PROCEEDINGS OF THE EIGHTH AUSTRALASIAN CONFERENCE

ON

MATHEMATICS AND COMPUTERS IN SPORT

edited by
John Hammond
and
Neville de Mestre

8M&CS

July 3rd to 5th 2006

Coolangatta, Queensland, Australia

Published by MathSport (ANZIAM)

ISBN: 0 7331 0022 8

The papers in these proceedings have gone through a full peer review refereeing process.

CONFERENCE DIRECTOR'S REPORT

Of the first seven conferences, five were held at Bond University on the Gold Coast, one at the University of Technology in Sydney during the year of the Sydney Olympics, and the last one was in New Zealand at Massey University. This eighth conference in the series returns to the Gold Coast, but is longer at Bond University because campus accommodation for conference participants is no longer available at that venue. Instead we gather at Greenmount Beach Resort, which has been used during the past decade for a number of Queensland and Australian Applied Mathematics Conferences. There will be 33 papers presented during the days across topics which cover a variety of individual and team sports

Participants are attending from the United Kingdom, France, Germany, India, New Zealand and Australia. Steve Clarke and Neville de Mestre have been to all eight conferences. Tony Lewis, of the Duckworth-Lewis formula for determining the winner in rain-affected international cricket matches, has been a regular attendee since 1996. It is great to see John Norman from Sheffield appear again, as he was at the first conference in 1992.

The Mathsport community wishes to thank His Worship the Gold Coast Mayor and former world record holder over middle distances, Ron Clarke, for giving up some of his valuable time to officially open the conference. In addition, we are grateful to our Keynote speakers, Roger Bartlett (Professor of Sports Science at the University of Otago) and Dr Keith Lyons (Head of Performance Analysis at the Australian Institute of Sport) in attending delivering their addresses. We hope that all those attending have an enjoyable and rewarding interaction in sports science over the next few days.

John Hammond and Neville de Mestre

LIST OF CONFERENCE PARTICIPANTS[1]

Michael Bailey	Michael.Bailey@med.monash.edu.au
Tristan Barnett	TBarnett@groupwise.swin.edu.au
Roger Bartlett	rbartlett@pooka.otago.ac.nz
Anthony Bedford	anthony.bedford@ems.rmit.edu.au
Alan Brown	abrown@labyrinth.net.au
Stephen Clarke	SClarke@groupwise.swin.edu.au
Uday Damodaran	uday@xlri.ac.in
Neville de Mestre	ndemestr@bond.edu.au
Don Forbes	dgforbes@groupwise.swin.edu.au
Sam Glasson	S3028040@student.rmit.edu.au
Richard Green	richard.green@canterbury.ac.nz
John Hammond	jhammond@scu.edu.au
Tim Heazlewood	t.heazlewood@mackillop.acu.edu.au
Ros Kerr	rfk15@student.canterbury.ac.nz
Martin Lames	martin.lames@sport.uni-augsburg.de
Tony Lewis	ajlewis@brookes.ac.uk
Keith Lyons	Keith.Lyons@ausport.gov.au
Denny Meyer	DMeyer@groupwise.swin.edu.au
Ken Noble	noblesoft@netspeed.com.au
John Norman	J.Norman@sheffield.ac.uk
Barbara O'Riley	barbara@eurekabeads.com.au
Darren O'Shaughnessy	darrenos@championdata.com.au
Matthew Ovens	matthew.ovens@its.monash.edu.au
Graham Pollard	graham@foulsham.com.au
Adrian Schembri	S3020239@student.rmit.edu.au
George Serrallach	george.serrallach@verusco.com
Didier Seyfried	didier_seyfried@yahoo.fr
Aaron Silk	aaronsilk@hotmail.com
Richard Smith	

[1] At time of going to press

PAPERS OF THE EIGHTH AUSTRALASIAN CONFERENCE ON
MATHEMATICS AND COMPUTERS IN SPORT
(in order of presentation – presenters are **bolded** if multiple authors)

CONTENTS (DAY ONE)

KEYNOTE ADDRESS: GREATER THAN THE SUM OF OUR PARTS? Keith Lyons……...1

Analysis and Modelling
PREDICTION VERSUS REALITY: THE USE OF MATHEMATICAL MODELS TO PREDICT ELITE PERFORMANCE IN SWIMMING AND ATHLETICS AT THE OLYMPIC GAMES. Tim Heazlewood…….2-11

THE USE OF MATHEMATICS AND COMPUTER SCIENCE IN THE ASSIGNMENT OF THE OPTIMAL INDIVIDUAL STROKE RATE PARAMETERS IN ELITE SWIMMERS. Didier Seyfried.……12-18

"RESTING TOUCHER": A TIME AND MOTION ANALYSIS OF ELITE LAWN BOWLS. **Aaron Silk**, Robert Weatherby and John Hammond……..19-28

MODELLING THE INTERACTION IN GAME SPORTS – RELATIVE PHASE AND MOVING CORRELATIONS. Martin Lames…….29-34

One-day cricket
IMPRESS YOUR FRIENDS AND PREDICT THE FINAL SCORE: AN ANALYSIS OF THE PSYCHIC ABILITY OF FOUR TARGET RESETTING METHODS USED IN ONE-DAY INTERNATIONAL CRICKET. **Barbara O'Riley** and Matthew Ovens……..35-43

FURTHER STEPS TOWARDS FAIRER PLAYER PERFORMANCE MEASURES IN ONE-DAY CRICKET. Tony Lewis……44-57

A MATHEMATICAL MODELLING APPROACH TO ONE-DAY CRICKET BATTING ORDERS. **Matthew Ovens** and Bruce Bukiet……58-67

Statistical Applications and Rules
NAISMITH'S RULE AND ITS VARIANTS. **John Norman** and Mike Hayes……68-72

STOCHASTIC DOMINANCE AND ANALYSIS OF ODI BATTING PERFORMANCE: THE INDIAN CRICKET TEAM, 1989-2005. Uday Damodaran……..73-80

STATISTICAL ANALYSIS OF NOTATIONAL AFL DATA USING CONTINUOUS TIME MARKOV CHAINS. **Denny Meyer**, Don Forbes and Stephen Clarke……..81-90

CONTENTS (DAY TWO)

KEYNOTE ADDRESS: ARTIFICIAL INTELLIGENCE IN SPORTS BIOMECHANICS: NEW DAWN OR FALSE HOPE? Roger Bartlett……..91-98

Computer Applications
THE IMPACT OF INTRODUCING REAL-TIME COMPUTER VISION TECHNOLOGY INTO AN ELITE SPORT TRAINING ENVIRONMENT. Ros Kerr……..99-108

BIOMECHANICALLY OPTIMISED REAL-TIME SYNCHRONISATION OF SKILLS IN NOVICE AND EXPERT VIDEO SEQUENCE PAIRS FOR AUTOMATIC FEEDBACK OF POSE SHAPE AND PHASE ERRORS. Kwangwu Lee and **Richard Green**……...109-116

Using Mathematics and Statistics in Sport
PUBLICISING MATHEMATICS AND STATISTICS THROUGH SPORT. Stephen Clarke……..117-121

USES AND MISUSES OF STATISTICS IN TEAM SPORTS: GOALS IN THE BANK? George Serrallach……...122-128

REDUCING THE LIKELIHOOD OF LONG TENNIS MATCHES. **Tristan Barnett**, Alan Brown and Graham Pollard…….129-140

Predictions
A COMPARISON OF COMPUTER AND HUMAN PREDICTIONS OF AFL. **Stephen Clarke** and Rod Clarke……...141-149

REAL-TIME PREDICTIVE 3D TRACKING OF CRICKET BALLS FOR KNOWLEDGE OF RESULTS AND STRATEGY TRAINING. Hongzhi Gao and **Richard Green**……...150-159

PREDICTING THE MATCH OUTCOME IN ONE DAY INTERNATIONAL CRICKET MATCHES, WHILE THE GAME IS IN PROGRESS. **Michael Bailey** and Stephen Clarke……...160-169

Tennis
COUNTBACK METHODS AND AUXILIARY SCORING SYSTEMS IN TENNIS. Graham Pollard and **Ken Noble**…….170-177

DEVELOPING A TENNIS MODEL THAT REFLECTS OUTCOMES OF TENNIS MATCHES. Tristan Barnett, **Alan Brown** and Stephen Clarke…….178-188

FAIRER SERVICE EXCHANGE MECHANISMS FOR TENNIS WHEN SOME PSYCHOLOGICAL FACTORS EXIST. Graham Pollard and **Tristan Barnett**…...189-198

CONTENTS (DAY THREE)

Australian Rules Football
AFL FOOTBALL – HOW MUCH IS SKILL AND HOW MUCH IS CHANCE? **Don Forbes**, Stephen Clarke, Denny Meyer, …….199-205

A PROBABILITY BASED APPROACH FOR THE ALLOCATION OF PLAYER DRAFT SELECTIONS IN AUSTRALIAN RULES FOOTBALL. **Anthony Bedford** and Adrian Schembri……..206-215

A BROWNIAN MOTION MODEL FOR THE PROGRESS OF AUSTRALIAN RULES FOOTBALL SCORES. Sam Glasson…….…….216-225

POSSESSION VERSUS POSITION: STRATEGIC EVALUATION IN AFL. Darren O'Shaughnessy…….……...226-238

Methods and Applications of Analysis
AN ANALYSIS OF TEN YEARS OF THE FOUR GRAND SLAM MEN'S SINGLES DATA FOR LACK OF INDEPENDENCE OF SET OUTCOMES. Graham Pollard, Rod Cross and **Denny Meyer**…….239-246

MAXIMISING HEIGHT, DISTANCE OR ROTATION FROM REAL-TIME ANALYSIS VISUALISATION OF TAKE-OFF ANGLES AND SPEED. Richard Green……...247-255

AN ANALYSIS OF GOAL-KICKING ACCURACY IN AUSTRALIAN RULES FOOTBALL. Anthony Bedford and **Adrian Schembri** …….……...256-265

GOLF PUTTING MODELS. Neville de Mestre…..……...266-272

THE APPLICATION OF AN EXPLORATORY FACTOR ANALYSIS TO INVESTIGATE THE INTER-RELATIONSHIPS AMONGST JOINT MOVEMENT DURING PERFORMANCE OF A FOOTBALL SKILL. Tina Smith, **Wendy Gilleard**, John Hammond and Lyndon Brooks……...……273-282

LOW COMPRESSION TENNIS BALLS AND SKILL DEVELOPMENT. **John Hammond** and Tina Smith.…….……..……..283-292

GREATER THAN THE SUM OF OUR PARTS?

Lyons, K.

Department of Performance Analysis, Australian Institute of Sport, Canberra

SUMMARY[1]

In this paper I explore a vision for the role that mathematics and computers will play in applied sports contexts. I propose to use examples from the Australian Institute of Sport to do this.

Work undertaken by members of the Mathematics and Computers in Sport community has, does and will make an important contribution to elite sport in Australia and New Zealand. I argue that there has never been a better opportunity for us to be greater than the sum of our parts.

The paper has five sections:

1. Introduction
2. The published papers of the Australasian Conferences of Mathematics and Computers in Sport.
3. Developments at the AIS
4. 'Interestingness'
5. Conclusion

[1] Copies of Keith's full paper will be available at the conference

PREDICTION VERSUS REALITY: THE USE OF MATHEMATICAL MODELS TO PREDICT ELITE PERFORMANCE IN SWIMMING AND ATHLETICS AT THE OLYMPIC GAMES

Heazlewood, T.

School of Exercise Science, Australian Catholic University, Strathfield, NSW, Australia

ABSTRACT

A number of studies have attempted to predict future Olympic performances in athletics and swimming based on trends displayed in previous Olympic Games. Some have utilised linear models to plot and predict change, whereas others have utilised multiple curve estimation methods based on inverse, sigmoidal, quadratic, cubic, compound, logistic, growth and exponential functions. The non linear models displayed closer fits to the actual data and were used to predict performance changes 10's, 100's and 1000's of years into the future. Some models predicted that in some events male and female times and distances would crossover and females would eventually display superior performance to males. Predictions using mathematical models based on pre-1996 athletics and pre-1998 swimming performances were evaluated based on how closely they predicted sprints and jumps, and freestyle swimming performances for both male and females at the 2000 and 2004 Olympic Games. The analyses revealed predictions were closer for the shorter swimming events where men's 50m and women's 50m and 100m actual times were almost identical to predicted times. For both men and women, as the swim distances increased the accuracy of the predictive model decreased, where predicted times were 4.5-7% faster than actual times achieved. The real trends in some events currently displaying performance declines were not foreseen by the mathematical models, which predicted consistent improvements across all athletic and swimming events selected for in this study.

KEY WORDS

swimming, athletics, olympic games, mathematical functions, extrapolation

INTRODUCTION

The prediction of future athletic performance by humans is a recurring theme during the Olympiad year, as well as forming the basis for some stimulating 'crystal ball gazing' in some of the learned sports science journals and in the mass media. Mathematics and science are based on the principles of description and more importantly prediction. The ability to make substantive and accurate predictions of future elite level sports performance indicates that such approaches reflect "good" science. Often these predictions are purely speculative and are not based upon any substantial evidence, rather they are based on the belief that records are made to be broken and that performances must continue to improve over time. The accessibility of data in the form of results from Olympic Games, world records and world best performances in a specific year allows the analysis of performances in any number of events. From these

analyses, changes in performance over time can be observed and predictions of future performance can be made utilising the process of mathematical extrapolation.

A number of researchers have attempted to predict future performances by deriving and applying a number of mathematical statistical models based on past performances in athletics. Prendergast (1990) applied the average speeds of world record times to determine a mathematical model for world records. The records or data used in the analysis spanned a 10 year period. Following his analysis, Prendergast (1990) raised the question of whether any further improvements can be expected or if the limits of human performance have been reached. The sports of athletics (Heazlewood and Lackey, 1996) and swimming (Lackey and Heazlewood, 1998) have been addressed in this manner and the knowledge of future levels of sporting performance has been identified by Banister and Calvert (1980) as beneficial in the areas of talent identification, both long and short term goal setting, and training program development. In addition, expected levels of future performance are often used in the selection of national representative teams where performance criteria are explicitly stated in terms of times and distances (Athletics Australia, 2004).

Some researchers such as Péronnet and Thibault (1989) postulate that some performances such as human male 100m sprinting is limited to the low 9 seconds, whereas Seiler (referred to by Hopkins, 2000) envisages no limits on improvements based on data reflecting progression of records over the last 50 years. According to Seiler improvements per decade have been approximately 1% for sprinting, 1.5% for distance running, 2-3% for jumping, 5% for pole vault, 5% for swimming and 10% for skiing for male athletes, whereas female sprint times may have already peaked. The differences for males and females it is thought to reflect the impact of successful drugs in sport testing on females.

The predictions of Heazlewood and Lackey (1996) paradoxically predicted the men's 100m to improve to zero by year 5038 and the women's 100m to reach zero by year 2429, which indicates a more rapid improvement over time for women sprinters. In their model (Heazlewood and Lackey, 1996), the women's times would be faster than men by 2060 where it was predicted the finalist at the Olympic Games would average 9.58s for men and 9.57 for women respectively. A similar crossover effect, where predicted female performances would exceed male performances, was noted for the 400m and high jump. The crossover effect was based on trends in athletic performances obtained prior to 1996; where in some events female improvements were more rapid than males.

In the sport of swimming (Lackey and Heazlewood, 1998), a similar crossover effect was observed for the 50m freestyle where predicted zero time was the year 2994 for men and 2700 for women. The concept that athletes will complete 100m sprints on land and 50m sprints in water in zero seconds appears unrealistic, however mathematical model based on actual data do derive these interesting predictions.

The curves that fit the data have also displayed interesting findings as no one curve fits all the data sets. Different events displayed different curves or mathematical functions (Lackey and Heazlewood, 1998) of best fit. In swimming the men's 50m freestyle was inverse, 100m freestyle compound, 200m sigmoidal, and the 400m and 1500m freestyle

cubic. For the women's freestyle events the 50m was inverse, 100m cubic, 200m sigmoidal, 400m cubic and 800m sigmoidal.

In athletics for the men's events the mathematical functions (Heazlewood and Lackey, 1996) were 100m inverse, 400m sigmoidal, long jump cubic and the high jump displayed four functions (compound, logistic, exponential and growth). In the women's events the mathematical functions were 100m cubic, 400m sigmoidal, long jump inverse and high jump displayed four functions (compound, logistic, exponential and growth). This may indicate that different events are dependent upon different factors that are being trained differently or factors underpinning performance evolving in slightly different ways. This has resulted in different curves or mathematical functions that reflect these improvements in training or phylogenetic changes over time.

However, at some point in time how accurately the predictive models reflect reality can be assessed. Since the models of Heazlewood and Lackey (1996) for athletics and Lackey and Heazlewood (1998) for swimming were derived, the 2000 and 2004 Olympic Games have occurred. Hindsight or real data can now enable the assessment of these models over a short timeframe, that is, 8-10 years. Assessing the accuracy of the models predicting performances hundreds or thousands of years into the future will be based on the research interests of future mathematicians, sports scientist and computer scientists.

The current research problem is how well the actual times and distances achieved by athletes at the 2000 and 2004 Olympic Games fit the predicted model for athletics and swimming based on the Heazlewood and Lackey (1996) and Lackey and Heazlewood (1998) prediction equations?

METHODS

The previous models were based on following model fit criteria used by Heazlewood and Lackey (1996) for athletics and Lackey and Heazlewood (1998) for swimming. The average time and distances for the finalist in each event were utilised to generate the data for the statistical analysis for curve estimation. Potentially both linear and non linear functions can be derived. The mean score of the actual performances from the 2000 and 2004 Olympic Games finalists were compared with the predicted values in the athletic events selected in this study (Wikipedia, 2006). Times for the 100m and 400m were in seconds and distances for the long and high jump were in metres.

The results for the finalists in current Olympic freestyle swimming events (50m, 100m, 200m, 400m, 800m women and 1500m men) were collected from internet based results (Wikipedia, 2006). Times were recorded to one hundredth of a second which is the recording method used by Federation Internationale de Natation Amateur or FINA (1997). These times were then converted from a minutes and seconds format to a seconds only format to facilitate calculations when applying the regression methods. The mean of the finalists in each event for each year in the study was then calculated. The mean was used as it is a measure that is representative of all scores in each group (Rothstein, 1986). The use of the mean of the finalists in this study may be more representative of the changes in human performance that world records as used by Jokl and Jokl (1976a, 1976b, 1977) and Edwards and Hopkins (1979). A world record

holder's performance may be far in advance of that of any other competitor and not be representative of overall performance in an event. For example, the women's 400m freestyle world record as set by Tracey Wickham in 1978 was not bettered until 1988 at the Seoul Olympic Games (Wallechinsky, 1996).

In the swimming pool the factor of wind resistance is not considered significant and as such wind readings are not required for swim records. In athletics assistive and resistive winds are thought to influence performance in events such as the 100m and long jump and the wind variable can be corrected to assess performance in still air conditions. The wind correction calculations are not presented in this paper just the times and distances reported for the athletic events, however correcting for the influence of wind may result in slightly different values for the original data.

The means were then included as a data set for each event for each Olympic Games for analysis using the Statistical Package for the Social Sciences (SPSS) program version 6.1 (Norušis, 1993 & 1994) to derive a number of possible regression equations. A number of criteria were used to evaluate the goodness of fit of each derived function for each individual event.

General Method of Determining the Appropriate Regressions Models

To investigate the hypotheses of model fit and prediction, the eleven regression models were individually applied to each of the athletic and swimming events. The regression equation that produced the best fit for each event, that is, produced the highest coefficient of determination (abbreviated as R^2), was then determined from these eleven equations. The specific criteria to select the regression equation of best were the magnitude of R^2, the significance of the analysis of variance alpha or p-value and the residuals.

The Coefficient of Determination

The coefficient of determination (R^2) is a measure of accuracy of the model used. A coefficient of determination of 1.00 indicates a perfectly fitting model where the predicted values match the actual values for each independent variable (Norušis, 1993 & 1994). Where more than one model was able to be selected due to an equal R^2, the simplest model was used under the principle of parsimony, that is, the avoidance of waste and following the simplest explanatory model.

Residuals

The residuals are the difference between the actual value and the predicted value for each case, using the regression equation (Norušis, 1993 & 1994) and the smaller the residual, the better the fit of the model. For each model the residuals were generated by the SPSS program. A large number of positive residuals indicate that the prediction is an over estimation (faster than the actual performance) and a large number of negative residuals indicates an underestimation (slower time than the actual performance).

Level of Significance

The level of significance, or p value, is a representation of the relationship between the model and the data. The smaller the p value, the higher the level of significance and the

greater the relationship where a small p value indicates a small possibility that the closeness of the predicted values to the actual values due to chance is small.

Logical Acceptance Based on Extrapolations

The ability of the model to generate extrapolations that appear to be reasonable when compared to previous means was also taken into consideration. When a model generated extrapolations that appear to be inconsistent with the actual results this model was discarded and the model with the next highest coefficient of determination was selected.

Applying the Model of Best Fit

After selection of the model to be used, according to the criteria previously stated, the equation of best fit was determined by applying the derived constants and coefficients to the generic formula for that model. Using this equation, a prediction of the mean result for the event at each Olympiad was calculated. At this stage, graphs representing the means of past and future performances for each event in each Olympiad were also generated in addition to predicted means using the appropriate regression equation.

Final Predictions for the Year 2000 and 2004

To predict the level of performance in the year 2000 and 2004, the data set that provided the greatest accuracy was chosen and the data from 1996 re-included in the data set, where appropriate. A series of regressions were made using the best fitting model and data set for each event. Using the constants and coefficients generated by regression models the future predictions were then calculated.

It is important to note that in some events the average for a complete field of competitors was not always possible due to disqualifications or injury. In the case of injury a competitor did not finish the event. This situation only occurred in a few events. At this point in time no attempt was made to re-evaluate the 1996 and 1998 models based on inclusion of 2000 and 2004 data.

RESULTS

The data for the predicted values and the actual values are provided in the table for each event. Table 1 indicates the events, mathematical functions, equations and R^2 values derived from the Heazlewood and Lackey (1996) for athletic events of the men's and women's 100m, 400m, long jump and high jump.

The trends in the mathematical functions indicate the men's and women's 400m and the high jump show identical trends in changes in performance over time, where the 400m was sigmoidal and the high jump was compound, logistic, exponential and growth. In the majority of events the explained variance or R^2 values were statistically very significant ($p<0.01$). It is interesting to note the R^2 values were highest for the men's and women's high jump (0.94) and lowest for the men's 100m (0.66) and men's long jump (0.78).

Table 1. Athletic events, type of mathematical function, equations and r^2 values.

Event	Regression Type	Equation *	R^{2**}
Men's 100m	Inverse	$Y = b_0 + (b_1/t)$	0.659
Women's 100m	Cubic	$Y = b_0 + b_3t^3$	0.902
Men's 400m	S	$Y = e^{(bo + b1/t)}$	0.907
Women's 400m	S	$Y = e^{(bo + b1/t)}$	0.843
Men's Long Jump	Cubic	$Y = b_0 + b_3t^3$	0.777
Women's Long Jump	Inverse	$Y = b_0 + (b_1/t)$	0.894
Men's High Jump	Compound Logistic Exponential Growth	$Y = b_0(b_1)^t$ $Y = b_0e^{b1t}$ $Y = e^{bob1t}$	0.944 0.944 0.944 0.944
Women's High Jump	Compound Logistic Exponential Growth	$Y = b_0(b_1)^t$ $Y = b_0e^{b1t}$ $Y = e^{bob1t}$	0.947 0.947 0.947 0.947

** All R^2 values in table 1 significant at $p<0.05$.

* Where b_0 = a constant
b_1, b_3 = regression coefficients
t = year
y = mean result for each event

Table 2 indicates the predicted performances and the actual performances achieved from the 2000 and 2004 Olympics Games. It can be observed that both 100m times for men exceeded the prediction, whereas the female 100m times were well below the predicted values. For the 400m, long jump and high jump both male and female athletes were below the predicted time and distances. However, the men's actual 400m times and long jump distances did show improvements from 2000 to 2004. In the women's 400m there was a performance decline and in the women's and men's high jump performances remained relatively static from 2000 to 2004.

Table 2. The predicted and actual performances for men's and women's 100m, 400m, long jump and high jump for 2000 and 20004 Olympics.

Year Predict	100M	100W	400M	400W	MLJ	WLJ	MHJ	WHJ
2000	10.07s	10.84s	43.84s	48.29s	8.36m	7.27m	2.42m	2.05m
2004	10.03s	10.76s	43.63s	47.78s	8.43m	7.39m	2.45m	2.09m
Year Actual	100M	100W	400M	400W	MLJ	WLJ	MHJ	WHJ
2000	10.05s	11.15s	44.92s	49.92s	8.26m	6.80m	2.32m	1.98m
2004	9.93s	11.04s	44.67s	50.00s	8.33m	6.92m	2.31m	1.98m

The mathematical models derived for swimming from pre-1998 data (Lackey and Heazlewood, 1998) for the men's and women's 50m, 100m, 200m, 400m, women's 800m and men's 1500m freestyle events are displayed in table 3. The functions for the 50m (inverse), 200m (sigmoidal) and 400m (cubic) are the same for both men and women, however the functions for the 100m (men compound and women cubic) and the longer distances displayed their own specific function (800m women sigmoidal and men 1500m cubic). The non significance for the men's and women's 50m freestyle equations and R^2 values is a result of the small degrees of freedom when calculating the level of significance, due to the 50m freestyle only being included in the Olympic swimming program from 1988.

Table 3. Men's and women's freestyle swimming events, type of mathematical function and predictive equations.

Gender	**Event**	**Type**	**Equation**	**R^2 value**
Men	50m	Inverse	$y = -44.77 + 134064.199 / year$	0.543 ns
	100m	Compound	$y = 42747.22 \times .99^{year}$	0.972 sig
	200m	Sigmoidal	$y = e^{-1.33+11971.98/year}$	0.862 sig
	400m	Cubic	$y = 12572.61 - 8.79 \times year + 6.53 \times 10^{-7} \times year^3$	0.977 sig
	1500m	Cubic	$y = 40457.13 - 27.31 \times year + 1.88 \times 10^{-6} \times year^3$	0.963 sig
Women	50m	Inverse	$y = -71.68 + 193578 / year$	0.732 ns
	100m	Cubic	$y = 7417.11 - 5.48 \times year + 4.5 \times 10^{-7} \times year^3$	0.959 sig
	200m	Sigmoidal	$y = e^{-0.54+8451.02/year}$	0.812 sig
	400m	Cubic	$y = 1565 - 1.68 \times 10^{-7} \times year^3$	0.976 sig
	800m	Sigmoidal	$y = e^{-1.26+14938.8/year}$	0.692 sig

ns due to the small degrees of freedom
sig. at $p<0.05$

The comparison of the predicted times with the actual times for each event and displayed in table 4, indicates congruence for the men's 50m and women's 50m and 100m. In all other events for both men and women the predicted times are faster than actual times, indicating rate of progress in these events appears to have slowed down based on data up to 1996. This indicates the prediction equations over estimated the rates of improvement.

The predictions were closer for the shorter swimming events where men's 50m and women's 50m and 100m, where actual times are almost identical to predicted times. In both men and women, as the swim distances increased, the accuracy of the predictive model decreased, where predicted times were 4.5-7% faster than actual times achieved. For example, the predicted men's 1500m of 489.03s for 2004 was 7% faster than the actual time of 509.06s.

Table 4. Actual 1996 means, predicted and actual means for 2000 and 2004 Olympics. All times are in seconds.

Event		Actual 1996	Predicted/**Actual** 2000	Predicted/**Actual** 2004
Men's	50m Freestyle	22.47	22.26/**22.19**	22.13/**22.11**
	100m Freestyle	49.31	48.17/**48.95**	47.53/**48.80**
	200m Freestyle	108.40	105.57/**107.43**	104.32/**106.47**
	400m Freestyle	230.57	219.03/**226.21**	215.29/**225.92**
	1500m Freestyle	910.29	854.72/**901.67**	835.76/**898.05**
Women's	50m Freestyle	25.37	25.11/**25.01**	24.91/**24.96**
	100m Freestyle	55.37	54.81/**54.76**	54.53/**54.58**
	200m Freestyle	119.95	117.90/**118.89**	116.91/**118.69**
	400m Freestyle	250.00	238.98/**249.46**	233.91/**248.05**
	800m Freestyle	514.89	496.38/**507.82**	489.03/**509.06**

DISCUSSION

The results indicate the ability to predict performances into the near future based on past performances is possible, however both sets of results derived from athletics and swimming indicate that in many events the improvements in performances were overestimated. Although improvement in the majority of athletic and swimming events studied did occur in absolute times or distances, the rate of improvement was slower than predicted by the numerous equations generated. The reality is, in some events performances have been static such as men's and women's high jump and women's 200m freestyle or actually declined such as women's 400m and women's 800m freestyle. The two events where actual performances exceeded, very slightly the predicted performances, were the men's 100m and men's 50m freestyle.

Recent emphasis on successful drug testing may have impacted more on women athletes than men, where the slowing down, and in some cases the declines in actual performances were noted for women.

Performances are expected to improve over a period of time due to a number of interacting sports scientific, ontogenetic (lifespan) and pharmacological factors, such as:

1. The use of more efficient running, jumping and swimming techniques, a biomechanical construct.
2. Improved training programs, which are exercise physiological and functional anatomical construct.

3. Enlarged population of athletes due to increased participation by more nations from which high performance athletes and swimmers are drawn. This will result in an increased sample from the human gene pool, a genetic construct.
4. Improved talent identification programs designed and implemented by national sporting organisations and sports institutes that will select and develop tomorrow's high performance athletes.
5. Changes in human physiology, such as the recent ontogenetic trends of increasing height and weight in Australia.
6. The use of performance enhancing drugs legal or illegal, especially androgenic-anabolic steroids and human growth hormone, which have a masculinising effect on women or the use of neutraceuticals (functional foods).

However, the exact mathematical trends these improvements and interactions take can be plotted mathematically to reveal future trends across athletic and swimming events. In effect, the actual performance is a summary of all these factors. In some cases the event can be predicted with a high degree of accuracy, such as the sprint swimming events for both men and women, whereas in other events such as 400m women and 800m freestyle are not predicted well as the events are currently displaying performance declines which were not identified by the mathematical models. It must be emphasised that all the models across all events indicated consistent improvements over time.

The primary purposes behind this type of predictive research are that we might understand the realistic limits to human improvement in many sports, to set new and realistic goals that athletes will have to achieve to make representative teams and Olympic finals, to provide a more coherent understanding of what performances of the past suggest about performances of the future, to understand if different events represent changes which reflect developments of human biomechanical, exercise physiological, motor learning and sport psychological functions as expressed in sport; and as a intellectual exercise to understand more completely the complex trends that underpin human evolution and training adaptation that are expressed in the sports arena.

CONCLUSION

As a heuristic exercise the derivation of mathematical-statistical models that predict changes in human sporting performance both in the near and distant future occupies definitely the minds of mathematicians and statisticians and if "we get it right" we will have a crystal ball into the future of sport. The problem is it just takes time to find out how good we are at solving such problems.

REFERENCES

Athletics Australia (2004) *Nomination criteria for 2004 Olympic Games Athens, Greece – August 2004 Athletics Australia*. Downloaded February-March, 2006. *http://www.athletics.org.au/events/selection/dp_details.cfm?ObjectID=108.*

Banister, E. and Calvert, T. (1980) Planning for future performance: Implications for long term training. *Canadian Journal of Applied Sports Science* **5**, 170-176.

Edwards, D. and Hopkins, W. (1979) Targets for the future. *Modern Athlete and Coach* **17**(3), 34-36.
Federation Internationale de Natation Amateur. (1997) Welcome to FINA. Downloaded February-March, 1998. *http://www.fina.org/.*

Heazlewood, I. and Lackey, G. (1996) The use of mathematical models to predict elite athletic performance at the Olympic Games. In N. deMestre (Ed). *Third Conference on Mathematics and Computers in Sport*. 30 September - 2 October, Bond University, Queensland. Conference Proceedings. 185-206.

Hopkins, W. (2000) Limits to performance. *SportScience.* Downloaded February-March, 2006. *http://www.sportsci.org/jour/0002/inbrief.html#limits*

Jokl, E., & Jokl, P. (1976a) Running and swimming world records. *Olympic Review* **107/108**, 536-543.

Jokl, E. and Jokl, P. (1976b) Running and swimming world records. *British Journal of Sports Medicine* **10**, 203-208.

Jokl, E. and Jokl, P. (1977) Running and swimming world records. *Journal of Sports Medicine and Physical Fitness* **17**, 213 - 229.

Lackey, G. and Heazlewood, I. (1998) The use of mathematical models to predict elite swimming performance. In N. deMestre and K. Kumar (Eds). *Fourth Conference on Mathematics and Computers in Sport.* 9-12 July, Bond University, Queensland. Conference Proceedings. 79-110.

Norušis, M. (1993) *SPSS for Windows: Base Systems User's Guide. Release 6.0.* SPSS Inc, Chicago, USA.

Péronnet, F. and G. Thibault, G. (1989) Mathematical analysis of running performance and world running records. *Journal of Applied Physiology* **67**, 453-465.

Prendergast, K. (1990) What do world running records tell us? *Modern Athlete and Coach* **28** (1) 33 -36.

Rothstein, A. (1985) *Research Design and Statistics for Physical Education*. Prentice Hall, Englewood Cliffs, USA.

SPSS Inc. (1994) *SPSS for Windows Release 6.1-Computer Software*. SPSS Inc, Sydney, Australia.

Wallechinsky, D. (1996) *The Complete Book of the Olympics 1996 edition*. Aurum Press, London, UK.

Wikipedia. (2006) Results of the 2000 and 2004 Olympic Games. Downloaded February-March, 2006. *http://en.wikipedia.org/*

THE USE OF MATHEMATICS AND COMPUTER SCIENCE IN THE ASSIGNMENT OF THE OPTIMAL INDIVIDUAL STROKE RATE PARAMETERS IN ELITE SWIMMERS

Seyfried, D.

Laboratoire d'Informatique Appliquée au Sport, INSEP, Paris, France

ABSTRACT

Looking to measure the spontaneous frequency of swimming to take advantage of it in an applied way for better learning (for the swimmer) and better teaching (for the coach) has been a method used since 1991 at the National Institute of Sport, Paris (INSEP) since it has appeared to be most effective in terms of swimming economy. A good number of mathematical relationships found between the speeds obtained on clean swimming but equally on non-swim areas (starts, turns, underwater part) can be a source of potential help to the individual swimmers if they can be calculated. Among them the stroke rate/velocity relationship is very interesting since its mastering has a lot of consequences on energy cost, delay of fatigue and better learning of the best stroke at any moment in training or competition. This starting point guided the development of several types of software beneficial to swimming training including NATAVIT (1995-2000) and AQUACYCLE (since 2000) but equally a concept of measurement automated in training (CHRONOCYCLE) and finally, a video-electronic system which is used in competition by the French Swimming Federation.

KEY WORDS

modelling, swimming, cadence, velocity, stroke length, feed back

INTRODUCTION

Looking to take the spontaneous frequency of activity to take advantage of it in an applied way for better learning (for the swimmer) and better teaching (for the coach) has been an objective since 1991. In swimming, the spontaneous stroke rate/velocity was at first an indicator then became a systemized test from 1995. It is used in many countries, notably in swimming to determine if the rate and by result, the distance per stroke, also called stroke length, used by a swimmer at a given moment in his swim (in training or competition) is the most effective in terms of economy.

MATHEMATICS AS A TRAINING TOOL

Transforming the observations obtained from swimmers during previous evaluations to create more specifically their future training sessions is the goal for most physiologists. A good number of mathematical relations found between the obtained speeds on clean swimming but equally on non-swim areas (starts, turns, underwater part) can be a source of potential help to the individual swimmers if they can be calculated. Among them, the stroke rate/velocity relationship is very interesting since its mastering has many consequences on energy cost, delay of fatigue and better learning of the best stroke at any moment in training sessions or competitions. This starting point guided the development of several types of software beneficial to swimming training including

NATAVIT (1995-2000) and AQUACYCLE (since 2000) but equally a concept of measurement automated in training (CHRONOCYCLE) and also a video-electronic system is used in competition by the French Swimming Federation.

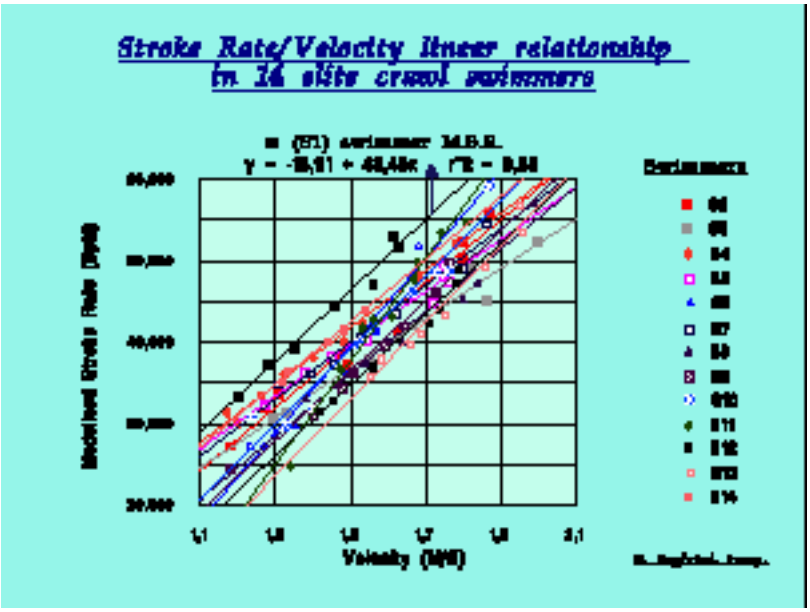


Figure 1

In France and at INSEP in particular, the choice of a kinetic study of cyclic parameters evolving within useful swim velocities was at the origin of a hypothesis to test the relationship between stroke rate and velocity. This was confirmed through the development of standards for this test and through the use of various methodologies and tools. The applied usage in training of these tests was carried out through the use of individual test results established using tables (Microsoft Excel), then on programmable machines (from 1992 to 1996), leading the French Swimming Federation (FFN) to ask INSEP to develop a specific software NATAVIT. This software allowed from data obtained in tests at the National Training Centre, the computerized mastering of the accurate individual velocities of swims, with indications for the best cyclic, physiological and technical methods.

THE SYSTEM OF EVALUATION AND FOLLOW UP IN TRAINING

The stroke rate/velocity incremental test involves observing 15 m of clean swimming during 8 progressive steps where 3 full cycles are measured considering the time and the spontaneous stroke rate of the swimmer. It allows every swimmer to be given theoretical cyclic parameters vs. observed ones during future sessions. Theoretical Stroke Length (TSL) vs. Observed Stroke Length (OSL), Theoretical Stroke Rates vs. Observed Stroke Rates (OSR) and even Theoretical Stroke index (TSI) vs. Observed Stroke index (OSI), are all good factors to take into account in terms of swimming economy during the observations of training sessions and races.

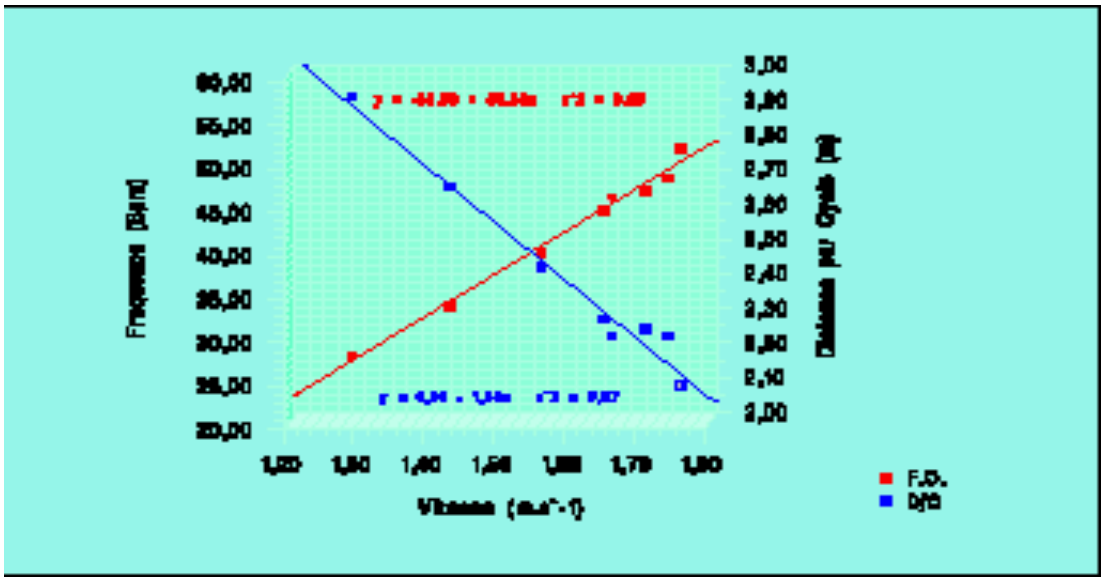


Figure 2: Individual linear relationships between velocity, stroke rate & stroke length

EXAMPLES OF APPLICATIONS IN TRAINING

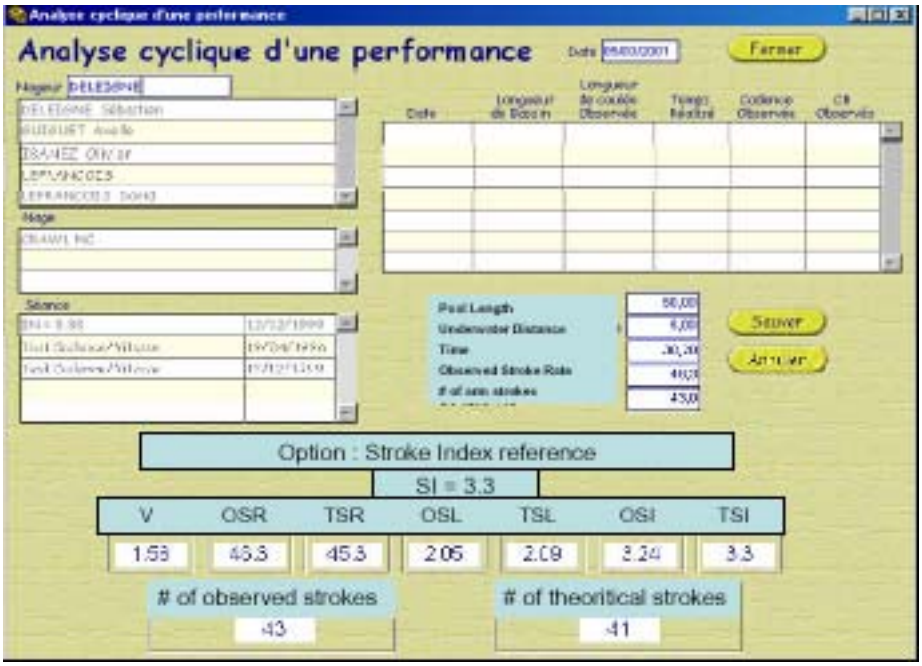


Figure 3: Observed Stroke rates, Stroke lengths and stroke indexes compared with theoretical individual model.

Different methods can help optimise the management of the cyclic fitting parameters to the velocity of the corresponding swims. The example of the relationship between stroke rate and the velocity of swims seems to be the most striking. After determining the equation rate velocity TSR (ideal individual stroke rate), this can be integrated into software or computer devices (programmable machine, software on computers or PAD) and will use the swimming economy index for the calculation of the ideal cyclic parameters for all the times, for all distances and size of pool and of given send-offs. This system provides a cyclic and technical example with the swimmer immediately

after each series. Recent developments (AQUACYCLE) allow as well the possibility for the swimmer to self-evaluate simply in terms of «arm hits» : checking the clock and counting the arm stokes is enough for better learning. These methods allow the swimmer to refine his or her mastery of spatial-mental skills, a critical step in the control of the cyclic parameters relating to velocity of swims. These situations where objective information is given during or immediately after the end of the exercise are essential for the education of swimmers. This is particularly useful in long distance swimming where the instruction or feedback during the exercise contributes to the learning curve necessary to maintain the stability of the energy cost.

Constant pace is a strategy that has been proven in terms of swimming economy and most notably in records surpassed; nevertheless, there are often other strategies that prove themselves essential to winning at championship meetings where the individual athlete is confronted with the primary objective of being on the podium. A faster start or accelerating at the right moment must have also been worked on in training; but to start faster does not necessarily mean a 'breakthrough' in attempting to go for broke, i.e. taking a strategic risk, using specific techniques, in a way adapted to the immediate needs of the swimmer. The mastery of the relationship between stroke rate and velocity allows working towards swimming at a constant speed and to a stable rhythm, to adapt with better effectiveness to all likely velocities used by the swimmer. The acquisition of biomechanical indicators by trial and error across the spectrum of velocities allowing the movement in water, with « feedback » in the fastest possible manner on execution will probably be a learning curve for the elite swimmer in the years to come.



