke bowled most of his overs very early in the innings (overs 2, 4, 6, 8, 10, 12, 14, 43 and 45) when the expectation on batsmen is not as great as later in the innings. R.Ratnayake bowled his overs closer to the end of the innings, when the expectation on batsmen (and therefore the reward to bowlers when runs are not scored) is greater. He bowled overs 11, 13, 15, 17, 19, 21, 40, 42, 47 and 49. These two examples show how the current statistics do not take into account the constraints involved in the game of one-day cricket.

TABLE 14.2. Performance measures for match on 26/12/89
First Innings - Australia batting

Batsman	Runs	Balls	Performance measure
M.Taylor	11	25	-10.1
G.Marsh	38	82	2.9
D.Boon	11	19	-7.8
D.Jones*	85	89	34.9
A.Border	11	15	-5.6
S.Waugh	5	7	-7.8
S.O'Donnell*	57	60	10.7
Extras	4		4.0

Bowler	Overs	Wickets	Runs	Performance measure
G.Labrooy	9	0	40	-11.0
J.Ratnayeke	9	1	47	-8.3
R.Ratnayake	9.5	1	43	1.7
A.Ranatunga	10	3	41	10.3
E.DeSilva	10	0	42	-11.1
A.Gurusinha	1	0	9	-2.9
Run Outs	-	0	-	0.0

* = not out

TABLE 14.2 (cont). Performance measures for match on 26/12/89
Second Innings - Sri Lanka batting

Batsman	Runs	Balls	Performance measure
R.Mahanama	36	67	-0.9
M.Samarasekera	30	50	0.8
A.Ranatunga	55	70	14.3
P.DeSilva	9	10	-6.2
S.Jayasuriya	3	5	-10.1
A.Gurusinha	22	40	-4.9
H.Tillekeratne	11	23	-12.2
J.Ratnayeke	0	2	-4.9
E.DeSilva	13	17	0.9
G.Labrooy *	6	2	-1.3
R.Ratnayake	0	1	-12.5
Extras	7		7.0

Bowler	Overs	Wickets	Runs	Performance measure
M.Hughes	9.2	2	41	6.1
G.Campbell	10	0	36	-8.1
S.O'Donnell	9	4	36	19.1
S.Waugh	6	1	26	1.1
P.Taylor	10	2	36	7.8
A.Border	3	0	17	-4.3
Run Outs	-	1	-	8.5

* = not out

The results from the match shown in Table 14.2 are presented in a summarised form in Table 14.3.

TABLE 14.3. Player performance summary for match on 26/12/89

Player	Batting performance measure	Bowling performance measure	Total performance measure
D.Jones	34.90		34.90
S.O'Donnell	10.73	19.05	29.78
A.Ranatunga	14.32	10.31	24.63
P.Taylor		7.79	7.79
M.Hughes		6.07	6.07
G.Marsh	2.86		2.86
M.Samarasekera	0.84		0.84
R.Mahanama	-0.93		-0.93
P.DeSilva	-6.24		-6.24
S.Waugh	-7.77	1.07	-6.70
A.Gurusinha	-4.91	-2.87	-7.78
D.Boon	-7.83		-7.83
G.Campbell		-8.14	-8.14
A.Border	-5.58	-4.31	-9.89
M.Taylor	-10.05		-10.05
S.Jayasuriya	-10.13		-10.13
E.DeSilva	0.92	-11.07	-10.15
R.Ratnayake	-12.45	1.65	-10.80
H.Tillekeratne	-12.21		-12.21
G.Labrooy	-1.31	-11.01	-12.32
J.Ratnayeke	-4.90	-8.28	-13.18

The performance measures suggest that there were three players whose performance was well above the others, Dean Jones (performance measure of 34.90), Simon O'Donnell (29.78) and A.Ranatunga (24.63). The 'man of the match' award which is presented to the player that the commentators feel was the best player for that match was given to Simon O'Donnell who was selected as the second best player by the performance measures. Of course the adjudicators may have included an allowance for fielding that the performance measure ignores. It is of interest that it was in fact O'Donnell who ran out the Sri Lankan batsman, which Table 14.2 shows cost Sri Lanka another 8.5 runs.

The measures for the complete 1989/90 Benson & Hedges Series have been calculated and Tables 14.4 and 14.5 show the top ten ranked batsmen and bowlers respectively, ranked in terms of average performance measure per innings. Only batsmen facing more than 100 balls and bowlers who have bowled more than 100 balls are included in these tables.

14.6. Further work

The current statistics kept on players do not allow for the varying dimensions of grounds or the ease or difficulty of batting on certain pitches. This is also the case with the performance measures. Currently the performance measures are all calculated using a single relationship between R and p_d. This relationship results in the expected score with 300 balls to go and 10 wickets in hand being 215. For a pitch that is very difficult (or very easy) to score on this may be unacceptable and produce measures where all the batsmen have high negative (positive) measures and all the bowlers have the reverse. This makes it difficult to compare performances in different matches. In common with current measures, we have here used the average measure over a series of matches to hopefully even out any injustices. This could be overcome at the expense of greater complication by using a different relationship for matches played on different pitches that reflects better the expected innings score on that pitch. Alternatively the deviation of each player's performance measure from the average measure of his team members for the innings would result in a performance measure that takes into account the standard of the pitch and the opposition and allow for a relative comparison of team members who played in different matches.