Figure4: Real Time display of Velocity, Stroke Rate, and Stroke Length as well as the Stroke Rate/Velocity Relation.

To learn how to manage precisely the cyclic parameters imposes on the coach an observation zone of 15 m., measured with precision, where three complete cycles of swims can be studied with a stopwatch with stroke rate function. To be pertinent, the

calculations of distance per stroke and of the swimming index must be carried out from the speed of swims measured in this zone. To facilitate these measurements, computer software was recently developed to immediately determine the relationship between stroke rate and velocity of the swimmer in practice. Software has been produced to record from a direct capture or of a video file that one can see in one of the windows of the application. The capture was tested with a click of the mouse, keyboard, and an external remote control to the PC radio HF but it can receive all types of triggers to the system (sensors, automatic recognition of pictures, light signals, sound cues).

DISCUSSION

The linear adjustments of the relationship of stroke rate/velocity and stroke length/velocity are more highly correlated as the level of expertise of the swimmers is increased. It is normal in physics because linked to drag that the stroke length diminishes in a manner inversely proportional to velocity according to a linear mathematical relationship. The swimmers seem insufficiently educated in this approach. The observation of swimmers in competition indicates unnecessarily high stroke rates in comparison with their individual demonstrated ideal economy of swims for the given velocity. The confusion between 'velocity and turn over' is all the more natural under the stresses of competition. This is the reason why it must be checked and practiced in training and in competition. From exchanges with T.M. Absaliamov, the long time expert in charge of the studies on Soviet then Russian swimmers, we reinforced the fact that pace regularity and stroke rates of Salnikov or of Sadovy, the technical management of race of Popov, had been obtained by practice and were the fruit of very elaborate feedback. In France, the number of repetitions at race velocity, while doing the report on the cyclic parameters is more and more used with the swimmers. Work at sub-maximal speed deserves equally to be given to swimmers with a reference to cyclic objectives, for example a number of arms strokes by length according to the speed of the swimmer.

CONCLUSION

Our evaluations show that for all the swimmers improving their times at INSEP over several years, the two values of stroke rate and race pace still evolve in the direction of an increase in the distance per stroke and of a decrease in the rate. Nevertheless, only the improvement of performance in times must remain the objective. The increase of the distance per stroke for a velocity of given swims (and therefore the proportional decrease of the rate) is only interesting if the swimmer preserves the same « potential of rate » throughout his or her career to attain a race pace which always improves in competition. The analysis taken from systematic measurements done in national and international competitions has also allowed the French team to be able to take advantage of the real-time evaluations from the last Athens Olympic Games and the Montréal World Championships since the French system was kept as the official analysis of FINA. For these two major competitions, the possibility to observe race strategies and capacities of the current best world class swimmers and to put them side by side with the French hopefuls in the same tests is rich in terms of immediate analysis and forecast. The comparisons then study of the stroke rates, on the data from both swimming and non-swimming parts, calibration of the velocities throughout the race

and so on, are important in the strategic aims of a federation that emerges from here onwards as being amongst the nations able to produce champions on a world level.

REFERENCES :

Argentin, S., Hausswirth, C., Bernard, T., Bieuzen, F., Leveque, J., Couturier, A. and Lepers R. (2006) Relation between preferred and optimal cadences during two hours of cycling in triathletes. *British Journal of Sports Medicine* **40**, 293-8.

Brisswalter, J. and Hausswirth, C. (2000) Energetically optimal cadence vs. freely chosen cadence during cycling : effect of exercise duration. *International Journal of Sports Medicine* **20**, 1-5.

Cavanagh, P. and Kram, R. The effect of stride length variation on oxygen uptake during distance running. *Medicine & Science in Sports & Exercise* **14,** 30-35.

Cavanagh P.R., Kram R. (1984) Mechanical and muscular factors affecting the efficiency of human movement. *Medicine & Science in Sports & Exercise* **17**, 326-331.

Chollet, D., Pelayo, P., Delaplace, C., Tourny, C. and Sydney M. (1997) Stroking characteristics variations in 100 m freestyle in different skill level swimmers. *Perceptual and Motor Skills* **85**, 167-177.

Coast, J.R. and Cox, R.H. (1986) Optimal pedalling rate in prolonged bouts of cycle ergometry. *Medicine & Science in Sports & Exercise* **18**, 225-230.

Foss, O. and Hallen, J. (2005) Cadence and performance in elite cyclists. *European Journal of Applied Physiology* **93**, 453-62.

Hagberg J.M., Mullin J.P., Giese M.D., Spitznagel E. (1981) Effects of pedalling rate on submaximal exercise responses of competitive cyclists. *Journal of Applied Physiology,* **51**, 447-451.

Marsh, A.P. and Martin, P.E. (1997) Effect of cycling experience, aerobic power and power output on preferred and most economical cycling cadences. *Medicine & Science in Sports & Exercise* **29**, 1225-1232.

Molen, N.H., Rozendal, R.H. and Boon W. (1972) Graphic representation of the relationship between oxygen consumption and characteristics of normal gait of the human male. *Proc. Kon. ned. Akad. Wet* **C**, 305-314.

Pai, Y., Hay, J.G. and Wilson B.D. (1984) Stroking techniques of elite swimmers. *Journal of Sports Science* **2** 225-239.

Seyfried, D. (1997) Evaluer et Analyser pour mieux entraîner : exemple du système développé à l'INSEP pour l'olympiade 1992-1996. *In L'Entraînement, Analyse de l'activité* **2**, 137-154, Atlantica.

Seyfried, D. and Van Hoecke, J. (1993) Validation et intérêts du test Fréquence-Vitesse pour l'évaluation et l'entraînement du nageur de haut-niveau. *In Actes du Congrès, 5° Journées de l'ACAPS,* 339-340, Caen.

Seyfried, D., Araar, N., Vallier, J.M. and Thomaidis, M. (1994) Analyse de paramètres biomécaniques et physiologiques au cours d'un protocole aérobie réalisé en crawl chez 18 nageurs de haut-niveau. *In Actes du Congrès, 1° Journées spécialisées de Natation, Lille.*

Seyfried, D. and Trepagny J.D. (1994) Conception et Développement du logiciel d'aide à l'entrainement NATAVIT pour les nageurs du CNE de l'INSEP. *In Actes du Congrès, 1° Journées spécialisées de Natation, Lille.*

Seyfried, D. (1994) Intérêts et Méthodologie du test Fréquence/Vitesse pour l'Evaluation et l'entraînement du nageur de 4 Nages. *In Actes du Congrès, 1° Journées spécialisées de Natation, Lille.*

Sidney, M., Pelayo, P., Chollet, D. and Tourny, C. (1997) Analyse des paramètres spatio-temporels en compétition et à l'entraînement. *In L'Entraînement, Analyse de l'activité* **2**, 95-127, Atlantica.

Swaine, I. and Reilly, T. (1983) The freely-chosen Swimming Stroke-rate at a maximal Swim on a Biokinetic Swim Bench. *Medicine and Science in Sports and Exercise* **15**, 371-375.

Tourny, C., Chollet, D., Pelayo, P. (1991) Variabilité de la distance par cycle et de la fréquence au cours d'épreuves compétitives en natation. *In Actes du Congrès, 3° Journées de l'ACAPS, Lille.*

Vercruyssen, F., Brisswalter, J., Hausswirth, C., Bernard, T., Bernard, O and, Vallier, J.M. (2002) Influence of cycling cadence on subsequent running performance in triathletes. *Medicine and Science in Sports and Exercise* **34**, 530-536.

Wakayoshi, K., Yoshida, T., Ikuta, Y., Mutoh, Y. and Myashita, M. (1993) Adaptations to Six months of Aerobic Swim Training: Changes in Velocity, Stroke Rate, Stroke Length and Blood Lactate. *International Journal of Sports Medicine* **14**, 368-372.

Wakayoshi, K., D'Acquisto, L.J., Cappaert, J. and Troup, J.P. (1995) Relationship between Oxygen Uptake, Stroke Rate and swimming Velocity in Competitive Swimming. *International Journal of Sports Medicine* **16**, 19-23.

Weissland, T., Pelayo, P., Vanvelcenaher, J., Marais, G., Lavoie, J.M. and Robin H. (1997) Physiological effects of variations of spontaneously chosen crank rate in incremental upper body exercise. *European Journal of Applied Physiology* **76**, 428-433.

‘RESTING TOUCHER’:
A TIME AND MOTION ANALYSIS OF ELITE LAWN BOWLS

Silk, A.J. [1], Hammond, J. [2] & Weatherby, R.P. [1]

[1] Department of Exercise Science and Sport Management, Southern Cross University.

[2] School of Education, Southern Cross University, Lismore, Australia.

ABSTRACT

Whilst numerous investigations have explored the physical demands placed upon competitive sportspeople from a wide array of sports little is known about the physical demands placed on lawn bowlers. The purpose of this study was to ascertain the movement activities of Australian representative singles and pairs players and to determine the frequency and duration of these activities. One match each of two male and two female players (one singles and one pairs player per gender) were videotaped during an international tournament. During playback of the videotaped matches (n=4), a single observer coded the players’ activities into five distinct categories (waiting, walking forward, walking backward, jogging and bowling) using a computerised video editing system (Gamebreaker™ Digital Video Analysis System). Field calibration of players over 30m for forward motions and 15m for the backward motion was performed to allow for the estimation of total distance covered during the match. Heart rate was monitored during each match. The duration of a match was found to be (mean±SD) 1hr 28±15mins. The total distance covered during each match was 2093±276m. The mean percentage of match time spent in each motion was: waiting, 61.8±9.3%; walking forward, 22.3±5.6%; walking backward, 2.0±0.4%; jogging, 1.1±0.5%; and bowling, 8.5±4.2%. Average heart rate was found to be 57±7% of age-predicted HR_{max} with a maximum of 78±9% of age-predicted HR_{max}. The results of this study suggest that playing lawn bowls at an international level requires light-moderate intensity activity similar to that reported for golf.

KEY WORDS

frequency, mean duration, heart rate, energy expenditure.

INTRODUCTION

The modern game of Lawn bowls has its origins in England early in the 13th Century. A similar game is reported to have been enjoyed by Ancient Egyptians, the Aztecs, North American Indians, and also in China and Polynesia. Lawn bowls is predominantly a game of skill but there is also a physical component necessary for its performance. This is also true for other so called “target sports” such as archery, golf, shooting and ten-pin bowling. An analysis of golf participation indicated that over the course of an 18-hole round mean heart rate was 108 beats/min, equating to 35-41% of maximal oxygen capacity (Murase et al., 1989) although this intensity has been reported to increase with the age of the player (Brohman et al., 2004). Numerous other competitive sports have been analysed for the physical component to investigate the demands placed on participants. Team sports such as rugby union (Duthie et al., 2005), soccer (Krustrup et al., 2005) and field hockey (Spencer et al., 2004) have been analysed by time and

motion studies to elucidate the physical demands of playing the sport. Sporting officials from rugby league (Kay and Gill, 2003), soccer (D'Ottavio and Castagna, 2001) and rugby union (Martin et al., 2001) have also been the target of time motion analyses. The information obtained from previous time and motion studies has been used to tailor specific conditioning programs for participants and also to develop sport-specific field tests. No such investigation has been carried out for Lawn Bowls.

The sport of lawn bowls is played on a bowling green, a level grass covered surface with dimensions approximately 40m x 40m (Judson, 2004). A game of lawn bowls is played on a rink, a demarcated strip of a bowling green, denoted by pegs at either end, usually between 5.5 and 5.8 meters wide (Judson, 2004) allowing several games of lawn bowls to be played concurrently on the same green. Singles, pairs, triples and fours games are played in accordance with the number of players constituting a team. Specific rules are followed concerning the number of deliveries allowed each player and when play is to end and a winner emerge. Players, and teams, alternately deliver their biased bowls down the green towards a small white non-biased bowl called the jack with the aim of having as many bowls as possible closer to the jack than their opponent. One point is awarded for each bowl closer to the jack than any opposing bowls. The direction of play reverses for successive ends, each of which follows the sequence of laying a protective mat, delivering the jack and delivering the allowable number of bowls, in turn. The current format for international level competitive bowls sees matches played over 2 sets of 9 ends with 4 bowls per player for singles matches and 2 bowls per player for pairs and triples.

The work reported here analysed the movement activities of Australian representative singles and pairs players and obtained some preliminary data on the physical demands on players at this level of competition.

METHOD

Subjects: Four Australian representative lawn bowlers gave their informed consent to participate in this research. Two males aged 29yrs and 26yrs with body masses of 113kg and 69kg respectively and two females aged 44yrs and 31yrs with body masses 70kg and 67kg participated.

Stride calibration: Prior to filming a stride calibration was conducted to ascertain stride length and movement velocity during the motion categories of walking (forward and backward) and jogging. This calibration was done over a 30m distance for walking forward and jogging and 15m for walking backward. This information was used to estimate distance covered during the match.

Games analysis: The matches that were analysed were played at the Tri-Nations Cup tournament between Australia, New Zealand and Malaysia that was held in Melbourne (Darebin International Sports Centre) Australia between the 31st January and 2nd February 2006. 2 singles matches (1 men's and 1 women's) and 2 pairs matches (1 men's and 1 women's) were recorded for analysis. The skips of the pairs combinations were filmed and used for analysis. During matches the on-green motion of players was captured by a video camera (Panasonic Model No.NV-GS150GN, Matsushita Electrical Industrial Co. Ltd. Japan) with a moveable field of vision. The camera was positioned at the side of the green that was closest to the rink on which the match was being played.

Player movement was followed for the duration of the match. Footage was recorded onto miniDV tapes (Panasonic DVM80, Matsushita Electrical Ind. Co. Ltd. Japan).

Player motion was subjectively characterised by an experienced operator while watching the video playback. A software analysis program (Gamebreaker™ Digital Video Analysis System, Sportstec Pty Ltd, Australia) was used to quantify each motion category. Video footage of the matches (n=4) was viewed and concurrently characterised by activating motion category buttons when a motion started and deactivating them when that motion ceased. This allowed the duration of each individual motion to be logged. At the completion of the analysis an output was obtained that detailed each motion category. This output included, for each of the five motion categories, the motion frequency, total time of motion during the match, motion as % of total match time and mean motion duration.

Motion categories: Player motions were coded into five different categories and were defined as follows:

1. Waiting: motionless or milling around the head or behind the mat (includes such activities as picking up bowls, shuffling feet, organising bowls, filling out scorecard, pacing, waiting to bowl, watching own bowl, inspecting the head, watching opponents bowl, and getting a drink).
2. Walking forward: forward motion with both feet in contact with the ground at same time during some point in the gait cycle.
3. Walking backward: backward motion with both feet in contact with the ground at same time during some point in the gait cycle.
4. Jogging: motion with an airborne phase.
5. Bowling: motion involved in delivering a bowl.

Heart rate monitoring: Players agreed to wear a heart rate monitor during the matches that they were filmed playing. The equipment used for recording heart rate constituted wearing an elasticised transmitting strap around the chest and a watch on the wrist (Polar Model No.S610, Polar Kempo, Austria). Average heart rate was recorded for each 15 second period of match time. The raw data that was collected was converted to the percentage of age-predicted maximum heart rate (HR_{max}). Age-predicted maximum heart rate was calculated using the equation $HR_{max} = 220 - age$.

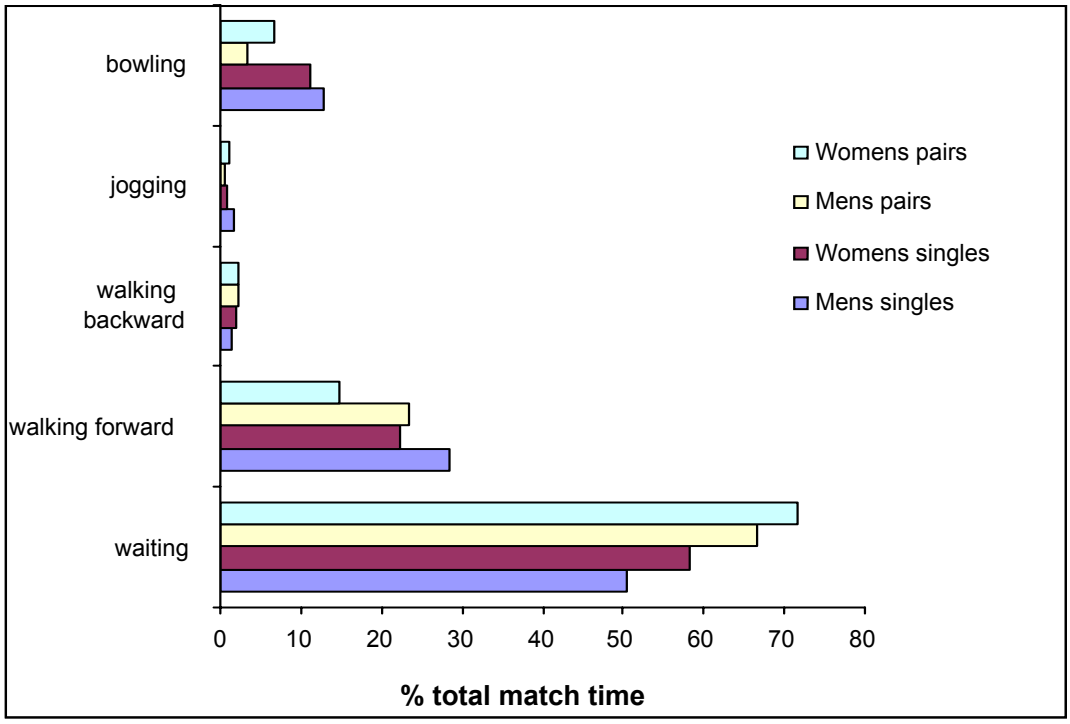
Reliability data: Player motion from the first set (9 ends) of each match (n=4) was analysed at three different times over a 10 day period to establish intra-tester reliability. The average duration of this footage was 44±2.5min. These times were Day 1 (1), Day 4 (2), and Day 10 (3). The viewing order of the four matches was randomised using a Latin Square design. A bivariate correlation (Pearson Correlation – 2 tailed) analysis was performed on the data (frequency and mean duration for all 5 motion categories). Intra-tester reliability was found to be within 95% confidence intervals.

Data analysis: The differences in motion frequency, duration and distance (if applicable) between gender and between singles and pairs were examined. All results are reported as mean±SD.

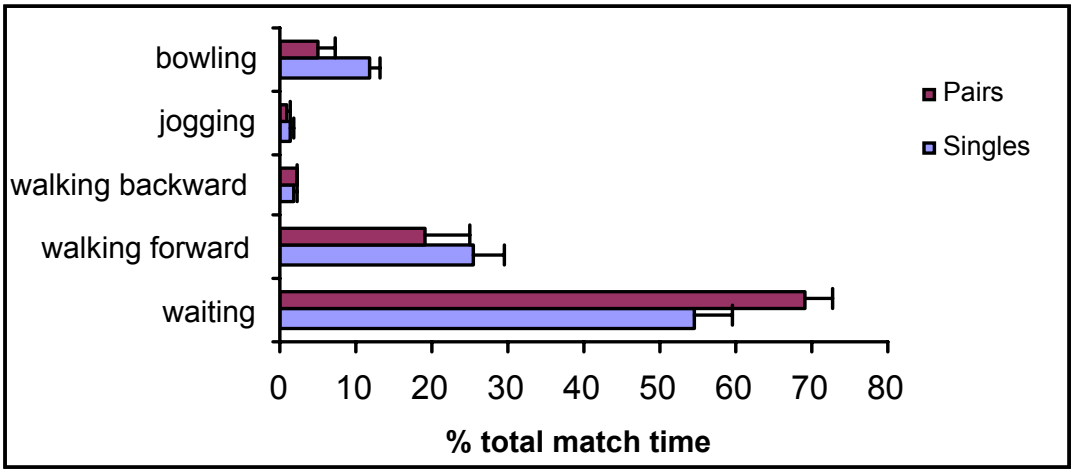
RESULTS

Match times and motion categories: Average match time was 1hr 28±15mins with 61.8±9.3% of this time spent waiting, 22.3±5.6% walking forward, 8.5±4.2% bowling,

2.0±0.4% walking backward and 1.1±0.5% jogging. Individual match times for men's singles, women's singles, men's pairs and women's pairs were 1hr 15mins, 1hr 39mins, 1hr 14mins and 1hr 42mins respectively. While the difference between average match times for gender was quite large (1hr 15mins for men vs. 1hr 41mins for women) when the average match time for singles and pairs matches were compared these were quite similar, being 1hr 27mins and 1hr 28mins respectively. The percentage of total match time within each motion category for each match (n=4) and also for combined singles (n=2) and pairs (n=2) matches is presented in Figure 1.



(a)



(b)

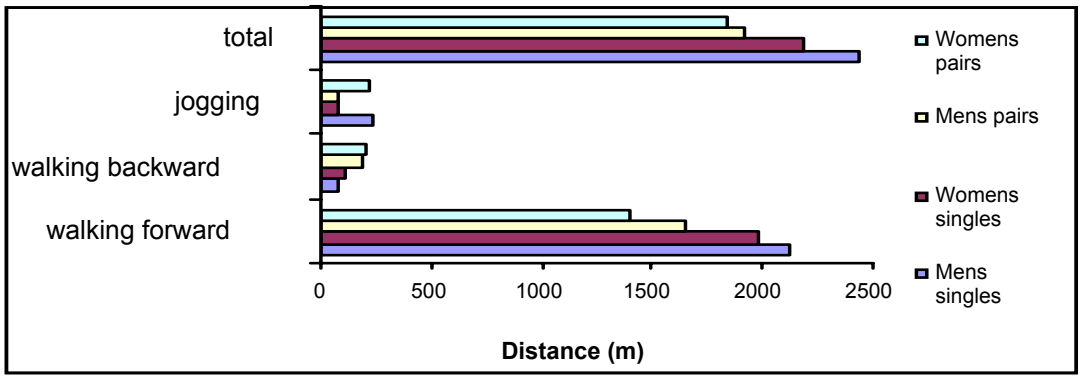
Figure 1: (a) Percentage of total match time spent within each motion category for each individual match; (b) Percentage of total match time spent within each motion category for combined singles (n=2) and pairs (n=2) matches. Displayed as mean±SD.

The frequency of motion occurrence varied greatly for some motion categories and very little for others. The frequency of waiting (143 vs. 90), walking forward (124 vs. 90) and bowling (79 vs. 33) motions were the standout differences between singles and pairs matches while walking backward (25 vs. 29) and jogging (9 vs. 10) motions were similar in number. Mean time spent in each motion was very similar when singles and pairs matches were compared with the exception of waiting (20s vs. 41s). These results are presented in Table 1.

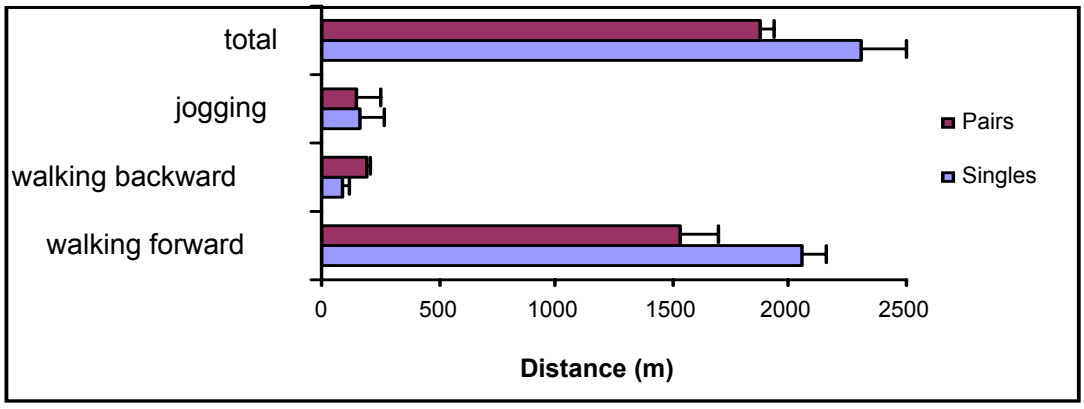
Table 1: Frequency and mean time for each motion category for all matches combined (n=4), singles (n=2) and pairs (n=2) matches. Results are presented as mean±SD.

Motion	**Frequency (n)**			**Mean duration (s)**		
	All	Singles	Pairs	All	Singles	Pairs
Waiting	117±39	143±35	90±20	30±12	20±1	41±2
Walking forward	107±20	124±6	90±8	11±1	11±0.5	11±2
Walking backward	27±8	25±6	29±12	4±1	4±0.5	5±1
Jogging	9±4	9±4	10±5	6±2	7±2	5±1
Bowling	56±28	79±16	33±4	8±3	8±1	8±5

Distance covered during the match: As seen in Figure 2, the average distance covered during a match was 2093±276m.



(a)



(b)

Figure 2: (a) Total distance and distance covered during separate motion categories for each individual match; (b) Total distance and distance covered during separate motion categories for the combined singles and pairs matches. Displayed as mean±SD.

The majority of this was in the form of forwards walking (1792±327m). Jogging contributed 7.5±4% of the total distance traveled, or 157±85m. Each jogging repetition was approximately 16m in distance. The distance covered in each motion category during each match (n=4) and also for combined singles (n=2) and pairs (n=2) is presented in Figure 2.

Heart rate during the match: The average heart rate recorded during the four matches was found to be 107±15bpm. This equated to 57±7% of age-predicted HR_{max}. The maximum heart rate that was observed during a match was 78±9% of age-predicted HR_{max} equating to 144±13bpm (highest recorded heart rate was 161bpm). Heart rate for each match (n=4) is presented in Figure 3. Figure 4 illustrates the percentage of match time spent in several percentages of age-predicted HR_{max} brackets during the four matches analysed.

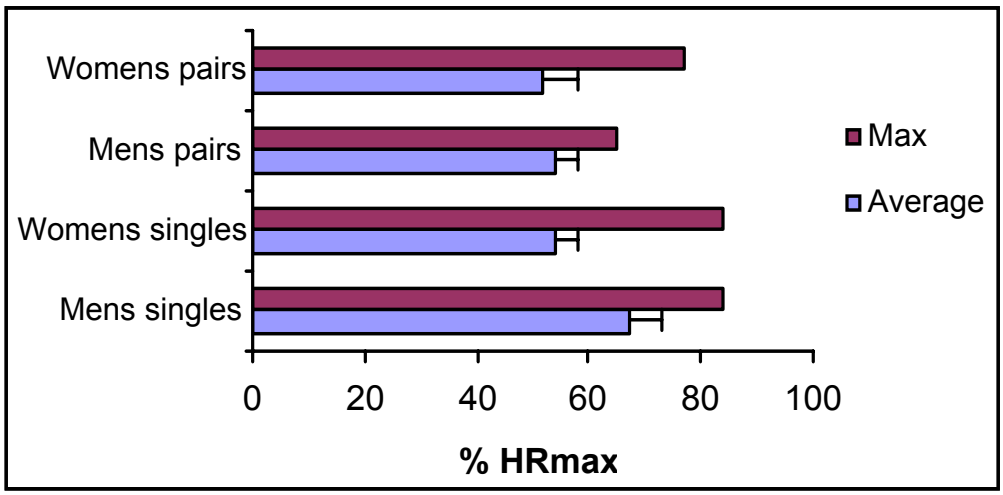


Figure 3: Average and maximum heart rates during each individual match. Average heart rates are presented as mean±SD.

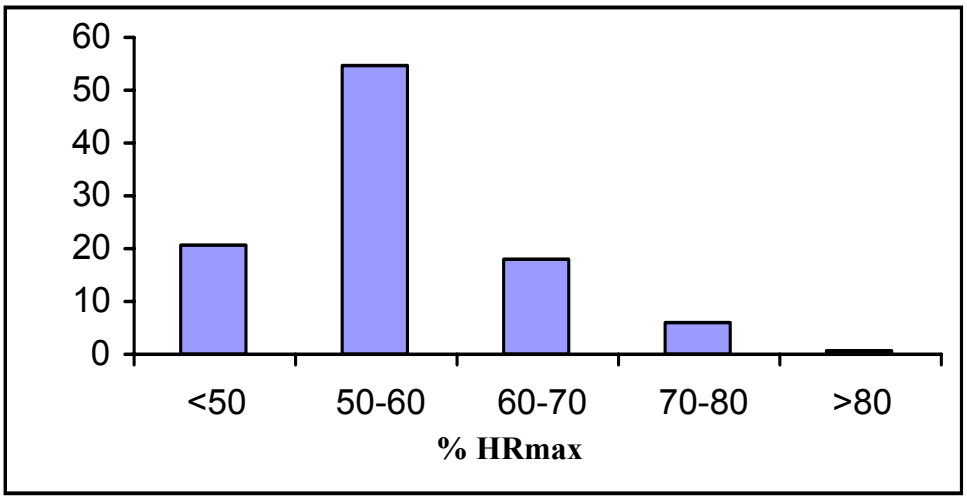


Figure 4: The average percentage of match time accounted for in several percentages of age-predicted HR_{max} brackets in the four matches analysed.

DISCUSSION

During the four matches analysed at the 2006 Tri-nations Cup Tournament in Melbourne it was found that the players spent most of their time engaged in what could be termed preparatory activities either behind the mat or at the head. In this study it was found that approximately 65% of match time was accounted for in this way. The remaining time was spent walking (≈25%), jogging (1%) and bowling (≈9%). The average duration of the matches was approximately 1½hr. On average players covered just over 2000m during the match. The majority of this distance can be attributed to walking forward (≈85%). Average heart rate during the matches was approximately 60% age-predicted HR_{max} with the maximum recorded heart rate being close to 80% age-predicted HR_{max}.

Although the total match times were very different when men's and women's matches are compared, the time taken to play the first set (9 ends) was not markedly different. Men took 43min to complete the first 9 ends while women required 45min. The explanation for this total match time difference is the nature of the games filmed. Both men's games were completed in 2 sets with the number of ends totaling 16 for singles and 15 for pairs whilst the women's matches comprised 2 sets and a tie-break for singles (total of 21 ends) and 2 sets for pairs (total of 18 ends). Therefore, men played 31 ends in a total time of 149min and women played 39 ends in 201min. This equates to 4:48min per end for men and 5:09min per end for women. To complete an 18 end match (2 sets of 9 ends each) it can be predicted that it would take women 6:18min longer than men. To put this in real terms, the men would finish while the women still had a little over an end to complete. Amongst bowling circles it often stated that women take longer to complete a match than men, tending to spend longer in discussion about tactics and inspecting the head. This closeness of predicted finishing times does not appear to support the belief that women are much slower players than men. Compared to women the men spent 30% more match time in the motions of walking and jogging. This is illustrated when distance covered per end is compared. Men covered an average of 70m per end and women 51.5m, a 35% differential.

Total match time was not found to be different when singles and pairs were compared however when the first 9 ends were analysed pairs took an average of 4min longer than singles. Singles players covered on average 500m more than their pairs counterpart (2 312m vs. 1 876m), this difference was primarily due to the amount of distance traveled in walking forward (2 057m vs. 1 528m). Interestingly, the skips of the pairs combinations covered more than twice the distance walking backward than the singles players, 194m vs. 95m. It was observed during the matches that after discussing tactics at the head or mid-rink that the skip of the pairs would take paces backward to either finish the conversation with their playing partner or to have a longer look at the formation of the head. This may explain the vast difference in distance covered in the motion of walking backward. Jogging was not found to be different (160m vs. 154m) and appeared to be very individual with singles players recording distances of 235m and 85m and pairs 82m and 225m (men and women respectively). Differences were observed in the % of match time spent in the motion categories of waiting (55% vs. 70%), bowling (12% vs. 5%), and walking forward (25% vs. 19%). Additionally, the frequencies of these same motions were also found to be different. Waiting (140 vs. 90), walking forward (124 vs. 90) and bowling (78 vs. 32). Due to the mean times of the

motion categories of bowling and walking forward being similar the difference in the frequency explains the difference in the % of match time each one accounted for. Although singles players had 50 more occurrences of waiting than the skips of the pairs combinations the difference in the mean time of this motion category (20s vs. 41s) explains the % match time difference observed for this motion.

During the matches analysed (n=4) a little more than 2000m was covered in 90 minutes. This gives an average movement velocity over the duration of the match of approximately 1.5km/h. However, 65% of this time was spent in the motion category of waiting. Walking, jogging and bowling accounted for the remainder of this time. Bowling is a relatively stationary event and seeing this accounted for approximately 9% of match time the approximate 2000m distance that was covered in a match was realistically done in 26% of total match time, or 23½min. This equates to an average movement velocity of approximately 5km/h. The motion category of jogging accounted for a minor portion of total distance covered (7.5%) with each bout lasting approximately 6s, covering 16m at an average velocity of 9.6km/h. Overall, a lawn bowls match would rate as a light-moderate intensity activity. Energy expenditure is estimated to be approximately 260kcal per match, roughly equivalent to the energy required to complete a 40min brisk walk (Whitney and Rolfes, 2002).

Heart rate was found to remain relatively stable throughout each match at an average of 60% age-predicted HR_{max}. In fact, 55% of match time (approximately 50min) was spent with heart rates between 50-60% age-predicted HR_{max}. A further 16min (or 18%) of match time was played at between 60-70% age-predicted HR_{max}. The combination of these two periods of match time (accounting for more than 70% of match time or approximately 60min) would roughly equate to exercising at 35-40% VO_2max (Swain et al., 1994). According to American College of Sports Medicine (2000) 60 minutes of activity at this exercise intensity would not be an effective weigh-loss orientated exercise session and would provide for only minimal cardiovascular and health related benefits. The maximum heart rate observed was on average close to 80% HR_{max}. Whilst this top end heart rate was only realised very briefly during each match it may pose problems to those suffering from cardiovascular complaints such as hypertension.

Golf participants covered 10km during an 18-hole round (Thériault and Lachance, 1998). Goalkeepers in soccer have been reported to cover around 4km per match (Stolen et al, 2005), rugby union referees 8.5km (Martin et al, 2001) and rugby league referees 6.7km (Kay and Gill, 2003). A total match distance of 10km was covered by elite female soccer players (Krustup et al, 2005). These values are all substantially greater than the 2km observed in this study.

Super 12 rugby union players spent approximately 40% of match time standing, 38% walking and 16% jogging with forwards engaging in static exertion 10% of total match time (Duthie et al, 2005). Elite female soccer players spent an average of 16%, 44% and 34% of match time in the motions of standing, walking and jogging respectively (Krustup et al, 2005). While the combination of the three motion categories (standing, walking and jogging) is approximately equivalent to this study the contribution that each makes to the total is vastly different. Interestingly the 10% match time spent in

static exertion reported by Duthie and co-workers (2005) in Super 12 forwards roughly parallels the 9% of match time spent bowling in this study, another static activity.

As could be predicted, average heart rates during play for all team sports are substantially higher than that observed during this study. Krustup and colleagues (2005) investigated 14 elite female soccer players and reported average and maximal heart rates of 87% and 97 % of HR_{max} respectively. During a round of golf it has been reported that younger players spend approximately 6% of playing time at a high intensity while 18% of playing time is at or below 50% HR_{max} (Broman et al, 2004). These values are similar to that observed in this study and with the average age of participants in the current study being similar to that of the work by Broman and co-workers (2004) it appears that the games of golf and lawn bowls may share a similar intensity.

CONCLUSION

The time-motion analysis for movements by players in the sport of Lawn Bowls has been investigated for the first time. The majority of match time was spent in a waiting or preparatory activity (65%) with the next most time consuming activity being forwards walking (≈25%). The average heart rate was 107±15bpm or 57±7% of age-predicted HR_{max}. A little more than 2000m was covered during the match with the vast majority (85%) of this total distance being forwards walking. Overall, Lawn Bowls appears to require light-moderate intensity activity and appears to be similar to the physical demands of golf.

Acknowledgements
The authors would like to thank the Australian Lawn Bowls Team that competed in the 2006 Tri Nations Cup and Bowls Australia Inc. for their cooperation during the data collection process. The authors appreciate the assistance during the analysis of footage given by Ms. Bec O'Connell.

REFERENCES

American College of Sports Medicine (2000) *ACSM's Guideline for Exercise Testing and Prescription.* 6 th edition. Lippincott Williams & Wilkins, Maryland.

Brohman, G., Johnsson, L. and Kaijser, L. (2004) Golf: a high intensity interval activity for elderly men. *Aging Clinical and Experimental Research* **16(5)**, 375-381.

Duthie, G., Pyne, D. and Hooper, S. (2005) Time motion analysis of 2001 and 2002 super 12 rugby. *Journal of Sport Sciences* **23(5)**, 523-530.

D'Ottavio, s. and Castagna, C. (2001) Physiological load imposed on elite soccer referees during actual match play. *Journal of Sports Medicine and Physical Fitness*. **41(1)**, 27-32.

Judson, R. (2004). Online Bowler Website. Retrieved January 16, 2006, from http://onlinebowler.com/bowlscoaching/index2.php?id=Credits&Lo=Home

Kay, B. and Gill, N.D. (2003) Physical demands of elite Rugby League referees: Part one – time and motion analysis. *Journal of Science and Medicine in Sport* **6(3)**, 339-342.

Krustup, P., Mohr, M., Ellingsgaard, H. and Bangsbo, J. (2005) Physical Demands during an Elite Female Soccer Game: Importance of Training Status. *Medicine and Science in Sport and Exercise* **37(7)**, 1242-1248.

Martin, J., Smith, N.C., Tolfrey, K. and Jones, A.M. (2001) Activity analysis of English Premiership rugby football union refereeing. *Ergonomics* **44(12)**, 1069-1075.

Murase, Y., Kamei, S. and Hoshikawa, T. (1989) Heart rate and metabolic responses to participation in golf. *Journal of Sports Medicine and Physical Fitness* **29(3)**, 269-272.

Stølen, T., Chamari, K., Castagna, C. and Wisølff, U. (2005) Physiology of Soccer: An Update. *Sports Medicine* **35(6)**, 501-536.

Swain, D.P., Abernathy, K.S., Smith, C.S., Lee, S.J. and Bunn, S.A. (1994) Target heart rates for the development of cardiorespiratory fitness. *Medicine and Science in Sport and Exercise* **26(1)**, 112-116.

Thériault, G. and Lachance, P. (1998) Golf Injuries: An Overview. *Sports Medicine* **26(1)**, 43-57.

Whitney, E.N. and Rolfes, S.R. (2002) *Understanding Nutrition.* 9th edition. Wadsworth/Thompson Learning, California.

MODELLING THE INTERACTION IN GAME SPORTS – RELATIVE PHASE AND MOVING CORRELATIONS

Lames, M.

Institute for Sports Science, Augsburg University, Bavaria, Germany

ABSTRACT

Model building in game sports should maintain the constitutive feature of this group of sports, the dynamic interaction process between the two parties. For single net/wall games relative phase is suggested to describe the positional interaction between the two players. 30 baseline rallies in tennis were examined and relative phase was calculated by Hilbert transform from the two time-series of lateral displacement and trajectory in the court respectively. Results showed that relative phase indicates some aspects of the tactical interaction in tennis. At a more abstract level the interaction between two teams in handball was studied by examining the relationship of the two scoring processes. Each process can be conceived as a random walk. Moving averages of the scoring probabilities indicate something like a momentary strength. A moving correlation (length = 20 ball possessions) describes the momentary relationship between the teams' strength. Evidence was found that this correlation is heavily time-dependent, in almost every single game among the 40 examined ones we found phases with a significant positive as well as significant negative relationship. This underlines the importance of a dynamic view on the interaction in these games.

KEY WORDS

game sports, model-building, relative phase, random walks

INTRODUCTION

Game sports may be defined as those sports, where two parties (teams, doubles or single) try to achieve their goal and to avoid that the opponent achieves his one (Lames, 1991). This constitutes an interaction process, and the observable performance is rather the emergent result of this interaction process than the display of skills and abilities of the two parties. The nature of game sports also implies that this interaction process is dynamic. It changes during the match due to the permanent search for successful behaviour, due to strategic considerations depending for example on the actual score or due to a reaction imposed by an action of the opponent. This constitutes a sharp contrast to other sports such as 100m dash or marathon where performance is largely determined by the (rather constant) skills and abilities of the athletes.

If this notion of game sports as dynamic interaction processes is accepted, two important consequences are to be drawn. First, some of the traditional methods of performance analysis in sports science become doubtful. For example, the search for behavioural norms becomes a futile endeavour if behaviour changes dynamically and emerges from the singular encounter of the two opponents. Also, assessing individual

skill in game sports will remain a problem as long as the measures used add up (weighted) frequencies of observed behaviour and do not respect the singularity and dynamics of an interaction process. The second consequence is that this notion stimulates the search for new models which are capable to describe the crucial properties of game sports, interaction and dynamics.

In this article, two approaches are outlined which tackle the challenges described above from different perspectives. First, the positional interaction between two players in single net and wall games is described by the relative phase between their trajectories. The second approach uses the random walk concept to assess the dynamical strength of the two teams in invasion games and studies the interaction between the two processes by moving correlations.

RELATIVE PHASE IN TENNIS

The idea of describing movements of two players with their relative phase was first introduced by McGarry et al. (1999) in squash. They were influenced by an interpretation of the players' moves as the moves of a dancing couple. Certainly, another source of this idea was the successful application of relative phase in order to describe coordinative patterns in movement science (Haken, Kelso & Bunz, 1985; Kelso, 1995). McGarry et al. examined the absolute distance of the players from mid-court and found dominantly an anti-phase behaviour. Palut and Zanone (2005) calculated relative phase for the first time with Hilbert transform. They used the lateral distance from mid-court in tennis and also found that most of the time, tennis players showed an anti-phase behaviour, but also in-phase values of relative phase showed a relative maximum.

Our own investigations were in tennis. We focused on methodological issues and addressed the question of the meaning of different values of relative phase for the status of the game.

Why is relative phase a promising approach to describe the spatial interactions in a net/wall game? From a systems point of view, the movements in tennis can be perceived as the movements of two subsystems, the players. These subsystems are strongly coupled by the nature of the game because they exchange strokes. While one player hits, the other tries to get in a "neutral" position, from where he has the best opportunities to arrive in time at the next stroke. As soon as he recognizes the direction of the stroke, he moves to the place of contact, while the other player moves to his "neutral" position. Figure 1 displays an idealised long-line and cross rally with the corresponding positions.

A very interesting hypothesis from a practical point of view is the relation between the relative phase and the state of the rally. One might assume that a stable relative phase indicates a stable game when no player has problems to arrive just in time for his stroke. The very nature of tennis demands, though, to use placement and speed of the strokes to create pressure and win the point at last. This should result in a perturbation of relative phase. So, the hypothesis is that in a stable phase of the rally the relative phase is stable, but in the final phase, when a winner is scored or the opponent is forced to commit an error, the relative phase becomes unstable. If this hypothesis could be proven it would allow to determine the pressure created during a rally which would in turn be a valuable instrument for practical analyses.

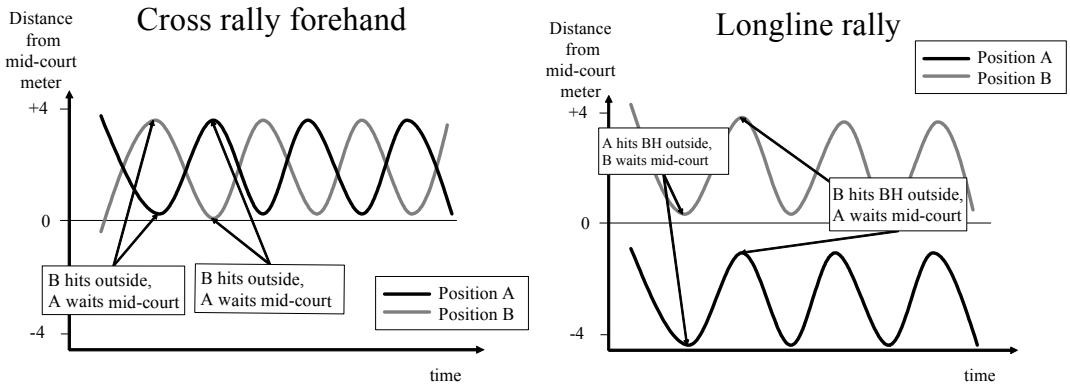
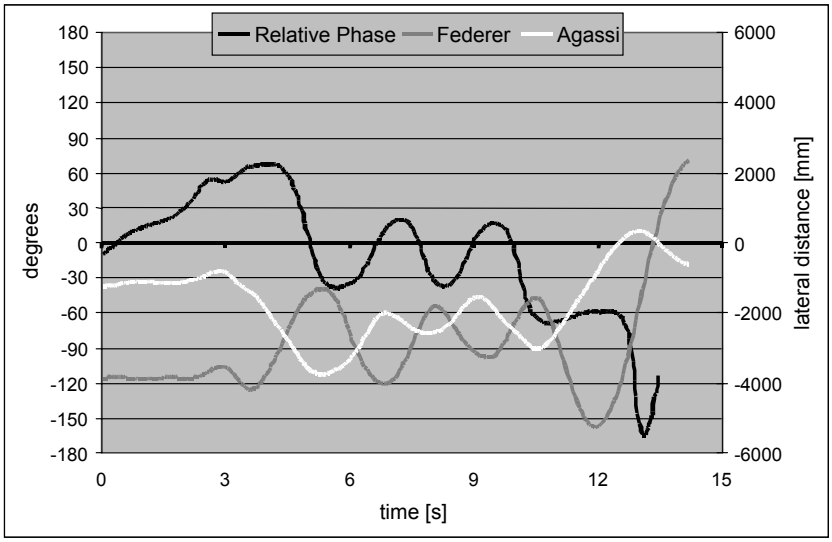


Figure 1: Idealised paths of lateral distance during a cross rally forehand and during a long-line rally in tennis. Relative Phase is constant in both cases, anti-phase (180°, left) and in-phase (0°, right).

We examined 30 rallies of top class athletes which we recorded from broadcasts of Grand Slam tournaments (Paris and Melbourne). The rallies were selected if they had a considerable length and if they were conducted and finished at the baseline. 18 rallies were played by female athletes. The positions of the players were obtained by image detection methods provided by the faculty of computer science, technical university Munich. Relative phase was calculated from the smoothed (1Hz filtering) time-series of positional data from the two players. The algorithm of Hilbert transform (MatLab) was used for the calculations. This procedure is well known in signal theory and allows to calculate continuous relative phase which is mandatory for we have comparatively few strokes in a rally (Pikovsky, Rosenblum & Kurths, 2001).

The first methodological issue we addressed was the optimal database for calculating relative phase. We found that the lateral displacements provide a good representation of the behaviour in the court, but have some weaknesses in their phase structure. This is due to the fact that even in baseline rallies the players move also perpendicular to the baseline in a considerable amount. As a result relative phase sometimes shows features that are hard to interpret when taking lateral displacements.

Figure 2: Lateral displacements and their relative phase of two male players.



The end of the rally is "announced" by a change in relative phase from in-phase to anti-phase. As an alternative we took the players' trajectory in the court from measurement to measurement (25 Hz). Actually these are speed data and relative phase now informs about the phase relation of moving speed of the players independent from their position on the court. With this data we usually get clear results for relative phase but we lack much of the understanding what is going on in the court. As a result, we suggest analysing lateral displacement as well as the two-dimensional trajectories in the court.

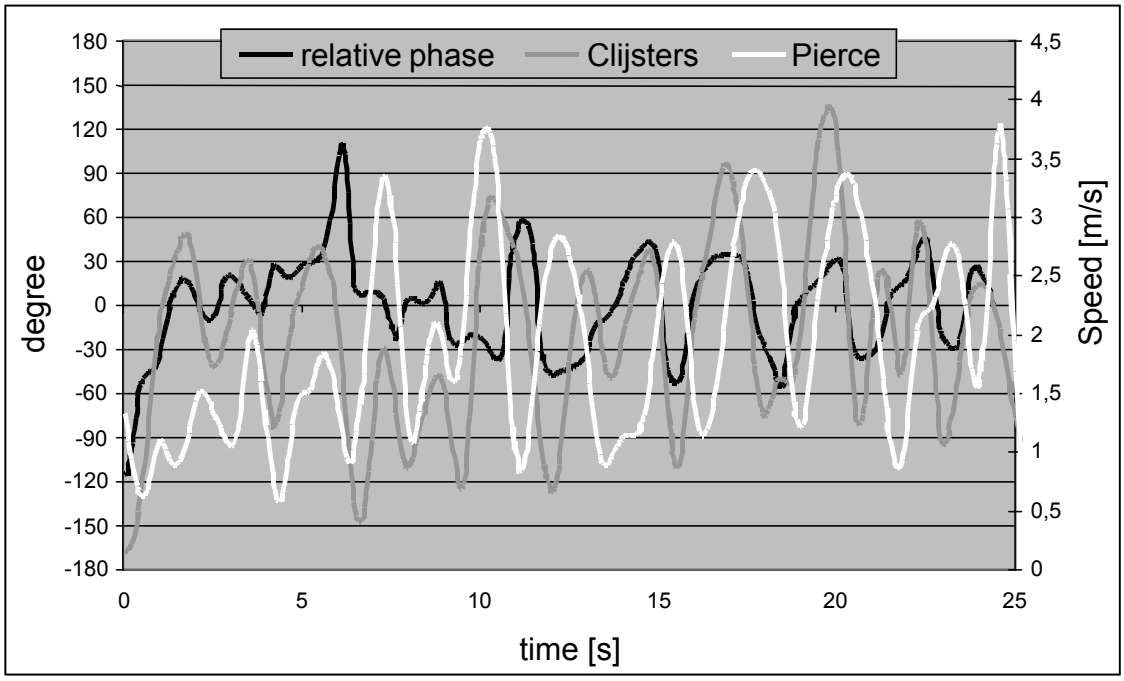


Figure 3: Speed of 2 female players during rally and relative phase between speed data.

The cyclical structure of the time series is evident, the rally ended with an unforced error of Clijsters which was not "announced" in relative phase which fluctuates around in-phase throughout the rally.

Results concerning the distribution of relative phase show that taking speed data we obtain a one-peak distribution indicating the dominance of in-phase. This is due to the fact that the rally synchronises the players in the sense that they alternate between two states: low speed while one player hits and the other orients for his next stroke, high speed while one player approaches the ball for his next stroke and the other comes back from his stroke towards a neutral position. This is in good agreement with the findings of Palut and Zanone (2005).

The dominant future task will be to link relative phase to tactical behaviour in the court. One way to achieve this will be a close examination of a larger sample of top-class rallies, but we will also instruct national-level tennis players to exhibit behaviour according to our instructions and study the provoked behaviour of relative phase.

RANDOM WALK MODELLING IN HANDBALL

The development of the score in handball for example may be perceived as two interlaced random walks. Each team has a probability p to score at ball possession,

P(1)=p, and a probability of q=1-p not to score, P(0)=q. Figure 4 left shows these two random walks for one example, a game between Germany and Croatia at the world championships in 2001. It becomes obvious that the processes are dynamic, we have phases where almost each ball possession leads to a goal but we find also periods with no goal scored. In some phases the two teams perform at the same level, in other phases there are differences. The local performance may be described by moving averages of the score. In figure 4 right the double backward moving average of length 4 is shown for each team. It reflects something like the momentary strength of a team and gives insight into the way the two teams interact.

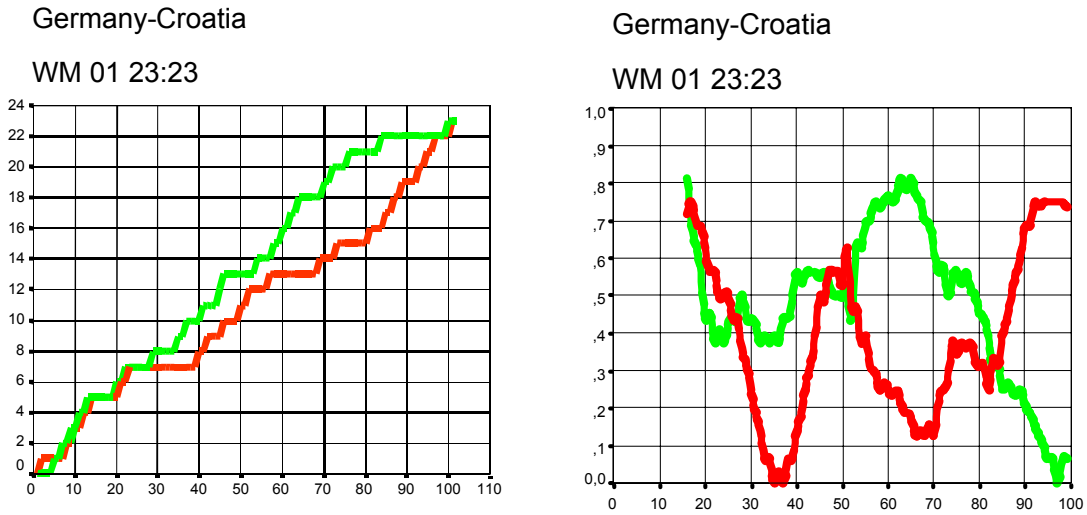


Figure 4: Left: development of score during a handball match
Right: Moving average of scoring probability

There is evidence for the hypothesis that a team's scoring rate is independent from the one of other team, but we see also phases with a seemingly strong dependence. Moreover, sometimes the momentary scoring probabilities seem to be negatively correlated (my team is good when the other is bad and vice versa, to be seen in the middle and the end of the example), but sometimes there is a positive relationship (my team performs well when the other does so, to be seen in the beginning).

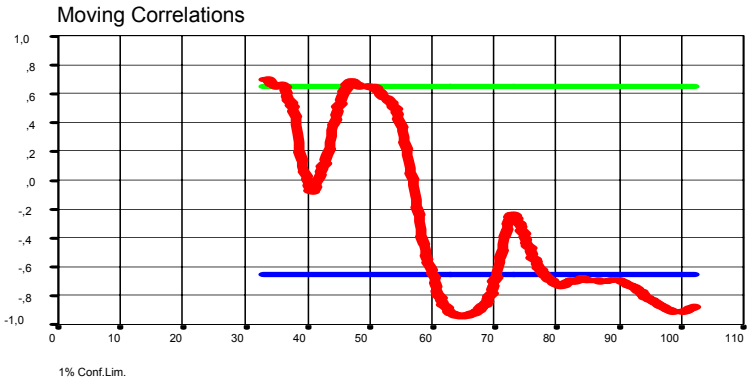


Figure 5: Moving correlations of length 20 between scoring probabilities in fig. 4 right

This lead to the idea of calculating moving correlations in order to study the relationship between the two scoring processes. Figure 5 shows that there are phases with significant positive and negative correlations. This behaviour is typical for most of the 30 games examined so far.

DISCUSSION

Modelling the positional interaction between tennis players by relative phase promises to reveal important insights into the nature of the game. Central aspects of game behaviour are described by relative phase. A certain limitation lies in the fact that relative phase is only apt to deal with longer rallies where the players try to create pressure by position play while scoring aces or unforced errors do not have an impact on relative phase. An interesting perspective is the description of other net/wall games by relative phase such as squash or badminton.

The examination of the scoring process as a random walk in handball provides theoretical as well as practical insights. For theorists it is fascinating to study the interaction dynamics during a game. For coaches it may be interesting to identify successful and less successful phases in a game as a starting point of a practical game analysis.

Acknowledgement

I'd like to thank my student Florian Walter for the calculations of relative phase with matlab.

REFERENCES

Haken, H., Kelso, J.A.S. and Bunz, H. (1985) A theoretical model of phase transitions in human hand movements. *Biological Cybernetics* **51**, 347-356.

Kelso, J.A.S. (1995) *Dynamic Patterns. The Self-Organization of Brain and Behavior*. Cambridge, Mass.: MIT-Press.

Lames, M. (1991) *Leistungsdiagnostik durch Computersimulation: Ein Beitrag zur Theorie der Sportspiele am Beispiel Tennis*. Frankfurt, Thun: Harry Deutsch.

McGarry, T., Anderson, D.I., Wallace, S.A., Hughes, M.D. and Franks, I.M. (2002) Sport competition as a dynamical self-organizing system. *Journal of Sport Sciences* **20**, 771-781.

McGarry, T., Khan, M.A. and Franks, I.M. (1999) On the presence and absence of behavioural traits in sport: an example from championship squash match-play. *Journal of Sport Sciences* **17**, 297-311.

Pikovsky, A., Rosenblum, M. and Kurths, J. (2001) *Synchronization: A universal concept in nonlinear sciences.* Cambridge: University Press.

Palut, Y. and Zanone, P.-S. (2005) A dynamical analysis of tennis players' motion: Concepts and data. *Journal of Sports Science* **23**, 1021-1032.

IMPRESS YOUR FRIENDS AND PREDICT THE FINAL SCORE: AN ANALYSIS OF THE PSYCHIC ABILITY OF FOUR TARGET RESETTING METHODS USED IN ONE-DAY INTERNATIONAL CRICKET

O'Riley, B. and Ovens, M.

School of Mathematical Sciences, Monash University, Australia

ABSTRACT

One-Day cricket's eternal problem is how to fairly account for an interruption that occurs during a team's innings. Several methods have been applied in the past, some more successfully than others. Numerous articles have been written about different target resetting methods applicable in one-day international cricket and how they "favour" one team over another. Gurram and Narayanan (2004) attempt to address this issue by quantifying the 'fairness' of five different methods,[1] using data from 377 uninterrupted matches[2]. In this paper we use an alternative approach looking at the psychic ability of four target resetting methods[3] and compare how well they predict the final score based on the present state of the first innings. We attempt to convert each of methods we investigate into a ball-by-ball predictive tool. We introduce a terminal interruption to the first innings at every ball and compute the predicted final score. We ascribe a nominal value to the difference between the final achieved score and the prediction given by each method. We compute our own 'Psychic Metric' to enable a comparison between the four methods. We also develop a computer package to manipulate the data from matches[4] in which the first innings was completed.

KEY WORDS

cricket, precicting scores, psychic abilities

INTRODUCTION

Many papers have been written about mathematics and cricket, on topics such as optimal batting strategies (Clarke, 1988), player performance measurements (Lewis, 2004, Allsopp and Clarke, 2000) and target resetting methods (Lewis, 1996, 1998, 2002, Gurram and Narayanan, 2004, Jayadevan, 2002, Armstrong, 1994, Bhogle, 1999). Several of the target resetting papers used historical data to build the method. We borrow this approach to evaluate the ability of four methods to predict the score achieved by the first batting team (Team 1), using data from 173 matches, some of which involved stoppages. Given that Team 2's target in an uninterrupted match is dependent on Team 1's final score, any target resetting method would need to have a reasonably good estimate of what Team 1 is likely to score. Our intention is to formulate a metric that could be used to assess the predictive ability of a target resetting method.

[1] The methods were: Duckworth/Lewis, Jayadevan, Most Productive Overs, Discounted Most Productive Overs and Average Run Rate

[2] Data extracted from the CricInfo website

[3] The methods chosen: Duckworth/Lewis, Jayadevan, Parab and Average Run Rate

[4] Data was kindly supplied by Champion Data

THE "FAIRNESS METRIC"

Gurram and Narayanan's 2004 paper addresses the fundamental issue of how "fair" some of the better-known target resetting methods are. Additionally, it provides some of the motivation for our work and as such, it is worthwhile considering issues raised by the paper. Firstly, in attempting to quantify how "fair" the chosen methods are, the paper fails to address some very important issues in relation to how the game is played. Gurram and Narayanan's fairness metric was only applied to games that went beyond the 25th over in the second innings[5]. Although the paper states that only games with "no interruption" were examined, "no interruption" was defined as a game consisting of two innings, where the innings concluded only when all balls had been bowled or all wickets had been lost. Consequently, this definition fails to acknowledge games that were interrupted, but resumed with no overs lost; and those games in which an interruption "threatened", but did not eventuate. By not taking these "interruptions" into account and given that Gurram and Narayanan only dealt with the second innings, one of the most important factors of one-day cricket is completely disregarded - the psychology of the game.

Secondly, Gurram and Narayanan originally found that the Average Run Rate method was "fairest" (with a fairness metric of 0.708). By their own admission (in section 2.1.1 of the paper), ARR has many downfalls, particularly the fact that the wickets remaining are not taken into account. Although ARR was once used in One Day International Cricket, it was eventually dismissed, due its many shortcomings, including the fact that it leads to an unfair advantage to Team 2 (as discussed in Lewis, 1996, Ovens, 2004). Consequently, it is worrying to see that ARR is still favourably viewed even when 20% of the "mismatches" are forgiven. It should also be noted that Gurram and Narayanan have no clearly defined rule to determine which mismatched overs should be removed; leading to the suspicion that one is able to take out particular overs in order to make one method perform better than another. On top of all this, their fairness metric asserts that of all the methods reviewed, Jayadevan's is the fairest, although the paper (like many others) states that one of the shortcomings of the Jayadevan method is that it fails to sufficiently address the issue of fallen wickets. The Duckworth/Lewis method appears to adequately take into account the effect of wicket loss, leading one to assume that this method would be "fairer". Over-all, Gurram and Narayanan's concept of computing a fairness metric as a way of comparing target resetting methods is laudable, but perhaps there are other factors that should be taken into account, e.g. the psychology of batting second. This provides motivation for our undertaking to formulate a metric that addresses these factors.

OUTLINE

The four methods chosen were Average Run Rate, PARAB, Duckworth/Lewis and Jayadevan. Average Run Rate (ARR) and PARAB (P) methods were chosen as they are easily adapted to predict a score achieved by the conclusion of an innings, using only the present runs scored and balls bowled. Duckworth/Lewis (D/L) was chosen as it is the current rain-rule used in One-Day international cricket. Jayadevan's method (J) was chosen as a potential alternative rain-rule that could be used to replace D/L (as

[5] According to ICC rules when the paper was written, a result would only be recorded if an interrupted match lasted for more than 25 overs per side.

discussed in Ovens, 2004). These last two methods both required manipulation to be able to be turned into predictive tools.

In order to use D/L as a predictive tool, we adapted the target formula (Equation 1) to the form shown in Equation 2.

$$T = \begin{cases} \left(S \times \dfrac{R2}{R1}\right) + 1 & if \quad R2 < R1 \\ S + 1 & if \quad R2 = R1 \\ S + \left(\dfrac{(R2 - R1) \times G50}{100}\right) + 1 & if \quad R2 > R1 \end{cases}$$

Equation 1: Standard Edition D/L Target Formula

Where T is the target for Team 2, S is the score achieved by Team 1, R1 is the resources available to Team 1, R2 is the resources available to Team 2 and G50 is the average score achieved in 50 overs in One-Day International Cricket (presently equal to 235).

$$T = S + \left(\frac{(100 - R1) \times 235}{100}\right) + 1$$

Equation 2: D/L "Predictive" Formula

In the case of a stoppage occurring during Team 1's innings, Jayadevan's method is applied as follows:

1. Determine the percentage of overs completed by Team 1.
2. Look up the corresponding normal score percentage in the normal table for the number of wickets fallen.
3. Determine the percentage of remaining overs after the stoppage with respect to the original number of overs remaining.
4. Look up the corresponding target score percentage in the target table.
5. Multiply the target score percentage by the difference between 100% and the normal score percentage.
6. Add this percentage to the normal score percentage obtained in step 2 to get the Effective Normal Score (ENS) in the total percentage of overs played.
7. Look up the target score percentage for the total percentage of overs played.
8. Multiply this target percentage by the ENS from step 6 to get the Multiplication Factor (MF).
9. Multiply the score made by Team 1 with MF to get the target for Team 2.

To convert Jayadevan's method into a predictive tool, we note that step 3 gives us 0% overs remaining, which in turn means that steps 4, 5 and 6 are unnecessary and the effective normal score is the normal score obtained in step 2. Thus, the result obtained in step 9 would be the target for Team 2, consequently the predicted score for Team 1 is one run less. It is worth noting that if Team 1's score is zero then this method results in a predicted score of zero. Further scrutiny of Jayadevan's method also indicates that, when using this as a predictive tool, the multiplication factor from step 8 will always be less than 1.

Software was written so that predictions using each method could be easily calculated from the present state of the match. Using the data provided by Champion Data we computed the predictions for each method on each ball of the first innings of the 173 matches. Figure 1 shows a screen shot of the predictions being computed ball by ball for ODI #1620.

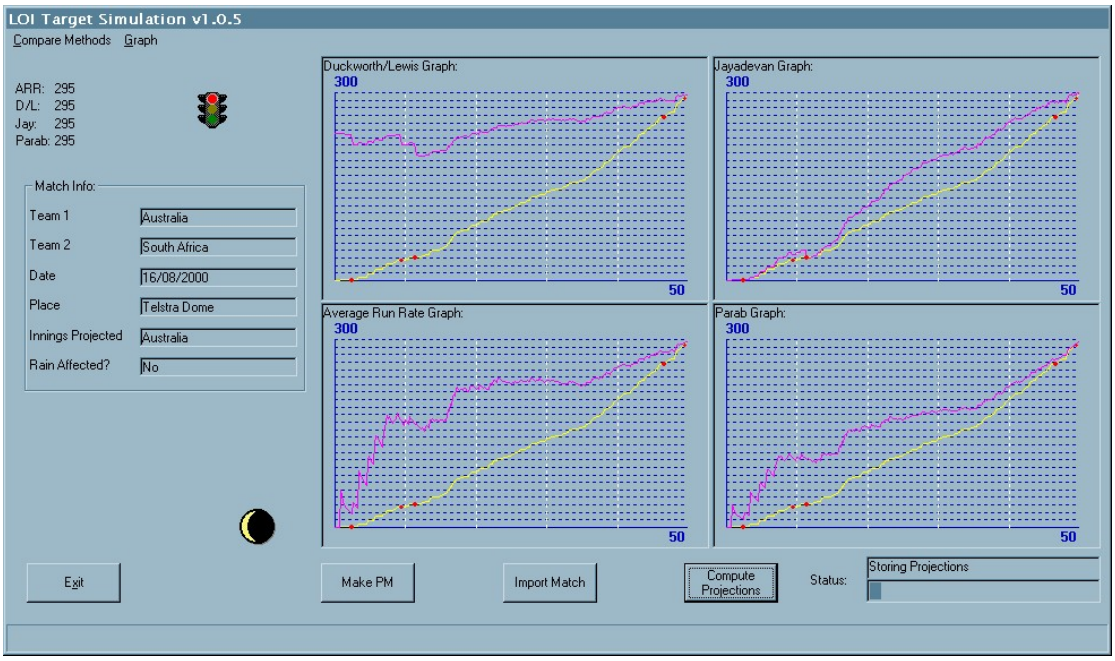


Figure 1: Software Screenshot

We then defined $OverProj_{ijk}$ as the difference between the prediction (on ball i, in match j, using method k) and the actual runs scored on ball i, match j. This $OverProj_{ijk}$ is then used to compute four different alternative Psychic Metrics; by ball, by delta, by delta/ball and by arbitrary. We define the four Psychic Metrics as follows:

$$PM_{ijk} = \left(OverProj_{ijk}\right)^2 \times i$$

Equation 3: Psychic Metric by ball

$$PM_{ijk} = \begin{cases} 1 - \left(\dfrac{OverProj_{ijk}}{Total_j}\right)^2 & if\ OverProj_{ijk} \le Total_j \\ 0 & if\ OverProj_{ijk} > Total_j \end{cases}$$

Equation 4: Psychic Metric by delta

$$PM_{ijk} = \begin{cases} \left[1 - \left(\frac{OverProj_{ijk}}{Total_j}\right)^2\right] \times i & if\ OverProj_{ijk} \le Total_j \\ 0 & if\ OverProj_{ijk} > Total_j \end{cases}$$

Equation 5: Psychic Metric by delta/ball

$$PM_{jk}(Raw) = \sum_{i=0}^{i=300} (300 - i) \times (S_i)$$

Equation 6: Psychic Metric by arbitrary

By observing equation 3, one can clearly see that the by ball method gives a squared difference, weighted by ball, an approach similar to that used to compute variance. Equation 4 (by delta) subtracts the squared proportion (of over projection divided by total) from 1 where a result closer to 1 indicates a better prediction. Equation 5 then weights this method by ball, such that a result closer to *i* indicates a better prediction. Equation 6 allows for a nominal value to be ascribed to a range of differences (which are able to be defined by the user).

Using the four psychic metrics, we were able to come up with scores scaled for each method.

RESULTS

We present examples using two of the psychic metrics to demonstrate the results obtained from the work undertaken in this paper.

1 An example of the Arbitrary Psychic Metric

To illustrate how the arbitrary psychic metric could be used, we have, for each ball, scored the absolute difference between the actual and predicted runs, according to the following scale.

Table 1: Psychic Scale

Difference	Score (S_i)
1	10
5	5
10	1
11+	0

This score was then multiplied by the balls remaining in the innings. Summing over the entire innings gives the following representation for Match j, Method k.

$$PM_{jk}(Raw) = \sum_{i=0}^{i=300} (300 - i) \times (S_i)$$

Equation 7: Raw Psychic Metric

From this representation, it is clear that the maximum possible raw score for any 50 over match would be 451,500 and thus:

$$PM_{jk} = \frac{PM_{jk}(Raw)}{451500}$$

Equation 8: Scaled Psychic Metric

As can be seen from Table 2, the mean for D/L is maximal over the four methods, although equally has the highest standard deviation. J is by far the most stable but also consistently less able to predict the final score. It is also interesting to note that none of the computed confidence intervals overlap that of the D/L method, indicating that there is a significant difference at the 5% level.

Table 2: Basic Statistics on Psychic Metric (Arbitrary)

	Mean	Std Dev	Lower CI	Upper CI
ARR	0.12	0.12	0.10	0.14
D/L	0.18	0.16	0.16	0.21
Jay	0.03	0.04	0.02	0.03
Parab	0.05	0.10	0.03	0.06

Figure 2 shows the average differences for each method, over the course of 300 balls.

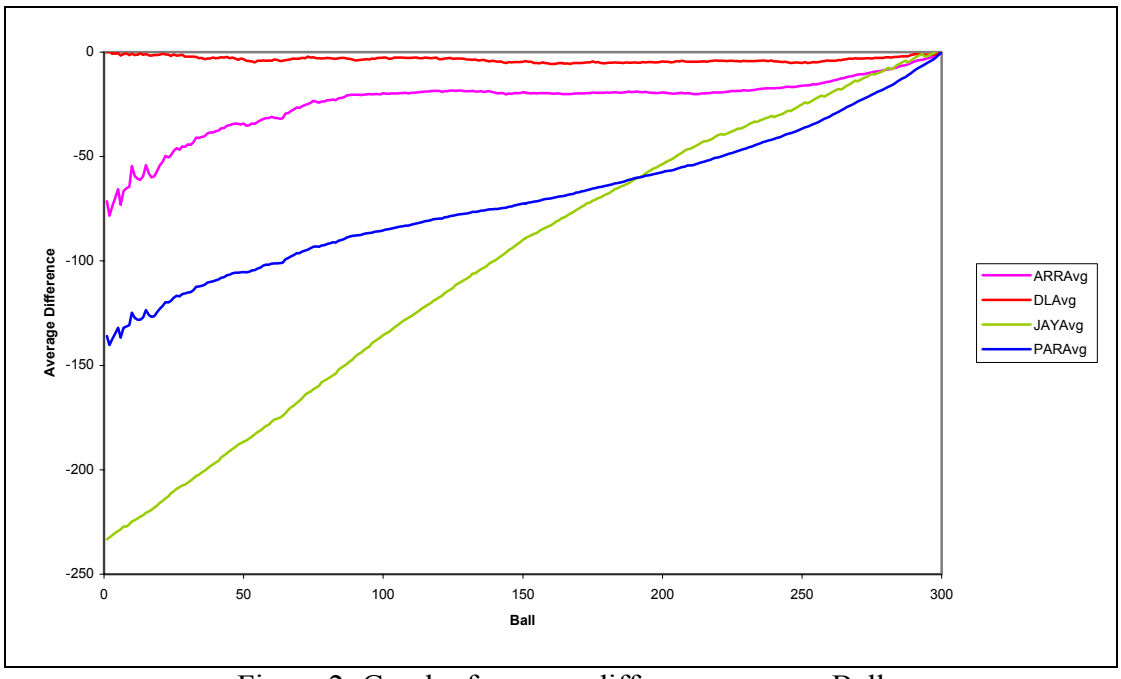


Figure 2: Graph of average differences versus Ball

As can be seen in the above graph, the Jayadevan, PARAB and Average Run Rate methods do poorly when compared to the Duckworth/Lewis method. One expects that as we get closer to the end of the game, the predictions will improve for all methods and this is evidenced in the above graph. The following table shows a comparison between

the four methods with 10 overs remaining. Table 3 shows that, at this point in the innings, ARR predicts, on average, 17.29 runs below the actual score, compared with D/L, 4.50 runs below, Jayadevan, 31.13 runs below and PARAB, 41.60 runs below. It is interesting to note that the minimum and maximums for D/L are almost symmetrical about zero, whereas the other three methods are asymmetrical about zero being further on the negative side.

Table 3: Method Comparison (Arbitrary)

	ARR	D/L	J	P
Average	-17.29	-4.50	-31.13	-41.60
Variance	635.06	475.68	584.47	798.24
Standard Deviation	25.20	21.81	24.18	28.25
Minimum	-91	-70	-95	-121
Maximum	62	72	39	26

2 An example using the By Delta/Ball Psychic Metric

As can be seen from Table 4, the means of ARR and D/L are both significantly close to one another and reasonably close to the 'ideal' average final score (150). We have not computed the confidence levels, as these values are bounded above.

Table 4: Basic Statistics on Psychic Metric (By Delta/Ball)

	Mean	Std Dev
ARR	144.59	88.28
D/L	146.34	88.23
Jay	128.57	97.61
Parab	136.98	89.51

Table 4 shows a comparison between the four methods when the last 10 overs are being played. At this stage of the game, the Duckworth/Lewis method's average predicted final score (267.66) is closest to the 'ideal' average score, closely followed by Average Run Rate (268.26), Jayadevan (264.24) and PARAB (258.94). The ideal average is 270.5 runs. Due to the definition of the metric, no minimum or maximum values have been given. The potential minimum is for all methods is 0 and the potential maximum is 300, therefore no further information would be gained by including these values.

Table 5: Method Comparison (By Delta/Ball)

	ARR	D/L	J	P
Average	268.26	267.77	264.24	258.94
Variance	389.12	660.89	1119.84	1720.73

The software used to obtain these results was adapted to allow us to check for any games that may have heavily influenced the results. This was done by visual inspection, using graphs plotted by the software.

DISCUSSION

We expected to find that both the Average Run Rate and PARAB methods would not perform well when compared to other methods, as it has been demonstrated time and time again that they have potentially serious shortcomings as target resetting methods. These shortcomings imply that the methods will not perform well over extended periods, however they were included to aid in comparison. Furthermore, we also expected that Jayadevan's method would not perform well, as it is a method that is not designed to predict scores but rather to reset the target, yet, as mentioned earlier, a target resetting method should have a reasonably accurate estimate of Team 1's expected final score. The intention of this work was to create a metric that could be used to assess the accuracy of a target resetting method in computing Team 1's expected final score. This leads us to conclude that Jayadevan's method would work best only when Team 1 has completed its innings. As the Duckworth/Lewis method is readily adaptable as both a predictive and a target resetting tool, it met our expectations to surpass the other methods. Assessing the four methods with our various psychic metrics, we conclude that the Duckworth/Lewis method is the most reliable in computing the expected final score of Team 1 and therefore should be chosen above other potential target resetting methods.

In attempting to create a metric to assess a target resetting method, we have inadvertently introduced potential weaknesses. Firstly, the data set we used, kindly supplied by Champion Data, consisted of only 173 matches, most of which came from series in which Australia was involved. Consequently, it was not a truly random sample. A more accurately representative data set, consisting of matches from various series and between various teams would allow us to address this problem. Secondly, in order to measure the methods against our psychic metric "ruler", we needed to adapt each of the methods to produce a 50 over score. For ARR, PARAB and D/L this is readily achieved, but for J this produces problems, as one of the assumptions underlying the method is that Team 2 cannot possibly face more overs than Team 1. Whilst this is true, it is a technical deficiency of the J method that restricts it from providing a prediction of Team 1's final score. Thirdly, in this work, we have only addressed terminal stoppages, which may have biased the results.

In the future, we plan to look at multiple stoppages, to see whether they affect a team's predicted final score. Another area that could be looked at would be to givesuch metrics the ability to be classified by both country and ICC ranking. This would allow one to deal with inefficiencies stemming from the issue of low scoring teams playing high scoring teams and the problems this causes when target resetting methods are applied. An additional aspect to consider for possible future research would be to investigate if suggested new rules affect how a team plays its innings and if this in turn affects the final score prediction. As an extension of this, one could look at if and how the batting order affects a final score prediction (like Bukiet et al, 1997 and to appear). Overall, in our opinion, D/L is presently the best target resetting method and is the most accurate at predicting Team 1's final score.

Acknowledgements

The authors would like to thank Champion Data and Mr David McKenzie for their assistance in gathering the data needed. Additionally, the authors would like to thank the anonymous reviewer for their insightful and helpful comments.

REFERENCES

Allsopp, P. and Clarke, S. (2000) Methods for quantifying performances in One-Day Cricket" *5th Australian Conference on Maths and Computers in Sport* (Eds, Cohen, G. and Langtry, T.) University of Technology, Sydney, Australia, pp 28-36

Armstrong, J. (1994) *Masters Thesis* Monash University, Victoria, Australia

Bhogle, S. (1999) "Rediff On The NeT" http://www.cmmacs.ernet.in/nal/pages/dlm.htm

Bukiet, B., Harold, E. and Palacios, J. (1997) *Operations Research* **45** (1)**,** 14-23.

Bukiet, B., and Ovens, M. (to appear) *A Mathematical Modelling Approach to One-Day Cricket Batting Order*s.

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

Duckworth, F. and Lewis, T. (1996) A fair method for resetting the target in interrupted one-day cricket matches. *Third Conference on Mathematics and Computers in Sport* (Ed. de Mestre, N.) Bond University, Queensland, Australia, pp51-68

Duckworth, F. and Lewis, T. (2002) Review of the Application of the Duckworth/Lewis Method of Target Resetting in One-Day Cricket. *MathSport* **Vol 6** (Eds. Cohen, G. and Langtry, T.) Bond University, Queensland, Australia, pp127-140

Duckworth, F. and Lewis, T. (2002) Can the Probability of Winning a One-day Cricket Match be Maintained Across a Stoppage? *MathSport* **Vol 6** (Eds. Cohen, G. and Langtry, T.) Bond University, Queensland, Australia, pp141-154

Gurram, M. and Narayanan, S. (2004) Comparison of the Methods to Reset Targets For Interrupted One-Day Cricket Matches. *7th Australasian Conference on Mathematics and Computers in Sport* (Eds, Morton, H. and Ganesalingam, S.) Massey University, Palmerston Norton, New Zealand, pp. 165-171.

Jayadevan, V. (2002) *Current Science* **83,** 577-586

Lewis, T. (2004) Steps Towards Fairer One-Day Cricketing Measures of Performance. *7th Australasian Conference on Mathematics and Computers in Sport* (Eds, Morton, H. and Ganesalingam, S.) Massey University, Palmerston Norton, New Zealand, pp. 198-213.

Lewis, T. and Duckworth, F. (1998) Developments in the Duckworth-Lewis (D/L) method of target-resetting in one-day cricket matches. *Fourth Conference on Mathematics and Computers in Sport* (Eds, de Mestre, N. and Kumar, K.) Bond University, Queensland, Australia, pp. 131-151.

Ovens, M (2004) If it rains, do you still have a sporting chance? *7th Australasian Conference on Mathematics and Computers in Sport* (Eds, Morton, H. and Ganesalingam, S.) Massey University, Palmerston Norton, New Zealand, pp. 242-252.

FURTHER STEPS TOWARDS FAIRER PLAYER PERFORMANCE MEASURES IN ONE-DAY CRICKET

Lewis, A.J.

Oxford Brookes University, UK

ABSTRACT

The author's paper at the 7th Conference on Maths and Computers in Sport on this topic showed that the Duckworth/Lewis methodology has considerable potential for providing more relevant and objective measures of performance in one-day cricket for single matches and over short series of matches such as the Australian VB Series. This paper evaluates how the methodology of these measures would apply on the longer term. This assessment is based upon several series of matches, involving the main players from most of the major cricketing nations. Although data used are not exhaustive they are of sufficient quantity that strong indications of the long term viability of the measures can be gleaned, leading to a realistic expectation of either inclusion of the methodology into the existing LG ICC rankings or the development of an independent set of performance indicators that more reliably reflect the value of players' inputs to matches. Such measures are expected to be of value to team management, to media commentators and, not least, to the public in assessing performances of one-day players both on current performance and longer term over their respective careers.

KEY WORDS

sports statistics, cricket, modelling

INTRODUCTION

This paper is concerned with further evaluating proposed improvements to measures of player performance in one-day cricket. In Lewis (2004, 2005) it was shown that the Duckworth/Lewis (D/L) methodology of target resetting in interrupted limited overs cricket matches can be used to assess, arguably more reliably, performances of batsmen and bowlers compared with the traditional measures of batting and bowling averages, and their subsidiaries of strike and economy rates. In addition, by using other aspects of the D/L methodology, Lewis (2004, 2005) and Beaudoin and Swartz (2003) have created different measures that assess 'all-round' performance, which previously was not possible in an objective way using the traditional measures of performance.

These measures were shown to assist in assessing contributions to the team effort from individual players, not only for a particular match, but also for a series of matches and to provide aids to adjudicators in deciding on 'man-of-the match' and 'player of the series' awards' .

Lewis (2004) concluded that the analysis of many more matches was necessary in order to investigate the reliability and viability of these measures in the longer term. He further suggested that if this were to be the case then their methodology could either be

incorporated into the popular player ratings or perhaps supplant the existing player-rating measures now run by the International Cricket Council (ICC) and promoted as the LG ICC ratings [http://www.cricketratings.com/ accessed 7 April 2006].

REVIEW OF D/L PERFORMANCE MEASURES

The performance measures proposed in Lewis (2004) are based on the D/L methodology (Duckworth and Lewis, 1998, 2004).

Net contribution

The measure called Net Batting (or Bowling) Contribution is a summation, on a ball-by-ball basis, of what a batsman scores (or a bowler concedes) relative to the runs expected from the D/L model (Lewis, 2004). The measure for each player is in the form of a positive or negative number of runs, which is usually non-integral, representing the overall extent to which the player has performed above or below expectation for the match situation for the balls received or delivered. With a common and additive scale it is possible to compare not only batsmen and bowlers amongst themselves but also batsmen with bowlers. Further, it is possible to combine a player's performance in batting and bowling in order to produce a measure of 'all-round' performance.

Resource (percentage) average

The second measure of player-performance proposed in Lewis (2004) was the Batting (or Bowling) Resource Percentage Average. According to the methodology developed, the delivering and receiving of a ball and the loss and taking of a wicket involve consumption of resources by the batsman and contribution of resources by the bowler. The respective Resource Percentage Averages are the average runs scored per unit of resource consumed by the batsman and the average runs conceded per unit of resource contributed by the bowler. For the latter this includes the number of wides and no-balls bowled as these forms of 'extras' count against the bowler but are not credited to the batsman.

Runs per match

Beaudoin and Swartz (2003) have produced a measure called Runs per Match (RM) both for batsmen and bowlers. These measures are a multiple of 100 of the batting and bowling resource percentage averages respectively. These figures are interpreted as the average number of runs scored or conceded were the player to do all of the batting or the bowling himself. According to the D/L model of one-day cricket (Duckworth and Lewis, 1998, 2004) the average runs in 50 overs, with all wickets available, is represented by $Z(50,0)$ and has a value of 235. For batsmen an RM measure above 235 represents an above average batsman with the converse true for a bowler.

Long-run performance measures

Batting and bowling averages are the traditional ways of evaluating player performance in the long term, although couched in qualifying terms because of the known weaknesses in these measures (Lewis 2004, 2005). Using their RM measures Beaudoin and Swartz (2003) undertook an analysis of the performances of several high-profile players between

1998 and September 2002 in order to achieve more meaningful measures of performance over those players' careers during that period. Some of these players also appear in the analysis that follows. Several interesting comparisons will be drawn.

In order to gain some indication of how Net Contributions and Resource Percentage Averages would work longer term, several one-day international (ODI) series have been analysed in exactly the same way as the Victoria Bitter (VB) Series for 2003 discussed in some detail in Lewis (2004, 2005). The series of matches summarised in subsequent analyses, in chronological order, are as follows. Several abbreviations subsequently employed in this paper are defined through this list.

(i) VB 2002 (VB02): between Australia (Aus), South Africa (RSA) and New Zealand (NZ) [14 matches]
(ii) NatWest Series 2002 (NW02): England (Eng), India (Ind) and Sri Lanka (SriL) [10]
(iii) VB 2003 (VB03): Australia, England, Sri Lanka [14]
(iv) NatWest Series 2003 (NW03S): England, Sth Africa, NZ [10]
(v) NatWest Challenge 2003 (NW03C): England, Pakistan (Pak) [3]
(vi) Tour 2003: West Indies (WI), Australia [7]
(vii) T V Sundaram Iyengar & Sons Cup 2003 (TVS03): India, Australia, NZ [10]
(viii) VB 2004 (VB04): Australia, India, Zimbabwe (Zim)[14]
(ix) Tour 2004: West Indies, England [5]
(x) Tour 2004: Sri Lanka, Australia [5]
(xi) NatWest Series 2004 (NW04S): England, NZ, West Indies [10]
(xii) NatWest Challenge 2004 (NW04C): England, India [3]
(xiii) Tour 2005: Sth. Africa, England [7]
(xiv) VB 2005 (VB05): Australia, NZ, Pakistan [14]
(xv) Tour 2005: NZ, Australia [5]
(xvi) NatWest Series 2005: England, Australia, Bangladesh (Bang)[10]

All major ODI countries have featured in at least three series, with the exceptions of Pakistan, Zimbabwe and Bangladesh. Consequently the principal players from these three countries are not included in subsequent analysis. Because the sources of the author's electronic data are located in the United Kingdom and Australia there is arguably a bias towards performances of players from these countries in subsequent analysis. However, such bias is believed not to invalidate the conclusions to be drawn.

Because of the turnover of players during the three-year period of this research, many players have played insufficiently for their overall performances to be regarded as representative. Players included in subsequent analyses are only those who have played a role in more than 12 innings in order to reduce bias due to small volumes of input. And in order that comparisons are as current as possible only players who have retained a ranking in the top 100 of the LG ICC ratings are included.

The LG ICC ratings

The LG ICC ratings [http://www.cricketratings.com/ accessed 7 April 2006] are generally regarded as the current market leaders of player performances in international cricket. There are two sets of rankings, one for Test Matches, and one for ODIs. This paper is concerned only with the latter set, which include ratings and rankings for batsmen and bowlers and more recently, rankings for all-round performance. The ratings are provided on a scale from 0 to 1000. The Internet website cited gives a general overview of the principal features of the ratings but does not provide full details of the calculations due to commercial sensitivity.

The mechanism of the LG ICC ratings uses only information on players' performances in a match that is provided by a standard summary scorecard as typically reproduced in Engels (2006). Consequently, the information captured by these ratings is likely to be more limited in scope than that used for analysis in this paper, although there is substantially less processing than is required by the D/L and RM performance measures.

Because of a discounting factor over time in the LG ICC ratings, players' more recent performances are given greater weight and players who are not selected for play quickly lose their status in the rankings

Whereas the information in the traditional scorecard provides the details for calculation of traditional averages, it is not sufficient for the analysis undertaken in this paper or in Lewis (2004). Substantial detail on the outcome of every ball is required. The author's data-sources, from Australia and the United Kingdom, provide detail in electronic form, to the level of who bowled to whom, how many runs were scored, or byes and leg-byes, wides and no-balls conceded and the mode of dismissal when the wicket falls.

Performance measures in the longer-term

It must be stressed that the analysis that follows is intended to indicate how the Net Contribution and Resource Percentage Average measures would operate over a series of matches. Because the number of matches included is just a sample from the players' careers there is no implication, at this stage, that the rankings produced are definitive rankings for the players involved. To some extent, however, comparison will be made with the current LG ICC international player ratings for the purpose of content validation of the D/L-based performance measures, and where possible, with the results of the study by Beaudoin and Swartz (2003).

In order to standardise the analysis, the net contributions and resource percentage averages will all be based on the D/L tables of September 2002 [http://www.icc-cricket.com/rules/d-l_table.pdf accessed 7 April 2006].

Batting rankings

Table 1 provides the performances of the eligible players (according to the criteria previously defined) over the one-day series mentioned above.