TABLE 14.4. Top 10 batsmen ranked by average batting performance measure

Batsman	Country	Matches	Average measure
D.Jones	Australia	9	15.42
A.Ranatunga	Sri Lanka	6	10.03
T.Moody	Australia	6	9.94
W.Akram	Pakistan	6	9.79
D.Boon	Australia	3	6.99
A.Border	Australia	9	4.22
M.Taylor	Australia	9	4.10
S.Malik	Pakistan	6	3.21
G.Marsh	Australia	4	2.96

TABLE 14.5. Top 10 bowlers ranked by average bowling performance measure

Bowler	Country	Matches	Average measure
S.O'Donnell	Australia	9	10.69
T.Alderman	Australia	8	9.94
C.Rackermann	Australia	7	4.72
P.Taylor	Australia	8	4.39
P.DeSilva	Sri Lanka	3	2.68
M.Hughes	Australia	4	2.65
W.Akram	Pakistan	7	1.24
G.Campbell	Australia	6	0.65
A.Border	Australia	5	-0.08
A.Gurusinha	Sri Lanka	4	-2.38

Although the model assumes that each team is made up of 10 batsmen of equal standard, this is never the case. Generally a team has specialist batsmen, a wicket keeper and specialist bowlers. Therefore when there are only specialist bowlers left to bat, batsmen place more emphasis on not losing a wicket (since these batsmen will not score many runs) than the dynamic programming formulation allows for, as it assumes that the batsmen still to bat are of equal standard with the batsmen at the crease. An adjustment to the model could be made to allow for this at the expense of increased complication.

14.7. Conclusion

This paper outlines a method of using a dynamic programming formulation to calculate performance measures for players involved in the game of one-day cricket. The measures represent a player's contribution to the expected score of the team, and better reflect the constraints involved than the measures currently used by automatically allowing for the stage and state of the game when runs are scored and wickets taken. They allow the performance of batsmen and bowlers to be compared directly.

While this measure is an improvement on the current statistics kept on one-day cricketers it does have its limitations that must be considered. The measure takes no account of the fielding ability of players which is a very important part of the one-day game. The measure treats all performance on an individual level whereas cricket is a team game and this must be considered when selecting the best representative team. However all of the limitations of the method outlined apply equally well to the current measures such as batting and bowling averages and run and strike rate.

It is hoped that this new method of calculating performance measures will be used by commentators to determine the 'man of the match', by selectors to help with selecting the strongest team and for fans of the game with listings of the best one-day players in the world.

CONCLUSION

What have I shown? Sport abounds with untested assertions. Carlton were unlucky not to make the finals. West Coast has a huge home advantage. Essendon had a bad draw this year. Geelong is a two to one chance tomorrow. The McIntyre final eight is a terrible play-off system. Wickets in hand are important at the end of one day innings. Runs should be sacrificed to protect a weak batsman. This thesis has shown how mathematical models can provide quantitative evidence relevant to such statements.

Performance measurement is best derived from a mathematical model. Whether we are measuring tactics, team performance, home advantage or the competition rules as a whole, measures should be based on a model of the system. Different performances can then be judged against a standard model. For example, a model that includes a performance measure and a home advantage allows the performance of two teams against a common opponent on different grounds to be compared. A model with set probabilities of teams in different positions winning facilitates the comparison of two different finals systems. Only through a model that allows for team goals can individual performance be correctly measured.

This thesis has investigated the application of mathematical models in sport. I have looked at applications in football and cricket, with particular emphasis on tactics and measuring performance. Performance measurement in these team sports has been applied at three levels, the individual player, the team, and the competition structure in which they compete. The importance of variability has been a recurring theme in my work, and this has resulted in an analysis of home advantage. The implementation of these ideas has been demonstrated by their use in a forecasting model which compares favourably to the expert tipsters and has received wide media coverage.

In broad terms the early chapters on football discuss measurement of team ratings and home advantage. For home ground advantage, the ad hoc measure of percentage of matches won by home team is shown to depend not only on home ground advantage but the spread of ability or average performance levels on a neutral ground of the teams. In addition, this measure is inappropriate for individual teams, as it makes no allowance for the quality of team and opposition. The measures demonstrated here depend on fitting a mathematical model to the match results, usually by least squares. It shows that models incorporating individual home advantages provide a significantly better fit over the common home advantage models previously used. The thesis has extended the measurement of a common home advantage to methods for measuring individual home advantages and paired home advantages. For the first time individual home advantage of Australian rules and English soccer clubs have been published. While the main purpose has been the calculation of home advantages, rather than the investigation of such causes, I have demonstrated a significant 'isolation' factor. The applicability of the suggested models has been shown by their use in forecasting Australian rules with an accuracy comparable to the expert tipsters.

All sports competitions take place under a framework of some overall competition rules. These are designed to produce a winner, or ladder order, or some other measure of overall success for the individual teams. The thesis investigates the performance of particular football and cricket competitions and finals systems, and show that in most cases the systems are not balanced for strength of opposition or home advantage, and that some teams are disadvantaged by competition rules.

The latter part of the thesis investigates the use of mathematical models for tactics. For the first time DP models in cricket are used to determine optimal strategies under a range of models and objective functions. The use of such models as the basis for player performance ratings that reward players for their contribution to team goals is demonstrated.

The major contributions of each chapter to the overall themes of the thesis are detailed below.

Chapter I gives an overview of the use of Operational Research methods in sport, with particular emphasis on the areas in which I have been involved. This places most of my work in a backdrop of the work being done at the time.