Table 1: Summary of batting performance measures for qualifying batsmen

Player	Coun-try	Batting Innings	Batting avge	Batting avge rank	Batting contrib	Batting contrib rank	Batting res% avge	Batting res% avge rank	LG ICC rating	LG ICC rank*
Pietersen	Eng	13	99.57	1	252.3	7	3.473	1	611	23
Lehmann	Aus	14	68.12	2	126.5	14	2.836	5	544	31
Kallis	RSA	20	54.87	3	40.6	26	2.306	31	708	8
Tendulkar	Ind	21	54.68	4	280.3	5	2.996	3	692	10
Sarwan	WI	23	50.82	5	156.1	11	2.661	10	672	13
Kaif	Ind	13	46.87	6	66.8	20	2.652	11	532	34
Clarke	Aus	33	46.66	7	243.9	8	2.649	12	705	9
Flintoff	Eng	31	43.96	8	309.8	3	2.998	2	711	5
Sangakkara	SriL	17	42.86	9	127.7	13	2.703	7	709	7
Laxman	Ind	21	41.66	10	57.8	22	2.389	24	532	35
Strauss	Eng	29	41.40	11	120.3	15	2.483	16	488	39
Lara	WI	22	41.15	12	158.5	10	2.681	8	646	20
Smith	RSA	14	41.00	13	40.0	27	2.368	26	722	3
Hayden	Aus	49	40.33	14	294.8	4	2.622	13	671	14
Chanderpaul	WI	16	39.73	15	69.3	19	2.444	20	625	21
Martyn	Aus	55	39.67	16	4.2	38	2.192	39	654	19
Ponting	Aus	61	39.52	17	253.3	6	2.478	18	759	1
Trescothick	Eng	53	38.57	18	396.3	2	2.747	6	658	18
Jayawardene	SriL	19	38.33	19	72.8	18	2.409	23	555	29
Symonds	Aus	47	37.75	20	241.0	9	2.596	14	660	17
Dravid	Ind	24	37.04	21	-1.5	40	2.184	40	711	5
Gilchrist	Aus	56	36.96	22	508.0	1	2.945	4	724	2
Jayasuriya	SriL	18	35.83	23	133.3	12	2.663	9	670	15
Gayle	WI	24	35.36	24	10.7	36	2.217	36	717	4
Gibbs	RSA	24	33.3	25	34.6	30	2.306	32	674	12
Y.Singh	Ind	25	32.63	26	82.9	17	2.477	19	546	30
Fleming	NZ	23	32.04	27	38.4	28	2.280	34	670	15
Solanki	Eng	22	31.85	28	49.8	23	2.370	25	490	38
McMillan	NZ	24	31.71	29	14.1	34	2.211	37	578	25
Cairns	NZ	17	31.64	30	58.9	21	2.501	15	588	24
Vaughan	Eng	41	31.64	31	0.7	39	2.198	38	569	26
Boucher	RSA	20	30.5	32	45.1	24	2.434	21	542	32
Collingwood	Eng	37	30.34	33	95.4	16	2.431	22	512	36
Sehwag	Ind	22	28.81	34	41.6	25	2.332	28	618	22
Atapattu	SriL	19	28.63	35	-24.3	43	2.043	43	675	11
Katich	Aus	14	28.54	36	37.9	29	2.480	17	348	46
JonesG	Eng	21	28.43	37	21.4	31	2.325	29	419	43
Astle	NZ	15	27.46	38	-54.8	47	1.921	46	569	27
Hogg	Aus	22	26.25	39	-44.8	45	1.928	45	407	44
Ganguly	Ind	21	25.85	40	-37.6	44	2.024	44	557	28
Harris	NZ	16	24.58	41	-50.4	46	1.844	47	426	42
Styris	NZ	15	24.15	42	-9.7	42	2.063	42	535	33
McCullum	NZ	14	24.08	43	-6.6	41	2.120	41	450	41
Powell	WI	19	23.93	44	20.0	32	2.315	30	503	37
Pollock	RSA	17	21.90	45	5.9	37	2.271	35	472	40
LeeB	Aus	23	18.93	46	13.9	35	2.287	33	346	47
Vettori	NZ	16	15.61	47	16.2	33	2.359	27	368	45

* The LG ICC ranking displayed is for the 47 players of this table having excluded players from the LG ICC lists not qualifying by the criteria of this paper.

Players are ranked in accordance with their traditional batting averages, but their ranks by the two measures suggested by this paper (called the D/L-based measures) are included for

comparative purposes. Beaudoin and Swartz (2003) RM measures are a multiple of 100 of the Batting Resource Percentage Averages

Table 1 also provides the number of innings in which each player batted, together with the LG ICC rating applicable to 31 May 2005. This was the date of their publication as near as possible to the final match of the final series included in the database, that of the NatWest Series of 2005 which concluded on 3rd July 2005.
Several features emerge from Table 1. One is the variation in the rankings of the players under the various methods of measuring batting performances. It is left to the reader to compare and contrast the rankings of players of interest.

Batting correlations

Table 2 summarises the correlations of the players' batting performance measures and with the number of innings in which they have contributed. These correlations are obtained from all of the 47 batsmen who have batted in more than 12 innings and have an LG ICC rating.

Table 2: Correlations of player's batting performances and number innings

Batting Correlations	*Batting Innings*	*Batting avge*	*Batting contrib*	*Batting res% avge*	*LG ICC rating*
Batting Innings	1				
Batting avge	0.004	1			
Batting contrib	0.590	0.479	1		
Batting res% avge	0.169	0.746	0.810	1	
LG ICC rating	0.400	0.437	0.499	0.358	1

The correlations between the three performance measures are moderate yet significant, the strongest being between the D/L based measures, as might be expected, and the weakest between standard batting average and net contribution. Beaudoin and Swartz (2003) found a correlation of 0.67 between batting average and their RM measure, which is not inconsistent with the 0.746 for this study.

The correlations of the performance measures with the LG ICC ratings are modest in all cases but slightly stronger for net batting contribution. It will be seen that the correlation of the number of innings with the standard batting average, 0.004, is insignificant, which probably justifies its use as a long term measure of evaluation of the particular aspect of batting that it captures. The number of innings by batsmen correlates weakly, at 0.169, with batting resource average but much more strongly, at 0.590, with net contribution. This

is suggesting that the more that batsmen play the greater the effect on their net contribution. This implies that the net contributions of the better batsmen will generally increase over time, and those of weaker batsmen decrease.

Table 3: Summary of ODI bowling performance measures

Player	Coun-try	Bowl-ing innings	Bowl-ing avge	Bowl-ing avge rank	Bowl-ing Contrib	Bowl-ing Contrib rank	Bowl-ing res% avge	Res % avge rank	LG ICC rating	LG ICC rank*
McGrath	Aus	34	19.30	1	388.8	1	1.686	1	865	1
LeeB	Aus	48	22.32	2	246.3	2	2.015	3	819	3
Kasprowicz	Aus	13	22.50	3	69.8	6	1.985	2	632	12
Ntini	RSA	22	22.82	4	90.8	4	2.222	8	757	6
Flintoff	Eng	31	24.50	5	55.7	7	2.163	5	727	8
Styris	NZ	14	25.53	6	-1.8	15	2.294	12	629	13
Khan	Ind	13	25.75	7	-63.5	28	2.509	20	625	16
Anderson	Eng	29	25.95	8	24.9	11	2.243	9	652	11
Lehmann	Aus	13	26.53	9	-37.9	19	2.592	24	436	30
Agarkar	Ind	17	26.96	10	-36.4	18	2.395	17	596	20
Gough	Eng	42	28.01	11	79.7	5	2.178	6	732	7
Harmison	Eng	27	28.27	12	-3.7	16	2.297	13	618	18
Cairns	NZ	18	29.45	13	-14.2	17	2.365	16	624	17
Gayle	WI	18	29.50	14	14.7	13	2.412	18	525	24
Hogg	Aus	37	29.64	15	18.2	12	2.249	10	627	15
Pollock	RSA	23	29.92	16	136.6	3	2.020	4	805	4
Bravo	WI	13	30.15	17	32.6	9	2.302	14	452	29
Vaas	SriL	18	31.33	18	-38.3	20	2.702	29	831	2
Clarke	Aus	33	32.61	19	-89.6	31	2.677	28	281	35
Boje	RSA	15	32.73	20	12.8	14	2.354	15	560	22
Gillespie	Aus	39	33.24	21	-133.7	34	2.503	19	786	5
Pathan	Ind	13	33.68	22	-60.7	26	2.520	21	629	14
Watson	Aus	17	34.11	23	-81.4	29	2.660	25	370	33
Kallis	RSA	19	34.85	24	-60.5	25	2.712	30	470	28
Vettori	NZ	25	35.36	25	26.4	10	2.263	11	722	9
Symonds	Aus	45	37.42	26	-164.6	35	2.528	23	515	26
Collingwood	Eng	31	38.83	27	-132.2	33	2.672	27	393	31
Nehra	Ind	13	40.61	28	-104.4	32	2.819	34	607	19
Ganguly	Ind	14	40.66	29	-63.2	27	2.733	32	314	34
Giles	Eng	31	43.03	30	41.5	8	2.196	7	675	10
Sehwag	Ind	14	45.50	31	-53.1	23	2.671	26	390	32
Vaughan	Eng	16	46.00	32	-57.5	24	2.71	31	279	36
Jayasuriya	SriL	18	48.30	33	-84.4	30	2.815	33	553	23
Harris	NZ	17	49.54	34	-46.3	22	2.525	22	474	27
Dillon	WI	15	57.83	35	-40.0	21	2.897	35	594	21
Kumble	Ind	16	59.33	36	-168.4	36	2.968	36	519	25

* The LG ICC rankings quoted are for the 36 players of this table, suitably adjusted for ODI players not qualifying for the table.

As a consequence, net contributions do not stabilise in order to be used to compare performances in the long run, nor when batsmen have different amounts of input. This characteristic implies that net contribution as a long-term measure of batting performance would be appropriate for comparisons of contributions over players' careers rather than for comparing players with differing volumes of input. On the other hand, the resource percentage average would appear to offer such facilities, once the effects of small amounts of input, as outlined in Lewis (2004, 2005), have been eliminated.

There would likely be no disagreement in the conclusion that the best bowler within the study of this paper is McGrath (Aus); there is a clear difference in all of his performance measures from the next-best bowler. Furthermore, there has been little change in McGrath's consistency since the Beaudoin and Swartz study, which found for him a bowling average of 20.1, an economy rate of 3.79 and an RM of 164.0.

Elsewhere in Table 3, there is arguably more consistency in the rankings than with the batsmen of Table 1. Again, the reader can compare and contrast various bowlers of interest.

Table 4: Correlations of player's bowling performances and number innings

Bowling Correlations	*Bowling innings*	*Bowling avge*	*Bowling Contrib*	*Bowling res% avge*	*LG ICC ratings*
Bowling innings	1				
Bowling avge	-0.272	1			
Bowling Contrib	0.300	-0.580	1		
Bowling res% avge	-0.422	0.746	-0.855	1	
LG ICC ratings	0.375	-0.455	0.595	-0.628	1

Bowling Correlations

Table 4 summarises the correlations between the bowling performance measures and the number of innings played that produced them.

Bowling average and Bowling resource (percentage) average are quite highly correlated (0.746) and in Table 3 it can be seen that there is only one name different in the top 12 of these two lists, with close agreement in their rankings. Beaudoin and Swartz (2003) found an even higher correlation of 0.93 but used only the top 12 bowlers active at the time.

In slight contrast to the situation relating to batting, the bowling contribution is only weakly correlated (0.300) with number of innings in which the bowler contributes. There is only a slightly higher correlation (in absolute terms) between bowling average and number of innings, and less than the correlation of resource percentage average with bowling innings. Comments similar to those in relation to batsmen are therefore appropriate on the applicability of net contribution for comparing bowlers in the longer term.

All-round performances

As the author has previously described (Lewis 2004, 2005), the net batting and net bowling contributions can be aggregated to provide a valid mechanism for comparing and aggregating batting and bowling performances. This was shown to provide a mechanism for the comparison of batsmen and of bowlers, and in determining all-round performance in single matches and over a series of matches. Beaudoin and Swartz (2003) suggest that their RM measures can be subtracted to yield an RM difference that assesses all-round ability.

Table 5 summarises the several suggested all-round measures using net contribution, RM difference and the-recently-introduced LG ICC all-round rankings. Table 5 provides the all-round information on the 18 players in this study who had batted and bowled in more than 12 innings and whose performances were sufficiently current that they retained an LG ICC rating for both batting and bowling. The ranking of these qualifying players is provided according to their RM difference over the 16 one-day series considered in this paper

Table 5: Summary of 'all-round' performance measures for qualifying players

Player	Country	Total contrib	Total contrib rank	Batting RM	Bowling RM	RM Diff	RM Diff rank	LG ICC rank
Flintoff	Eng	365.4	1	299.8	216.3	83.5	1	1
LeeB	Aus	260.2	2	228.7	201.5	27.2	2	
Pollock	RSA	142.5	4	227.1	202.0	25.1	3	2
Lehmann	Aus	88.6	5	283.6	259.2	24.4	4	
Cairns	NZ	44.7	8	250.1	236.5	13.7	5	
Vettori	NZ	42.6	9	235.9	226.3	9.6	6	
Symonds	Aus	76.3	6	259.6	252.8	6.8	7	
Clarke	Aus	154.3	3	264.9	267.7	-2.7	8	
Jayasuriya	SriL	48.9	7	266.3	281.5	-15.2	9	
Gayle	WI	25.4	10	221.7	241.2	-19.6	10	4
Styris	NZ	-11.5	11	206.3	229.4	-23.1	11	
Collingwood	Eng	-36.8	15	243.1	267.2	-24.1	12	
Hogg	Aus	-26.6	14	192.8	224.9	-32.2	13	
Sehwag	Ind	-11.5	12	233.2	267.1	-33.8	14	
Kallis	RSA	-19.9	13	230.6	271.2	-40.6	15	
Vaughan	Eng	-56.8	16	219.8	271.4	-51.6	16	
Harris	NZ	-96.7	17	184.4	252.5	-68.1	17	
Ganguly	Ind	-100.9	18	202.4	273.3	-70.9	18	

Again, several interesting aspects of their rankings are raised. By a considerable margin, Flintoff is rated the best all-rounder by all of the measures.

As all bowlers get some opportunities to bat, in the modern game it is becoming more and more important for bowlers to improve their batting skills. This is not only to score runs, but also towards the end of the innings to support the other batsman who may be a higher-order player capable of scoring many runs and guiding the team to victory. Lee would be

regarded as a specialist bowler but he has acquired batting expertise of the nature mentioned and exhibited in the VB series of 2003 (Lewis 2004, 2005) and in subsequent series. Pollock has also acquired a similar reputation.

TOWARDS DUCKWORTH/LEWIS PLAYER PERFORMANCE INDICES

This paper now proposes indices of player-performance.

The DL Batting or Bowling Index

The LG ICC ratings of teams, introduced in 2002 [http://www.cricketratings.com/ accessed 5 Dec 2005], rank the major cricket playing countries using an index based upon 100 as the 'average'. Since such indices are now familiar to cricket enthusiasts, equivalent performance measures for players are now proposed.

In Lewis (2004, 2005) a player's resource percentage average for a match is defined as $\sum s_i / \sum p_i$ where the s_i are the runs scored on the i^{th} ball remaining. The p_i are the resources consumed for the i^{th} ball remaining including the resource percentage consumed for the loss of the wicket at the terminal ball. These are then summed over the balls received during his innings. The resource percentage average P_a for the batsman's career to date would be $P_a = \sum_j \sum_i s_{ij} / \sum_j \sum_i p_{ij}$ where s_{ij} now represents the runs scored from the i^{th} ball remaining in the j^{th} innings in his career. There would be a similar quotient P_b for a bowler where the s_{ij} of runs conceded, would be replaced by $s_{ij} + h_{ij}$ which includes the number of extras, that is, wides and no balls h_{ij}, which are charged against the bowler. The p_{ij} are the resource percentages contributed by the bowling of balls and the taking of wickets.

The average value of $P_{(.)}$ is *Z(50,0)/100,* namely 2.35 in the Standard Edition of the D/L model (Duckworth and Lewis, 2004) that is used in the analysis and applicable for the majority of the series of matches identified above.

The proposed batting index I_a is defined as $I_a = 100 P_a / Z(50,0)$ and expressed as a percentage; the higher the figure the better the evaluation of the batsman. For a batsman such as Lehman (Aus) from Table 1, his batting index, I_a, would be 100*2.836/2.35 = 120.7. This would be interpreted as assessing Lehman as a batsman performing at 20.7% above that of an average ODI batsman.

The problem of the equivalent bowler's index is that good performance would generate a low index. In order to preserve the nature of high indices representing good performance and high ranking, the proposal for the bowler's index is $I_b = Z(50,0) / P_b$.

As an example, Lehmann from Table 3 would have a bowling index of 235/2.592 = 90.7. His bowling performances would be 90.7% of an ODI average bowler, which is 9.3% below average.

All round Index

An all-round index I_r could be defined as the geometric mean of the batting and bowling indices so that $I_r = \sqrt{I_a * I_b}$. For Lehman, his all-round index would have the value of $\sqrt{120.7 * 90.7} = 104.6$ representing an all-round performance 4.6% above average.

It will be seen that in I_r, the term $Z(50,0)$ is eliminated so that the index is merely $100\sqrt{P_a / P_b}$. The removal of the D/L reference point means that an index of, say, 104.6 may have arisen from many combinations of batting resource percentage average P_a and bowling resource percentage average P_b. For Lehman these were 2.836 and 2.592 respectively. Another combination of say 1.845 and 1.686 respectively would yield the same all-round index of 104.6 but with a much stronger bowling than batting performance. Far from being a disadvantage to the proposed index it is suggested that the endless combinations in obtaining a given value of the index would be appreciated by *aficionados* of cricket statistics as providing quantitative support to discussion on who is, or was, the best all-round cricketer!

Although the geometric mean is a recognised mechanism for the averaging of indices, there is arguably a very practical cricketing interpretation to its square; or simply the product of their batting and bowling indices and re-expressed as a percentage. This interpretation would be that of the compound effect of the player's performance in the two disciplines. For Lehman, his all-round index would then be 109.4 providing a compounded all-round performance measure 9.4% above that of an 'average' ODI cricketer.

Table 6 summarises the various indices of all the players in this study who had contributed with bat and ball in more than 12 one-day innings and held LG ICC ratings that were current for both of cricket's main disciplines. These players' ratings and rankings include their DL, RM and LG ICC ratings for batting *and* bowling, and their rankings as all-rounders. Clearly, for ranking purposes, there is no difference between using the geometric mean or its square. Table 6, however, records the square of the geometric mean as the DL all-round index.

The LG ICC all-round rankings are also based on the product of the player's LG ICC ratings. The website [http://www.cricketratings.com/ accessed 7 April 2006] ranks just the top five players deemed to be international all-rounders.

The rankings of batsmen by the DL index are, of course, identical to that using the batting RM since the former is constant divisor of the latter. The bowling rankings by the DL index are slightly different from the RM bowling rankings due to the reciprocal nature of the two contributing measures.

The DL all-round index provides identical rankings to the RM difference for the top ten players. Thereafter there are minor differences in the rankings of no more than one position

Table 6: DL indices for batting, bowling and all-round performance measures for qualifying players

Player	Country	RM diff	RM diff rank	LG-ICC Batting rating	LG ICC Bowling rating	LG ICC All-round rank	DL Batting Index	DL Bowling Index	DL All-round Index	DL All round rank
Flintoff	Eng	83.5	1	711	727	1	127.6	108.7	138.6	1
LeeB	Aus	27.2	2	346	819		97.3	116.6	113.5	2
Pollock	RSA	25.1	3	472	805	2	96.6	116.3	112.4	3
Lehmann	Aus	24.4	4	544	436		120.7	90.7	109.4	4
Cairns	NZ	13.7	5	588	624		106.4	99.4	105.8	5
Vettori	NZ	9.6	6	368	722		100.4	103.8	104.3	6
Symonds	Aus	6.8	7	660	515		110.4	93.0	102.7	7
Clarke	Aus	-2.7	8	705	281		112.7	87.8	99.0	8
Jayasuriya	SriL	-15.2	9	670	553		113.3	83.5	94.6	9
Gayle	WI	-19.6	10	717	525	4	94.3	97.4	91.9	10
Collingwood	Eng	-24.1	12	512	393		103.4	88.0	91.0	11
Styris	NZ	-23.1	11	535	629		87.8	102.5	89.9	12
Sehwag	Ind	-33.8	14	618	390		99.3	88.0	87.3	13
Hogg	Aus	-32.2	13	407	627		82.0	104.5	85.7	14
Kallis	RSA	-40.6	15	708	470		98.1	86.6	85.0	15
Vaughan	Eng	-51.6	16	569	279		93.5	86.6	81.0	16
Ganguly	Ind	-70.9	18	557	314		86.1	86.0	74.1	17
Harris	NZ	-68.1	17	426	474		78.5	93.1	73.0	18

It is felt, however, that Flintoff's clear first-place ranking, with an index of 138.6, would have more meaning in terms of a player whose combined performances rate him 38.6% above average than the equivalent RM difference of 83.5 runs. To rate him as scoring on average 83.5 runs per match more than the opposition were he to do all of the batting and bowling himself, might stretch the abilities of interpreting this statistic rather more than the proposed DL all-round index.

The LG ICC rankings provided in Table 6, as applicable on 31/05/05, are consistent amongst themselves but omit two all-rounders from Pakistan who have not qualified for inclusion in the analysis of this paper due to insufficient contributions with either bat or ball or both. The LG ICC ratings are also unclear on the criteria by which a player is deemed an all-rounder for consideration within the rankings. Indeed the site only ranks five players suggesting a dearth of players purporting to be international all-rounders. Using the criteria above, the style of which could be easily applied in the LG ICC all-rounder rankings, there is no need for a subjective judgement as to whether the player is deemed to be an all-rounder. A substantial proportion of batsmen do not bowl and would therefore not meet the minimum bowling contribution requirement for inclusion. But all bowlers will be required to bat at some stage. Subject to the minimum participation requirement, bowlers will thus acquire a batting index and hence an all-round index in due course. It would then be ability in both disciplines of cricket that would indicate, or otherwise, that a bowler could be regarded as an all-round player. Subjective judgment in this respect would therefore be removed. For example, Lee (Aus) is rated highly by all measures in Table 3 as a bowler.

From Table 1, he has also made a small but positive net contribution of 13.9 runs with the bat even though his batting average is below 20 and his DL batting index is just below 100. In this study his combined performance has placed him in second place of the all-rounder list of Table 7. Compared with countryman Lehman, who is a better batsman but a poorer bowler, Lee has performed slightly better overall.

FURTHER STUDY

The LG ICC ratings and rankings are designed to assess current form and contributions. As a consequence they have an inbuilt discounting procedure that weights recent performances greater than performances earlier in player's careers and also depreciates players' contributions more quickly when they are not selected, or are not available possibly through injury. The LG ICC website, previously cited, provides an overview of the process of evaluating players but does not provide the detail of their calculations due to commercial confidentiality.

In order to compare the D/L-based performance measures with the LG ICC ratings, it will be desirable to define a mechanism that discounts performances over time. This will be the focus for continuing research toward fairer measures of player performance in one-day cricket

SUMMARY

This paper has looked at whether the measures of player performance called net contribution and resource percentage average for batsmen and bowlers are viable for the purposes of evaluating ODI cricket players in the longer term and over players' careers. Through the analysis of many series of matches covering a three-year time span and involving several of the world's regular and leading players it has been shown that there are limitations to the net contribution measure. Whereas this measure was shown to be very effective in evaluating performances in single matches and for a series of matches with a major advantage of being able to assess all-round ability, the analysis in this paper reveals that the measure is unstable and depends very much on frequency of appearance. The net batting contribution measure in particular is shown to depend clearly on the quantity of cricket that the batsman plays. Data show that this is a problem also with the net bowling contribution but is not quite as severe.

On the contrary, whereas the resource percentage average was shown to be unstable when applied in small numbers of matches, it becomes a more reliable and stable measure over the longer term. But there is a problem over interpretation of resource percentage averages. Beaudoin and Swartz (2003) solve this by multiplying by 100 and calling the result Runs per Match that they interpret as the average performance of the batsman, or bowler, were he to do all of the batting, or bowling, himself.

This paper proposes an index based on the ODI average of 235 runs per match giving 2.35 as the average runs per percentage of resource. As international country rankings are given in index form, it is suggested that the cricketing public might more easily accept this measure of performance.

The all-round measure of performance suggested in this paper is the product of a player's batting and bowling indices, which has the virtue of reflecting the compounding effect of performance in the two disciplines rather than through the geometric mean as a measure of an average all-round performance. Beaudoin and Swartz suggest an RM difference with a 'hard to swallow' interpretation of the average performance of the player were he to do all of the batting and bowling himself.

Comparisons of rankings have been made with the LG ICC ratings, and show reasonable comparability for the players qualifying for the analysis summarised in this paper. There are major differences in the methodology of these ratings in comparison with the methods proposed in this paper. One is the limited information in a standard scorecard that is used to create the LG ICC ratings. A second difference is the use of a discounting factor in the LG ICC ratings that attempt to assess current rather than career levels of performance.

Further work on this aspect is suggested if the methodology of the measures proposed in this and the author's previous papers (Lewis, 2004, 2005) were to be either incorporated into the LG ICC ratings or to create a separate set of DL performance measures.

Acknowledgements
The author would like to acknowledge, with thanks, assistance received in preparing this paper: colleague Frank Duckworth for valuable comments on the draft manuscripts; and, Darren O'Shaughnessy and his colleagues at Champion Data (Australia), and also the staff of Press Association, UK (Sports) for the provision, in electronic form, of the ball-by-ball details of the matches absolutely essential for undertaking the analysis summarised in this paper.

REFERENCES

Beaudoin, D. and Swartz T.B., (2003). The best batsmen and bowlers in one-day cricket. *South African Statistical Journal* **37,** 203-222

Duckworth, F.C. and Lewis, A.J., (1998). A fair method for resetting the target in interrupted one-day cricket matches. *Journal of the Operational Research Society* **29**, 220-227.

Duckworth, F.C. and Lewis, A.J., (2004). A successful Operational Research intervention in one-day cricket. *Journal of the Operational Research Society* **55**: 749-759.

Engels, M., (2006). *Wisden Cricketers' Almanack.* John Wisden & Co. Ltd.: Hampshire

Lewis, A.J., (2004). Steps towards fairer measures of performance in one-day cricket. In: Morton, R.H and Ganesalingham, S., (Eds), *Proceedings of 7th Conference Mathematics and Computers in Sport,* Palmerston North, New Zealand.198-213.

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

A MATHEMATICAL MODELLING APPROACH TO ONE-DAY CRICKET BATTING ORDERS

Ovens, M.[1] and Bukiet, B.[2]

[1] School of Mathematical Sciences, Faculty of Science, Monash University, Australia

[2] Center for Applied Mathematics and Statistics, Department of Mathematical Sciences

New Jersey Institute of Technology

ABSTRACT

While scoring strategies and player performance in cricket have been studied, there has been little published work about the influence of batting order with respect to One-Day cricket. We apply a mathematical modelling approach to compute efficiently the expected performance (runs distribution) of a cricket batting order in an innings. Among other applications, our method enables one to solve for the probability of one team beating another or to find the optimal batting order for a set of 11 players. The influence of defence and bowling ability can be taken into account in a straightforward manner. In this presentation, we outline how we develop our Markov Chain approach to studying the progress of runs for a batting order of non-identical players along the lines of work in baseball modelling by Bukiet et al. (1997). We describe the issues that arise in applying such methods to cricket, discuss ideas for addressing these difficulties and note limitations on modelling batting order for One-Day cricket. By performing our analysis on a selected subset of the possible batting orders, we apply the model to quantify the influence of batting order in a game of One Day cricket using available real-world data for current players.

INTRODUCTION

Many cricket commentators will suggest that certain players perform best as "Number Three" in the batting line-up. Listen to almost any commentary team during the course of a One-Day game and you will hear statements based upon rules of thumb like, "Ricky Ponting is a genuine number three." This raises the question, is there really a "Batting Order Effect" and assuming there is, how would you test this? Suppose that a cricket coach decided to test every possible batting order for a team of 11 players, how many games would they have to play? With a team of 11 players, there are nearly 40 million possible line-ups, thus if they could play 1 game every day, it would take a little more than 109286 years (assuming players lived and could play for that long). However, if one has data for the ability of each of the batsmen in a cricket lineup, one can apply techniques of mathematical modelling to ascertain how well a set batting order (against a specified set of bowlers) should perform.

PREVIOUS RESEARCH

There has been little published work about the influence of batting order with respect to One-Day cricket. The earliest work on mathematical analysis of cricket was by Wood (1945) and Elderton (1945) who studied whether cricketers' scores follow geometric progressions. Starting in the late 1980's and continuing up to the present, Clarke's and

his group (e.g. Clarke (1988), Johnston et al. (1993) and Norman and Clarke (2004)) have studied cricket scoring strategies and player performance, among other cricket issues, by applying dynamic programming methods. Clarke (1988) addressed cricket strategies in terms of optimal run rates using dynamic programming techniques. Johnston et al. (1993) assessed player performance by dynamic programming and developed a ranking system to aid in assessing a player's performance. Cohen (2002) studied the probability of dismissing a team before 50 overs and the geometric nature of scoring strokes. Norman and Clarke (2004) investigated the effect of a sticky wicket and how a batting team should adjust its line-up in the longer form of the game. Although Swartz et al. (2006) did study the problem of finding optimal cricket batting orders; their work has been from a statistical simulation, rather than a mathematical modelling perspective. Swartz, et al. (2006) devised a statistical method to compute the probability of each outcome for each batsman with corrections based on wickets and balls remaining. They then simulated games with various batting orders (10,000 games per batting order) and used simulated annealing to reduce number of choices of batting orders to consider.

In the current work, we apply a mathematical modelling approach to compute efficiently the runs distribution of a cricket batting order in an innings. The approach, which will be described in the following section, uses Markov Chains and is based on the method developed by Bukiet et al (1997) that has been used over a number of years for the modelling of the run production in baseball. Among other applications, our method enables one to solve for the expected number of runs a batting order should score and the probability of one team beating another. By considering all 11! or 39,916,800 batting orders, one could potentially find the optimal batting order for a set of 11 players, i.e., the batting order that can be expected to attain the most runs. The model is set up such that the influence of defence and bowling ability can be taken into account in a straightforward manner. As we note in later sections of this paper, the time it takes to analyse a single lineup using our technique is such that evaluating all possible lineups (full enumeration) would take a prohibitively long amount of time. Thus, presently, the model is mainly for theoretical purposes in terms of finding *the* best batting order, but it could be applied to address some questions. For example, by considering a selected subset of the possible batting orders for 11 Australian players and applying our model, we demonstrate that there is a difference between best and worst batting orders and that this difference is significant.

METHODS

In this paper, we outline how we develop our Markov Chain approach to studying the progress of runs for a batting order of non-identical players along the lines of work in baseball modelling by Bukiet, et al. (1997). We describe the issues that arise in applying such methods to cricket and how we have addressed the difficulties particular to cricket.

In a Markov process, it is not important to know how a given situation arose, just that you are in a particular situation. The probability of going from one situation to any other is known. There are a finite number of situations.

In the context of cricket, the dynamics of run production depends mainly on the interaction of the bowler and the batsman. So the game can be modelled as a sequence

of one-on-one interactions. A batsman takes a turn and then we stop and have a new situation. The probability of any occurrence depends only on the current situation (who is the facing batsman, who is the bowler, who is the batsman at the other wicket (the non-striker), how many balls are left) and possibly only a small subset of that. For the most part, there are only 7 states to which a given situation will commonly transition on a single bowl of the ball.

- A batsman is dismissed and no runs score with probability P_d
- Zero runs score (no dismissal) with probability P_0
- One run is scored (no dismissal) with probability P_1
- Two runs are scored (no dismissal) with probability P_2
- Three runs are scored (no dismissal) with probability P_3
- Four runs are scored (no dismissal) with probability P_4
- Six runs are scored (no dismissal) with probability P_6

Making the situation slightly more complex is the effect of the batsman switching places. If an odd number of runs are scored, the batsman will have switched ends. Similarly, if a multiple of six balls have been bowled, the bowler has finished the over and a new bowler will begin bowling from the other end, resulting in the non-striker becoming the facing batsman. (We note that it is possible, but uncommon to score 5 runs. Similarly, it is possible to score runs when a batsman is runout. We disregard these events, other rare offensive possibilities as well as rules concerning fielding restrictions. Our method could handle most of these at the cost of greatly increased computational time. It appears, at least in the case of modelling baseball that ignoring many rare situations makes little difference in the results as there is much cancellation between including positive events (e.g. fives) and negative events (e.g., runouts on run scoring balls)).

Let the multidimensional Matrix M have entries $M(b,r,w,b_1,b_2)$ represent the number of balls bowled, runs scored and wickets down, the next batsman and the batsman at the other wicket, respectively. For each number of balls bowled, we can compute the probability of being in a given situation by multiplying the (multidimensional) matrix representing the set of probabilities after the b-1 balls by the probability of each of the events listed above occurring. For example, the game begins with 0 balls bowled, 0 runs scored, 0 wickets gone and batsman number 1 about to hit, with batsman number 2 at the other wicket. Thus, $M(0,0,0,1,2)=1$ and all other entries of $M(0,r,w,b_1,b_2)$ are zero. After the first ball, $M(1,0,1,3,2) = P_d$, $M(1,0,0,1,2) = P_0$, $M(1,1,0,2,1)=P_1$, $M(1,2,0,1,2) = P_2$ and so forth. These values are obtained in the general case (b balls) by multiplying each non-zero entry of $M(b-1,r,w, b_1,b_2)$ by each of the probabilities P_d, P_0-P_6 and placing the result in the appropriate location in the $M(b,r,w,b_1,b_2)$. After the computation has considered 300 balls (with 10 wickets down causing no future balls to be bowled) we end up with the probability of any given number of runs having been scored (the runs distribution). The computation is actually simplified by looping through the number of balls and saving only the situation after b-1 balls to compute the situation after b balls. Also, one need not keep track of wickets dismissed since the batsmen currently in the game provide that information. (If batsman number 6 in the order is in the game, but 7-11 are not, then 4 batsmen have been dismissed). Thus, an 11 X 11 X 1800 (batsmen X batsmen X runs) matrix needs be maintained and updated. We

note that the method automatically takes into account that the batsmen early in the lineup, if they are the best batters will face more balls than the later batsmen. Summing the product of each possible number of runs and its probability of being the result in the game gives the expected number of runs for the batting order considered.

This strategy is the same in philosophy as that of Bukiet, et al (1997) only the details are modified. The strategy involves mathematically only addition, multiplication and some logic. The method makes sense only because cricket has the following properties:

- Dynamics of run production depends mainly on interaction of bowler and batsman so the game can be modelled as a sequence of one-on-one interactions
- There are a finite number of states in the game (batsman, outs, runs)
- The probability of an occurrence depends only on current situation (to a reasonable approximation) (One can also implement run or score dependence).

Some aspects of One-Day Cricket make it more complicated and lengthy to model than for baseball.

- There are 11 players who bat on a team in an innings. Thus there are over 39 million batting orders to consider in a full enumeration.
- Up to 300 balls are bowled in an innings, in groups of 6 (an over). After each over has been completed, a different bowler bowls the ball from the opposite side of the field. (In baseball one can consider the situations as batter by batter and there are only about 50 batters up in a game for each team).
- An odd number of runs scored on a ball results in the other active batsman batting next (unless this is the end of an over).
- Bowlers switch sides at the end of an over and can only bowl a maximum of 10 overs in a match. This makes some of the logic and bookkeeping more complicated.
- Typical one-day cricket matches result in each team scoring 200-300 runs (much more than the 0-10 runs common in baseball, with extreme cases running up to 20 runs for a team). In cricket, in theory a team could get up to 1800 runs, although the present record is 438 runs[1]. This increases run time and storage requirements.
- P_d is very small for most batsmen, typically about 0.03 or less, which means outs are very rare when compared to baseball.
- By comparison to baseball, players play in fewer games per year. As a result, small errors can have a larger effect on an individual player's probabilities.
- When the innings is nearing its end, batsmen begin taking more risks and score at a higher rate. Players who come in to bat with very few overs remaining will usually score at a higher rate than they normally would if they came in earlier in the innings. Only the first batsman has a non-zero probability of facing all 300 balls in a match. This could potentially skew the probability distribution (P_0-P_6) obtained from the data set for a later order batsman.
- For one line-up a simply written code for 1-day cricket takes ~15 seconds on an average PC (up to 600 runs considered). To evaluate 11! ~ 40,000,000 line-ups such a code would take about 1000 days.

[1] ODI #2349 South Africa v Australia at New Wanderers Stadium, Johannesburg on 12 March 2006, South Africa scored 9/438 in reply to Australia's 4/434. Previous record was 5/398 scored by Sri Lanka in ODI #1074 Sri Lanka v Kenya at Asgiriya Stadium, Kandy during the 1995/96 Wills World Cup. Source: CricInfo

The large computational time involved using this straightforward approach (referred to later on as "the straightforward method" makes it unattractive as a planning tool for coaches, however some streamlining improves the performance. Instead of considering each batsman and each ball individually, we consider (in what we call our "streamlined method") each pair of batsmen (11 x 10 pairs) and each over individually (50 overs). As a further simplification, we assume that a maximum of 1 wicket can fall in any given over; the result is about a 1 second improvement in processing time per line-up studied.

To include bowling and defensive performance, one could scale the offensive characteristics in an appropriate way. For example, if a given bowler has performance level, say, 2% worse, than the average bowler, by some measure, then opposing players would have their offensive performance (P_1-P_6) increased by 2% and P_0 decreased accordingly. Ideally, one would like to have enough data on how well each bowler performs against each batsman (and vice versa), but that is not likely to be the case. Another method of scaling batsman performance might take into account his "handedness", that of the bowler, and/or the type of bowler (e.g., a spin bowler) bowling. One of the authors has looked into various methods of considering pitcher ability in baseball and found that considering such complications did not lead to improved results.

RESULTS

Using data gathered from the CricInfo website and that kindly supplied by Champion Data, we were able to find estimates (for various players) of the probability of scoring 0, 1, 2, 3, 4 or 6 runs. Treating being not out at the end of a game as the same as being run out on the last ball, we can find an estimate for the probability of being dismissed. Table 1 shows these estimates for 11 Australian players. These players have enough experience such that the data used takes into account at least 100 balls being bowled to each of the players and each player's performance in at least 15 matches (except for Michael Slater[2] and Glen McGrath[3] as shown in Table 2).

Table 1

P_0	P_1	P_2	P_3	P_4	P_6	P_d	Player
0.513	0.311	0.071	0.006	0.083	0.012	0.027	Symonds
0.561	0.230	0.055	0.013	0.128	0.011	0.027	Gilchrist
0.546	0.297	0.067	0.014	0.070	0.004	0.019	Waugh, M.
0.563	0.273	0.058	0.012	0.080	0.011	0.021	Hayden
0.545	0.295	0.059	0.012	0.072	0.013	0.022	Ponting
0.512	0.294	0.115	0.000	0.038	0.038	0.025	Slater
0.533	0.319	0.060	0.004	0.060	0.020	0.051	Bichel
0.526	0.332	0.071	0.005	0.026	0.032	0.083	Lee
0.515	0.387	0.054	0.018	0.018	0.006	0.109	Gillespie
0.687	0.234	0.046	0.000	0.031	0.000	0.250	McGrath
0.555	0.324	0.072	0.003	0.039	0.004	0.051	Warne

[2] The data collected did not have many matches including Michael Slater, although he has played 42 matches for Australia

[3] Glen McGrath normally is the last batsman in the Australian batting lineup. This limits the amount of data available on his performance.

The probabilities P_0-P_6 above sum to approximately 1 and are the distribution of 0-6 runs scored by the players whilst not being dismissed. The probability of dismissal, P_d, is calculated as the number of innings played divided by the number of balls faced. Whilst this does not account for innings in which the player does not get dismissed, this error in the data is small when a player has played many innings and unlikely to be as large as the error caused by using small data sets (e.g. a single innings). The data from Table 1 can also be used to determine the expected runs if the team were made up of only 1 player occupying all 11 places, which then enables us to rank the players. This ranking, computed when limiting runs to 600, is shown in Table 2.

Table 2

Player	Runs per innings	Rank	Balls	Innings
Symonds	244.04	3	2096	57
Gilchrist	263.89	1	4203	117
Waugh, M	202.15	5	2477	49
Hayden	216.39	4	3401	73
Ponting	177.64	6	5056	112
Slater	254.57	2	78	2
Bichel	148.66	7	444	23
Lee	61.31	9	334	28
Gillespie	53.95	10	165	18
McGrath	13.60	11	64	16
Warne	120.13	8	635	33

Using this ranking, we compute the expected runs using the "straightforward method" when the players are ordered in Best-Worst ranked order (Gilchrist bats first, Slater second, Symonds third and so forth), the given order and Worst-Best ranked order. These expectations are shown in Table 3.

Table 3

Best-Worst Ranked Order	210.66 runs
Given Order	208.75 runs
Worst-Best Ranked Order	194.16 runs

To enable us to consider a greater number of batting orders, we used our "streamlined method" to evaluate about 10,000 line-ups, permuting only the last 8 players. we find that the minimum number of expected runs is approximately 219 compared with a maximum of almost 229. These results are higher than those achieved before the streamlining and is most likely due to the restriction on the number of wickets than can be lost per over. Table 4 shows a comparison of results achieved using the straightforward and streamlined methods. The line-up used in generating tables 1-4 uses players (like Mark Waugh) whom no longer play for Australia, but each player has played at least 10 games at the international level. Suppose, however, that we wish to replace 4 players with current or new players, like Hussey and Hodge. Table 5 shows the estimated probability distributions using the same source data. (Here, each player's data only includes at least 2 games and 50 balls bowled to him, to allow for newer players, as shown in Table 6 except for Stuart MacGill[4].)

[4] Stuart MacGill normally bats after Glen McGrath when both are playing for Australia (see Footnote 2). This also limits the amount of data available on his performance.

Table 4

Player	Runs/Inn Straightforward	Runs/Inn Streamlined
Symonds	244.04	249.29
Gilchrist	263.89	269.91
Waugh, M	202.15	203.06
Hayden	216.39	217.92
Ponting	177.64	178.88
Slater	254.57	258.88
Bichel	148.66	165.98
Lee	61.31	69.63
Gillespie	53.95	68.96
McGrath	13.60	23.68
Warne	120.13	134.19

Table 5

P_0	P_1	P_2	P_3	P_4	P_6	P_d	Name
0.553	0.313	0.048	0.010	0.075	0.003	0.015	JL Langer
0.563	0.273	0.059	0.012	0.080	0.011	0.021	ML Hayden
0.545	0.295	0.060	0.012	0.073	0.013	0.022	RT Ponting
0.412	0.471	0.039	0.039	0.039	0.000	0.039	BJ Hodge
0.522	0.317	0.064	0.016	0.074	0.006	0.022	MEK Hussey
0.514	0.312	0.072	0.007	0.083	0.013	0.027	A Symonds
0.561	0.231	0.055	0.014	0.128	0.011	0.028	AC Gilchrist
0.556	0.324	0.072	0.003	0.039	0.005	0.052	SK Warne
0.527	0.332	0.072	0.006	0.027	0.033	0.084	B Lee
0.714	0.286	0.000	0.000	0.000	0.000	0.286	SCG MacGill
0.688	0.234	0.047	0.000	0.031	0.000	0.250	GD McGrath

Table 6 shows, like Table 2, the expected runs if these players made up the entire line-up (using the streamlined method with a maximum of 1800 runs allowed).

Using this data we compute the expected runs from a convenient subset (163,724 samples at this time) of batting line-ups, namely, those line-ups with the openers and some with the third batsman already determined. Table 7 shows a brief analysis of the data. It is interesting to note that the mean is almost exactly 235, which is also the current value of the G50 constant in the Duckworth/Lewis method for target resetting. Figure 1 shows a histogram of these results. We note that the distribution is unimodal, has small variance, but is highly skewed with large number of outliers. The line-ups that produced the minimum and maximum number of expected runs (among the 163,724 lineups studied) are shown in Table 8.