Chapter II introduces Australian rules football, and investigates the problem of estimating team performance measures and home advantage in a competition which is neither balanced for strength of opposition nor home grounds. It discusses the drawbacks of the usual measures, and calculates home advantages of individual clubs by using various techniques. It shows:

- The quality of opposition can be allowed for by looking at the paired HA. These clearly show an isolation factor.
- The individual HAs for all clubs are calculated in alternative ways. An ordering of the clubs by HA clearly shows an isolation effect with interstate clubs heading the table and inner city Melbourne clubs bringing up the rear.
- There was evidence for MCG teams and teams playing for the first season on a new ground having a lower than average HA.
- Investigation of models of varying complexity shows that the use of models more detailed than those incorporating only a common home advantage is justified.

Chapter III investigates special techniques for calculating team ratings and HAs in balanced competitions by investigating 15 years of English Association football data. It shows:

- The previous method used in the literature of calculating HA for competitions is unsuitable for individual teams.
- The existence of a spurious HA, due to the HA of all other teams, when home and away performances of a particular team are compared.
- Fitting a model allowing for ability and individual HA by least squares to individual match results is equivalent to a simple method based on the end of year ladder.
- The individual HAs in goals per match of all 94 clubs for the 15 years 1980-81 to 1994-95 are calculated. Analysis of these show no division effects but significant year effects and reasonable evidence that clubs do not have a common HA. There was evidence of an 'isolation effect' with three of the top 10 clubs being particularly isolated, and many London clubs with low HAs.
- The isolation hypothesis was confirmed by showing a definite linear relationship between a pair of clubs paired HA and their distance apart.

Chapter IV gives a case study showing an implementation of the above methods for computer forecasting. It shows:

- Automated forecasting systems using variations of the models discussed in earlier chapters can be successfully applied to Australian rules football.
- A simple program using a pre calculated common home advantage and exponential smoothing to produce team ratings was used to provide forecasts for the media for several years.
- An 'improved' program used an extension of individual home advantages and a power method to measure error.

Chapter V looks at the absolute and relative performance of the computer model of Chapter IV by analysing in detail its performance in 1991. It shows

- The computer performed in 1991 better than most expert tipsters in predicting winners and margins.
- Unlike the computer, humans are biased toward their own team, and generally do not select enough close margins.

Chapter VI combines the previous chapters by looking at some international comparisons for both forecasting and home advantage.

- A comparison of HA in several international and national competitions including soccer, American and Australian football and baseball, showed HA varied with the amounts of the three factors normally suggested as causes of HA.
- The performance of radically different automated systems applied to the same data tend to be similar. This suggests that the accuracy is limited by the data content, and more data rather than improved methods is necessary to make progress.

Chapter VII looks at the measurement of the fairness of competition rules under which teams compete. By assuming all teams are equal, or have probability of victory and HA as developed by the computer models of Chapter IV it shows:

- The computers estimates of ground effects and errors of prediction can be used to evaluate the effect of a change of venue on a team's chances of winning .
- The AFL draw is unfair, and the imbalance does not appear to even out over the years.
- Ladder positions at the end of home and away matches are affected by up to 10 places by randomness.
- The Macintyre final eight system passes most tests of fairness. However the chance of two teams meeting in the grand final is not in order of their combined ladder position. Moreover a lack of consistency from year to year on a team's path to the finals increases the effects of randomness.
- Bookmakers odds do not reflect the intricacies of the AFL finals draw.
- The knockout structure of the finals increases the effect of home advantage.

Chapter VIII uses a DP approach to investigate the correct tactics near the end of a game in a commonly occurring situation in Australian rules. Since individual players should make decisions that maximise some goal for his team, a knowledge of the correct strategy is sometimes important in assessing the effect of competition rules.

- A simple Markov model is set up by dividing the ground into seven areas, with strategies depending on the area. The model is checked by investigating steady state probabilities and estimated scores during a match.
- It is often advantageous, depending on the score difference and the time remaining in the match, to concede a behind.
- The effects on the optimal strategy due to a proposed rule change are shown.

Chapter IX gives a comprehensive literature survey of cricket.

Chapter X shows the importance of HA and competition rules applies equally in cricket. By analysing the Sheffield Shield results it shows:

- HA exists in cricket. Most teams win most of the points allocated on their home grounds.
- The competition rules create an unfair home advantage. Some teams have many more points allocated on their home grounds.
- Two suggested scoring methods of partly overcoming the problem always produce the same rank ordering of teams at the end of the year.

Chapter XI continues the theme of tactical evaluation and applies a DP formulation to one day cricket to investigate optimal run rate. In the first recorded application of Operational Research methods to cricket, both a first and second innings formulation are solved numerically, and the implications for competition rules and player performance measures are noted. The model assumes all batsmen are of the same ability. It shows

- Contrary to current practice the optimal run rate is always faster than the expected average rate for the remainder of the innings.
- A heuristic gives a good outcome under a range of relationships between probability of dismissal and scoring rates. Batsmen should score at the rate which would see their last wicket fall at the end of the final over.
- The results can be applied to selection decisions and player evaluation.
- The current rain interruption rules are unfair. The model could be used to develop fairer rules.

Chapter XII investigates a DP model applicable to a common situation in test cricket with batsmen of two different abilities. Should they refuse a possible run in order to protect the weaker batsmen? It shows:

- A model minimising the chance of the weaker batsman being on strike at the start of the next over is solvable analytically.
- Trying to put the better batsman on strike at the start of the over does not necessarily maximise the expected duration of the partnership.
- To maximise the expected duration of the partnership the better batsman should generally run if possible off the second last or last ball of the over.

Chapter XIII develops further DP models that extend those of Chapters XI and XII to allow for batsmen of different ability and scoring profiles where the objective function is to maximise the number of runs. It shows

- Various relationships and symmetry can reduce the complexity of models where the states are the facing and non facing batsmen and the number of runs on offer.
- Numerical solutions of such models can be used to advise batsmen on tactics near the end of an innings.
- The practice of refusing runs to protect the weaker batsman is sensible under certain conditions.

Chapter XIV shows how the models of the previous three chapters could be used for player performance measures that reflect a players contribution to the team goals.

- The first innings dynamic programming formulation of Chapter XI is used to develop an innovative method of calculating the contribution, in runs, made by batsmen and bowlers to the team's score in a game of one-day cricket.
- Data from a one day series was used to apply the method to all players. Tables of the measure are given and compared with the usual statistics.

Where to from here? Beginning from 1996, I have been a collaborator in the collection of player statistics from AFL matches. The statistics have been collected from video and live, and consist of better quality data than has been collected previously. Rather than just tally the number of possessions and disposals, they have been rated for quality. Possessions are rated as to difficulty of obtaining, and kicks and handballs have been categorised as long or short, effective or not. The quality and power of the statistics is such they have obtained widespread publicity on both radio and television, (Hopkins (1996), Wright (1996) and are purchased by most clubs to assist in training and planning tactics against opponents. The data is being analysed to determine the contribution of each performance statistic to winning performance, with a view to developing a player rating based on the statistics. Further work will involve analysis of the data to determine if home ground advantage manifests itself in the number of possessions or quality of possession and disposal. It will also be of interest to determine if incorporation of this data can improve the computer predictions. Alternatively the methods demonstrated for predicting final scores could be applied to the secondary data itself, to predict the style of game that will be played. Since the data has some coding which places the action in broad areas of the field, it will also find a use in improving the Markov decision model of football. Late in 1997, discussions began with the Victorian Institute of Sport on extending this data collection into other sports, particularly cricket. This will overcome the lack of ball by ball data which hampers much of the work in this area. A closer look at HA in cricket can be undertaken, and many of the possible investigations outlined in Chapter IX may become feasible.

Research results need to be brought to the attention of possible users. In sport these are players, administrators and fans. The publication in the popular media of the computer predictions for Australian rules football has been discussed in the thesis. During 1997, the analysis of the football statistics discussed above were the basis of a weekly article in the Australian Financial Review (e.g., Hopkins, 1997a). Late in 1997, with the end of the football season, these articles were extended to other sports. Hopkins (1997b, 1997c, 1997d, 1997e, 1997f) are examples that have discussed my research work in cricket, golf and soccer. This coverage will continue and hopefully extend to other sports and other researchers.

In the last 20 years the applications of mathematical modelling to sport has increased significantly. However there are still many problems which will yield to contributions from Operational Researchers and Statisticians. A continuing challenge is to demonstrate to players, administrators and supporters that modellers have a continuing role in measuring and improving sports performance.

BIBLIOGRAPHY

Alexander, D., McClements, K., & Simmons, J. (1988). Calculating to win. *New Scientist*, 10 December, 30-33.

Barnett, V., & Hilditch, S. (1993). The effect of an artificial pitch surface on home team performance in football (soccer). *Journal of the Royal Statistical Society (Series A)*, 156 (1), 39-50.

Baumeister, R. F., & Steinhilber, A. (1984). Paradoxical effects of supportive audiences on performance under pressure: The home field disadvantage in sports championships. *Journal of Personality and Social Psychology*, 47, 85-93.

Bellman, R. (1977). Dynamic Programming and Markovian decision processes, with application to baseball. In S. P. Ladany & R. E. Machol (Eds.), *Optimal Strategies in Sports* (pp. 77-85). Amsterdam: North Holland.

Benjafield, J., & Liddel, W. W. (1989). Is there a home field disadvantage in professional sports championships. *Social Behavior and Personality*, 17 (1), 45-50.

Berkmann, M. (1990). *Deloitte Ratings: The Complete Guide to Test Cricket In the Eighties*. London: Transworld.

Bland, N. D., & Bland, J. M. (1996). Correspondence: Home ground advantage and average home attendance. *The Statistician*, 45 (3), 381.

Bosi, A. (1976). The effect of bonus points on the Cricket County Championships. *Mathematical Spectrum*, 9 (3), 75-82.

Brooks, W. J., & Hughston, L. P. (1988). A problem in squash strategy. *The Mathematical Gazette*, 72 (460), 93-95.

Clarke, S. R. (1979). Tie point strategy in American and International squash and badminton. *Research Quarterly* , 50 (4), 729-734.

Clarke, S. R. (1981). Lou versus the computer. Paper presented at Australian Mathematics Society one day meeting on Mathematical Problems in Competitive Sport, Pt Cook.

Clarke, S. R. (1984). Mathematics in sport. In A. Maurer (Ed.), *Conflicts in Education* (pp. 130-135). Melbourne: Mathematics Association of Victoria.

Clarke, S. R. (1985a). Mathematics in tennis. *Function*, 9 (2), 21-24.

Clarke, S. R. (1985b). Squash mathematics. *Function*, 9 (3), 25-28.

Clarke, S. R. (1986a). Another look at the 1985/86 Sheffield Shield competition cricket results. *Sports Coach*, 10 (3), 16-19.