Table 6

Name	Exp_Runs	Rank	Balls	Innings
AC Gilchrist	281.34	1	4203	117
A Symonds	259.20	2	2096	57
MEK Hussey	242.79	3	312	7
RT Ponting	241.75	4	5056	112
ML Hayden	240.17	5	3401	73
BJ Hodge	239.58	6	51	2
B Lee	232.00	7	334	28
JL Langer	222.43	8	400	6
SK Warne	192.78	9	635	33
GD McGrath	128.37	10	64	16
SCG MacGill	80.06	11	7	2

Table 7

Mean	235.1
Sample Variance	95.2
Mode	209.8
Count	163724
Minimum	187.6
First Quartile	234.1
Median	236.9
Third Quartile	239. 5
Maximum	257.6
Range	70.0
IQR	5.4

Table 8

Batsman	Minimum	Maximum
1	JL Langer	JL Langer
2	ML Hayden	ML Hayden
3	BJ Hodge	A Symonds
4	GD McGrath	AC Gilchrist
5	SK Warne	SK Warne
6	SCG MacGill	MEK Hussey
7	AC Gilchrist	B Lee
8	B Lee	SCG MacGill
9	MEK Hussey	GD McGrath
10	RT Ponting	RT Ponting
11	A Symonds	BJ Hodge

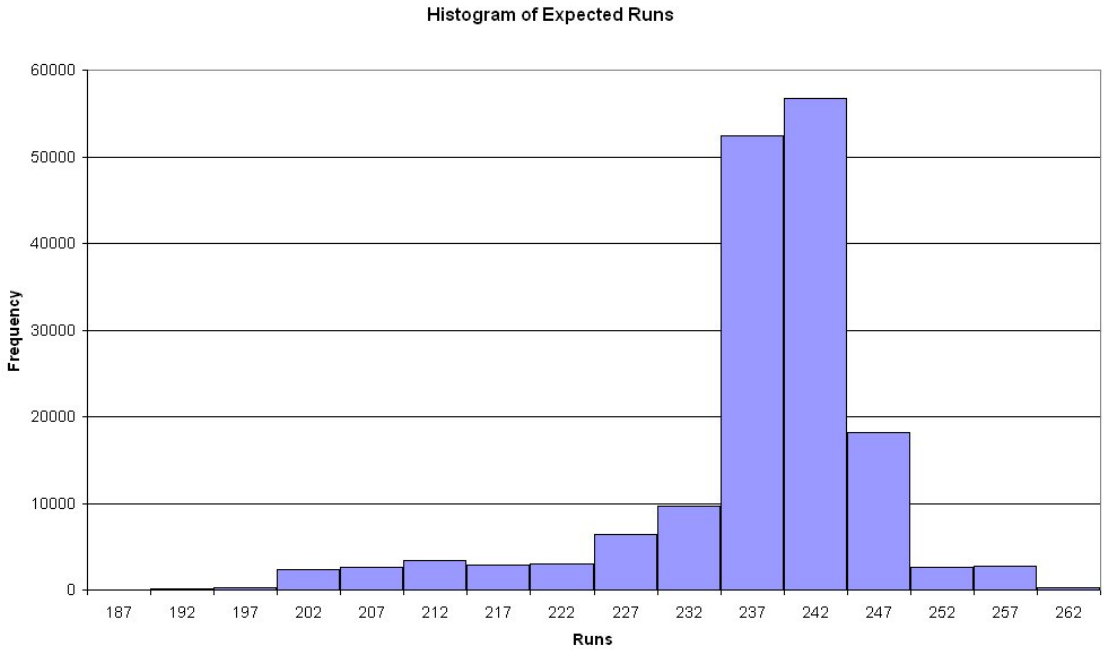


Figure 1: Histogram of expected number of runs for 164,724 line-ups taken from the players listed in Table 6

DISCUSSION AND CONCLUSION

Cricket is a game with many variables affecting the outcome; the weather, the pitch, the players and even the spectators at the match. Attempting to model a cricket match requires reducing these many variables down to a manageable and quantifiable subset. In this paper we have attempted to simplify these many variables down to the batsman's average offensive ability versus every bowler they have faced (within the available data set). We have developed a straightforward and a streamlined approach for evaluating the distribution and thus, the expected number of runs a line-up should produce against average bowling. The large number of possible line-ups makes a "straightforward" approach to finding the optimal batting order virtually impossible to achieve in a reasonable time frame, however smaller subsets could be calculated rather quickly. This means that the most practical use of our work would be in determining the order of three or four batsmen with the rest of the line-up fixed. Our work could also be used to quantify the effect of the "super-sub" under the new laws of the One-Day game.

We find it interesting that the results of our model show a mean expected number of runs scored of almost exactly 235. Since we were not able to study all 40 million line-ups, we have shown for the set of players considered that the best batting order can expect to produce at least 70 runs more than the worst possible line-up. Figure 1 suggests that it is easier to find a very poor line-up than it is to find a very good one. We expect the result of allowing for slight variations in player ability will have a similar effect to such variations in baseball player ability as studied by Sokol (2003). That is, that while our technique will find (given enough computation time) the line-up with the greatest expected number of runs, slight variations in player performance ability would

result in a different line-up being "better". In other words, the best line-up is not robust. However, any of a set of nearly optimal line-ups would be indistinguishable within the limit of accuracy of the input probabilities.

Whilst we have done a large number of calculations and streamlined the code, we see that it is unlikely that the calculations could be finished (on a single computer) within a reasonable timeframe. However, we see other opportunities for future work and applications, for example: 1) We could study optimal batting order and impact of batting order on probability of winning a game. Although we may not perform a full enumeration, there are ideas which would allow us to study likely subsets (e.g. Swartz); 2) The data collection in One-Day Cricket is more difficult than baseball. We would like to obtain more data and more team information, but perhaps we could investigate ways of interpolating using available data; 3) We could expand our model to include bowling ability (defence) as was done for baseball (e.g. Bukiet, 1997); 4) A further extension of the model would be to compute the probability of winning a game and the effect of using one player instead of another (e.g. the "rotation" policy effect).

Acknowledgements

The authors would like to thank Champion Data and Mr David McKenzie for their assistance in gathering the data needed. Additionally, the authors would like to thank the anonymous reviewer for their insightful and helpful comments.

REFERENCES

Bukiet, B., Harold, E. and Palacios, J. (1997) *Operations Research,* **45** (1)**,** 14-23.

Clarke, S. R. (1988) *Journal of Operational Research Society,* **39** (4)**,** 331-337.

Cohen, G. L. (2002) Cricketing chances. In G. Cohen and T. Langtry (Eds). *Proceedings of the 6th Australasian Conference on Mathematics and Computers in Sport*. Bond University, Queensland, Australia, pp. 1-13.

Elderton, W. E. (1945) *Journal of the Royal Statistical Society, Series A,* **108,** 1-11.

Johnston, M. I., Clarke, S. R. and Noble, D. H. (1993) *Asia-Pacific Journal of Operational Research,* **10,** 45-55.

Norman, J. and Clarke, S. R. (2004) Dynamic programming in cricket: batting on a sticky wicket. In H. Morton and S. Ganesalingam (Eds). *Proceedings of the 7th Australasian Conference on Mathematics and Computers in Sport.* Massey University, Palmerston North, New Zealand, pp. 226-232.

Sokol, J. (2003) *Journal of Heuristics,* **9** (4)**,** 353-370.

Swartz, T. B., Gill, P. S., Beaudoin, D. and de Silva, B. M. (2006) *Computers and Operations Research,* **33** (7)**,** 1939-1950.

Wood, G. H. (1945) *Journal of the Royal Statistical Society, Series A,* **108,** 12-22.

NAISMITH'S RULE AND ITS VARIANTS

Norman, J. and Hayes, M.

Sheffield University Management School, Sheffield, England

ABSTRACT

Naismith's Rule, a guide for people walking in mountainous country, suggests allowing 20 minutes for each mile as the crow flies and 30 minutes for each thousand feet of ascent. The rule thus suggests that one foot of ascent is equivalent to 7.92 feet on level ground. In this paper, we review results from various sources to estimate the equivalence of height climb and distance as the crow flies - the "N-value", signifying that one foot of climb is equivalent (in time taken) to N feet on the level. Naismith's N-value is thus 7.92. The sources include results from fell races in the UK, the Catskill mountain road race, the Barr Trail mountain race, treadmill experiments and a paced running event conducted at a constant work rate. The N-values vary from 2.8 for the treadmill to 8.6 for winners of UK fell races. It is suggested that the following factors may give rise to higher N-values: greater difficulty of terrain (decreasing "runability"); increased congestion, making it difficult to pass other runners, especially on ascents; the tendency of competitive runners to work harder on ascents than on the flat. Taken with results of previous work, these conclusions suggest that, if he has a choice, a runner does well to run directly up slopes of less than 1 in N, but to zigzag up slopes greater than that, keeping to a gradient of 1 in N.

KEY WORDS
Naismith's rule, running uphill

INTRODUCTION

Naismith's Rule, a guide for people walking in mountainous country, suggests allowing 20 minutes for each mile as the crow flies and 30 minutes for each thousand feet of ascent. Variations of the Rule have been suggested: an example (Hayes and Norman 1994) is Hayes's rule for runners attempting the arduous Bob Graham round (72 miles and 27,000 feet of ascent in 24 hours). Hayes suggests:

(1) Flat or uphill. Allow 10 minutes per mile + 10 minutes per 1000' of ascent
(2) Gentle downhill (under 500' per mile). Allow eight minutes per mile
(3) Steep downhill (over 500' per mile) Allow 10 minutes per mile + two minutes per 1000' of descent
(4) Rough ground on flat, uphill or gently downhill sections. Allow 15 minutes per mile + 15 minutes per 1000' of ascent
(5) Rough, steep downhill. Allow 10 minutes per mile + 10 minutes per 1000' of descent.

Naismith's rule suggests that one foot of ascent is equivalent to 7.92 feet on level ground. Hayes's rule suggests that one foot of ascent is equivalent to 5.28 feet on the level. Scarf (in press) has analysed the record times for the 300 or so races in the 1994 fell racing calendar and derived formulas for the expected times for men and women shown below:

For men, $E(t) = 2.09(x + 8.6y)^{1.14}$ (1)
and for women $E(t) = 2.14(x + 10.6y)^{1.16}$ (2)
where x is the distance as the crow flies and y the climb, both measured in kilometres, and E(t) the expected time, measured in minutes. The power term is sometimes interpreted as a fatigue factor: essentially, it takes account of the fact that runners run slower in longer races. Scarf calls the term inside the brackets the "equivalent distance". He thus suggests that for men, one foot of ascent is equivalent to 8.6 feet on level ground, while for women, the ratio is 1:10.6.

Other work has used Naismith's idea of a proportional allowance for climb directly, without a power term, in effect using a formula

$$T = (x + Ny)/v \quad (3)$$

where x and y are defined as before, v is the runner's speed (km/minute) and T is the runner's estimated time (minutes). Analysis of treadmill experiments (Davey et al 1994,1995) and the records of a mountain road race (Norman 2004) have resulted in estimates of N = 2.8 for the former and N = 4.4 for the latter. We may wonder why these two values are different and whether the simple approach giving rise to equation (3) is compatible with the sophisticated statistical approach which results in equations (1) and (2). In this note, we analyse results from two running events, the first a mountain race held annually in the United States and the second a paced endurance event which takes place around midsummer in the Lake District in England. The data for the first event have been obtained from the Internet (http://www.runpikespeak.com accessed 24.8.2005).

THE BARR TRAIL MOUNTAIN RACE

This race takes place annually at Pikes Peak National Park in Colorado. It "starts at the Cog Railway Depot at an elevation of 6,570'. The race proceeds 6 miles up the Barr Trail gaining 3,630' where it turns around at Barr Camp at 10,200'. Runners then head back down the Barr Trail to Hydro Street where they turn left and up with the finish being in the pullout to the right for a total of 12 miles". Apart from four short "gentle downhills", the first half of the race is uphill all the way and apart from these and the "nasty little hill where you finish your race at 6,650' ", the second half of the race is downhill almost all the way. We shall lose little in accuracy and gain much in simplicity if we disregard the nasty little hill at the end and regard the race as 6 miles gaining 3,630' followed by the same 6 miles losing 3.630'.

Using equation (3) and defining T_u and T_d as the estimated uphill and downhill times, then we may write

$$T_u = (x + Ny)/v \quad (4)$$

and

$$T_d = x/v \quad (5)$$

so that

$$T_u/T_d = 1 + Ny/x \quad (6)$$

Split times are given for all runners so this ratio may be computed for individuals. In the 2004 race, the 231 male finishers had a mean ratio of 1.70, with a standard deviation of 0.15. Substituting in equation (6) with x = 9.6 and y =1.10 (measured in km) gives a value for N of N = 6.0. A similar calculation for the 95 female finishers results in a mean ratio of 1.73 with a standard deviation of 0.15 resulting in a value for N of N = 6.4.

DISCUSSION

These N-values are higher than those given previously for running on a treadmill and in a mountain road race and it may be helpful to set them out with some notes on the running conditions.

Treadmill running N = 2.8. No wind resistance and some "give" underfoot. Similar experiments (Margaria et al 1938,1963,1976) suggest N = 2.5. Obviously, there is no congestion, but it may be that a different style of running from that used in the open is used. Some writers have questioned whether treadmill training is useful for the serious athlete (http://www.pponline.co.uk/encyc/treadmill.html accessed 24.8.2005)

Road running in the Catskill Mountain Road Race. N = 4.4. Unchanging surface underfoot and no congestion, and thus no difficulty in passing other runners. This is a feature of the Catskill race not normally found in road races, or in races generally.

Trail running in the Barr Trail Mountain Race. N = 6.0. Varied surface underfoot and some congestion. There is some asphalt surface at the start and finish and though much of the surface has been described as "smooth", it is certainly not as smooth as an asphalt road surface..

These N-values are all lower than the value of N = 8, suggested by Scarf for winners of fell races. The notes above, however, suggest that the values of N increase with the severity of the terrain and perhaps with the degree of congestion as runners try to pass each other. As Scarf remarks, "in fell races, steeper paths will in general be in poorer condition than more moderately inclined paths. Consequently, steeper paths may be slower not just as a result of the extra climb but because of the reduced runability".

PACED RUNNING IN THE BOB GRAHAM ROUND

Point (1) of Hayes's rule suggests an N-value of 5.3, but runners attempting the Bob Graham round are generally supported by other runners who pace them. There is virtually no congestion, though some complicating factors arise not found in other events. For example, most successful runners complete the course in little under 24 hours and consequently run some of it in darkness. The data for the second event relate to a successful completion of the Bob Graham round and were derived by M.H.

Hayes acted as a pacer for two fell runners attempting the Bob Graham Round, who were supported by other fell runners from the Dark Peak Fell Runners Club on different sections of the Round. Hayes's section was the 22.5 km stretch in the Hellvellyn range from Dunmell Raise to Threkeld, The two runners, on arriving at Dunmell Raise, had already been running for 13 hours. Accompanied by Hayes, they then ran 22.5km, with ascents totalling 1679 metres and descents totalling 1664 metres. Hayes, an experienced fell runner, knew the route well, and made only one navigational error, losing about 5 minutes between Great Dodd and Calfhow. Hayes wore a heart rate monitor and used it to check that his pace never deviated much from 140 beats per minute, so as to maintain a constant work rate. To quote Hayes's account, "no action was taken if the heart rate was between 135 and 145. If it exceeded 145, pretty strong measures were taken to rein in the leading contestant and get him to slow down. If it

fell below 135, constant encouragement was given to the trailing contestant in an effort to get him to increase his pace".

Hayes timed the contestants at the beginning and end of each of 21 sections of the stage and the four rest stops. He had previously recorded the distance on the ground and the vertical height gains and losses, all in metres, in each of the 21 stages. This formed the basis of a later statistical analysis. In brief, discarding the Great Dodd to Calfhow section and taking no account of descent, estimated time was given by:
Estimated time = .041 +.009*distance + .052*height gain
with an adjusted R^2 = .980. The coefficients of distance and height gain had p-values <.001. The p-value of the constant term was .956. Forcing the regression equation through the origin resulted in the following equation:
Estimated time = .0086*distance + .0522*height gain.
The corresponding N-value =.0522/.0086 = 6.1.

This N-value is again less than the N-value of 8 found by Scarf in his analysis of fell race winners. On the Bob Graham Round section, there was no congestion, but parts were run in darkness. The main identifying characteristic, however, shared only with the treadmill experiments, is that of a constant work rate, measured by heart rate. Perhaps fell runners in a race, competing for the lead, exert themselves more on ascents than on descents and on the level, increasing the time they take over what they might accomplish in even-paced running.

CONCLUSION

In this note, we have reviewed results from several sources to estimate the equivalence of height climb and distance as the crow flies – the N-value, signifying that one foot of climb is equivalent (in time taken) to N feet on the level. Considerable variation in the N-values was noted, ranging from N = 2.8 or the treadmill to N = 8.6 for male winners of fell races. Some possible reasons for the various values have been suggested.

The remark that follows can only be conjectural, though the commentary above shows that there is support for it. The following factors may give rise to higher N-values:

- greater difficulty (more difficult runability) of terrain
- increased congestion, making it difficult to pass other runners, especially on ascents
- the tendency of competitive runners to work harder on ascents than on the level

A final note. A recreational runner may be concerned with only the first of these factors. If the conclusions of Davy et al (1994,1995), based on running on a treadmill, can be applied to running outdoors, then the results of this paper suggest that, if he has a choice, a runner does well to run directly up slopes of less than 1 in N, but to zigzag up slopes greater than that, keeping to a gradient of 1 in N.

REFERENCES

Davey, R., Hayes, M. and Norman, J. (1994) Running uphill: an experimental result and its applications. *Journal of the Operational Research Society* **45**, 25-29

Davey, R., Hayes, M. and Norman, J. (1995) Speed, gradient and workrate in uphill running. *Journal of the Operational Research Society* **46**, 43-49

Hayes, M. and Norman, J. (1994) Strategy in fell running: an analysis of the Bob Graham Round. *Journal of the Operational Research Society* **45** 1123-1130

Margaria, R. (1938) Sulla fisiologia, e specialmente sul consumo energetico, della marcia e della corsa a varie velocita ed inclinazioni del terreno. *Atti Accademia Nazionale dei Lincei Memorie serie VI*, **7**, 299-368

Margaria, R., Cerretelli, P., Aghemo, P. and Sassi, G. (1963) Energy cost of running. *Journal of Applied Physiology* **18**, 367-370

Margaria. R. (1976) *Biomechanics and Energetics of Muscular Exercise*. Clarendon Press, Oxford

Norman, J. (2004) Running uphill: Energy needs and Naismith's rule. *Journal of the Operational Research Society* **55**, 308-311

Scarf, P. (1998) An empirical basis for Naismith's rule. *Mathematics Today* 149-151

Scarf, P. (in press) On the equivalence of distance and climb: Naismith's rule for walkers, runners and cyclists. *Journal of Sports Sciences.*

For the Catskill Mountain Road Race;
http://www.angelfire.com/ny5/corprace/wildbunch.html

For the Barr Trail Mountain Race;
http://www.runpikespeak.com/course.htm

STOCHASTIC DOMINANCE AND ANALYSIS OF ODI BATTING PERFORMANCE: THE INDIAN CRICKET TEAM, 1989-2005

Damodaran, U.

XLRI Jamshedpur School of Management, Jamshedpur, Jharkand, India

ABSTRACT

Relative to other team games, the contribution of individual team members to the overall team performance is more easily quantifiable in cricket. Viewing players as securities and the team as a portfolio, cricket thus lends itself better to the use of analytical methods usually employed in the analysis of securities and portfolios. This paper demonstrates the use of stochastic dominance rules, normally used in investment management, to analyze the One Day International (ODI) batting performance of Indian cricketers. The data used span the years 1989 to 2005.

In dealing with cricketing data the existence of 'not out' scores poses a problem while processing the data. In this paper, using a Bayesian approach, the 'not-out' scores are first replaced with a conditional average. The conditional average that is used represents an estimate of the score that the player would have gone on to score, if the 'not out' innings had been completed.

The data thus treated are then used in the stochastic dominance analysis. To use stochastic dominance rules we need to characterize the 'utility' of a batsman. The first derivative of the utility function, with respect to runs scored, of an ODI batsman can safely be assumed to be positive (more runs scored are preferred to less). However, the second derivative need not be negative (no diminishing marginal utility for runs scored). This means that we cannot clearly specify whether the value attached to an additional run scored is lesser at higher levels of scores. Because of this, only first-order stochastic dominance is used to analyze the performance of the players under consideration. While this has its limitation (specifically, we cannot arrive at a complete utility value for each batsman), the approach does well in describing player performance. Moreover, the results have intuitive appeal.

KEY WORDS

Bayesian, utility function, batting average, conditional average, geometric distribution

INTRODUCTION

As a game, cricket is a statistician's delight. Each game of cricket throws up a huge amount of performance related statistics. As other games have evolved and developed, they too have become richer in the use of performance statistics. For example, use of statistics like 'unforced errors' in lawn tennis or 'assists' in basketball is increasingly becoming popular. However in cricket these statistics have *always* been part and parcel of the game. Cricket is one of the few games in which a 'scorer' is required to continuously maintain statistical

data on key game/player-specific performance statistics. It is one of the few games that have detailed 'scoring sheets'. These scoring sheets were maintained manually in the pre-digital age and are maintained electronically today.

In spite of this legacy and long history of maintaining statistical data, two aspects associated with cricketing data are striking. The first is the idiosyncrasy that has persisted in the treatment of the 'not out' scores of a player. The second is the lack of effort in exploiting the richness of data to improve the representation of player performance.

The batting average of player *i*, $\overline{R}_i$, is computed as:

$$\overline{R}_i = \sum_{t=1}^{n} Rit / (n-k) \qquad (1)$$

Where:

R_{it} is the number of runs scored by the *i* th player in the *t* th innings;
n is the total number of innings in which the *i* th player has batted and
k is the number of innings in which the *i* th player has remained 'not out'.

Equation (1) introduces an upward bias in the average. This bias is caused because the numerator is the total runs scored over all innings while the denominator excludes the innings in which the player has remained 'not out'. This bias cannot seemingly be avoided. Taking the denominator to be *n* instead of *n-k* would instead introduce a downward bias in the average. A similar problem arises while preparing the input data required for the stochastic dominance rules developed later in the paper. The input data that is required is the innings-by-innings runs scored by the player. What should be done with the scores for the innings in which the player has remained 'not out'? This paper first proposes a method to deal with this problem.

The second aspect of cricketing data is the scant attention that has been focused by researchers on certain aspects of cricket. A substantial portion of the work has focused on devising optimal playing strategies. The strategies studied have either focused on batting strategies (Clarke, 1988; Clarke and Norman, 1999; Preston and Thomas, 2000; Swartz et al., 2006) or on bowling strategies (Rajadhyaksha and Arapostathis, 2000). A fair amount of work has also focused on the problem of arriving at a fair result when a game has to be prematurely terminated due to weather conditions or other disturbances (Duckworth and Lewis, 1998; Preston and Thomas, 2002; Carter and Guthrie, 2004).

The third stream of work, on the understanding and development of player-specific performance statistics, (Wood, 1945; Kimber and Hansford, 1993; Lemmer, 2004; Lewis, 2005), has received little attention. Cricket, with its slow pace and non-continuous nature, is a very television-friendly game. It allows viewers the leisure of watching replays without impinging on real-time action. It thus allows for the presentation of a vast amount of descriptive statistics during the course of a game. In spite of this feature of the game and the long history of the game, cricket commentators sometimes seem to feel constrained by

the inability of performance statistics to really describe player performance. Comments like "Statistics don't say everything" are very commonly heard. The attention devoted by researchers to this aspect of cricket, therefore, seems surprisingly scant in relation to its importance and relevance.

This paper seeks to develop methods to assess the performance of batsmen in cricket that (i) makes use of more information than current methods do and (ii) can be converted into visually appealing graphics for the television medium. The method is demonstrated using player statistics for the some of the key members of the Indian One Day International (ODI) cricket team between 1989 and 2005. The names of the players included in the study are listed in Table 1.

METHODS

The primary measure of a batsman's performance in cricket today is the player's batting average defined as in Equation (1). This measure suffers from the shortcoming that it is a one-dimensional number and does not capture the richness of the underlying data. Though cited very often, this measure fails to capture the various facets of a batsman. It does not provide answers to many questions that arise during the course of a game. These un-addressed concerns or questions feature often in the comments of cricket commentators. For example, commentators of the game are found to say "Player X is a dangerous player once he is set". Or "Player X has the ability to convert a good start to a big score". Or "Though his average does not reflect it, Player X is a more consistent performer than Player Y".

Adjusting the raw data

The raw data used in the development of any method for representing a batsman's performance are the innings-by-innings runs scored by the player. However, using this raw data poses a problem. In some of the innings the batsman would not have been dismissed. In such cases the score would not reflect the number of runs the player could potentially have gone on to score. The scores for these innings (the 'not out' situations) have thus to be replaced by a number that is a good estimate of the number of runs the player would have scored had he batted on.

In an early work Wood (1945) had provided empirical support to support the claim that a batsman's scores follows a geometric distribution. Under this assumption, because of the memoryless property of the geometric distribution, a batsman's chance of getting out is independent of the score he is on. This assumption can be used to arrive at an estimate of the number of runs a 'not-out' player would have scored had he batted on. However, the assumption of a geometric distribution for a batsman's scores might not hold for all players. There may be some players who are 'slow starters' and who therefore do better as they progress. There may be other players who become more adventurous as their score increases. For such adventurous players their chances of getting out might increase as their score increases.

Kimber and Hansford (1993) did consider deviations from the geometric distribution, but their focus was on arriving at an optimal estimator for the population mean. On the other hand, we need a method to arrive at an estimate of the number of runs a 'not out' batsman would have gone on to score. A Bayesian approach has been adopted in this paper to arrive at this estimate. This is achieved in the following manner.

Assume that in his *j* th innings player *i* remains 'not out' on a score of R_{ij}. Define a binary variable G_{rik} such that:

$G_{rik} = 0$ if $R_{ik} < R_{ij}$ and $= 1$ if $R_{ik} >= R_{ij}$ for $k = 1, 2, \ldots j-1$ (2)

Define $n_{ij} = \sum_{k=1}^{j-1} G_{rik}$ (3)

Define $C_{ik} = 0$ if $R_{ik} < R_{ij}$ and $= R_{ik}$ if $R_{ik} >= R_{ij}$ for $k = 1, 2, \ldots j-1$ (4)

The estimate of the number of runs that the 'not-out' batsman would have gone on to score is then given by:

$$E_{ij} = \sum_{k=1}^{j-1} C_{ik} / n_{ij} \quad (5)$$

In other words, the estimator used for the runs that the 'not out' batsman would have gone on to score is the conditional average of the batsman at that point of time, given that he has already scored a certain number of runs. In every instance of a 'not out', the batsman's score in that innings *j* is replaced by the estimate E_{ij}. This approach has the advantage of handling deviations from the geometric distribution assumption. It is also information efficient, with the posterior values of the conditional average incorporating more information on the batsman's performance. Table 2 gives an example of the computational procedure used for finding the replacement values for the first two 'not outs' in the career of one member of the Indian ODI team, Sachin Tendulkar.

Table 1: Names of Players Included in the Study

Serial Number	Player	Serial Number	Player
1	S. Tendulkar	11	A. Nehra
2	A. Kumble	12	M. Kaif
3	S. Ganguly	13	M. Karthik
4	R. Dravid	14	S. Bangar
5	A. Agarkar		
6	V.V.S. Laxman		
7	H. Singh		
8	V. Sehwag		
9	Z. Khan		
10	Y. Singh		

Table 2**:** Replacing 'Not Out' Scores with Estimates of Runs Likely to Have Been Scored by a Batsman: Sachin Tendulkar's first 15 ODI Innings

Sl. No.	Runs Scored	Whether Dismissed	Adjusted Score	Remarks
1	0	Y	0	
2	0	Y	0	
3	36	Y	36	
4	10	Y	10	
5	20	Y	20	
6	19	Y	19	
7	31	Y	31	
8	36	Y	36	
9	53	Y	53	
10	30	Y	30	
11	4	Y	4	
12	53	Y	53	
13	52	N	53	Average of Sl Nos. 9,12
14	22	Y	22	
15	11	N	35	Average of Sl Nos 3,5,6,7,8,9,10,12, 13 (adjusted), 14

Stochastic Dominance

The adjusted raw data is now used to arrive at an analytical representation of the player's batting performance. The approach adopted draws from methods normally used for the analysis of securities and portfolios in investment management.

The focus in investment management is on wealth creation. The problem of portfolio choice is that of selecting a portfolio that maximizes the utility for the investor. The utility function for the investor attaches a utility to various levels of wealth. The utility function can be constrained to have certain properties like non-satiation (more wealth is always preferred to less) or risk aversion (diminishing marginal utility for incremental units of wealth). In mathematical terms the first constraint requires the first derivative of the utility function to be positive. Again, in mathematical terms the second constraint requires the second derivative of the utility function to be negative.

Consistent with some of the above-listed features of utility functions, the traditional approach to the portfolio selection problem has been the mean-variance approach. Amongst the alternative approaches to the portfolio selection problem suggested in the investment management literature is the set of stochastic dominance rules (Levy, 1973; Ali, 1975; Bawa, 1978). To use stochastic dominance rules we need to characterize the utility function of the investor. According to the first-order stochastic dominance rules a portfolio A is preferred to another portfolio B if, for any level of return, the cumulative probability of portfolio A giving a return lesser than the given level of return is never greater, and

sometimes less, than the cumulative probability of portfolio B giving a return lesser than that given level of return. This rule is consistent with the assumption that in the investors' utility function more wealth is preferred to less. (Elton and Gruber, 1997)

Analogous to the portfolio selection problem, a similar approach is adopted in this paper to represent the batting performance of cricketers. Using this approach we can say that a batsman A's performance is better than another batsman B's if, for any level of score, the probability of batsman A getting a score greater than the given score is never lesser, and sometimes greater, than the probability of batsman B getting a score greater than that given score. This rule corresponds to the first-order stochastic dominance rules and assumes that more runs are always preferred to less.

The cumulative probability charts of various batsmen can now be charted with runs on the X-axis (with the origin as zero) and the probability of scoring more runs than the X-axis value of the score (that is one minus the cumulative probabilities of scoring lesser than the X-axis value of score) on the Y-axis. Visually this would mean that a batsman whose stochastic dominance curve envelops another's curve stochastically dominates the other batsman.

RESULTS

The method is demonstrated using data for the Indian ODI cricket team spanning the years 1989 (the year one of India's most highly rated players, Sachin Tendulkar, made his debut) to 2005. This period was chosen because this was a period during which the compositional changes in the Indian ODI team were very few. A sample batting performance stochastic dominance chart output for five Indian players is given in Figure 1.

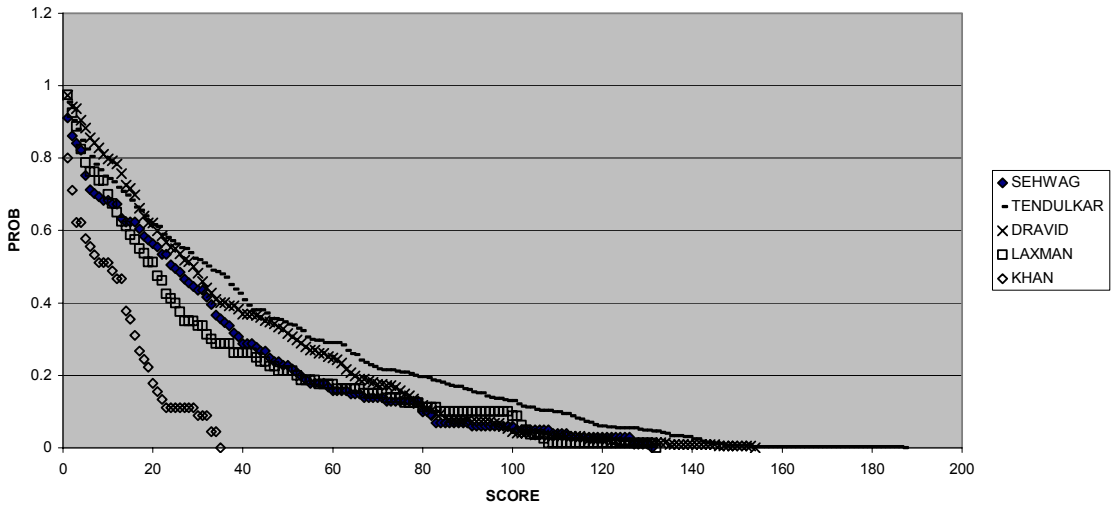


Figure 1: Sample Stochastic Dominance Curves

Four of the five players represented are essentially specialist batsmen (Tendulkar, Dravid, Sehwag and Laxman) and one a specialist bowler (Khan). The results are interesting and have intuitive appeal. They are consistent with popular notions regarding the batsmen whose performances were studied. For example, the curve for Sachin Tendulkar, who is considered an icon of Indian cricket, almost completely envelops the curves for other players. And the curve for Rahul Dravid, who is referred to as 'the wall' because of his perceived consistency, does indeed dominate the curves for other players till the 20 run point. In other words, the chances of Rahul Dravid getting a score less than 20 is lesser than the chances for any other player in the Indian team getting a score lesser than 20. Finally, the curves for the specialist batsmen very clearly dominate the curves for the specialist bowlers, as should be the case.

DISCUSSION

The method that has been developed only provides an alternative approach to represent the batting performance of cricket players. This alternative approach is visually and intuitively appealing. The attempt in this paper is not to arrive at a model to rank the utility of players. Nor is the goal to develop a model to assist in team selection. The utility of a player goes far beyond the runs scored by him. Factors like tactical skills, passive support to the partner batsmen, etc. cannot be gauged by looking at the runs scored. Even if we use runs scored as the sole measure of utility, first-order stochastic dominance rules alone cannot be used to rank players in terms of their utility. And if we go on to second-order stochastic dominance rules the utility function might not have a negative second derivative. In other words, there could be potentially match winning situations in which a batsman who is batting on a very high score (say, 108) has to score one more run in order for the team to win the match. In this situation the incremental one run (from 108 to 109) might be much more valuable than the incremental one run the batsman scored while he was on a lower score (say, 23) during the same innings.

CONCLUSION

Within the limits of this study, the paper seeks to highlight the tremendous scope that exists to improve and develop on the measures currently used to describe the performances of cricket players in general, and batsmen in particular. The measures used today do not adequately capture the richness of the underlying data. Similar approaches can be adopted to represent the performances of bowlers too.

REFERENCES

Ali, M. (1975) Stochastic Dominance and Portfolio Analysis. *Journal of Financial Economics* **2**, 205-230

Bawa, V. (1978) Safety-First, Stochastic Dominance, and Optimal Portfolio Choice. *Journal of Financial and Quantitative Analysis* **XIII**, 225-271

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

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

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

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

Elton, E.J. and Gruber, M.J. (1997) *Modern Portfolio Theory and Investment Analysis*. 5th edition. John Wiley and Sons.

Kimber, A.C. and Hansford, A.R. (1993) A statistical analysis of batting in cricket. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **156**, 443-455.

Lemmer, H.H. (2004) A measure for the batting performance of cricket players. *South African Journal for Research in Sport, Physical Education and Recreation* **26** (1), 55-64

Levy, H. (1973) Stochastic Dominance, Efficiency Criteria, and Efficient Portfolios: The Multi-Period Case. *American Economic Review* **LXIII**, 986-994

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

Preston, I. and Thomas, J. (2000) Batting Strategy in Limited Overs Cricket. *Journal of the Royal Statistical Society: Series D (The Statistician)* **49**, 95-106.

Preston, I. and Thomas, J. (2002) Batting Strategy in Limited Overs Cricket. *Journal of the Royal Statistical Society: Series D (The Statistician)* **51**, 189-202.

Rajadhyaksha, G. and Arapostathis, A. (2000) Using a Bayesian network to recommend the best bowling action. Available from URL:http://webspace.utexas.edu/gaureshr/pubs/cricket.pdf

Swartz, T.B., Gill, P.S., Beaudoin, D. and de Silva, B.M. (2006) Optimal batting orders in one-day cricket. *Computers and Operations Research* **33** 1939-1950

Wood, G.H. (1945) Cricket scores and geometric progressions. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **108**, 12-22.

STATISTICAL ANALYSIS OF NOTATIONAL AFL DATA USING CONTINUOUS TIME MARKOV CHAINS

Meyer D., Forbes D. and Clarke S.
Swinburne University of Technology

ABSTRACT

Animal biologists commonly use continuous time Markov chain models to describe patterns of animal behaviour. In this paper we consider the use of these models for describing AFL football. In particular we test the assumptions for continuous time Markov chain models (CTMCs), with time, distance and speed values associated with each transition. Using a simple event categorisation it is found that a semi-Markov chain model is appropriate for this data. This validates the use of Markov Chains for future studies in which the outcomes of AFL matches are simulated.

KEY WORDS
homogeneity in time, sequential dependency, semi-Markov process, football

INTRODUCTION

Animal biologists frequently perform ethological studies creating models in order to provide an accurate description of animal behaviour. The effects of various factors can then be studied in terms of the parameters of these models. The data often consists of continuous time records of behaviour which can be described using Markov chain models which take into account both the duration and the sequence of acts. Using these models it is possible to determine whether behaviour is homogeneous during an observation period, and, when behaviour is not homogeneous, changes in the model parameters can be used to determine when and how behaviour changes. Sports can be studied in a similar manner, using notational analysis to collect the data as described in Forbes and Clarke (2004) and Forbes (2006). If changes in behaviour can be linked to successful outcomes we will have a valuable tool for player development.

Markov chains have been previously used to model sports events (Bellman, 1977, Bukiet and Harold et al. 1997, Forbes and Clarke 2004, Hirotsu 2002, Hirotsu and Wright 2003a and 2003b, Forbes 2006). Forbes and Clarke (2004) and Forbes (2006) created discrete Markov chains for AFL football, but this is probably the first time that an attempt has been made to model AFL football using CTMCs, because the data was not previously available. The model is similar to the above rat model; however, in addition to associating times with events we also have distances and speeds.

METHODS

CTMC Assumptions
There are a number of assumptions associated with a continuous time Markov chain. The Markov property implies that transitions are independent of the time for previous transitions as well as the type of previous transitions. In addition it is assumed that the characteristics of the transitions have exponential distributions for each state. In animal

behaviour it is commonly found that the times for behaviours (bouts) do have an exponential distribution.

In our analysis of AFL football we refer to the states of the Markov chain as events, such as a Kick. We have the times for each transition between events as well as the distance and speed associated with each transition, so we shall endeavour to include all three of these transition characteristics into our CTMC model. It is unlikely that these variables will have exponential distributions because these dimensions are confined by field size and shape and because there is a grouping of behaviours under each of the events (e.g. a Kick may be long, short, a ground kick, a clanger, a kick to advantage or an ineffective kick). Also it may be that we do not have a first-order Markov model in that the transition probabilities may not be independent of the previous sequence of events. This paper will investigate these issues in detail.

Processes which do not have exponentially distributed transition times are called Semi-Markov chains. A common distribution in the animal behaviour literature is a displaced exponential distribution which allows for a non-zero minimum value. The gamma distribution has also been used to describe the duration of animal behaviours, allowing for a mixture of exponential distributions. Log-normal distributions are also used and even normal distributions which have been censored at zero. All these possible distributions can be tested for our time, distance and speed variables. Of course, a multivariate distribution allowing for correlations between these three variables should also be considered.

Haccou and Meelis (1992) recommend the following process for analysing behavioural data in animals.

1) Search for homogeneous periods in order to reduce the error variances in the model. In particular it is possible to determine whether changes are abrupt or gradual.
2) Analysis in the presence of gradual changes require special modelling but where there are abrupt changes the data should be divided into homogeneous subphases and analysed as indicated below for each subphase.
3) Determine suitable distributions for the time, distance and speed variables and test the sequential dependence properties of the process. Also search for outliers.
4) If the distributions are exponential and there is first-order sequential dependency a standard Markov chain analysis is possible.
5) If the distributions are a mixture of exponential or gamma distributions and there is first-order sequential dependency the behavioural categories (Markov states) may have to be subdivided before a standard Markov chain analysis is possible.
6) If the distributions are not exponential or mixtures of exponential or gamma distributions, but there is first-order sequential dependency, a semi-Markov chain analysis is possible.
7) If there is higher order sequential dependency ad hoc analysis methods are required.

We will follow this process, in the analysis below, using a data set derived from four AFL matches during the 2004 season.

The Data

The data was collected by Champion Data, the official provider of AFL statistics, for four matches during the 2004 season. These matches, the venues and the results are described in Table 1. In this paper we apply the ideas of Haccou and Meelis (1992) in this context using the event definitions shown in Table 2. These event definitions are over-simplistic in that they do not identify the teams involved in each transition. However, this simple event definition does allow us to test the assumptions of the CTMC. Forbes (2006) and Forbes and Clarke (2004) use a different set of event definitions in their work. Their definitions identify the teams involved in each transition; however, they do not differentiate between handballs and kicks, making the modelling of distances, times and speeds problematic for these events.

Table 1: Description of Data: 4 AFL matches in the 2004 season

Venue	Home Team	Away Team	Winner	Home Team Score	Away Team Score
Kardinia Park, Geelong	Geelong	St Kilda	Geelong	101	94
Subiaco Oval, Perth	West Coast	Western Bulldogs	West Coast	106	57
Melbourne Cricket Ground	Melbourne	Hawthorn	Melbourne	107	63
Sydney Cricket Ground	Sydney	Kangaroos	Kangaroos	112	118

Table 2: Description of event codes

Event code	Event Description
BEHI	Behind
BUBO	Ball up bounce
CEBO	Centre ball up
HB	Hand ball
KI	Kick in
KK	Kick
THIN	Throw in

In the following analysis we start with an exploratory analysis in which we examine the assumption of an exponential distribution for time, distance and speed for each type of event. Thereafter we test for time inhomogeneity in our data and then test the nature of any time dependencies.

RESULTS

Exploratory Data Analysis

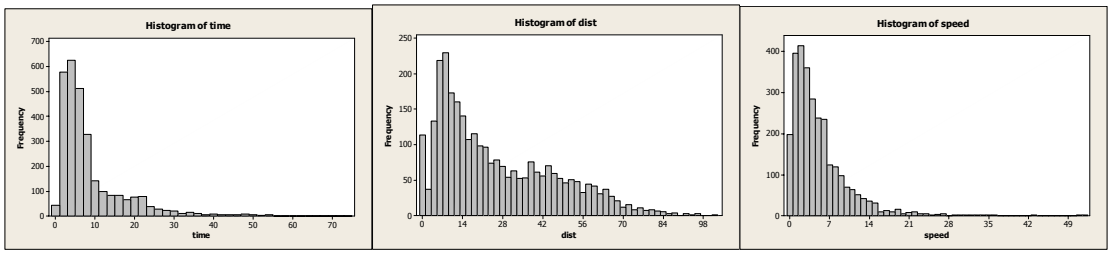
A transition matrix was derived using the above event codes and the average times (sec), distances (m) and speeds (m/sec) were calculated for each event as shown in Table 3. Clearly kicks (KK) are the most common event followed by handballs (HB).

The mean times, distances and speeds vary markedly for the different types of events as expected.

Table 3: Transition matrix and average event statistics

	To Events								
From Events	BEHI	BUBO	CEBO	HB	KI	KK	THIN	Total	Pct % (St.Dev.)
BEHI	0	0	0	0	85	0	0	85	2.93
BUBO	1	14	0	35	0	41	3	94	3.24
CEBO	0	16	0	61	0	49	0	126	4.34
HB	0	16	0	385	0	511	29	941	32.44
KI	0	0	0	20	0	64	0	84	2.90
KK	85	38	112	381	0	738	90	1444	49.78
THIN	0	11	0	61	0	48	7	127	4.38
Total	86	95	112	943	85	1451	129	2901	100.00
Mean Time	18.41	11.85	9.16	4.48	6.73	9.51	11.87	8.22	(8.66)
Mean Dist.	36.22	9.33	9.45	13.37	7.37	37.51	18.40	25.80	(20.35)
Mean Speed	2.29	1.10	1.28	4.60	2.68	6.54	2.23	5.06	(4.87)

Figure 1: Distributions for event values: Time, Distance and Speed



The histograms in Figure 1 show the distributions for time, distance and speed when all event types are combined. A right skew distribution is exhibited in all cases with skewness coefficients of 2.40 for time, 0.85 for distance and 2.51 for speed. In particular, the lumpiness of the distance distribution demonstrates the effect of the different events. Figure 2 shows a time plot for the events for each of the four matches. This plot shows obvious differences between the four matches. The Geelong match shows relatively few behinds except in the last quarter. The number of behinds peaked in the middle of the match for the Perth and in the first half of the Melbourne game. The number of goals is related to the number of centre bounces (CEBO), with three of the four games showing relatively few goals in the last few minutes of the match. Kicks were relatively rare for the Sydney game while Kick-Ins were relatively rare for the Geelong game.