Clarke, S. R. (1986b). Mathematics in cricket. In N. Ellerton (Ed.), *Mathematics, Who Needs What?* (pp. 299-302). Melbourne: Mathematics Association of Victoria.

Clarke, S. R. (1988a). Computer tipping of football (and other sports). In D. Firth (Ed.), *Maths Counts - Who Cares?* (pp. 243-247). Melbourne: Mathematics Association of Victoria.

Clarke, S. R. (1988b). Dynamic Programming in one day cricket - Optimal scoring rates. *Journal of the Operational Research Society*, 39, 331-337.

Clarke, S. R. (1988c). Tinhead the tipster. *OR Insight*, 1 (1), 18-20.

Clarke, S. R. (1991a). Australian student projects on sport result in a USA research connection. *ASOR Bulletin*, 10 (2), 2-5.

Clarke, S. R. (1991b). Consistency in sport - with particular reference to cricket. *Proceedings 27th NZ OR conference*, 30-35.

Clarke, S. R. (1992a). Computer and human tipping of AFL football - a comparison of 1991 results. In N. de Mestre (Ed.), *Mathematics and Computers in Sport* (pp. 81-93). Gold Coast, Qld: Bond University.

Clarke, S. R. (1992b). *How mathematics, with a little help from computers, can solve problems in sport* [Audio Tape]. Padstow, NSW: Quickcopy.

Clarke, S. R. (1993a). Analysing football finals with a spreadsheet. *Australian Senior Mathematics Journal*, 7 (1), 35-44.

Clarke, S. R. (1993b). Computer forecasting of Australian rules football for a daily newspaper. *Journal of the Operational Research Society*, 44, 753-759.

Clarke, S. R. (1994a). An adjustive rating system for tennis and squash players. In N. de Mestre (Ed.), *Mathematics and Computers in Sport* (pp. 43-50). Gold Coast, Qld.: Bond University.

Clarke, S. R. (1994b). Analysing 1993 AFL football data with SAS/JMP. In C. Beesey & D. Rasmussen (Eds.), *Mathematics Without Limits* (pp. 156-162). Brunswick: Mathematics Association of Victoria.

Clarke, S. R. (1994c). *Variability of scores and consistency in sport.* In N. de Mestre (Ed.), *Mathematics and Computers in Sport* (pp. 69-81). Gold Coast, Qld.: Bond University.

Clarke, S. R. (1995). Analysing 1994 AFL football data with JMP. In A. Richards (Ed.), *FLAIR Forging Links and Integrating Resources* (pp. 93-100). Darwin: Australian Association of Mathematics Teachers.

Clarke, S. R. (1996a). Calculating premiership odds by computer - an analysis of the AFL final eight playoff system. *Asia Pacific Journal of Operational Research*, 13 (1), 89-104.

Clarke, S. R. (1996b). Home advantages in balanced competitions - English soccer 1990-1996. In N. de Mestre (Ed.), *Mathematics and Computers in Sport* (pp. 111-116). Gold Coast, Qld.: Bond University.

Clarke, S. R. (1997a). *Home ground advantage in the Australian Football League, 1980-85.* Paper presented at Asia Pacific Operational Research Society Conference, Melbourne.

Clarke, S. R. (1997b). Test Statistics. In J. Bennet (Ed.), *Statistics in Sport*. Edward Arnold. (to appear)

Clarke, S. R., & Handley, D. (1994). What do mathematicians do? Some projects of Swinburne mathematics students. In C. Beesey & D. Rasmussen (Eds.), *Mathematics Without Limits* (pp. 163-168). Brunswick: Mathematics Association of Victoria.

Clarke, S. R., & Norman, J. M. (1978). Playing up to 10 - What are the chances? *The Squash Player* (July), 50-51.

Clarke, S. R., & Norman, J. M. (1979). Comparison of North American and International squash scoring systems - Analytical results. *Research Quarterly*, 50 (4), 723-728.

Clarke, S. R., & Norman, J. M. (1994). Home ground advantage of individual clubs in English soccer. *Swinburne University of Technology School of Mathematics Technical Report series*.

Clarke, S. R., & Norman, J. M. (1995a). Home ground advantage of individual clubs in English soccer. *The Statistician*, 44 (4), 509-521.

Clarke, S. R., & Norman, J. M. (1995b). Protecting a weak batsman from taking strike at the start of a new over - A simple Dynamic Programming model in cricket. In D. L. Hoffman (Ed.), *Network Conferencing, proceedings of ASOR: The 13th National Conference* (pp. 45-55). Canberra: Australian Society of Operations Research.

Clarke, S. R., & Norman, J. M. (1996). Authors' reply. *The Statistician*, 45 (3), 381-383.

Clarke, S. R., & Norman, J. M. (1997a). Authors' replies. *The Statistician*, 46 (3), 443-444.

Clarke, S. R., & Norman, J. M. (1997b). Dynamic Programming in Australian rules football: When to rush a behind. In R. L. Jenson & I. R. Johnson (Eds.), *Proceedings of the Twenty-Sixth Annual Meeting of the Western Decision Sciences Institute* (pp. 741-743). Hawaii: The Decision Sciences Institute.

Clarke, S. R., & Norman, J. M. (1997c). To run or not? Some Dynamic Programming models in cricket. In R. L. Jenson & I. R. Johnson (Eds.), *Proceedings of the Twenty-Sixth Annual Meeting of the Western Decision Sciences Institute* (pp. 744-746). Hawaii: The Decision Sciences Institute.

Clarke, S. R., & Norman, J. M. (1998a). Dynamic Programming in cricket: Protecting the weaker batsman. *Asia Pacific Journal of Operational Research*, in press.

Clarke, S. R., & Norman, J. M. (1998b). When to rush a behind in Australian rules football: a Dynamic Programming approach. *Journal of the Operational Research Society*, in press.

Clarke, S. R., & Rice, J. M. (1995). How well do golf courses measure golf ability? An application of test reliability procedures to golf tournament scores. *ASOR Bulletin*, 14 (4), 2-11.

Clemen, R. T., & Winkler, R. L. (1985). Limits for the precision and value of information from dependent sources. *Operations Research*, 33, 427-442.

Courneya, K. S., & Carron, A. V. (1991). Effects of travel and length of home stand/road trip on the home advantage. *Journal of Sport & Exercise Psychology*, 13, 42-49.

Courneya, K. S., & Carron, A. V. (1992). The home advantage in sport competitions: A literature review. *Journal of Sport & Exercise Psychology*, 14, 13-27.

Croucher, J. S. (1979). The battle for the ashes - a statistical analysis of post-war Australia-England cricket, *Proceedings of the 4th National Conference of the Australian Society of Operations Research* (pp. 48-58): Australian Society of Operations Research.

Croucher, J. S. (1982a). Anglo-Australian test cricket dismissals 1946-1980. *Bulletin in Applied Statistics*, 9 (2), 179-193.

Croucher, J. S. (1982b). Australia - England test cricket. *Teaching Statistics*, 4 (1), 21-23.

Croucher, J. S. (1982c). The effect of the tennis tie-breaker. *Research Quarterly*, 53 (4), 124-127.

Croucher, J. S. (1984). The effect of changing competition points in the English Football League. *Teaching Statistics*, 6 (2), 39-42.

Croucher, J. S. (1986). The conditional probability of winning games of tennis. *Research Quarterly*, 57 (1), 23-26.

Croucher, J. S. (1987). Scientific method in sports. *Proceedings of the 8th National Conference of the Australian Society of Operations Research* (pp. 44-54): Australian Society of Operations Research.

Croucher, J. S. (1994). Ranking soccer teams. *Teaching Statistics*, 16 (3), 90-92.

Crowe, S. M., & Middeldorp. (1996). A comparison of leg before wicket rates between

Australians and their visiting teams for test cricket series played in Australia, 1977-94. *The Statistician*, 45 (2), 255-262.

Deloittes. (1988). *Deloittes ratings performance review of test batsmen and test bowlers over 1987/88 winter series* : Test and County Cricket Board.

Dixon, M. J., & Coles, S. G. (1996). Modelling Association football scores and inefficiencies in the UK football betting market. *Journal of the Royal Statistical Society, Series C*, (To appear).

Dixon, M. J., & Robinson, M. E. (1996). A birth process model for Association football matches. *Private communication*.

Dowe, D. L., Farr, G. E., Hurst, A. J., & Lentin, K. (1996). Information theoretic football tipping. In N. de Mestre (Ed.), *Mathematics & Computers in Sport conference* (pp. 233-241). Gold coast, Qld.: Bond University.

Dowie, J. (1982). Why Spain should win the World cup. *New Scientist*, 94, 693-695.

Duckworth, F. (1997). Dealing satisfactorily with interruptions in one-day matches. *The Cricket Statistician* (98), 10-15.

Duckworth, F., & Lewis, T. (1996). A fair method for resetting the target in interrupted one-day cricket matches. In N. de Mestre (Ed.), *Mathematics and Computers in Sport* (pp. 51-68). Gold Coast, Qld: Bond University.

Dundas, R. (1991). Benson and Hedges World Series Statistics. In G. Thomas (Ed.), *World Series 1991/92 Guide* (pp. 73-74): PBL Marketing.

Elderton, W. E. (1927). *Frequency Curves and Correlation*. (2nd ed.). London: Layton.

Elderton, W. E. (1945). Cricket scores and some skew correlation distributions. *Journal of the Royal Statistical Society (Series A)*, 108, 1-11.

Elderton, W. P., & Elderton, E. M. (1909). *Primer of Statistics*. London: Black.

Elo, A. E. (1978). *The rating of chess players, past and present*. London: Batsford.

Engel, M. (Ed.). (1997). *Wisden Cricketers' Almanack*. (Vol. 134). Guilford: John Wisden.

Ganesalingam, S., Kumar, K., & Ganeshanandam, S. (1994). *A statistical look at cricket data*. In N. de Mestre (Ed.), *Mathematics and Computers in Sport* (pp. 89-104). Gold Coast, Qld.: Bond University.

Gayton, W. F., Mathews, G. R., & Nickless, C. J. (1987). The home field disadvantage in sports championships: does it exist in hockey? *Journal of Sport Psychology*, 9, 183-185.

Harville, D. (1977). The use of linear model methodology to rate high school or college football teams. *Journal of the American Statistical Society*, 73 (358), 278-289

Harville, D. A. (1980). Predictions for National Football League games via linear-model methodology. *Journal of the American Statistical Association*, 75, 516-524.

Harville, D. A., & Smith, M. H. (1994). The home-court advantage: How large is it, and does it vary from team to team? *The American Statistician*, 48 (1), 22-28.

Hastings, N. A. J. (1973). *Dynamic Programming with Management Applications*. London: Butterworths.