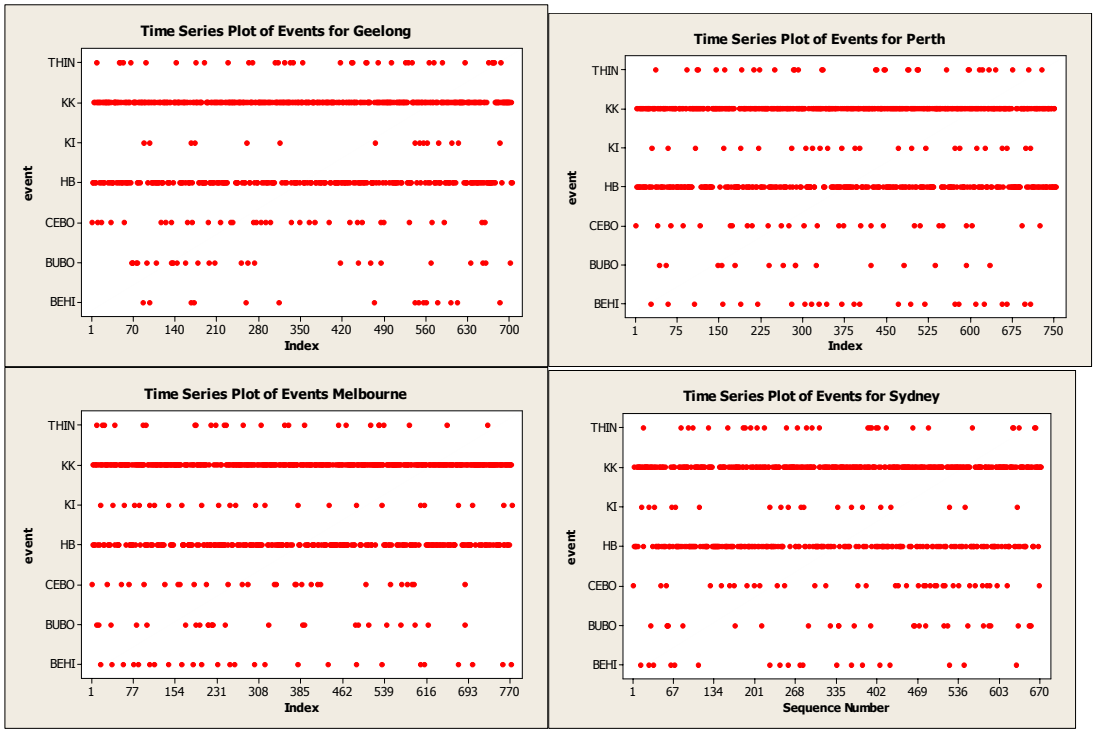


Figure 2: Event Sequences for each of the four venues

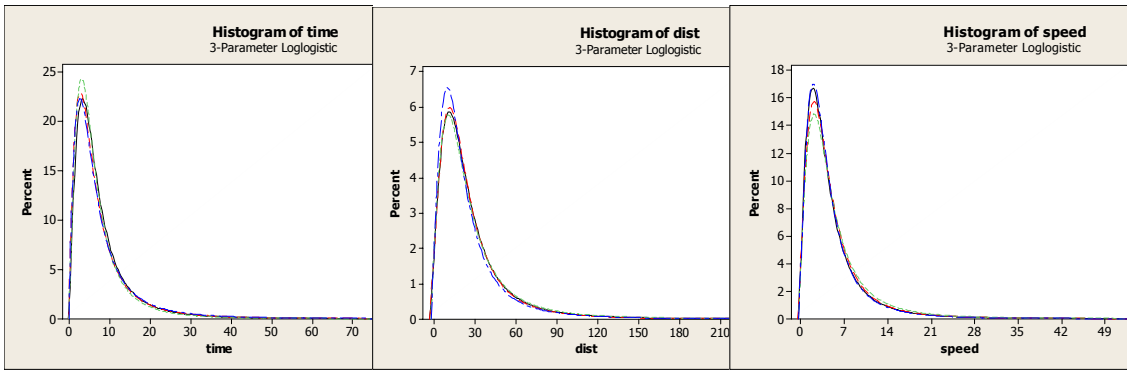


Figure 3: Distribution of event times, distances and speeds for the four venues.

Figure 3 compares the time, distance and speed distributions for each match using a 3 parameter LogLogistic distribution to describe each distribution. This is a versatile distribution shown below to describe the data well. There are obviously quite small differences between the matches, and the nonparametric Kruskal-Wallis tests in Table 4 suggest that there is a significant difference only for speed, with a slower game played in Sydney than in Melbourne.

Table 4: Mean Values and Standard Deviations for each venue

	Venue				
Variable	Geelong	Melbourne	Perth	Sydney	p-value
Time	8.37 (0.32)	7.70 (0.28)	8.46 (0.33)	8.84 (0.37)	0.349
Distance	26.06 (0.76)	27.29 (0.77)	25.49 (0.72)	24.59 (0.77)	0.190
Speed	4.99 (0.19)	5.58 (0.18)	5.14 (0.17)	4.88 (4.76)	0.048

Figure 4 compares the distribution of event times, durations and speeds for each type of event, again using a 3-parameter LogLogistic distribution. There are clearly very significant differences between the various types of events (Kruskal Wallis: $p<0.001$). Moreover it is clear that an exponential distribution is not appropriate in most cases.

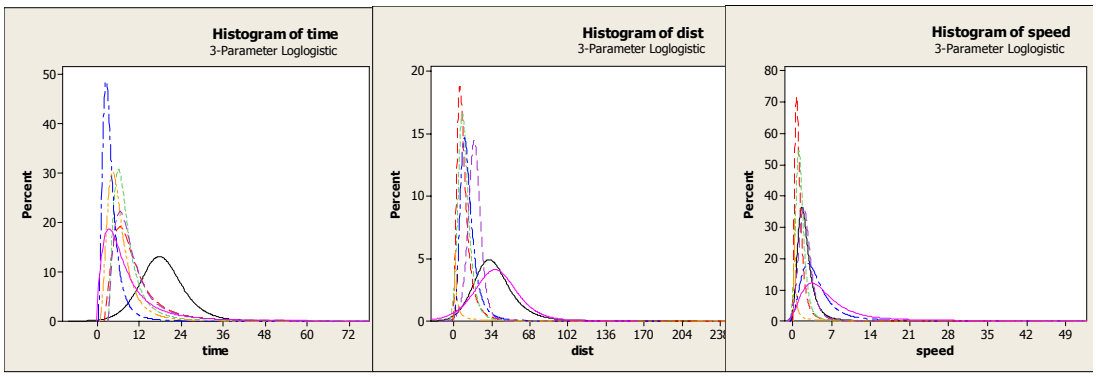


Figure 4: Distribution of event times, distances and speeds for the seven events.

The goodness of fit for a set of four common survival distributions was studied using the Anderson-Darling statistic. This statistic measures the area between the fitted distribution function and the nonparametric empirical distribution function. As shown in Table 5, the 3-parameter LogNormal distribution (LN) and the 3-parameter Loglogistic (LL) distributions gave the most consistently good results. The 3-parameter Gamma (G) and the 3-parameter Weibull (W) distributions were less appropriate in most instances.

Table 5: Ranking of Distributional Fit using the Andersen-Darling Statistic

Event	Time				Distance				Speed			
	LL	G	W	LN	LL	G	W	LN	LL	G	W	LN
THIN	1	3	4	2	1	3	4	2	1	3	4	2
KK	1	4	3	2	3	2	4	1	4	1	3	2
KI	2	3	4	1	3	2	1	4	3	2	1	4
HB	1	3	4	2	1	3	4	2	2	3	4	1
CEBO	1	3	4	2	4	2	1	3	4	3	1	2
BUBO	1	3	4	2	1	3	4	2	1	3	4	2
BEHI	4	2	1	3	1	3	4	2	1	3	4	2
Mean Rank	1.6	3.0	3.4	2.0	2.0	2.6	3.1	2.3	2.3	2.6	3.0	2.1

Table 6: Spearman Rank Correlations for Transition Values (** $p<0.001$)

	Correlations		
Event	Time*Distance	Time*Speed	Distance*Speed
BEHI	0.030	-0.604 (**)	0.717 (**)
BUBO	0.072	-0.577 (**)	0.697 (**)
CEBO	0.129	-0.435 (**)	0.798 (**)
HB	0.213 (**)	-0.659 (**)	0.533 (**)
KI	-0.248 (**)	-0.265 (**)	0.998 (**)
KK	0.438 (**)	-0.629 (**)	0.356 (**)
THIN	0.069	-0.633 (**)	0.415 (**)
All	0.414 (**)	-0.520 (**)	0.520 (**)

Table 6 shows the correlations between our time, distance and speed variables for each type of event. Spearman rank correlations were used on account of the lack of normality in most cases. The correlations are particularly interesting for Kick-Ins, with longer kicks apparently associated with shorter times, resulting in a much quicker speed. This is expected since a run with the ball is more likely before a short kick than before a long kick.

Analysis for time inhomogeneity in the case of abrupt changes
Visual methods can be used for detecting inhomogeneity in time. Our time plots give some indication of inhomogeneity in time and between matches in that Figure 2 suggests that the frequency of events varies over time and between matches. However, Table 4 showed that the mean time and duration were similar for all the matches with a barely significant difference in the case of speed. This suggests that any inhomogeneity in our CTMCs will be confined to the transition probabilities. Hypothesis tests can be used to confirm whether this is true, using changes in mean termination rates or in the sequence of transitions to detect any time inhomogeneity.

If the number of change points is known a Kruskal-Wallis test can be used to test whether the distribution of values for a specific event differs between the differing periods. This test makes no assumption about the distribution of values for a specific event. We compared the time, distance and speed distributions for each event between the quarters in any match and found no significant differences for any event when the Bonferroni correction was applied ($\alpha = 0.05/28$). This confirms that there is no time inhomogeneity in the time, distance and speed distributions.

Change points in the transition matrix can be tested using the following likelihood function for the Poisson distribution, with A and B denoting two consecutive states in a Markov process. In animal behaviour studies it is not usual to allow a transition from a state to itself, however, we shall allow this in AFL football so that we can track the passage of the ball from player to player. On the other hand there are some transitions that are not possible in AFL football (e.g. a Kick-In is the only event that can follow a Behind), so we will ignore all transitions with a frequency of zero in Table 3.

For the sake of simplicity we again consider the end of each quarter as possible change points for each of the four matches. Our multinomial logistic regression analysis shows no significant match or quarter effect, suggesting that the transition matrix, like the transition variables, is homogeneous in time. As a result we shall use our complete data set for all four matches to test for sequential dependency.

Tests of sequential dependency
In a continuous time Markov chain (CTMC) a first-order dependency in the sequence of states is assumed. This means that the transition probability for states A and B in time Δ is independent of the sequence of preceding states. This implies that the transition durations are independent for a given sequence of states. Dependencies may be short-term, long-term, or periodic in nature. They may relate to the sequence of states or dependencies between transition values and preceding and/or following states, or they may relate to correlations with transition values in subsequent transitions. In the case of animal behaviour transitions from state A to itself cannot occur, but as mentioned above

this is not true in the case of AFL football. Instead there are several other transitions that are impossible as exhibited in table 3.

Deviation from first-order dependency in a sequence of states is commonly tested with a chi-squared test. This test has reasonable power, however, it does not necessarily detect dependencies of higher than second order. Multinomial logistic regression was therefore used to model the occurrence of event Y based on the two previous events (X and A). It was found that only the most recent event had a significant influence ($\chi^2(36) = 762.0$, $p<0.001$) while the effect of the previous event was not significant ($\chi^2(42) = 44.6$, $p = 0.384$).

The next form of dependency occurs when the transition value distributions depend on the preceding state. This can be tested using a Kruskal-Wallis test, making no assumptions regarding the nature of the value distributions. Not unsurprisingly there was a strong relationship between the type of previous event and the values for time ($\chi^2(6) = 20.7$, $p = 0.002$), distance ($\chi^2(6) = 186.7$, $p<0.001$) and speed ($\chi^2(6) = 210.6$, $p<0.001$).

Relations between subsequent transitions for the same and for different states produce a further form of dependency found in (semi-)Markov models, which can be measured using autocorrelation. Autocorrelations were initially calculated for all types of events simultaneously. For the time variable there was a very weak but significant positive autocorrelation of 0.05 for every second transition, suggesting that shorter events, such as handballs, would alternate with other types of event such as kicks. This theory is supported by the transition matrix in table 3. For the distance variable there was a weak but significant negative autocorrelation of 0.07 for successive events (lag one), again suggesting a tendency to alternate handballs and kicks, while for the speed variable there was a weak but significant negative autocorrelation of 0.05 for successive events and a stronger positive autocorrelation of 0.12 for every second event. Although all these correlations are weak they do tell us something interesting about the game. It is expected that these correlations will be automatically incorporated in the model through the transition matrix.

When autocorrelations are considered for each type of event separately, only in the case of Kicks do we obtain any significant autocorrelations. The time taken for consecutive kicks has a weak but significant negative correlation of 0.15, suggesting that short duration kicks alternate with longer duration kicks. However, the speed for consecutive kicks has a weak but significant positive correlation of 0.10. Although weak, these correlations probably need to be incorporate in the modelling process.

DISCUSSION AND CONCLUSION

Our analysis of four 2004 AFL football matches has shown that inhomogeneity is unlikely to be a problem within an AFL football match. In our case there were similar processes for all four matches, perhaps on account of the similar scores for the four matches. However, for our definition of events there were marked differences in time, distance and speed requiring a separate analysis for each type of event. The distributions for the time, distance and speed variables varied for the different types of event, however, the 3-parameter LogLogistic and the 3-parameter LogNormal distributions

tended to give the best fit. There were strong correlations between these variables for most of the events. Finally, it was confirmed that a first-order sequential dependency existed for the events, and that for successive kicks there was a weak correlation for the speed and time variables.

These results suggest that a semi-Markov model is appropriate since the distributions are not usually exponential or mixtures of exponential or gamma distributions, but there is first-order sequential dependency. This model could be used for simulation purposes. An initial centre bounce (CEBO) would result in Ball-up Bounce (BUBO) a handball (HB) or a kick (KK) with respective probabilities of 13%, 48% and 39%. The associated time, distance and speed could be generated using the appropriate CEBO three-parameter log-normal distributions, allowing appropriate correlations between the times, distances and speeds. Similarly, results for all ensuing game events could be simulated. Through changes to the transition matrix and/or other model parameters, the resulting model could be used in order to predict the effect of rules changes and changes in play strategy.

However, although the total number of goals and behinds would be known, the final score and the winner would not be known. In order to develop a more useful model all that is needed is a split of the events to identify the teams involved in each transition. The current work suggests that a semi-Markov model would be appropriate for this extended model, allowing a simulation similar to that described above, from which scores and the winning team could be determined for each simulated game.

In the above analysis we have associated distances, speeds and times with each transition in time. The addition of directions for each transition would make it possible for a spatial simulation to be performed. In this case it would make sense to define the events according to spatial zone (within the field) as well as activity. An alternative approach would have been to use the quarters of the field as the events, again using time, distance, speed and direction to describe each transition. This approach would also not allow the simulation of match outcomes but it would help coaches and players to better understand the spatial patterns of play. A further extension to this work could allow continuous changes in the model parameters over time with the possible inclusion of covariates in the models for the transition probabilities.

Acknowledgements

The authors wish to thank Champion Data for access to the data on which this paper is based and they wish to thank a referee for helpful comments.

REFERENCES

Bellman, R. (1977) Dynamic Programming and Markovian decision processes, with application to baseball. *Optimal Strategies in Sports*. Ed: Ladany, S.P. and Machol, R. E. Amsterdam: North Holland. 77-85.

Bukiet, B., Harold, E. and Palacios, J. (1997) A Markov chain approach to baseball. *Operations Research* **45**(1), 14 - 23.

Forbes, D. and Clarke, S. (2004) A seven state Markov process for modeling Australian rules football. *Seventh Conference on Mathematics and Computers in Sport*. Ed: Morton, H. Palmerston North: Massey University. 148-158.

Forbes, D. (2006) *Dynamic Prediction Models In Australian Rules Football Using Real-time Performance Statistics*. Submitted Doctoral Thesis, Melbourne: Swinburne University of Technology.

Haccou, P. and Meelis, E. (1992) *Statistical Analysis of Behavioural Data: An Approach Based on Time-Structured Models*. New York: Oxford University Press.

Hirotsu, N. (2002) A Formulation of Optimal Substitution Strategies Using a Markov Process Model in Baseball and Soccer. *Management Science* **48**(2), 306.

Hirotsu, N. and Wright, M. (2003) An evaluation of characteristics of teams in association football by using a Markov process model. *Journal of the Royal Statistical Society: Series D* **52**(4), 591 - 602.

Hirotsu, N. and Wright, M. (2003) A Markov chain approach to optimal pinch-hitting strategies in a designated hitter rule baseball game. *Journal of the Operations Research Society of Japan* **46**(3), 353 - 371.

ARTIFICIAL INTELLIGENCE IN SPORTS BIOMECHANICS: NEW DAWN OR FALSE HOPE?

Bartlett, R.

School of Physical Education, University of Otago, Dunedin, New Zealand

ABSTRACT

This article reviews developments in the use of Artificial Intelligence (AI) in sports biomechanics over the last decade. It outlines possible uses of Expert Systems as diagnostic tools for evaluating faults in sports movements ('techniques') and presents some example knowledge rules for such an expert system. It then compares the analysis of sports techniques, in which Expert Systems have found little place to date, with gait analysis, in which they are routinely used. Consideration is then given to the use of Artificial Neural Networks (ANNs) in sports biomechanics, focusing on Kohonen self-organizing maps, which have been the most widely used in technique analysis, and multi-layer networks, which have been far more widely used in biomechanics in general. Examples of the use of ANNs in sports biomechanics are presented for javelin and discus throwing, shot putting and football kicking. I also present an example of the use of Evolutionary Computation in movement optimization in the soccer throw in, which predicted an optimal technique close to that in the coaching literature. After briefly overviewing the use of AI in both sports science and biomechanics in general, the article concludes with some speculations about future uses of AI in sports biomechanics.

KEY WORDS

artificial intelligence, artificial neural networks, evolutionary computation, expert systems, Kohonen self-organizing maps, sports biomechanics.

INTRODUCTION – WHERE WE WERE IN 1995

Lapham and Bartlett (1995) published a review of the use of AI in sports biomechanics. In this, we reported no evidence of the use of AI in sports biomechanics, although Expert Systems and ANNs were being used in gait analysis. We did, however, predict a bright future for the use, in particular, of Expert Systems in sports biomechanics. So what has happened in the decade since?

EXPERT SYSTEMS

Expert Systems are, effectively, a database combined with a knowledge base, 'reasoning' and a user interface. The knowledge base contains specific knowledge, or facts, for the 'domain'. The knowledge rules can also include logic operations, managed by probability theory, as in this example from a hypothetical Expert System for the analysis of fast bowling in cricket: IF 'shoulder-axis counter-rotation' is high; THEN 'technique' is mixed (p=0.8).

This example was chosen to illustrate that much information is vague – 'high' in the above example has varied from 10 to 20 to 30 to 40º in the scientific literature on fast bowling (see, for example, Bartlett, 2003), showing that much information is 'fuzzy'. The difference between 'crisp' and 'fuzzy' knowledge is shown in Figure 1 for fast bowling. Note that in the fuzzy representation, side-on and mixed techniques overlap as do mixed and front-on. These fuzzy overlaps are supported by the division of the mixed technique into side-on-mixed and front-on-mixed.

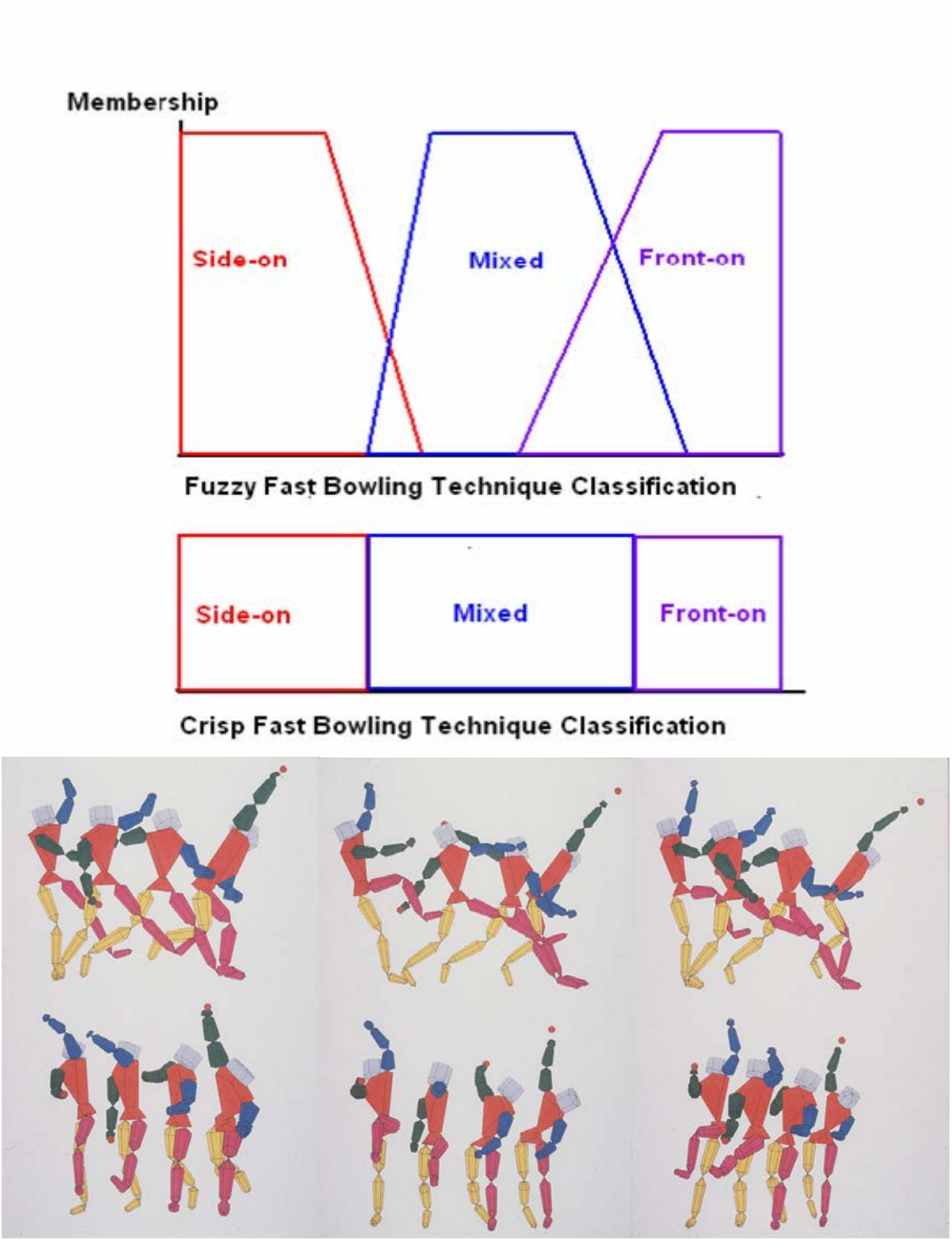


Figure 1: Classification of cricket fast bowling techniques

So, as Expert Systems are good diagnostic tools and system 'shells' are readily available, it is surprising that they are rare in sports science. The closest thing to Expert Systems in sports biomechanics at present is found within qualitative video analysis packages, such as SiliconCOACH's 'wizards'. Although not, strictly speaking, Expert Systems, these wizards do provide a formula engine that could be used by wizard developers to arrive at decisions by taking into account one or more responses to other data entered into the wizard; whether this provision is used is up to the wizard developer. This reality conflicts with the positive view of the utility of Expert Systems by Lapham and Bartlett (1995).

The use of Expert Systems in gait analysis (e.g. Bekey et al., 1992) suggests an extension to the analysis of sports techniques; both are branches of biomechanics. In gait analysis, however, there is a strong developmental motivation – patient health – which helps to attract funding. Clinicians are expensive, making investment in complex software development worthwhile financially. Gait analysis is a confined expert domain - gait and its variants with many experts. It is laboratory-based, so automatic marker tracking systems are commonplace and data are abundant. Analysis of sports techniques is more complex than gait analysis and there is a weak developmental motivation: research into sport performance is not well funded. Coaches and sport scientists are not expensive; technique analysis is often field-based, preventing the automatic tracking of markers; and it is a broad expert domain, involving many sports. There is little data for technique analysis Expert Systems and there are fewer experts than for gait analysis.

ARTIFICIAL NEURAL NETWORKS

Artificial Neural Networks (ANNs) allow computers to learn from experience and by analogy. They are computer programs that try to create a mathematical model of neurons in the brain. An ANN is an interconnection of simple adaptable processing elements or nodes. They are non-linear programs that represent non-linear systems, such as the human movement system, and, from a notational analysis perspective, games. Artificial Neural Networks have nodes, which are simplified models of brain neurons, inputs, outputs and weights. The network stores experiential knowledge as a pattern of connected nodes and synaptic weights between them. Multi-layer ANNs have several 'hidden' layers and normally learn using the 'back-propagation learning law'.

Kohonen self-organizing maps have one hidden layer and using 'competitive learning' – only one neuron is selected for weight adjustment each iteration, based on the minimum 'distance' between the sums of its inputs and its weight. These networks require lots of 'training' data and, once trained, can only be used for testing, not further learning.

ARTIFICIAL NEURAL NETWORKS IN SPORTS BIOMECHANICS

Given their usefulness for classification, clustering and prediction, and that they are easily available, how widespread is the use of ANN in sports biomechanics? Well, unlike Expert Systems, they have been used, as well as in notational analysis and elsewhere in sport and exercise science (see, for example, Perl, 2001, 2005). Perl (2005) and Perl and Weber (2004) highlighted the importance of pattern recognition using ANNs; the patterns can be tactical ones from a game, performance patterns in training, or – the focus of the rest of this paper – movement patterns of sports performers. In this

last application, the ANN is normally used to transform a high-dimensional vector space of biomechanical time series into a low-dimensional output map.

Kohonen self-organizing maps were used to analyze discus throws by Bauer and Schöllhorn (1997). They used 53 throws (45 of a decathlete, 8 of a specialist) recorded using semi-automated marker tracking over a one-year training period. Each throw had 34 kinematic time series, for 51 normalised times; these complex, multi-dimensional time series were mapped on to a simple 11x11 neuron output space (Figure 2). Each sequence was then expressed as the mean deviation (*d* in Figure 2) of the output map – the continuous line - from that of one of the throws by the specialist thrower, shown by the dashed line.

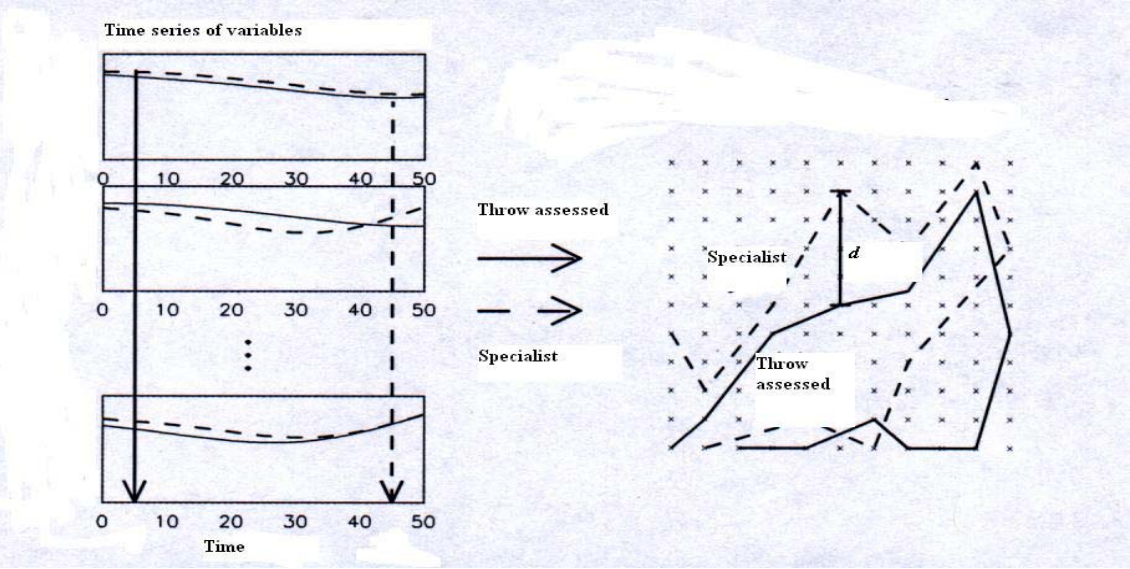


Figure 2: Use of Kohonen self-organizing maps in discus technique analysis (adapted from Bauer and Schöllhorn, 1997).

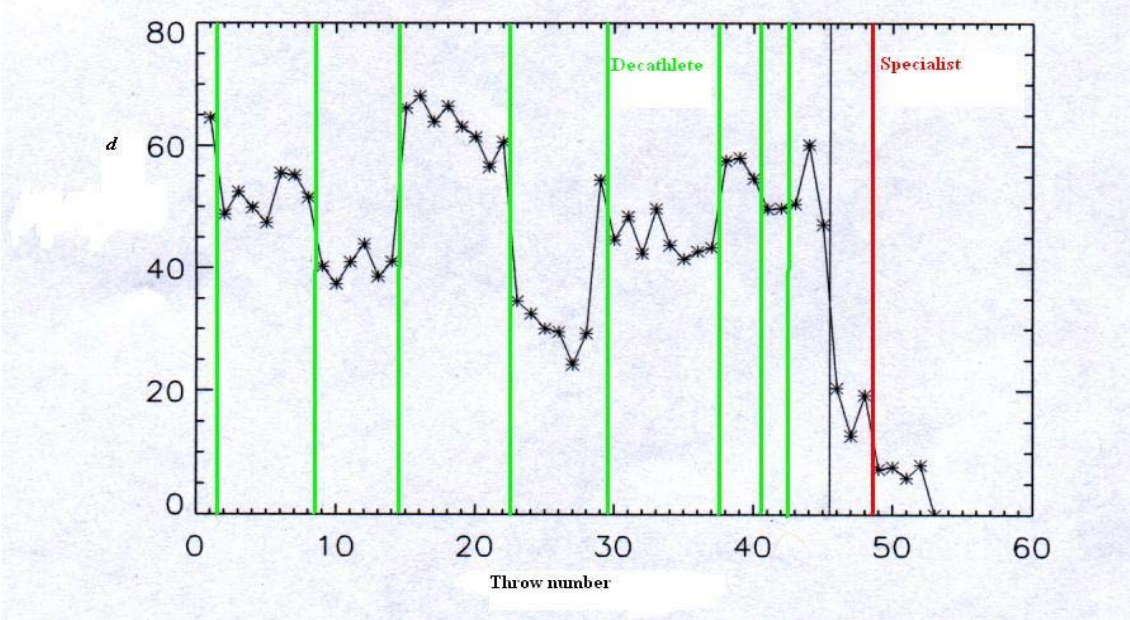


Figure 3: Grouping of throws within sessions, between vertical lines (adapted from Bauer and Schöllhorn, 1997).

The deviations for the eight specialist throws are shown on the right of Figure 3, the decathlete's 45 throws on the left. The 'distances' are less for the specialist thrower as the comparator was one of his throws. Note the clustering of groups of throws, between the vertical lines, within training or competition sessions. There was more variability between than within sessions; for five groups of five trials, the authors computed inter- and intra-cluster variances, giving an inter-to-intra variance ratio of 3.3±0.6. This shows that even elite throwers cannot reproduce invariant movement patterns between sessions. The supposed existence of such invariant patterns – which arises from the motor programs of cognitive motor control - has often been used, explicitly or implicitly, to justify the use of a 'representative trial' in sports biomechanics.

Bauer and Schöllhorn (1997) claimed that the map output reveals information about the whole movement that is not discernable from the detailed kinematics. It is, undoubtedly, simpler and different. What we have here is, in effect, the detection and recognition of a pattern that is obscured by the enormous fine detail of the multiple time series.

Schöllhorn and Bauer (1998) reported a similar approach to analyse 49 javelin throws from eight elite males, nine elite females and ten heptathletes. This time, manual digitising of estimated joint centre locations was used. Clustering was found for the male throwers – as a group - and for the two females for whom multiple trials were recorded. Variations in the cluster for international male athletes were held to contradict any existence of an 'optimal movement pattern'. This view was supported by an analysis at the 1995 World Athletics championships, reported by Morriss et al. (2000), with a focus on arm contributions to release speed. The large shoulder angular velocity for the silver medalist suggested reliance on shoulder extension and horizontal flexion to accelerate the javelin, suiting his linear throwing technique. In contrast, the gold medalist used medial rotation of the shoulder to accelerate the javelin; this movement, plus an elbow extension angular velocity at least 18% faster than for any other finalist, was the reason he was able to achieve the greatest release speed. However, some scepticism about the results of both these studies is warranted in the light of recent research by Bartlett et al. (2006). We found, in a two-dimensional laboratory study of treadmill running, that it is impossible to distinguish movement variability between trials from variability within and between operators who manually digitized joint centres without the use of markers. This would be far worse for a field-based three-dimensional study.

Lees et al. (2003) reported the results of a study that used Kohonen maps to analyse instep kicks by two soccer players for distance or accuracy. Joint angles were obtained from the three-dimensional coordinates of automatically-tracked markers. These were then mapped on to a 12x8 output matrix and showed differences between tasks and players; these output patterns were repeatable for the same task for one player. The authors claimed that the output map 'nodes' were related to characteristics of the movement technique, although what these characteristics are remains to be determined. Lees and Barton (2005) used a similar approach for several kicks by six soccer players, three right- and three left-footed. In this study, 14 joint angles were obtained from the three-dimensional coordinates of automatically-tracked markers for 80 equispaced time instants from take-off for the last stride to the end of the follow through of the kick. The output maps distinguished well between the right- and left-footed groups, which the authors stated was a non-trivial problem using just the joint kinematics. Again, intra-player differences were small.

Adopting a different approach from that of the previous studies, Yan and Wu (2000) used a multi-layer ANN with one hidden layer to analyse the shot putts of 155 throws by 31 national-standard Chinese females. The network was 'trained' using values of 20 global and 33 local technique parameters from manually-digitized coordinates, to predict release angle and speed from 134 throws of all throwers; it was then tested with data from 21 throws of 11 throwers. The errors between the network outputs and the measured release parameters were then compared to those obtained using regression analysis. The ANN errors were typically 25-35% less than those from regression analysis, e.g. 0.20 compared to 0.31 $m \cdot s^{-1}$ for release speed and 0.91 compared to 1.26º for release angle. Whether such an improvement merits the use of a more complicated approach is a matter of judgment, although it is worth noting that regression models cannot learn. What might need emphasizing is that the errors from both methods are smaller than the uncertainties in release parameter values that occur using manual digitizing, as in this study, for which errors in release angle of ±1.5º and in release speed of ±0.5 $m \cdot s^{-1}$ are common. This network was then used by Yan and Li (2000) to analyse the shot putting techniques. The authors claimed that this showed weaknesses of technique compared with those of the elite putters, although this was not well substantiated by the paper, possibly because the Chinese authors were writing in English.

Artificial Neural Networks have been more widely used than Expert Systems in sports biomechanics. In technique analysis, Kohonen self-organising maps have been claimed to reveal the 'forest' rather than the 'trees'. Simplification is undoubtedly an important feature of ANN, although the ways in which we can best use the outputs of these mappings remains to be determined. If the mapping rules within these opaque and very non-linear networks never become transparent, as some ANN experts predict, then explicit mappings between specific features of the kinematic time series and the output maps may never emerge. Even under these circumstances, however, this novel approach to the analysis of sports movements might still prove to be a powerful tool in the analysis of human movement in sport, such as by possibly providing a non-linear measure of movement variability. Artificial Neural Networks represent an important link to non-linear dynamical systems theory; for example, Kelso (1995) reported the use of ANNs in studies of perception and noted that the networks model hysteresis, stimulus bias, and adaptation effects, all key tenets of non-linear dynamical systems theory.

EVOLUTIONARY COMPUTATION

Evolutionary Computation includes genetic algorithms, genetic programs and evolutionary strategies, and uses artificial – numerical – 'chromosomes' to simulate evolution. Bächle (2003) used an evolutionary strategy to optimize the joint torques at hip, shoulder and elbow to maximize distance thrown in a soccer throw in. This study predicted an optimal throwing technique close to that described in the coaching literature, with the initially passive torque of the hip accelerating the trunk forwards while the negative elbow torque kept the forearm back. Then, 30 ms before release, the trunk was decelerated by a negative hip torque, while a positive elbow torque accelerated the forearm forwards. Seifriz and Mester (2002) used genetic algorithms to calculate the optimum trajectory of a skier, but this was only published as an abstract.

CONCLUSION: A ROSY FUTURE FOR AI IN SPORTS BIOMECHANICS?

Automatic marker-tracking systems allow more, and more accurate, human movement data to be collected. This could lead to the use of fuzzy Expert Systems for diagnosis of faults in sports techniques, a substantial development of the rudimentary Expert Systems currently embedded in some video analysis packages. Kohonen mapping will become commonplace in sports biomechanics, particularly if the technique elements captured by the mapping can be identified. Dynamically controlled networks will become more widely used in studying learning of movement patterns. Multi-layer ANNs will have an important role in technique analysis, a view supported by their use elsewhere in biomechanics, including the closely related domain of gait analysis. Other AI applications – particularly Evolutionary Computation and hybrid systems - will feature in future developments in the optimization of sports techniques and skill learning. Finally, the links with dynamical systems theory will become even more apparent, leading, for example, to an enhanced understanding of movement coordination and the role of movement variability. But Lapham and Bartlett were equally optimistic in 1995 and, so far, their expectations have not been fully realised.

REFERENCES

Bächle, F. (2003) The optimisation of throwing movements with evolutionary algorithms on the basis of multi-body systems. *International Journal of Computer Science in Sport* **Special Edition 1**, 6-11.

Bartlett, R. (2003) The science and medicine of cricket: an overview and update. *Journal of Sports Sciences* **21**, 733-752.

Bartlett, R., Bussey, M. and Flyger, N. (2006) Movement variability cannot be determined reliably in no-marker conditions. *Journal of Biomechanics.* In press.

Bauer, H. and Schöllhorn, W. (1997) Self-organizing maps for the analysis of complex movement patterns. *Neural Processing Letters* **5**, 193-199.

Bekey, G., Kim, J., Gronley, J., Bontrager, E. and Perry, J. (1992) GAIT-ER-AID: An expert system for diagnosis of human gait. *Artificial Intelligence in Medicine* **4**, 293-308.

Edelmann-Nusser, J., Hohmann, A. and Henneberg, B. (2002) Modelling and prediction of competitive swimming performance in swimming upon neural networks. *European Journal of Sport Science* **2(2)**, 1-10.

Kelso, J. (1995) *Dynamic Patterns: The Self-organization of Brain and Behavior*. MIT Press, Cambridge, MS, USA.

Lapham, A. and Bartlett, R. (1995) The use of artificial intelligence in the analysis of sports performance: A review of applications in human gait analysis and future directions for sports biomechanics, *Journal of Sports Sciences* **13**, 229-237.

Lees, A. and Barton, G. (2005) A characterisation of technique in the soccer kick using

a Kohonen neural network analysis. In T. Reilly, J. Cabri and D. Araùjo (Eds), *Science and Football V*. Routledge, London, UK, 83-88.

Lees, A., Barton, G. and Kershaw, L. (2003) The use of Kohonen neural network analysis to qualitatively characterize technique in soccer kicking. *Journal of Sports Sciences* **21**, 243-244.

Perl, J. (2001) Artificial neural networks in sports: New concepts and approaches. *International Journal of Performance Analysis in Sport* **1** 106-121.

Perl, J. (2005) A [sic] computer Science in sport: An overview of present fields and future applications (Part II). *International Journal of Computer Science in Sport* **4(1)**, 35-45.

Perl, J. and Weber, K. (2004) A Neural Network approach to pattern learning in sport. *International Journal of Computer Science in Sport* **3(1)**, 67-70.

Schöllhorn, W. and Bauer, H. (1998) Identifying individual movement styles in high performance sports by means of self-organizing Kohonen maps. In H.J. Riehle and M. Vieten (Eds), *Proceedings of the XVI ISBS 98 Konstanz*, ISBS, Konstanz, Germany, 574-577.

Seifriz, F. and Mester, J. (2002) Modelling in sports: from mathematical fundamentals to applied use in mass media. *International Journal of Computer Science in Sport* **2**, 135.

Yan, B. and Li, M. (2000) Shot put technique using an ANN AMT model. In Y. Hong and D. Johns (Eds). *Proceedings of the XVIII International Symposium on Biomechanics in Sports, Volume 2*. The Chinese University of Hong Kong, Hong Kong SAR, China, 580-584.

Yan, B. and Wu, Y. (2000) The ANN-based analysis model of the sports techniques. In Y. Hong and D. Johns (Eds). *Proceedings of the XVIII International Symposium on Biomechanics in Sports, Volume 2*. The Chinese University of Hong Kong, Hong Kong SAR, China, 585-589.

THE IMPACT OF INTRODUCING REAL-TIME COMPUTER VISION TECHNOLOGY INTO AN ELITE SPORT TRAINING ENVIRONMENT

Kerr, R.F.

University of Canterbury, New Zealand

ABSTRACT

This paper describes the introduction of a computer vision system into a sports training environment. It is generally assumed that because a computer application is designed specifically to enhance sporting performance, the coaches and athletes will find it useful and will incorporate it into their training regime. In this paper, the extent and way in which a real-time computer vision system was used by coaches and athletes, who had no experience using computer systems before, is examined. The environment chosen was a high level gymnastics club that introduced a computer vision replay system to assist their high performance trampolinists and rhythmic gymnasts. This system allowed athletes to view their training continuously on the screen in real time or seconds after their "turn" enabling coaches and athletes to review their performances throughout each training session. Six weeks of detailed field observations and interviews with the coaches were undertaken with a view to describing how the system was integrated into their training sessions. The results show that the system was adopted by the two groups in quite different ways over the six week period, which were different again to the way in which the designer envisioned it being used.

KEYWORDS

sport technology, video feedback, gymnastics, trampolining, actor network theory.

INTRODUCTION

Sociologists of sport are increasing concerned with the importance of science and technology in the production of sport, particularly at the elite level. Ethical debates discuss whether doping, genetic manipulation and technology have now reached a level where sport is about the best athlete's technology as opposed to the best athlete (McCrory 2003, MacCarthur and North 2005, Miah 2004, Patel and Graydanus 2002, Randerson 2004, Tenner 1995: 30). Equipment and clothing design are shown to be highly influential in increasing sporting performance (Robinson et al 2002, Gunston 2005, Randerson 2004, Jenkins 2002: 65) as are high tech devices such as heart monitors, pedometers and complex software (Earnest 2005, Schnirring 2004).

Sociologists of sport discuss the large amount of time spent training by elite athletes who use this equipment. Researchers have examined the effects of training on such areas as adult social relationships (Drummond 2002, Hemery 1986), youth peer relationships (Patrick 1999, Donnelly 1993, Duncan 1997, Coakley 1993, Weiss 2000) and post-competition life (Schmitt and Leonard 1986). Yet, as Butryn points out, "…there has been little attempt on the part of sport sociologists to investigate athletes' interactions with perceptions of technology." (Butryn 2003: 19). Technology appears to be unquestioningly accepted as an integral part of the production of elite sport with little explanation on how it is used.

As Butryn notes, the question of how technologies and equipment are integrated into and used in the competitive training environment is rarely discussed. The scientists and designers of sporting technology often work in laboratories, quite separate from the gymnasiums, courts or fields where sport training generally occurs. How then, do the new ideas developed by scientists, reach the sport training ground? As Sands points out, it can be difficult for coaches to establish whether a new piece of technology will be beneficial to their athletes (Sands 1999). It may or may not be something that will assist with that particular training programme. Moreover, even once introduced, the coaches must then be taught how to use the new technology to its maximum effectiveness (Lim 2005).

In this study, a real time computer vision system providing instant video feedback was introduced to a sports training environment. The effectiveness of video feedback in sports coaching is well documented and therefore not specifically examined here (Darden 1999, Holcomb 2002, Jambor 1995, Seifried 2005). Instead, the focus is on how was this technology integrated into a sports training environment and how, why and whether it became a regular part of the training programme.

METHOD

The environment chosen for this research was a gymnastics club which had high performance competitive programmes for several types of gymnastics and was about to introduce a computerised replay system into the gymnasium to assist with training. The new system was specifically being introduced for the trampolinists and the rhythmic gymnasts, who were the highest level of gymnasts in the club and, club officials reasoned, would benefit most effectively from the new technology (Interview with the software designer, July 2005).