Hayes, M., & Norman, J. M. (1984). Dynamic programming in orienteering: Route choice and the siting of controls. *Journal of the Operational Research Society,* 35, 791-796.

Haygarth, A. (1862-1885, 1925). *Cricket scores and biographies of celebrated cricketers*: Vols1-IV, Lilywhite, Vols V-XV, Longman.

Hopkins, T. (1996). A new standard in stats. *The Herald Sun*, (April 2), pp. 69.

Hopkins, T. (1997a). Kicking matters most. *The Australian Financial Review*, (August 22), Weekend Review, pp. 4.

Hopkins, T. (1997b). Shield points system is full of bias. *The Australian Financial Review*, (October 11-12), pp. 62.

Hopkins, T. (1997c). What a difference a day makes - if the programming is dynamic. *The Australian Financial Review*, (November 8-9), pp. 62.

Hopkins, T. (1997d). Success came back when spoiled brats turned into focussed athletes. *The Australian Financial Review*, (November 15-16), pp. 62.

Hopkins, T. (1997e). Amen to management of Botany Bay breezes on PGA holes 13-17. *The Australian Financial Review*, (November 22-23), pp. 62.

Hopkins, T. (1997f). Scientists hoping to score in the hunt for soccer success. *The Australian Financial Review*, (November 8-9), pp. 62.

James, W., & Stein, C. (1981). Estimation with quadratic loss. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1, 361-379.

Johnston, M. I. (1992). *Assessing team and player performance in one day cricket using Dynamic Programming.* Unpublished M.App.Sc, Swinburne University.

Johnston, M. I., Clarke, S. R., & Noble, D. H. (1992). An analysis of scoring policies in one day cricket. In N. de Mestre (Ed.), *Mathematics and Computers in Sport* (pp. 71-80). Gold Coast, Qld: Bond University.

Johnston, M. I., Clarke, S. R., & Noble, D. H. (1993). Assessing player performance in One Day Cricket using Dynamic Programming. *Asia-Pacific Journal of Operational Research*, 10, 45-55.

Kimber, A. (1993). A graphical display for comparing bowlers in cricket. *Teaching Statistics*, 15 (3), 84-86.

Kimber, A. C., & Hansford, A. R. (1993). A statistical analysis of batting in cricket. *Journal of the Royal Statistical Society A*, 156 (3), 443-445.

Kohler, D. A. (1982). Optimal strategies for the game of darts. *Journal of the Operational Research Society*, 33, 871-884.

Kuk, A. Y. C. (1995). Modelling paired comparison data with large numbers of draws and large variability of draw percentages among players. *The Statistician*, 44 (4), 523-528.

Ladany, S. P., & Machol, R. E. (1977). *Optimal strategies in sports*. Amsterdam: North Holland.

Leake, R. J. (1976). A method for ranking teams with an application to college football. In R. E. Machol, S. P. Ladany, & D. G. Morrison (Eds.), *Management Science in sports* (pp. 27-46).Amsterdam: North Holland.

Longford, N. (1997). Correspondence: Home ground advantage and average home attendance. *The Statistician*, 46 (3), 443.

Lunn, H. (1993). The legendary cricket draw. In R. Cashman, D. Headon, & G. Kinross-Smith (Eds.), *The Oxford Book of Australian Sporting Anecdotes* (pp. 194-195). Melbourne: OUP.

Machol, R. E., Ladany, S. P., & Morrison, D. G. (1976). *Management Science in Sports*:.Amsterdam: North Holland.

Miles, R. E. (1984). Symmetric sequential analysis: the efficiencies of sports scoring systems (with particular reference to those of tennis). *Journal of the Royal Statistical Society, Series B*, 46, 93-108.

Monahan, J. P., & Berger, P. D. (1977). Play-off structures in the National Hockey League. In S. P. Ladany & R. E. Machol (Eds.), *Optimal Strategies in Sports* (pp. 123-128). Amsterdam: North Holland.

Morris, C. (1977). The most important points in tennis. In S. P. Ladany & R. E. Machol (Eds.), *Optimal strategies in sports* (pp. 131-140). Amsterdam: North Holland.

Morrison, D. G., & Schmittlein, D. S. (1991). How many forecasters do you really have? Mahalanobis provides the intuition for the surprising Clemen and Winkler result. *Operations Research*, 39, 519-523.

Mottley, C. M. (1954). The application of Operations Research methods to athletic games. *Operations Research*, 2, 335-338.

Norman, J. M. (1985). Dynamic programming in tennis - when to use a fast serve. *Journal of the Operational Research Society*, 36, 75-77.

Norman, J. M. (1995). Dynamic programming in sport: a survey of applications. *IMA Journal of Mathematics Applied to Business and Industry*, 6 (December), 171-176.

Pace, A., & Carron, A. V. (1992). Travel and the home advantage. *Canadian Journal of*

Sport Sciences, 17, 60-64.

Pollard, G. H. (1983). An analysis of classical and tie-breaker tennis. *Australian Journal of Statistics*, 25, 496-505.

Pollard, G. H. (1986a). *A stochastic analysis of scoring systems.* Unpublished Ph.D., Australian National University.

Pollard, G. H. (1992). The optimal test for selecting the greater of two binomial probabilities. *Australian Journal of Statistics*, 34 (2), 273-284.

Pollard, R. (1977). Cricket and statistics. In S. P. Ladany & R. E. Machol (Eds.), *Optimal Strategies in Sports* (pp. 129-130). Amsterdam: North Holland.