The methodology chosen for this project was actor network theory, which involved extended fieldwork, observations and interviews. Following authors such as Latour and Woolgar, the aim was to "learn from the actors". However, for these authors, the word “actor” refers to both humans and non-humans and both are understood together to create their own networks, their own language and metalanguage (Miller 1997: 360). They argue that fieldwork, making close observations, makes it possible to illuminate how things are assembled or disassembled through following the actors through the labyrinths or networks they construct (Austrin and Farnsworth 2005).

Due to the emphasis on the significance of both human and non-human entities, these methods have become very effective in technology studies. Yet despite the increased importance of "non-humans" such as equipment and sport science technologies, the methodology has never been applied to sport. This research uses these methods to examine the impact and integration of a new piece of video related software into an elite training environment. Field observations of training in the club occurred during several weeks prior to the introduction of the system and for six weeks afterwards. Interviews were undertaken with the trampolining and rhythmic coaches and the software designer, as well as some informal discussions with the athletes. The observations between the two areas were compared.

RESULTS

Before we can examine how the introduction of technology altered training, it is necessary to first understand how training worked in the different disciplines prior to the system being integrated.

Trampolining



Figure 1: The trampolining area of the gym.

There are between six to eight females in the group at each session, aged between 14 and 29. It is crucial to note that at all times, only four trampolinists are actually on the trampolines, as it is only possible to have one person on each of the four trampolines at a time. This means that at all times, there are two to four trampolinists doing something other than bouncing on the trampoline. When not on the trampolines, the trampolinists tended to go to the bathroom, talk to each other, write in their books or, "throw in" one of the blue or green safety mats.[1] At all times, the coach stood on the floor area, a little bit back from the trampolines and watched the two trampolines closest to the wall.

In a typical training session, the trampolinists performed stretching and strengthening exercises[2] on the floor area beside the trampolines[3] for about 20 minutes. They then began bouncing on the trampolines, first doing simple moves like jumping up and down keeping their bodies straight, then moving to more complex somersaults with multiple twists. Each turn on the trampoline takes about five minutes. They spent some turns on one or two skills, and some turns doing their whole ten skill routine.

---

[1] The mat is "thrown in" during the difficult move so that the trampolinist performing the move lands on the mat, not on the trampoline, cushioning their bounce and ensuring they land more safely.

[2] These consist of stretches for the hamstrings, shoulders, splits, quadriceps, back and stomach. Strengthening exercises are generally sit-up type exercises for the abdominal muscles, and push up type exercises to build arm muscles.

[3] The "floor area" is the official artistic gymnastics floor which is shared with artistic gymnasts at this time.

Between each five minute turn, they got off the trampoline, did something else, then returned for another turn usually about ten minutes later. In interviews, it was revealed that trampolinists generally need to use the bathroom every ten minutes because bouncing puts pressure on the bladder, so it is quite normal for a trampolinist to leave the trampolining area and use the bathroom after almost every turn on the trampoline (Interview trampoline coach, May 2005).

Rhythmic Gymnastics



Figure 2: The rhythmic gymnastics floor

Similar to trampolining, there were about six to eight females in the high performance group. They are a little younger than the trampolinists, aged between 11 and 16. In a typical training session, the gymnasts arrived, placed their bags and apparatus in the storeroom and then spent one hour on warm up exercises in the ballet room, followed by stretching[4].

After stretching, they moved to the rhythmic floor and spent half an hour to an hour on specific skills. They usually lined up on the edge of the rhythmic floor and performed rows of specific balances, turns and leaps or walkovers. The coach stood watching and continuously issued verbal corrections. Alternatively they used their apparatus (hoop, ball, rope or ribbon) to practice throw skills on the mat. Then they spent the remaining two hours or so practicing routines.

In contrast to the trampolining program, in the rhythmic training there was often more than one person working with the girls. There were always one or two coaches present, and often a third person, a judge, who although she did not direct the girls, continually pointed out corrections and clarified judging rules with the coaches[5]. Routines were

[4] The rhythmic gymnasts spend a great deal of time practicing the splits and bridges. They practice splits with one foot placed up on a chair and hold for several minutes, plus many types of exercises to increase their back flexibility.

[5] In an interview, it was explained that rhythmic coaches must submit a detailed judging sheet describing exactly what their gymnasts will be performing in their routines prior to competing. These means the coach needs a substantial amount of help from judges to plan routines and complete these sheets.

done with one girl performing a routine to music on the floor with her apparatus, while the others practiced parts of their routines around her on the floor, moving out of her way if she comes near them. They each took turns (for about 1.5 minutes) to do one routine, and then returned to practicing individual skills and skill sequences. Unlike the trampoline, the rhythmic floor is able to have many people on it at once and all the girls are constantly on the floor and all practicing rhythmic gymnastics at all times.

The introduction of the Computer Replay System

In May 2005, two computer vision replay systems were placed in the gymnasium to provide constant real time video feedback to the athletes and their coaches on a continual basis. The systems could be set to show a continuous replay of the athlete from several seconds to several minutes ago[6]. Coaches at all levels of sport regularly use videotaped performances and practices to assess, demonstrate to, and motivate their players (Seifried 2005:36). In this case, the designer, as a club official, believed that creating a permanently set up video system would assist athletes in improving performance. He placed one monitoring the trampolining area and another monitoring the rhythmic gymnastics floor.

Trampolining with the system

In trampolining, during the first week, the new system was of great interest to the trampolinists. The trampolinists used it by simply watching the screen as soon as their turn on the trampoline was finished, and only then going to the bathroom or throwing the mat in as normal. They rushed with great interest to the screen to see themselves, however what they saw wasn't always what they expected to see. One of the older trampolinists remarked after her turn "I can't wait to see that move on the computer, it was perfect!" However, upon watching herself on the computer, discovered that she was very far from perfect, expressed great disappointment with herself and surprise that it "felt" perfect to her in the air yet was clearly not so to the eye (Field notes May 2005).

During the next six weeks, the computer system became a "normal" part of training for the trampolinists, though not all athletes used it identically. The older trampolinists, who were all also trained as coaches, used it to coach themselves. They would get off the trampoline after their turn and watch themselves without first speaking to the coach for feedback, often selecting slow-motion replay or pausing at certain frames. They might also discuss what they are seeing with another trampolinist or the coach, discussing ideas on how to fix what they could see was wrong on the monitor. One of the younger trampolinists, not at all trained in coaching, would not only watch herself. She would watch one of the better trampolinists and compare how they looked with herself. She would move the footage forward and backward frame-by-frame, flicking between the other trampolinist and herself (Field notes June 2005).

The trampolining coach described that since the new system went in:

> "The girls are taking more interest in self evaluation. They're more likely now to look at their skills from a judging perspective and compare them to other people's

[6] A webcam was placed on the wall next to the trampolines and connected to the computer. This system is less time consuming than setting up a hand held video camera as it needs no one to work it, but simply records and provides feedback on a continual basis.

> skills… They can make adjustments based on what they know to be correct." (Interview with trampolining coach, May 2005).

The coach, although at first showing some resistance to the new system, became gradually more used to it as the weeks went by and began referring the trampolinists to it more often.

The trampolining coach also noted that the older trampolinists are the ones who use the system the most because they have:

> "…the knowledge and experience to know how to use it better… the fact that all the trampolinists are at least level one coaches does give them an ability to self coach." (Interview with trampolining coach, May 2005).

The coach inadvertently points to the shifting of assemblages, coming about due to the older athletes' knowledge of coaching. The assemblages of trampolinist and coach were previously reliant on each other to produce results. With the computer system, the assemblage of trampolinist/computer system/coaching knowledge works to create high performance trampolining as effectively as the previous assemblage of trampolinist/coach.

This is confirmed by the coach's initial comment "if you girls don't get back to work, I'm going to turn that machine off forever!" Although the designer of the technology intended it to "augment" the coaching process, the coach instead immediately recognised that this new system had the ability to substitute for himself and was therefore threatened, although in time, he realised its value.

After six weeks, trampolining training had stabilised with the new system accepted as an important actor in the production of trampolining. The trampolinists believed the system to be allied to them in showing them faults in their body shapes which the coach may or may not point out. The coach believed the system to be his ally as it is able to prove to the trampolinists that what he is saying about their performance was true. Thus, although initially the new actor disrupted the trampolining network through being perceived as a competitor to the coach, the network soon reassembled and stabilised.

Rhythmic Gymnastics with the system

The rhythmic gymnasts, unlike the trampolinists, had no "off time" away from the rhythmic floor in which to watch themselves on the computer. Therefore, the coaches incorporated the computer system in the gym in a different way.

Sessions worked with one gymnast doing a routine, which one coach watched. At the end of the routine, the coach would take the gymnast over to the corner of the mat where the computer was to watch the routine on the screen together. In the meantime, the other coach or the judge would watch the next girl's routine. So at all times, there would be one coach and one gymnast at the computer, while another gymnast did a routine and was watched by the other coach. In between their turns with the coaches, the gymnasts were almost always on the rhythmic floor practising parts of their routines, often watched by the judge. However, as the weeks since the introduction of the new

system went by, the system was used less and less. One of the coaches also took a holiday, and while she was away the system was not turned on. After six week, the system was rarely used.

DISCUSSION

There are several possible explanations for why the computer system was so easily adapted into the trampolining environment as opposed to the rhythmic gymnastics environment.

The most obvious explanation is that in trampolining, there is always "downtime" which allowed for the computer system to be used by the athletes. The shape of the trampoline means that only one trampolinist can work on each trampoline at a time, and with more than four people in the training group, this necessitates that there are always trampolinists not on the trampoline. In this particular case, the trampolinists had no specific tasks to fulfil when not on the trampoline and therefore, had time to use the computer system. Training did not need to be altered in order for the system to be effective. By contrast, the rhythmic gymnasts lacked the time to look at the computer screen. They were required to be training more or less constantly, allowing no time during the training session for them to watch the screen.

Another possible explanation is the age and ability of the athletes. The trampolinists were generally older than the rhythmic gymnasts. As a result, several of these trampolinists also had experience in coaching. Therefore, it was easier for these athletes to use the computer system to coach themselves. The rhythmic gymnasts, being younger and without any coaching experience, may not have known what exactly to look for when watching themselves on the screen, therefore the coach felt the need to watch the routine on the computer with them. Other studies suggest however, that it is the ability of the performers that impacts the effectiveness rather than the age (Darden 1999), and in this study, it appears to be related to both coaching ability and age.

The attitude of the coaches is also significant, which is consistent with other findings (Darden 1999). In trampolining, although the coach was originally not enthusiastic about the system, he was prepared to try it and by the end of six weeks had concluded that it was a vital coaching aid. By contrast, in an interview with the rhythmic coach, it was revealed that she actually had quite strong negative opinions about the use of video feedback:

> "I found in the past that gymnasts, especially when they first begin watching themselves on video, are very harsh and only see the worst parts. And it can be, for some girls, particularly those perfectionists, can actually be a negative thing, because they don't like what they're seeing." (Interview with rhythmic gymnastics coach)

Her belief that video feedback tends to be negative would undoubtedly have influenced the amount of use the system was given. Darden (1999) discusses how the negativity experienced when first using video must be overcome in order for video feedback to be effective. It is possible that if this coach were to use the system over a longer period, she may see her athletes overcome this feeling and find the system more useful.

CONCLUSIONS

The results of this study suggest that when creating or purchasing technology to enhance sporting performance, it is not enough to consider only whether the technology will be effective technically. The way that training sessions operate, the views of the coaches and the age and abilities of the athletes appear to be highly relevant in predicting whether a new piece of technology will be useful. It is likely that a closer relationship between the designers of technology, coaches and athletes would be beneficial to creating technology that can be incorporated smoothly into the sports training environment.

REFERENCES

Austrin, T. and Farnsworth, J. (2005) Hybrid genres: fieldwork, detection and the method of Bruno Latour. *Qualitative Research* **5,** 147-165.

Butryn, T. (2003) Posthuman podiums: cyborg narratives of elite track and field athletes. *Sociology of Sport journal* **20,** 17-39.

Butryn, T. and Masucci, M. A. (2003) It's not about the book: A cyborg counternarrative of Lance Armstrong. *Journal of Sport and Social Issues* **27,** 124-144.

Callon, M. (1991) Techno-economic networks and irreversibility. In: *A Sociology of Monsters: Essays on Power, Technology and Domination.* Ed: Law, J. London and New York: Routledge, 132 – 164.

Coakley, J. (1993) Social Dimensions of intensive training and participation in youth sports. In: *Intensive participation in children's sports.* Ed: Cahill, B. A. Champaign: Human Kinetics Books.

Darden, G. (1999) Videotape Feedback for Student Learning and Performance: A Learning-Stages Approach. *JOPERD--The Journal of Physical Education, Recreation & Dance* **70**, 40.

Donnelly, P. (1993) Problems associated with youth involvement in high performance sport. In: *Intensive participation in children's sports.* Ed: Cahill, B. A. Champaign: Human Kinetics Books. Champaign, Human Kinetics Books.

Drummond, M. (2002) Sport and images of masculinity: the meaning of relationships in the life course of "elite" male athletes. *Journal of men's studies* **10,** 129.

Duncan, J. (1997) Focus groups interviews with elite young athletes, coaches and parents. In: *Young people's involvement in sport.* Ed: Kremer, J., Trew, K., and Ogle, S. London: Routledge.

Earnest, C. P., Wharton, R. P., Church, T. S. and Lucia, A. (2005) Reliability of the load excalibur sport ergometer and application to computrainer electromagnetically braked training devices. *Journal of strength and conditioning* **19,** 344-349.

Guadagnoli, M., Holcomb, W. and Davis, M. (2002) The efficacy of video feedback for learning the golf swing. *Journal of Sports Sciences* **20**, 615.

Gunston, R. (2005) Play ball! How sports will change in the 21st century. *The Futurist,* Jan/Feb 2005**,** 31-36.

Haraway, D. (1991) A cyborg manifesto: science, technology, and socialist-feminism in the late twentieth century. In: *Simians, Cyborgs and Women: The Reinvention of Nature.* Ed: Haraway, D. New York City: Routledge.
Hemery, D. (1986) *The pursuit of sporting excellence: a study of sport's highest achievers.* Champaign, Human Kinetics Books.

Jambor, E. and Weekes, E. (1995) Videotape feedback: make it more effective *JOPERD--The Journal of Physical Education, Recreation & Dance* **66**, 48.

Jenkins, M. J. (2002) Advanced materials and sporting performance. *Interdisciplinary Science Reviews* **27,** 61-65.

Johnson, J. (1988) Mixing humans and non-humans together: the sociology of a door closer. *Social problems* **35,** 298-310.

Latour, B. (1991) Technology is society made durable. In: *Sociology of monsters: essays on power, technology and domination.* Ed: Law, J. London: Routledge.

Latour, B. (1995) A door must be either open or shut: a little philosophy of techniques. In: *Technology and the politics of knowledge.* Ed: Feenberg, G.A. and

Hannay, A. Bloomington: Indiana University Press.
Law, J. (2004) *After method: mess in social science research,* London, Routledge.

Liebermann, D., Katz, L., Hughes, M., Bartlett, R., McClements, J. and Franks, I. (2002) Advances in the application of information technology to sport performance. *Journal of sport science* **20,** 755-769.

Liebermann, D., Katz, L., Sorrentino, R. (2005) Experienced coaches attitudes towards technology *International journal of computer science and sport* **4,** 21.

Lim, J. (2005) Preparing preservice teachers to effectively use technology as teaching tools. *Research quarterly for exercise and sport* **76,** 81-83.

Macarthur, D. G. and North, K. N. (2005) Genes and human elite athletic performance. *Human Genetics* **116,** 331.

McCrory, P. (2003) Super athletes or gene cheats? The threat of gene transfer technology to elite sport. *British Journal of sports medicine* **37,** 192-194.

Miah, A. (2004) *Genetically modified athletes: biomedical ethics, gene doping and sport,* London, Routledge.

Miller, P. (1997) The multiplying machine. *Accounting, organisations and society* **22,** 355-364.

Patel, D. R. and Graydanus, D. E. (2002) Genes and athletes. *Adolescent Medicine* **13,** 249-257.

Patrick, H., Ryan, A., Alfeld-Liro, C., Fredericks, J., Hruda, L. and Eccles, J. (1999) Adolescents' commitment to developing talent: the role of peers in continuing motivation for sports and the arts. *Journal of youth and adolescence* **28,** 741.

Randerson, J. (2004) Triumph of the machines. *New Scientist* **183,** 31-34.

Robinson, M. G., Holt, L. E. and Pelham, T. W. (2002) The technology of sprint racing canoe and kayak hull and paddle designs. *International Sports Journal* **6,** 68-86.

Sands, W. (1999) Communicating with coaches: envisioning data. In: *International symposium on biomechanics in sports.* Ed: Sanders, R. and Gibson, B. Perth: Edith Cowan University,.

Schmitt, R. and Leonard, W. (1986) Immortalising the self through sport. *The American journal of sociology* **91,** 1088 - 1111.

Schnirring, L. (2004) Core temperature goes high. *The Physician and sports medicine,* **32,** 10.

Seifreid, C. (2005) Using videotaped athletic contests within Mosston's teaching methods. *Journal of Physical Education, Recreation and Dance* **76,** 36-39.

Tenner, E. (1995) The technological imperative. *The Wilson Quarterly* **19,** 26.

Varney, W. (2002) Tumbling into gendered territory: gymnastics and its technologies. In: *Sport technology: history, philosophy and technology.* Ed: Miah, A. and Eassom, S. London, Elsevier Science.

Weiss, M. (2000) Culture, context and content analysis: an exploration of elite women gymnasts in the high school world. In: *Games, sports and cultures.* Ed: Dyck, N. Oxford, Berg.

Woolgar, S. (1996) Science and technology studies and renewal of social theory. In: *Social theory and sociology: the classics and beyond.* Ed: Turner, S. Cambridge: Blackwells.

BIOMECHANICALLY OPTIMISED REAL-TIME SYNCHRONISATION OF SKILLS IN NOVICE AND EXPERT VIDEO SEQUENCE PAIRS FOR AUTOMATIC FEEDBACK OF POSE SHAPE AND PHASE ERRORS

Lee, K.[1] and **Green, R**.[1]

[1] Department of Computer Science, University of Canterbury, New Zealand

ABSTRACT

Synchronising video of a movement skill with that of an expert currently requires manual alignment and scaling of two video sequences. This is problematic because different athletes perform skills at different speeds and even body position phase relationships vary during execution. This paper introduces a method for temporally and spatially synchronising pairs of image sequences of human body movement in real-time while normalising for athlete size. Real-time image sequences summersaults are used in this case study. Based on biomechanical features calculated from movement sequences, a detailed implementation is described together with measurement of the difference between two moving objects, performance, and limitations of the proposed algorithms. Skill pairs are either temporally synchronised using motion segmentation or spatially synchronised based on pose alignment to enable feedback of both pose phase errors and pose shape errors respectively. The results indicate that the real-time temporal and spatial synchronising algorithms introduced here automatically synchronise movement in novice and expert video sequence pairs for useful pose shape and pose phase feedback to athletes.

KEY WORDS

computer vision, motion analysis, sport coaching, human body, temporal synchronisation

INTRODUCTION

Given image sequences of a novice and expert performing the same skill, one of the most basic and useful tasks would be spatially and temporally matching the videos by synchronising their motion. However, although this operation is cognitively easy for humans, implementing it on a computer system is a challenging problem. In general sequences of motion, synchronising two similar moving objects, which are varying in their speed and the direction could be a difficult task, if it has to be done manually by a simple video editing tool. Having a robust method to automatically track objects and synchronise videos based on some invariant features is the goal of this research.

In computer vision literature, modelling the human body and its motion is an area of great interest to many researchers, as a survey by Moeslund and Granum (2001) shows. The research can be roughly grouped into two parts: tracking method with model reconstruction, and movement recognition. Already impressive results have been reported in the recognition area alone. Yamato et al. (1992) reported successful recognition of six different tennis strokes among three subjects. Starner and Pentland

(1995) also worked on hand gesture system, American Sign Language, with a 99% successful recognition rate. These are based on the use of Hidden Markov Models (HMM). Recently, Green and Guan (2004a, 2004b), proposed a more unified and general framework for recognising human movement patterns which are also based on HMM.

For tracking the human body, a popular method is CONDENSATION (CONditional DENSity propagATION) algorithm (Isard and Blake, 1998), or also known as particle filter (Deutscher et al., 2000). Because of multiple predictions for each variable being tracked, this algorithm is more robust to heavy background clutter than the traditional Kalman filter (Kalman, 1960).

In this paper, we demonstrate the experiment with "salto" as our case study example, which is a somersault with a full 360° rotation about the transverse axis of the body, performed on a trampoline. With video to be synchronised consisting of relatively simple and elegant movement sequences, the salto is an ideal choice for the study of human motion in a controlled indoor environment. This research proposes a computationally efficient temporal synchronising algorithm using the centre of mass and principal axis calculated in real-time from a sequence of images.

BACKGROUND

It is assumed that only one main moving object is captured by a fixed camera environment. This assumption is necessary when applying a Double Difference Algorithm (Aggarwal and Cai, 1999) to obtain an object motion mask.

The entire region of the moving object (i.e., the gymnast performing a salto) stays within the image boundary, so that the centre of mass can be always computed correctly.

Overview of the Procedure

An outline of the general procedure of this experiment is:

1. Apply double difference algorithm (DDA) to isolate the moving object.
2. Apply morphological closing operation to obtain smoothed object (Shapiro and Stockman, 2001).
3. Calculate the centre of mass and principal axis angle.
4. Apply heuristics to adjust principal axis angle.
5. Calculate the alignment error of two salto actions frame by frame, spatially and temporally aligned on the centre of mass.
6. Calculate the alignment error of two salto actions frame by frame, spatially aligned on the centre of mass and temporally aligned by the principal axis.

Kinematics of Motion

We exploit the kinematics of saltos by observing that the following two features are the most important parameters: (1) centre of mass and (2) principal axis.

Centre of Mass

According to the law of Newtonian physics, a system of particles can be regarded as a single particle at the position of the centre of mass $\vec{r}_{cm}$ with the mass M, which is the sum of all the mass of the system of particles ($M = \sum_{1}^{k=n} m_k$). The following equation establishes the position of the centre of the mass by:

$$M\vec{r}_{cm} = m_1\vec{r}_1 + m_2\vec{r}_2 + \cdots + m_n\vec{r}_n$$

The path of its centre of mass can be replaced as a single particle with the sum of mass at the centre and determined by the initial velocity. As shown Figure 1, the lower graph line represents the vertical component of the centre of mass position over the horizontal axis which is the timeline based on frame index (33mS/frame @ 30fps). Like any other object thrown into air, the familiar parabolic path of motion is followed using physics based model of flight under gravity (neglecting drag) (Halliday et al., 1993).

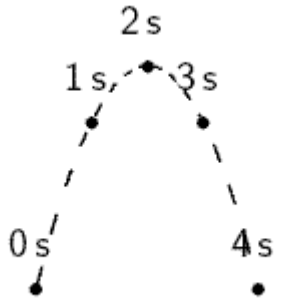


Figure 1: An example of parabolic flight path of the centre of mass.

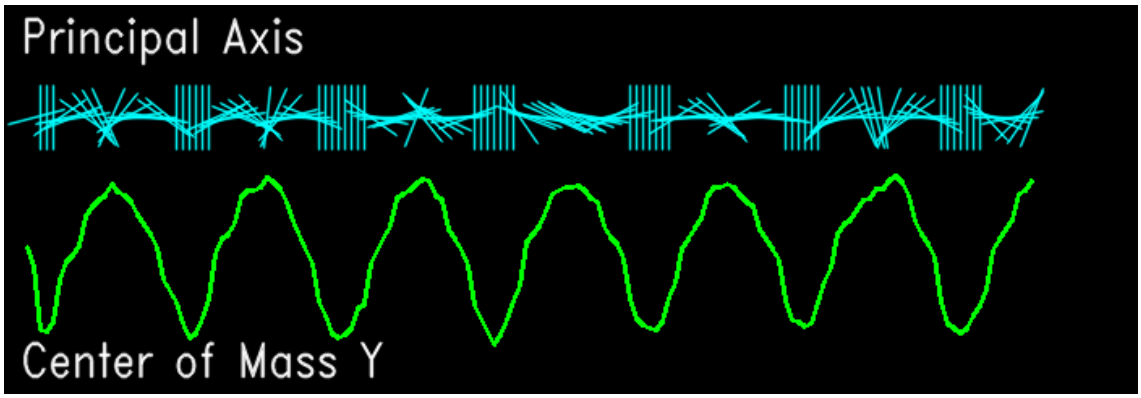


Figure 2: Plot of multiple saltos. Centre of mass *Y* and principal axis angle is represented on frame index (33mS/frame).

Principal Axis

The angle of the principal axis of an object is based on the rotational inertia and is calculated as follows using a special case equation (Jain et al., 1995) of 2D binary image:

$$\tan^2\theta + \frac{\mu_{20} - \mu_{02}}{\mu_{11}} \tan\theta - 1 = 0 \tag{1}$$

where the second order of moments μ_{20}, μ_{11} and μ_{02} are defined as:

$$\mu_{20} = \sum_{i=1}^{m} \sum_{j=1}^{n} B[i,j](x_{ij} - \bar{x})^2 \tag{2}$$

$$\mu_{11} = \sum_{i=1}^{m} \sum_{j=1}^{n} B[i,j](x_{ij} - \bar{x})(y_{ij} - \bar{y}) \tag{3}$$

$$\mu_{02} = \sum_{i=1}^{m} \sum_{j=1}^{n} B[i,j](y_{ij} - \bar{y})^2 \tag{4}$$

and $B[i,j]$ is intensity value (0 or 1) of an m-by-n binary image at (i,j), x_{ij} and y_{ij} are the vertical and horizontal coordinates of the image, respectively. $\bar{x}$ and $\bar{y}$ are the centre of the mass of horizontal and vertical spatial position, respectively.

Solving Equation 1 yields:

$$\theta = \frac{1}{2} \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \tag{5}$$

Calculating the angle of the principal axis has a number of issues. Unless it is masked out, the motion of the trampoline bed can generate horizontal noise as a part of the moving object causing errors of the angle as high as 90˚, even though the body posture on the bed is close to the straight upright position.

Another problem is the range of the angle θ ($0^\circ < \theta \leq 180^\circ$) from Equation 5, which corresponds to the upper or lower half of the full circle. This causes an 180˚ ambiguity of upside-down between the two videos. To solve this problem, a constant angular velocity is assumed (or at least the conservation of angular momentum) so any sudden change of angle greater than 90˚ is assumed to be a shifting to the upper or lower range of circle.

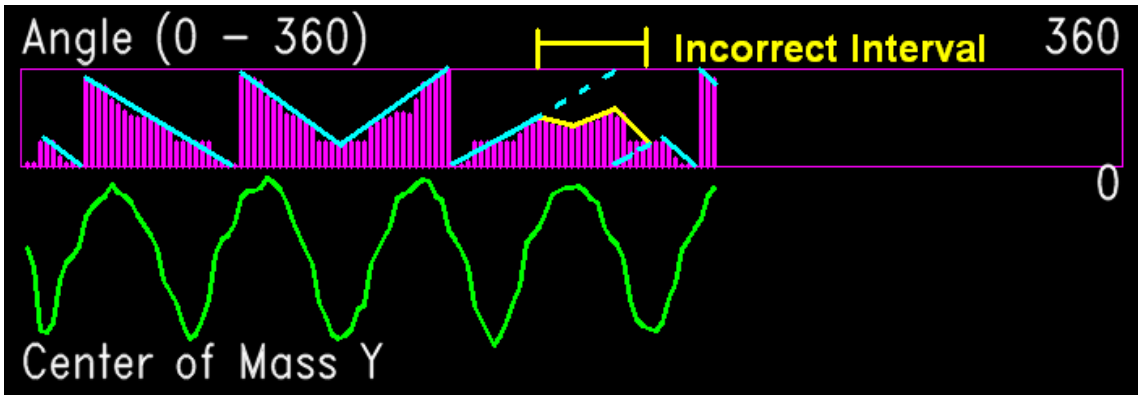


Figure 3: An example of principal axis angular velocity flipping 180˚ angle. The dotted-line represents the projection of the expected angle.

Finally, a tucked body shape in the salto action can cause the minimum and maximum axes to have similar values, which results in a poor estimate of the principal axis. As shown in Figure 3, at the fourth peak of parabolic graph, the angle begins to decrease. This eventually loses the tracking angle and can flip to an opposite angle causing an error of 180˚. The conservation of angular momentum L can be described as:

$$L = I\omega = \text{a constant}, \tag{6}$$

or

$$I_i\omega_i = I_f\omega_f\,. \tag{7}$$

where I is the rotational inertia and ω is the angular velocity.

While an object in flight may change the rotational inertia from its initial value I_i to a smaller or larger value I_f, the rotational inertia cannot become negative value. Therefore, while the angular speed can be increased or decreased (Figure 4), the direction of rotation cannot reverse while in flight.

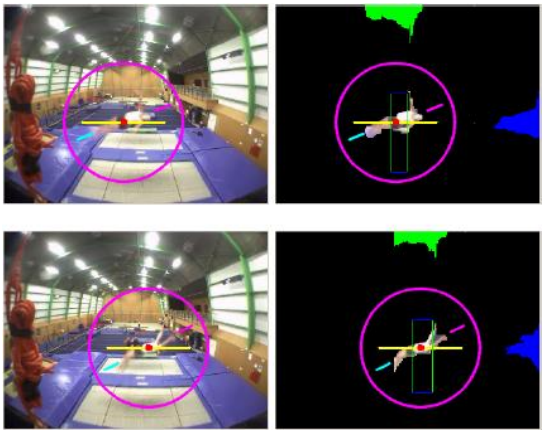


Figure 4: Tracking of centre of mass and principal axis through two video sequences.

In other words, the sign of the angular velocity (i.e. the tangent of the upper graph in Figure 3) cannot be changed, and the angle should be either monotonically increasing or decreasing. The angular velocity direction is allowed to change only at the lowest point of the centre of mass while in contact with the trampoline bed.

This problem is overcome by (1) maintaining the sign of the angular velocity until next landing, or by (2) calculating the angular momentum. When the rotational momentum becomes as small as a certain threshold value, we interpolate the angle based on the current angular velocity.

RESULTS

Performance

Intel OpenCV Library and Microsoft MFC were used on a 2GHz PC with 512MB memory. The 30fps video clips were adjusted below 20 frames per second to avoid dropped frames returned by a slow OpenCV frame grabbing function.

Experimental Result

Figure 5 shows a number of features of the given frame that may be useful for synchronisation. The histogram on the right and the histogram at the top are the projections of images onto a line obtained by summing rows of pixels and summing columns of pixels. The thick white horizontal line inside the circle is a visual cue for 0˚. From the two line segments in the circle, one can see the principal axis. A projection image histogram is used to identify sudden noise increases such as movement of the trampoline bed.

The box represents the maximum and the minimum centre of mass over all the frames. To compute these values, all frames need to be passed once. During the first pass, the frame indices of the local extremum – the highest and the lowest point – of the centre of mass, are calculated.

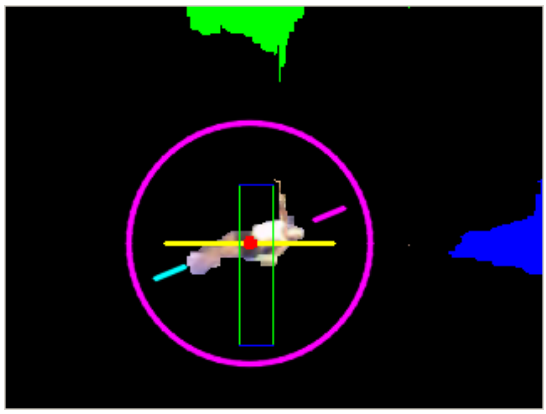


Figure 5: An analytical version including: the maximum and the minimum centre of mass, principal axis, and histograms of projection.

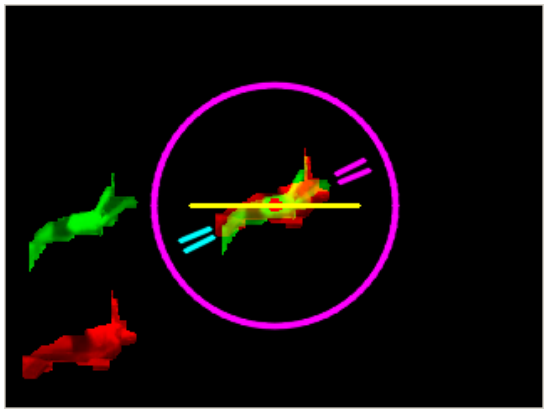


Figure 6: An image with two gymnasts overlapped.

Measurement of Difference

In this experiment, the measure of how closely the two human body images match is calculated as the ratio of the number of pixels overlapped (intersection) to the total number of non-zero intensity valued pixels (union). In Figure 7, the last third column images show overlap of the first and the second column images. The first row is

overlapped at the centre of mass, while the second row images are aligned on the principal axis as well. Clearly, in this synchronisation, the overlapped image with both centre of mass and the principal axes aligned together produce an improvement of 45% that without alignment of the principal axis gained only a 30% to 15% synchronisation improvement.

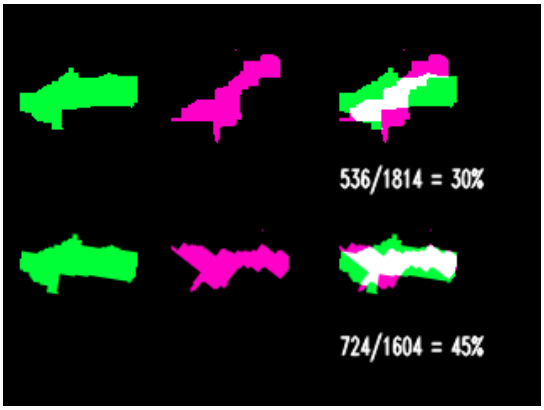


Figure 7: The first row illustrates the alignment error using centre of mass synchronisation. The second row illustrates the alignment error using principal axis synchronisation.

CONCLUSION

The experiment demonstrates the detailed discussion of computing kinematics features of motion in image sequences. The results indicate that the proposed algorithm is a useful method for matching two different motion sequences for comparison and analysis.

We have introduced methods for calculating principal axis angles, and a set of novel methods to measure the difference of two image sequences. This temporal synchronisation algorithm could also be applied to a number of other applications such as video retrieval.

However, the lack of an explicit human model increases the alignment error since an articulated body model is required for more accurate body alignment.

In future work, this method will be extended to include approaches using particle filters (Deutscher et al., 2000) with an articulated human body model (Plänkers and Fua, 2001, Rehg and Kanade, 1995).

REFERENCES

Aggarwal, J. and Cai, Q. (1999) Human Motion Analysis: A Review. *Computer Vision and Image Understanding: CVIU.* **73(3),** 428-440.

Deutscher, J., Blake, A., and Reid, I. (2000) Articulated Body Motion Capture by Annealed Particle Filtering. In *CVPR*. 126-133.

Green, R. and Guan, L. (2004a) Quantifying and Recognizing Human Movement Patterns From Monocular Video Images – Part I: A New Framework for Modelling

Human Motion. *IEEE Transactions on Circuits and Systems for Video Technology,* **14(2),** 179-190.

Green, R. and Guan, L. (2004b) Quantifying and Recognizing Human Movement Patterns From Monocular Video Images – Part II: A New Framework for Modelling Human Motion. *IEEE Transactions on Circuits and Systems for Video Technology,* **14(2),** 191-198.

Halliday, D., Resnick, R., and Walker, J. (1993) *Fundamentals of Physics.* John Wiley & Sons, 4th edition.

Isard, M. and Blake, A. (1998) CONDENSATION – conditional density propagation for visual tracking. *International Journal of Computer Vision.* **29(2)**, 5-28.

Jain, R., Kasturi, R., and Schuck, B. (1995) *Machine Vision.* McGraw-Hill.

Kalman, R. (1960) A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME—Journal of Basic Engineering.* **82**(Series D), 35-45.

Moeslund, T. and Granum, E. (2001) A Survey of Computer Vision-Based Human Motion Capture. *Computer Vision and Image Understanding: CVIU.* **81(3),** 231-268.

Plänkers, R. and Fua, P. (2001) Articulated Soft Objects for Video-based Body Modeling. In *ICCV*. Vancouver, Canada.

Rehg, J. and Kanade, T. (1995) Model-Based Tracking of Self-Occluding Articulated Objects. In *ICCV*. 612-617.

Shapiro, L. and Stockman, G. (2001) *Computer Vision.* Prentice-Hall.

Starner, T. and Pentland, A. (1995). Real-Time American Sign Language Recognition From Video Using Hidden Markov Models. In *SCV95.* page 5B Systems and Applications.

Yamato, J., Ohya, J., and Ishii, K. (1992) Recognizing Human Action in Time-Sequential Images using Hidden Markov Model. In *Proc. Conf. on Computer Vision and Pattern Recognition.* 379-385.

PUBLICISING MATHEMATICS AND STATISTICS THROUGH SPORT

Clarke, S.

Faculty of Life and Social Sciences, Swinburne University of Technology, Hawthorn, Vic., Australia

ABSTRACT

With Australian students shunning the mathematical sciences, it is necessary to take every opportunity to publicise interesting mathematical applications. Publicising research in the popular media can also have important spin-offs for the researcher as well as his department and institution. It will always be important to publish in the relevant scientific journal, to be read by the few colleagues who work in your area, and gain some research points for your University. But it may also be advantageous to get WWW, newspaper, radio or TV coverage. Among the millions of readers or watchers there will be colleagues, superiors, future students, possible consulting or research project clients who would otherwise not learn about your work.

In Australia, interest in sport is almost universal, and the anxiousness of the media to publish sporting material can be used to overcome their usual avoidance of anything scientific. Working in the application area of sport gives us a huge advantage in gaining media coverage for our research. Using personal experiences, this paper will discuss some of the ways we can realise that advantage. The use of personal contacts, press releases, web sites, media guides and special events will be discussed, along with some of the necessary characteristics and pitfalls of being a media junkie.

KEY WORDS

media, publicity

INTRODUCTION

A month after Don Forbes joined Swinburne Sports Statistics to start a PhD on Australian Rules football, he was appearing on national television discussing results of his research. However he was not discussing football, but his first love of horse racing. This followed an enquiry from Today/Tonight on whether we did any research on horse racing, and could supply predictions for the Melbourne Cup. While at that stage we had no relevant prediction models, Don was happy to spend a couple of days developing some simple regression models to forecast the race. This example illustrates our willingness to satisfy media requests if possible.

In contrast to many researchers, we actively seek media coverage of our research results. Don has subsequently had television coverage of his PhD work. Michael Bailey has also appeared on national television, and along with all my other students has done many radio interviews. We believe such coverage aids the individual researcher in their career, and enhances the reputation of the researcher's institution. It is a strategy we would like other contributors to this conference to follow.

METHODS

Working in sport in Australia gives us a huge advantage. Newspapers devote a large proportion of their paper and often special lift out sections to sport; news bulletins always have a special section on sport; both radio and TV have live coverage of important sporting events; and there are even radio stations devoted entirely to sport. The emphasis Australians place on sport is illustrated by an example from the Melbourne based daily The Herald/Sun, the Monday after one of the most tumultuous weekends in Victorian politics. Over the weekend the Australian government had sent troops overseas to East Timor, which took 4 pages of coverage. In a huge upset, the seemingly invincible Kennett Government had fallen in a Victorian election loss, which took 17 pages of coverage. The space given to these events meant the Herald/Sun's coverage of sport was reduced to a mere 45 pages.

The exposure given to sport, generated no doubt by the interest of readers, listeners and watchers, means journalists are on the lookout for possible articles. This gives us the opportunity to demonstrate scientific methods to the general public.

Clarke (2002) details some of the ways of creating media interest for scientific research. Here we discuss these in a more parochial context.

www

The easiest way to publicise your work is through personal pages on the internet. This allows you complete control over content and timing, but may lack the flair and professional look that a journalist or regular media outlet could contribute. We have made great mileage out of our regular weekly sports predictions. These pages were among the most popular at Swinburne, with nearly half the hits coming from overseas. Some of these hits move on to other Swinburne pages on research or courses. Several of our Graduate students first learned of our work through these pages.
Most media outlets now also have associated web pages. These can create a more lasting record than a mere radio, TV or even newspaper interview. They are often more accessible to the general public, particularly those overseas. Allsopp and Clarke (2004) showed it was better for Test Cricket teams to bat second, rather than first as the so called expert commentators would have us believe. Before publication, we sent the article to the ABC Science website, which featured it in their Science News of the day. This in turn created interest from other outlets and subsequent newspaper articles. Moreover, the web pages are still obtainable through their archive system at http://www.abc.net.au/science/news/stories/s1074814.htm.

Press releases

A press release is a simple way to get the message out to a range of media outlets. It must be short and have an angle that will appeal to the media. Editors and producers get literally hundreds of these releases every day, and you need to attract their attention. However we rarely send out a release without getting some resultant article or request for interview. A press release prior to the Sydney Olympics based on a paper given at MCS5 on home advantage in the Olympics, resulted in a couple of Newspaper articles, a TV news item, and over 15 radio interviews. Events which occur yearly are a good candidate for releases. Michael Bailey sent out a press release on the likely winner of the Brownlow medal each year, while Tristan Barnett regularly sends out a release before the Melbourne tennis open, on some aspect such as likely winner or court

surface. Swinburne always send out a press release about my AFL computer tips prior to the season. In 2005 this resulted in a Melbourne radio station SEN, Campus review, and a couple of local papers taking on the regular publication of the tips. With my partial retirement, it is good to see the loose ball being picked up by Michael Bailey. In 2006 he had several radio interviews following a press release on his pre-season AFL ladder predictions. Unfortunately, following his successful PhD completion, it is now Monash receiving the media credit.

Because most institutions value media exposure, there will usually be a unit controlling this aspect, and there will be some expert who will assist in preparation of a media release. To my knowledge, this conference has never appointed media people to coordinate publicity. However the 2003 ICIAM conference in Sydney did so. This conference had a one day session on sport, organised by Neville de Mestre, and the resulting press release created much media interest. Neville de mestre, Eliot Tonkes, Darren O'Shaughnessy and myself all had radio interviews and Newspaper articles.

Personal contacts

Once a journalist has interviewed or published an article on your work, it is a good idea to try and build up a personal relationship. These people will be predisposed to publish future work. I have found journalists who have the onerous task of doing a regular column or spot to be particularly amenable to ideas for possible articles. We built up a great relationship with Ted Hopkins, when he was writing for the Australian Financial review. He published articles detailing virtually all of our past and current research work, and we often carried out special analysis to complement an article he was writing. That relationship developed into an ARC Linkage grant using Ted's company as the industry partner. We have also built up a good rapport with Justin Kemp who has a regular sports program, 'Run like you stole something', on community radio. He has interviewed me several times and all my PhD students at least once.

One outlet statisticians might like to consider is Robert Bolton's regular article, 'Lies and Statistics', in the Australian Financial Review. Each week this publishes a short article signed by the authors on some aspect of statistics that may be different to public perception. While often on politics, Robert is quite happy to have articles on sport. We have had published letters for Allsopp and Clarke on batting first in Test Cricket, Barnett and Clarke on Simpson's paradox in tennis, Barnett and Rod Cross on court surface speed and Barnett and Meyer on comparing men's and women's tennis. When the queue of future articles gets dangerously short, we now often get an email asking for a contribution.

Reputation and Media Guides

Once you build up a reputation for media work, you will be approached for comment. For example, at the start of every footy season I am usually asked for tips on tipping. At some stage in the season a journalist will do an article on how everyone is going in the tipping competitions, or home advantage, and I will be contacted for comment. Several years ago we were approached by the Media unit for the Australian Grand Prix. They knew of our work in football, and wanted similar predictions for motor cycling and F1 racing. Jonathan Lowe took on the task, and via a simulation we provided predictions for the Melbourne races and World Championships. These were included as part of the F1 Grand Prix media release, and resulted in several (some overseas) magazine and newspaper articles. Swinburne University has a media guide, which lists staff willing to

comment to the media on current issues. Lately my listing under gambling has resulted in a couple of television interviews. You will also find that media exposure rarely comes in ones. In particular, a newspaper article will invariably result in radio interviews. Sometimes the reverse occurs.
With my approaching retirement, for the last few years I have not been actively seeking media exposure, yet it seems as busy as ever, driven by approaches from the media rather than vice versa. Thankfully most of my previous students have learnt the advantages of extended coverage, and are taking over the load.

DISCUSSION

It should be recognised that publication in the popular press is different to peer publications in international journals. It is speedy but less permanent. It does not have to be the latest research, but may be relatively simple work. It doesn't have to be new, so it can be recycled. It is less serious and often less accurate. It is consumed less carefully, but by more people. To get exposure, it probably has to be interesting or have an angle, and almost certainly be current.