Pollard, R. (1986b). Home advantage in soccer: A retrospective analysis. *Journal of Sport Sciences*, 4, 237-248.

Pollard, R., Benjamin, B., & Reep, C. (1977). Sport and the negative binomial distribution. In S. P. Ladany & R. E. Machol (Eds.), *Optimal Strategies in Sports* (pp. 188-195). Amsterdam: North Holland.

Reep, C., & Benjamin, B. (1968). Skill and chance in association football. *Journal of the Royal Statistical Society A*, 131, 581-585.

Reep, C., Pollard, R., & Benjamin, B. (1971). Skill and chance in ball games. *Journal of the Royal Statistical Society A*, 134, 623-629.

Riddle, L. H. (1988). Probability models for tennis scoring systems. *Applied Statistics*, 37, 63-75.

Rodgers, S., & Browne, A. (1996). *Every Game Ever Played*. (Fifth ed.). Ringwood: Penguin.

Russel, K. G. (1980). Balancing carry-over effects in round robin tournaments. *Biometrika*, 67 (1), 127-131.

Schutz, R. W. (1970). A mathematical model for evaluating scoring systems with specific reference to tennis. *Research Quarterly,* 41 (552-561).

Schutz, R. W., & Kinsey, W. K. (1977). Comparison of North American and International squash scoring systems - a computer simulation. *Research Quarterly*, 48 (248-251).

Schwartz, B., & Barsky, S. E. (1977). The home advantage. *Social Forces*, 55, 641-661.

Schwertman, N. C., & Howard, L. (1989). A probability model for the Victorian Football League finals series. *Australian Mathematics Teacher*, 45 (2), 2-3.

Schwertman, N. C., & Howard, L. (1990). Probability models for the Australian Football League. *The Australian mathematical Society Gazette*, 17 (4), 89-94.

Sheahan, M. (1996). *Stop the 'own goal'*. Herald Sun, Melbourne, 31 May.

Simmons, J. (1989). A probabilistic model of squash: strategies and applications. *Journal of the Royal Statistical Society (Series C),* 38 (1), 95-110.

Snyder, E. E., & Purdy, D. A. (1985). The home advantage in collegiate basketball. *Sociology of Sport Journal*, 2, 352-356.

Sphicas, G. P., Humes, J. W., & Ladany, S. P. (1975). The optimal aiming line. *Operational Research Quarterly*, 26 (3, i), 495-506.

Sphicas, G. P., & Ladany, S. P. (1976). Dynamic policies in the long jump. In R. E. Machol, S. P. Ladany, & D. G. Morrison (Eds.), *Management Science in sports* (pp. 113-124).Amsterdam: North Holland.

Stefani, R. T. (1977). Football and basketball predictions using least squares. *IEEE Transactions on systems, man, and cybernetics*, 7, 117-121.

Stefani, R. T. (1980). Improved least squares football, basketball and soccer predictions. *IEEE Transactions on Systems, Man and Cybernetics*, SMC -10 (2), 116-123.

Stefani, R. T. (1983). Observed betting tendencies and suggested betting strategies for European football pools. *The Statistician*, 32, 319-329.

Stefani, R. T. (1987). Applications of statistical methods to American football. *Journal of Applied Statistics*, 14, 61-73.

Stefani, R. T., & Clarke, S. R. (1991). Australian rules football during the 1980s. *ASOR Bulletin*, 10 (3), 11-15.

Stefani, R. T., & Clarke, S. R. (1992). Predictions and home advantage for Australian rules football. *Journal of Applied Statistics*, 19 (2), 251-261.

Sumner, J., & Mobley, M. (1981). Are cricket umpires biased. *New Scientist* (2 July), 29-31.

Thomas, L. C. (1978). Student exercise : Department of Decision Theory, University of Manchester.

Tobin, P. C., & Clarke, S. R. (1993). Some statistics of the triathlon. In J. Mousley & M. Rice (Eds.), *Mathematics: Of Primary Importance* (pp. 433-440). Brunswick: Mathematics Association of Victoria.

Townend, M. S. (1984). *Mathematics in Sport*. Ellis Horwood: Chichester.

Walsh, G. R. (1975). *Methods of Optimization*. Wiley: London.

Washburn, A. (1991). Still more on pulling the goalie. *Interfaces*, 21, 59-64.

White, D. J. (1993). *Markov Decision Processes*. Chichester: Wiley.

Willis, R. J. (1994). Scheduling the Australian state cricket season using simulated annealing. *Journal of the Operational Research Society*, 45, 276-280.

Willis, R. J., & Armstrong, J. (1993). Scheduling the World Cup of cricket - a case study. *Journal of the Operational Research Society*, 44, 1067-1072.

Wood, G. H. (1941). What do we mean by consistency? *The Cricketer Annual*, 22-28.

Wood, G. H. (1945). Cricket scores and geometrical progression. *Journal of the Royal Statistical Society (Series A)*, 108, 12-22.

Wright, C. (1996). Boot-up there, Cazaly. *The Australian Financial Review*, (20 September), pp. 1,10,11.

Wright, M. B. (1988). Probabilities and decision rules for the game of squash rackets. *Journal of the Operational Research Society*, 39, 91-99.

Wright, M. B. (1991). Scheduling English cricket umpires. *Journal of the Operational Research Society*, 42, 447-452.

Wright, M. B. (1992). A fair allocation of county cricket opponents. *Journal of the Operational Research Society*, 43, 195-201.