If you wish to use the media, you have to accept its limitations. They will often highlight the controversial, and often get things wrong. This once worked in our favour. When Michael Bailey did his Brownlow tips (Bailey and Clarke, 2002) he had to introduce a dummy variable for fair headed or dark skinned players, as these seemed to attract more votes. The Age picked up on this, with an article entitled "Hair's to all those Brownlow hopefuls". We were a bit embarrassed when the ultimate winner was only placed 10th in Michael's table, but bemused when a newspaper article trumpeted that Swinburne had got the result correct, as the winner had blonde hair.

Working in sport you must also be prepared for your research to be treated a bit tongue in cheek. I am often publicised as having a degree in Footy tipping, and a newspaper cartoon showed the chancellor presenting my degree via a drop kick. When Don was asked to produce tips for the Melbourne Cup, he didn't have 3 years in which to produce a proper study. It would have been easy to say we couldn't satisfy the media request. However within a few days Don produced predictions that we could say had some basis in Science. He was also prepared to be treated with some degree of irreverence. In addition to Don's scientific tips, the program also asked an astrologer and a dog. For years, my computer tips on Adelaide TV regularly had a dog as a competitor, and on Adelaide radio we now compete against a guinea pig.

You also need to try and keep technical jargon to a minimum, and not get hung up on the more subtle details – well most of the time. Before Don appeared on TV, we sent him down for a crash media course. They stressed the golden rule "Keep it simple". The first thing the interviewer told Don was to make it sound as complicated as possible, so he could feign going to sleep while Don discussed the intricacies of regression models.

CONCLUSION

There are many opportunities for gaining exposure: regular weekly team events in your sport, such as football, soccer and cricket; related within competition events such as best player awards; other regular events such as tennis majors, Commonwealth and Olympic

Games, World Championships. Don't be restricted to elite competition; local papers are interested in local sports events.

While there are some disadvantages, I believe they are outweighed by the many advantages of increased media exposure for your research. Your research becomes better known by your colleagues and superiors. Among the millions of readers or watchers there will be colleagues, superiors, future students, possible consulting or research project clients who might become involved and would otherwise not learn about your work. In addition I believe it subtly educates the general public and schoolchildren on the pervasiveness and usefulness of mathematics.

All the persons mentioned in this paper are past attendees at this conference. Many other attendees have obtained exposure for their work. I am sure most of them would have positive feedback from the experience. Most of the papers at this conference we enjoy not because they are within our particular mathematical expertise, but because they are about sport. This demonstrates the popular appeal of research work in sport. The general public are interested in your research – so tell them about it.

REFERENCES

Allsopp, P. and Clarke, S. (2004) Rating cricket teams and analysing outcomes in one-day and test cricket. *Journal of the Royal Statistical Society (Series A)* **167**, 657-667.

Bailey, M. and Clarke, S. (2002) Predicting the Brownlow medal winner. In: *Proceedings of the sixth Australian conference on mathematics and computers in sport.* Eds: Cohen, G. and Langtry, T. Gold Coast: University of Technology, Sydney**:** 56-62.

Clarke, S. (2002) Publicising OR projects in the media. *OR Insight* **15**(3), 4-14.

USES AND MISUSES OF STATISTICS IN TEAM SPORTS: GOALS IN THE BANK?

Serrallach G.

Verusco Technologies Ltd, Palmerston North, New Zealand

ABSTRACT

Statistics are seen by some people involved in the coaching of sport as a magic potion for improving performance, but it is the information extracted from them and how this information is applied that makes the difference in improving sporting performance. This paper offers examples of the way statistics are misused by coaches in sport and how they can be used correctly to positively impact on future performance.

KEY WORDS

soccer, shots, kpi, uses, misuses, team

INTRODUCTION

Simple statistics in themselves cannot be used as a tool to change future outcomes as what they show is merely the distribution of events that have occurred in the past. What they do provide is information which, when extracted correctly and applied with the right tools, can provide coaches with the ability to take steps to alter future performance.

As they come from a sporting rather than scientific background, coaches generally have little understanding of statistical tools such as standard deviation and correlation, or how to actually use these tools and other statistical data correctly. Due to this lack of understanding, many coaches try to forecast the future through merely looking at historical simple statistics. And as Albert Einstein once said, the height of "stupidity" is doing the same thing over and over again, and expecting different results.

The most instructive way to show how statistics can be misused and how they can be correctly used in sport is through the presentation of an example of each of these. The information used for these examples is drawn from Verusco's GoalSeeker Premier Analysis System and the football information database held by the company.

To provide some background, the GoalSeeker Premier Analysis System is a unique and comprehensive football analysis tool that delivers an array of statistical data, all linked to video footage, in a format that allows for the practical application of the information. Every action that occurs during the game and the quality of its execution is inputted during the "coding" process, which can take around 40 working hours per game. Inputting this amount of information, rather than selectively analysing only some actions, allows for a series of Key Performance Indicators (KPIs) to be generated by the system. These KPIs highlight the crucial areas of the game and the performance of players and teams in these areas.

While the technological advances of the last ten years have rapidly progressed the functionality and output capability of GoalSeeker Premier, the concept of the system and its development was started by company founder Dr. Serrallach around 50 years ago.

MISUSE OF STATISTICS

Looking firstly at the misuse of statistics, during a recent football season Verusco presented GoalSeeker Premier to a Premier League team in England. The team was at the time performing poorly, and though it was only halfway through the season, they were already in danger of relegation. Prior to the presentation Verusco had 'coded', down to the level of each individual task, a number of this team's games in order to gather relevant statistical data.

From this raw data, the system was able to generate and output a series of KPIs on the team. A KPI is a statistical figure that shows the performance of a team or player either in a specific area of the game or in a summary of several areas. Some KPIs have two values, one based on option taking (decision making) abilities and the other on the execution of the techniques used in that specific or summary area.

There are four KPIs that are very important in football and represent or show:

- **Rating**: the weighted average of all KPIs
- **Delivery:** the ability of players to communicate, pass the ball to team mates or put the ball into the net. It also shows the loss of ball possession and recovery by the other team allowing them to initiate a counterattack
- **Moves Influence**: the ability of a player/team to build attack and create goal scoring chances
- **Playing Profile**: the mental attitude of the player/team towards the game.

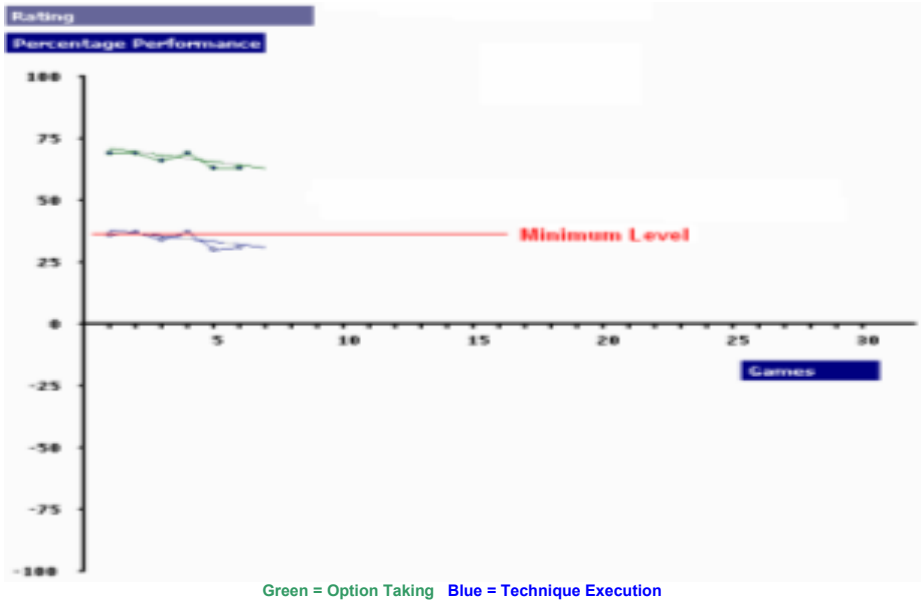


Figure 1 – Rating KPI for a series of games

The trend of the fundamental KPIs showed that the performance of the team was below the minimum level required to stay in the Premier League. Three of these KPIs, Rating, Delivery and Moves Influence, were outlined in depth to highlight the areas of the game in which they were seriously underperforming (see Rating KPI Figure 1). The obvious conclusion from the data was that if they did not investigate these areas of concern and change on field performance levels, they would be demoted at the end of the season.

The feedback from the people attending the presentation, which included the team's manager and a very high ranking official of the English Football Association, was that while all Verusco's statistics were extremely good, they did not focus on the "real and important statistics". When asked to expand on what they meant by this, the high ranking official stated that "real statistics" should include shots at goal, penetration in the penalty area, passes into the penalty area and a only few other events. And the cornerstone of these "real statistics" was relying on simple historical data such as the competition's averages. For example, it was mentioned that for every 100 shots at goal 10 goals are scored.

As it could be demonstrated that the team was only scoring at a rate of 5 goals per 100 shots in the current season, the question of where the other 5 goals were was put to them. The reply was that the "goals were in the bank". When this reply was challenged the response was that the GoalSeeker Premier KPI trends were wrong and that the "missing goals" would be scored during the rest of the season (balancing out to the competition average) and therefore they would win the necessary games to stay in the Premier League. No reply was received when it was asked that as one of the leading teams in the Premier League was scoring 15 goals every 100 shots, were they borrowing the "goals from the bank" or taking them from the underperforming teams.

It was later learnt that strikers and midfielders of this club were given a quota of shots at goal that they were expected to achieve. No mention or expectation was placed on quality of these shots at goal. It was thought that if enough shots were sent in the direction of the goal, then 10 out of every 100 would somehow end up in the back of the net, and that is exactly the flawed thought processes the players took onto the field.

There were two main points reached from the meeting and discussions with this Premier League team. First, the way they looked at historical statistics as a direct basis for future performance was fundamentally flawed from a statistical standpoint. As chance has no memory, goals not scored cannot be recovered at a future time. An increase in goal scoring is linked to the quality of performance rather than a straight increase in quantity of shots at goal taken.

The second point related to the presentation of complex statistical data from sport in ways that coaches can easily access and understand it. The coaches of this team could not comprehend the information presented to them and this was partially the result of presenting the statistics in a way, using the jargon of professional statisticians and not simple language, which made it difficult for a lay person to understand. To be able to change the future statistics have to be presented in such a way that the coach is made aware, through directing attention to data on trends and the quality of performance, of aspects of the game that if they are improved the performance of the team will increase.

As a postscript to this case, the team involved was demoted at the end of the season. While the players reached their targets on shot quantity, there was no improvement in the quality through the rest of the season. At the time of demotion the team still had 31 "goals in the bank".

CORRECT USE OF STATISTICS

When used correctly statistics can be a powerful weapon in a coach's armoury. This example of the correct use of statistics looks at the delivery (passing) tasks of a player and shows the links between the quality of tasks performed and successful outcomes.

From coding a number of games, GoalSeeker Premier is able to generate a Playing Profile on individual players. This profile breaks the actions performed by a player into 4 key task areas so that coaches can examine the separate aspects of a player's game. The Playing Profile of a well known player is shown in Figure 2.

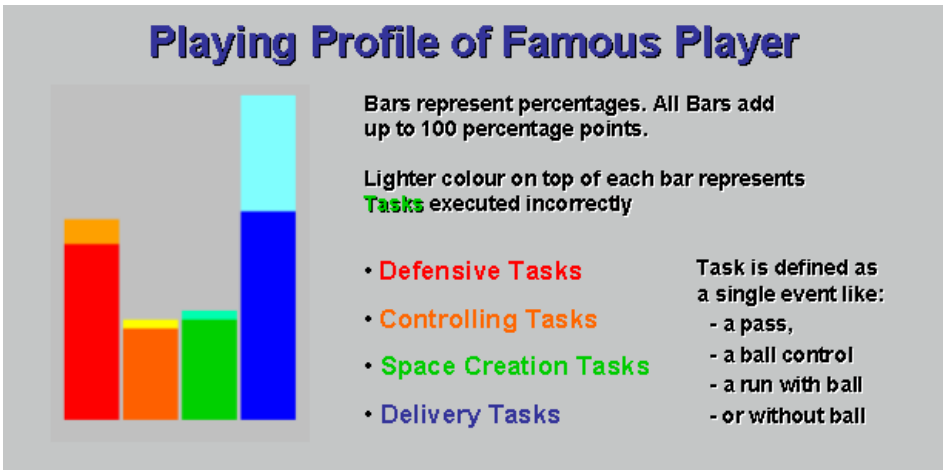
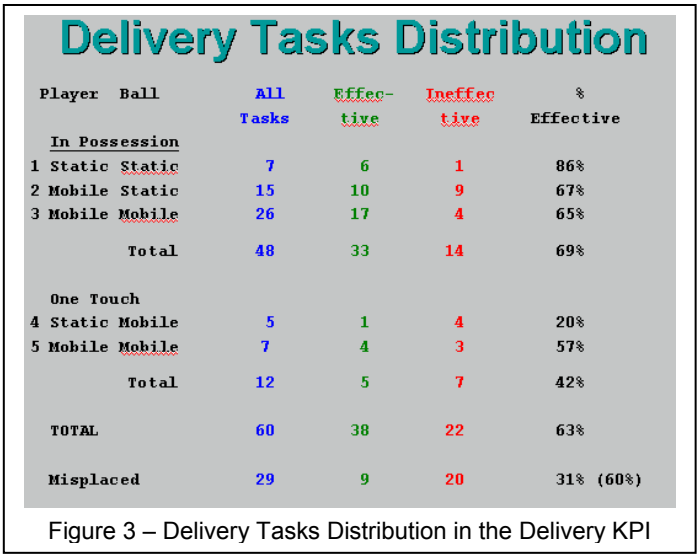


Figure 2 – Playing Profile

Looking at his Playing Profile, it can be seen that the player performs a higher proportion of delivery Task than any of the other technical skill areas, and that there is a high level of poorly executed tasks in the delivery area. This indicates that this area needs further investigation in order to improve the player's and team's performance.

To look into this further we can look at a series of games and break down the delivery tasks into subsets for closer examination. Delivery techniques (see Figure 3) can be broken down by ball movement at the pass (mobile or static), player movement at the pass (mobile or static), the player's control of the ball (In Possession or One Touch), and then examined to find the areas of poor performance.

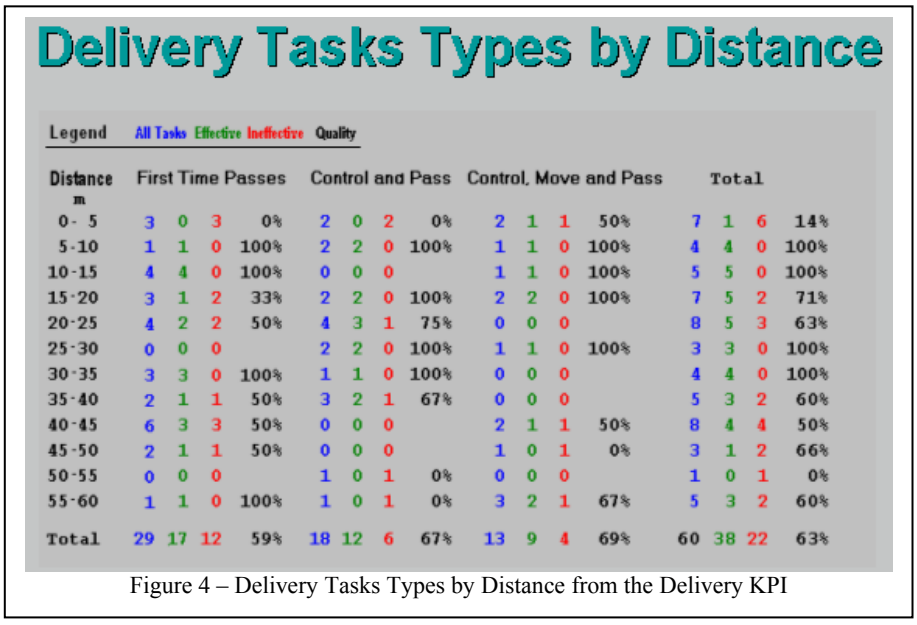


Delivery Tasks Distribution

Player	Ball	All Tasks	Effective	Ineffective	% Effective
In Possession					
1 Static	Static	7	6	1	86%
2 Mobile	Static	15	10	9	67%
3 Mobile	Mobile	26	17	4	65%
	Total	48	33	14	69%
One Touch					
4 Static	Mobile	5	1	4	20%
5 Mobile	Mobile	7	4	3	57%
	Total	12	5	7	42%
TOTAL		60	38	22	63%
Misplaced		29	9	20	31% (60%)

Figure 3 – Delivery Tasks Distribution in the Delivery KPI

The GoalSeeker Premier Analysis System also allows the user to break down the delivery tasks by the likes of angle on the field and distance of the delivery (see Figure 4). From looking at all these areas the patterns of play can be seen, and then checked against the video to confirm the analysis and to ascertain why there are defects in the player's performance.



Delivery Tasks Types by Distance

Legend: All Tasks | Effective | Ineffective | Quality

Distance m	First Time Passes				Control and Pass				Control, Move and Pass				Total			
0- 5	3	0	3	0%	2	0	2	0%	2	1	1	50%	7	1	6	14%
5-10	1	1	0	100%	2	2	0	100%	1	1	0	100%	4	4	0	100%
10-15	4	4	0	100%	0	0	0		1	1	0	100%	5	5	0	100%
15-20	3	1	2	33%	2	2	0	100%	2	2	0	100%	7	5	2	71%
20-25	4	2	2	50%	4	3	1	75%	0	0	0		8	5	3	63%
25-30	0	0	0		2	2	0	100%	1	1	0	100%	3	3	0	100%
30-35	3	3	0	100%	1	1	0	100%	0	0	0		4	4	0	100%
35-40	2	1	1	50%	3	2	1	67%	0	0	0		5	3	2	60%
40-45	6	3	3	50%	0	0	0		2	1	1	50%	8	4	4	50%
45-50	2	1	1	50%	0	0	0		1	0	1	0%	3	1	2	66%
50-55	0	0	0		1	0	1	0%	0	0	0		1	0	1	0%
55-60	1	1	0	100%	1	0	1	0%	3	2	1	67%	5	3	2	60%
Total	29	17	12	59%	18	12	6	67%	13	9	4	69%	60	38	22	63%

Figure 4 – Delivery Tasks Types by Distance from the Delivery KPI

From analysing all the information linked to the Playing Profile, I was able to ascertain that the player has issues in the following areas of his play:

- Delivery quality decreases with the increase in difficulty of the delivery type
- Difficultly in using instep or inside of foot when delivering ball directly ahead
- Poor delivery over short distance when under pressure
- Delivery problems when applying power to the ball

With this information a coach can then look to work on the quality of the individual player's skills in these areas with the aim of improving the quality and outcome of the actions performed. The emphasis is placed on quality as opposed to quantity and all goals set for players to achieve need to be understandable and able to be measured.

This information can be taken a step further to look at the player's contribution to the team's overall in delivery tasks. As can be seen in Figure 5, this player had an overly high proportion of total ineffective delivery tasks in comparison with his share of total delivery tasks. He is therefore having in effect a negative influence in the performance of the team in this area.

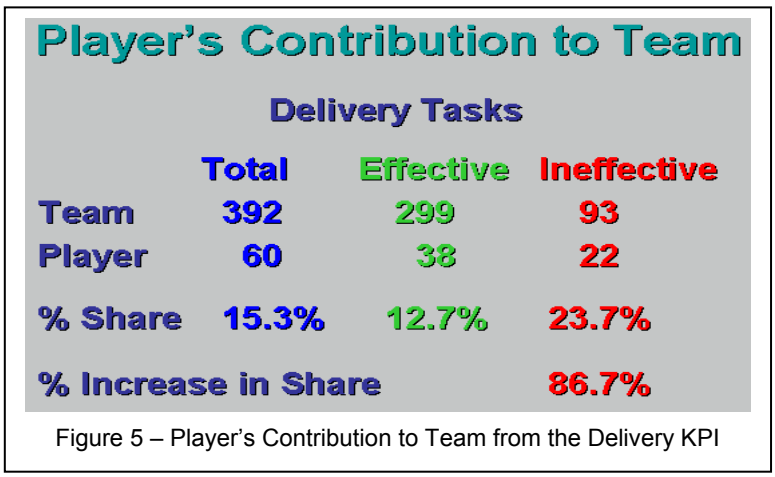


Figure 5 – Player's Contribution to Team from the Delivery KPI

From this, and looking at his contribution in the other areas of the game, the value of a player to the team can be ascertained. This information can also be viewed from an overall club management perspective, linking it to questions such as what was paid for the player, was the price paid justified, can weak areas of play be corrected to increase the player's value, and can this help in driving revenue for the club?

STATISTICAL TOOLS

The usefulness of using statistics in this way will be limited if once technical weaknesses in a player are identified there is no clear system for correcting them. Having the player, the coach and the team merely knowing where there is a weakness will not in itself improve performance. The technical weakness needs to be addressed in order to improve the quality of tasks and therefore team performance.

The natural progression then is to build into the analysis system a techniques structure that can identify, isolate and index all the tasks and techniques used by a player during a game. By doing this the "perfect" execution of each task and technique can be stored in the system, and this can be compared with the execution of tasks by a player in match or practice situations. By making this comparison deficiencies in technique can be identified and correct technique practised until the task execution of the player meets the required level of quality.

This comparison needs to be made using tools that the player can readily comprehend and identify with. Here the use of video footage has proved invaluable as direct comparisons can be made between correct task execution, the individual player's task execution, and the task executions of top players at international level. Now being added to this is the ability to generate 3D models of tasks and techniques to replicate the actions seen in video footage. This allows coaches and players to even more closely examine the tasks that are being executed and isolate areas of technique deficiency.

These types of video and 3D analysis tools can then be taken beyond looking at task and technique execution and can be applied to other parts of the game such as option taking. A player making the correct decision at a crucial point of the game often proves the difference between winning and losing. Comprehensive analysis tools incorporating video and 3D can provide the functionality of allowing players to "relive" past situations or to look at other games involving top level players to learn from the correct options taken in similar situations.

CONCLUSION

Even at the highest levels of professional sport, examples can be found of statistical data being used poorly by coaches who do not understand how to interpret and apply the information it provides them with. So what tends to happen is that statistics are used in such a simple manner that they have no impact in actually changing future team or player performance.

Using statistics correctly requires extracting key information and trends from the data and applying it using sound statistical principles. But to do these things you need more than just simple statistics and a basic analysis system, you need a complete analysis system that:

- Is outcome based
- Measures tasks by quality not just quantity
- Delivers in depth statistics to coaches in a way they can understand and use
- Allows the user to database a large number of games to provide trends
- Incorporates structured teaching tools
- Presents information to players in formats they can easily relate to (e.g. linked to video and 3D)

Without this type of system coaches will have such limited information that they may guess what will be delivered in future performances, but they will powerless to actually change the outcome of future performances.

REDUCING THE LIKELIHOOD OF LONG TENNIS MATCHES

Barnett, T.[1], Brown, A.[1] and Pollard, G.[2]
1 Swinburne University, Melbourne, VIC, Australia
2 University of Canberra, Canberra, ACT, Australia

ABSTRACT

Long matches can cause problems for tournaments. For example, the starting times of subsequent matches can be substantially delayed causing inconvenience to players, spectators, officials and television scheduling. They can even be seen as unfair in the tournament setting when the winner of a very long match, who may have negative aftereffects from such a match, plays the winner of an average or shorter length match in the next round. Long matches can also lead to injuries to the participating players.

One factor that can lead to long matches is the use of the advantage set as the fifth set, as in the Australian Open, the French Open and Wimbledon. Another factor is long rallies and a greater than average number of points per game. This tends to occur more frequently on the slower surfaces such as at the French Open. The mathematical method of generating functions is used to show that the likelihood of long matches can be substantially reduced by using the tiebreak game in the fifth set, or more effectively by using a new type of game, the 50-40 game, throughout the match.

KEY WORDS

tennis, scoring systems, sport, generating functions, long tennis matches

INTRODUCTION

In recent years there have been a number of grand slam matches decided in long fifth sets. In the third round of the 2000 Wimbledon mens singles, Philippoussis defeated Schalken 20-18 in the fifth set. Ivanisevic defeated Krajicek 15-13 in the semi-finals of Wimbledon in 1998. In the quarter-finals of the 2003 Australian Open mens singles, Andy Roddick defeated Younes El Aynaoui 21-19 in the fifth set, a match taking 83 games to complete and lasting a total duration of 5 h. The night session containing this long match required the following match to start at 1 am. Long matches require rescheduling of following matches, and also create scheduling problems for media broadcasters. They arise because of the advantage set, which gives more chance of winning to the better player (Pollard and Noble, 2002), but has no upper bound on the number of games played. It may be in the interests of broadcasters and tournament organizers to decrease the likelihood of long tennis matches occurring.

Pollard (1983) calculated the mean and variance of the duration of a best-of-three sets match of classical and tiebreaker tennis by using the probability generating function. It is well established that the mean and standard deviation completely describe the normal distribution. When a distribution is not symmetrical about the mean, the coefficients of skewness and kurtosis, as defined in Stuart and Ord (1987), are important to graphically

interpret the shape of the distribution. This commonly has been done by using the probability or moment generating function. The cumulant generating function (taking the natural logarithm of the moment generating function), can also be used to calculate the parameters of the distribution in a tennis match. The cumulant generating function is particularly useful for calculating the parameters of distributions for the number of points in a tiebreaker match, since the critical property of cumulant generating functions is that they are additive for linear combinations of independent random variables.

The layout of this paper is as follows. For convenience of the less mathematically inclined we defer the presentation of the mathematics of generating functions applied to tennis till Section 3. Instead we will begin in Section 2 with a discussion on several aspects of long matches, relying on graphical results to advance our arguments as to how they might be curtailed. We aim to show that the likelihood of long matches can be substantially reduced by using the tiebreak game in the fifth set, or more effectively by the use of a new type of game, the 50-40 game (Pollard and Noble, 2004), throughout the match. In Section 4 we make some concluding remarks.

DISCUSSION OF THE PROBLEM (using graphical results)

Up until 1970 (approx), all tennis sets were played as advantage sets, where to win a set a player must reach at least 6 games and be ahead by at least 2 games. The tiebreaker game was introduced to shorten the length of matches. A tiebreaker game is played when the set score reaches 6-games all. However in three of the four grand slams (Australian Open, French Open and Wimbledon), an advantage set is still played in the deciding fifth set. Figure 1 represents a comparison of a match with 5 advantage sets (5adv), 5 tiebreaker sets (5tie) and 4 tiebreaker sets with a deciding advantage set (4tie1adv). The probability of each player winning a point on serve is given as 0.6 to represent averages in men's tennis. The long tail given by the match with 5adv gives an indication as to why the tiebreaker game was introduced to the tennis scoring system.

It is well known that the dominance of serve in men's tennis has increased since the introduction of the tiebreaker game. This creates a problem when two big servers meet in a grand slam event where the deciding fifth set is played as an advantage set. Figure 2 represents a match with 4tie1adv for different values of players winning points on serve. It shows that for two strong servers winning 0.7 of points on serve, there is a long tail in the number of points played. In comparison with Figure 3, which represents a match with 5tie, the tail is substantially reduced for two players winning 0.7 of points on serve.

Figure 4 represents a match with 5 tiebreaker sets, where a standard 'deuce' game is replaced by a 50-40 game. It shows an even greater improvement to reducing the number of points played in a match compared to Figure 3. In the 50-40 game the server has to win the standard 4 points, while the receiver only has to win 3 points. Such a game requires at most 6 points.

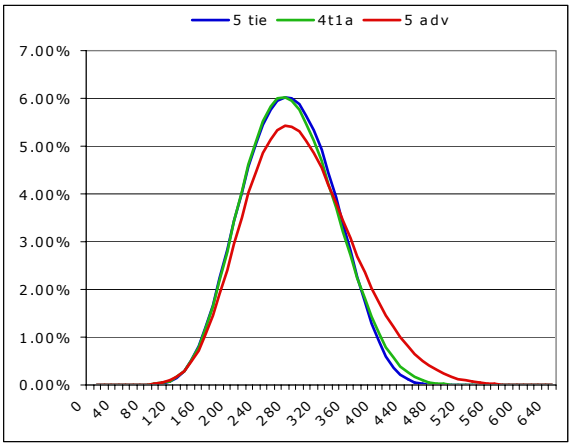


Figure 1: Distribution of a match with different scoring systems

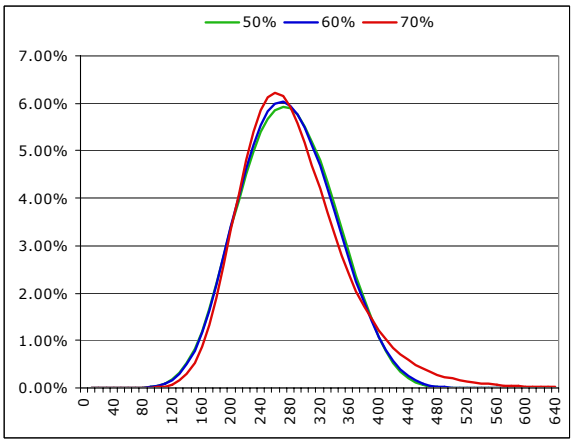


Figure 2: Distribution of an advantage match (4tie1adv) for different values of players winning points on serve

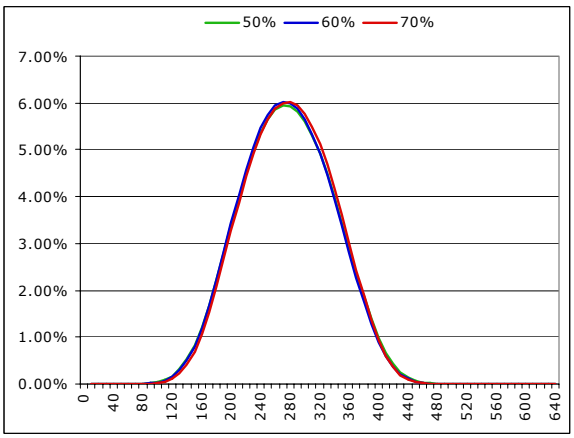


Figure 3: Distribution of a tiebreaker match (5tie) for different values of players winning points on serve

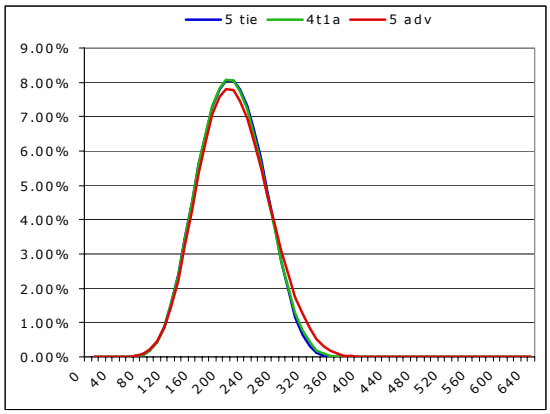


Figure 4: Distribution of a tiebreaker match (5tie) for different values of players winning points on serve, by using 50-40 games instead of standard 'deuce' games

THE MATHEMATICS OF GENERATING FUNCTIONS

Modelling a tennis match

Forward recursion

The state of a tennis match between two players is represented by a scoreboard. The scoreboard shows the points, games and sets won by each player, and is updated after each point has been played. It is assumed that the conditional probability of the server winning the point depends only on the data shown on the scoreboard. This enables the progress of the match to be modelled using forward recursion. An additional assumption is that the probabilities of each player winning a point on his own service remain constant throughout the match.

Development of generating functions of distributions

The forward recursion enables the probabilities of various possible scoreboards to be calculated. These probabilities can be collected in the form of probability generating functions, or equivalently, moment generating functions (using the transformation $v = e^u$).

Lemma: If X and Y are independent random variables and $Z = X + Y$ then:

$$m_Z(t) = m_X(t) * m_Y(t).$$

It becomes convenient at times to take logarithms, and work in terms of cumulant generating functions, since $K_Z(t) = K_X(t) + K_Y(t)$.

The higher order cumulants depend on powers of the scale for the random variable, and for the purposes of communication it is useful to transform them into non-dimensional statistics (i.e. numbers) such as the coefficients of variation, skewness and kurtosis.

The inversion of the cumulants using normal power approximation

This gives a continuous approximation to a discrete distribution (Pesonen, 1975). The formula is asymptotic and works reasonably well for unimodal distributions with the

coefficient of skewness less than 2 and the coefficient of kurtosis less than 6. i.e. tails die off at least as fast as the exponential distribution.

The number of points in a game

Let X be a random variable of the number of points played in a game. Let $f^{pg}_A(x)$ represent the distribution of the number of points played in a game for player A serving, where $f^{pg}_A(x) = P(X = x)$. This gives the following:

$f^{pg}_A(4) = N^{pg}_A(4, 0) + N^{pg}_A(0, 4)$
$f^{pg}_A(5) = N^{pg}_A(4, 1) + N^{pg}_A(1, 4)$
$f^{pg}_A(6) = N^{pg}_A(4,2) + N^{pg}_A(2, 4)$
$f^{pg}_A(x) = N^{pg}_A(3, 3)[p^2_A + (1 - p_A)^2][2p_A(1 - p_A)]^{(x-8)/2}$, if $x = 8, 10, 12, ...$

where:
$N^{pg}_A(a,b)$ represents the probability of reaching point score (a,b) in a game for player A serving
p_A represents the probability of player A winning a point on serve

Croucher (1986) gives algebraic expressions for calculating $N^{pg}_A(a,b)$.

Let $m(t)$ denote the moment generating function X. Generating functions can be used to describe a distribution, such as $f^{pg}_A(x)$ for all x. It is well established (Stuart and Ord,1987) that the mean, variance, coefficient of skewness and coefficient of kurtosis of X can be obtained from generating functions.

The moment generating function for the number of points in a game for player A serving, $m^{pg}_A(t)$, becomes:

$$\sum_x e^{tx} f^{pg}_A(x) = e^{4t} f^{pg}_A(4) + e^{5t} f^{pg}_A(5) + e^{6t} f^{pg}_A(6) + [N^{pg}_A(3,3)(1 - N^{pg}_A(1,1))e^{8t}] / [1 - N^{pg}_A(1,1)e^{2t}]$$

The mean number of points in a game M^{pg}_A , with the associated variance V^{pg}_A are calculated from the moment generating function using *Mathematica* and given as:

$$M^{pg}_A = \frac{4\{p_A(1 - p_A)[6p^2_A(1 - p_A)^2 - 1] - 1\}}{\{1 - 2p_A(1 - p_A)\}}$$

$$V^{pg}_A = \frac{4p_A(1 - p_A)[1 - p_A(1 - p_A)(1 - 12p_A(1 - p_A)(3 - p_A(1 - p_A)(5 + 12p^2_A(1 - p_A)^2)))]}{[1 - 2p_A(1 - p_A)]^2}$$

Similar expressions can be obtained for the coefficient of skewness S^{pg}_A, and the coefficient of kurtosis K^{pg}_A .
Let U^{pg}_A represent the standard deviation of the number of points in a game for player A serving. Let C^{pg}_A represent the coefficient of variation of the number of points in a game for player A serving. It follows that $U^{pg}_A = \sqrt{V^{pg}_A}$ and $C^{pg}_A = U^{pg}_A / M^{pg}_A$.

Table 1: The parameters of the distributions of points in a game for different values of p_A

p_A	$M^{pg}{}_A$	$U^{pg}{}_A$	$C^{pg}{}_A$	$S^{pg}{}_A$	$K^{pg}{}_A$
0.50	6.75	2.77	0.41	2.16	6.95
0.55	6.68	2.73	0.41	2.17	7.01
0.60	6.48	2.59	0.40	2.20	7.21
0.65	6.19	2.37	0.38	2.25	7.59
0.70	5.83	2.10	0.36	2.34	8.25
0.75	5.45	1.78	0.33	2.46	9.27

Table 1 represents $M^{pg}{}_A$, $U^{pg}{}_A$, $C^{pg}{}_A$, $S^{pg}{}_A$ and $K^{pg}{}_A$ for different values of p_A. Notice that the mean and standard deviation are greatest when $p_A = 0.50$, but the coefficients of skewness and kurtosis are greatest when p_A approaches 1 or 0. The generating functions to follow are for player A serving first in the tiebreaker game or set.

The moment generating function for the number of points in a tiebreaker game, $m^{pgT}{}_A(t)$ becomes:

$$m^{pgT}{}_A(t) = e^{7t}f^{pgT}{}_A(7)+e^{8t}f^{pgT}{}_A(8)+e^{9t}f^{pgT}{}_A(9)+e^{10t}f^{pgT}{}_A(10)+e^{11t}f^{pgT}{}_A(11)+e^{12t}f^{pgT}{}_A(12)+ N^{pgT}{}_A(6,6)(1-N^{pgT}{}_A(1,1))e^{14t}/[1-N^{pgT}{}_A(1,1)e^{2t}]$$

where:
$f^{pgT}{}_A(x)$ represents the distribution of the number of points played in a tiebreaker game
$N^{pgT}{}_A(a,b)$ represents the probability of reaching point score (a,b) in a tiebreaker game

The moment generating functions for the number of games in a tiebreaker set, $m^{gsT}{}_A(t)$ and advantage set, $m^{gs}{}_A(t)$ become:

$$m^{gsT}{}_A(t)=e^{6t}f^{gsT}{}_A(6)+e^{7t}f^{gsT}{}_A(7)+e^{8t}f^{gsT}{}_A(8)+e^{9t}f^{gsT}{}_A(9)+e^{10t}f^{gsT}{}_A(10)+e^{12t}f^{gsT}{}_A(12)+ e^{13t}f^{gsT}{}_A(13)$$

$$m^{gs}{}_A(t) = e^{6t}f^{gs}{}_A(6)+e^{7t}f^{gs}{}_A(7)+e^{8t}f^{gs}{}_A(8)+e^{9t}f^{gs}{}_A(9)+e^{10t}f^{gs}{}_A(10)+ N^{gs}{}_A(5,5)(1-N^{gs}{}_A(1,1))e^{12t}/[1-N^{gs}{}_A(1,1)e^{2t}]$$

where:
$f^{gsT}{}_A(x)$ represents the distribution of the number of games played in a tiebreaker set
$f^{gs}{}_A(x)$ represents the distribution of the number of games played in an advantage set
$N^{gs}{}_A(c, d)$ represents the probability of reaching (c,d) in an advantage set

The number of points in a set

The parameters of distributions of the number of points in a set

Let $m^{pg}{}_{A+}(t)$ and $m^{pg}{}_{A-}(t)$ be the moment generating functions of the number of points in a game when player A wins and loses a game on serve respectively. Let $m^{pg}{}_{B+}(t)$ and

$m^{pg}{}_{B-}(t)$ be the moment generating functions of the number of points in a game when player B wins and loses a game on serve respectively. Let $s(c, d)$ be the moment generating function of the number of points in a set conditioned on reaching game score (c,d). It can be shown that
$s(6,1) = 3[m^{pg}{}_{A+}(t)]^3[m^{pg}{}_{B-}(t)]^2[m^{pg}{}_{A+}(t)m^{pg}{}_{B+}(t) + m^{pg}{}_{A-}(t)m^{pg}{}_{B-}(t)]$ and
$s(1,6) = 3[m^{pg}{}_{A-}(t)]^3[m^{pg}{}_{B+}(t)]^2[m^{pg}{}_{A+}(t)m^{pg}{}_{B+}(t)+m^{pg}{}_{A-}(t)m^{pg}{}_{B-}(t)]$. Similar conditional moment generating functions can be obtained for reaching all score lines (c, d) in a set. The moment generating function for the number of points in a tiebreaker set becomes:

$m^{psT}{}_A(t) = N^{gsT}{}_A(6, 0)s(6,0) + N^{gsT}{}_A(6,1)s(6,1) + N^{gsT}{}_A(6,2)s(6,2) + N^{gsT}{}_A(6,3)s(6,3) + N^{gsT}{}_A(6,4)s(6,4)+N^{gsT}{}_A(7,5)s(7,5)+N^{gsT}{}_A(0,6)s(0,6)+N^{gsT}{}_A(1,6)s(1,6)+N^{gsT}{}_A(2,6)s(2,6)+ N^{gsT}{}_A(3,6)s(3,6) + N^{gsT}{}_A(4,6)s(4,6) + N^{gsT}{}_A(5,7)s(5,7) + N^{gsT}{}_A(6,6)s(6,6)m^{pgT}{}_A(t)$

A similar moment generating function can be obtained for the number of points in an advantage set.

Let $M^{ps}{}_A$, $U^{ps}{}_A$, $C^{ps}{}_A$, $S^{ps}{}_A$ and $K^{ps}{}_A$ represent the mean, standard deviation, and coefficientsof variation, skewness and kurtosis for the number of points in an advantage set. Let $M^{psT}{}_A$, $U^{psT}{}_A$, $C^{psT}{}_A$, $S^{psT}{}_A$ and $K^{psT}{}_A$ represent the mean, standard deviation, and coefficients of variation, skewness and kurtosis for the number of points in a tiebreaker set. Table 2 represents $M^{ps}{}_A$, $U^{ps}{}_A$, $C^{ps}{}_A$, $S^{ps}{}_A$, $K^{ps}{}_A$, $M^{psT}{}_A$, $U^{psT}{}_A$, $C^{psT}{}_A$, $S^{psT}{}_A$ and $K^{psT}{}_A$ for different values of p_A and p_B. The table covers values in the interval $0.50 \leq pA \leq pB \leq 0.75$ as this is the main area of interest for men's professional tennis. It can be observed that: $M^{ps}{}_A > M^{psT}{}_A$, $U^{ps}{}_A > U^{psT}{}_A$, $C^{ps}{}_A > C^{psT}{}_A$, $S^{ps}{}_A > S^{psT}{}_A$ and $K^{ps}{}_A > K^{psT}{}_A$.

The mean number of points in a set is affected by the mean number of points in a game and the mean number of games in a set. The mean number of points in a game is greatest when p_A or $p_B = 0.50$. For a tiebreaker set, when $p_A = p_B = 0.50$, $M^{pg}{}_A = M^{pg}{}_B = 6.75$, $M^{gsT}{}_A = 9.66$ and $M^{psT}{}_A = 65.83$. When $p_A = p_B = 0.70$, $M^{pg}{}_A = M^{pg}{}_B = 5.83$, $M^{gsT}{}_A = 10.94$ and $M^{psT}{}_A = 66.22$. For this latter case, even though the mean length of games is shorter, the mean number of points in a tiebreaker set overall is greater since more games are expected to be played. Both players have a 0.90 probability of holding serve, which means that very few breaks of serve will occur and there is a 0.38 probability of reaching a tiebreaker. This is further exemplified in an advantage set, where for $p_A = p_B = 0.70$, $M^{ps}{}_A = 86.43$. This is also highlighted by the coefficients of variation, skewness and kurtosis being much greater for an advantage set, compared to a tiebreaker set, when p_A and p_B are both "large".

Approximating the parameters of distributions of the number of points in a set

The moment generating function for the number of points in an advantage set $m^{ps}{}_A(t)$, when $p_A = 1 - p_B$, becomes:

$m^{ps}{}_A(t)=[f^{gs}{}_A(6)](m^{pg}{}_{AB})^6+[f^{gs}{}_A(7)](m^{pg}{}_{AB})^7+[f^{gs}{}_A(8)](m^{pg}{}_{AB})^8+[f^{gs}{}_A(9)](m^{pg}{}_{AB})^9+ [f^{gs}{}_A(10)](m^{pg}{}_{AB})^{10}+N^{gs}{}_A(5,5)(1-N^{gs}{}_A(1,1))(m^{pg}{}_{AB})^{12} / [1-N^{gs}{}_A(1,1)(m^{pg}{}_{AB})^2]$

where: $m^{pg}{}_{AB}(t) = [m^{pg}{}_{A}(t)+m^{pg}{}_{B}(t)]/2$ is the average (in this case equal) of two moment generating functions.

Table 2: The parameters of the distributions of points in a tiebreaker and advantage set for different values of p_A and p_B

p_A	p_B	$M^{psT}{}_{A}$	$U^{psT}{}_{A}$	$C^{psT}{}_{A}$	$S^{psT}{}_{A}$	$K^{psT}{}_{A}$	$M^{ps}{}_{A}$	$U^{ps}{}_{A}$	$C^{ps}{}_{A}$	$S^{ps}{}_{A}$	$K^{ps}{}_{A}$
0.50	0.50	65.83	16.54	0.25	0.55	-0.04	67.71	21.15	0.31	1.62	4.75
0.50	0.60	61.99	15.97	0.26	0.65	0.18	63.39	19.76	0.31	1.71	5.39
0.50	0.70	54.73	13.75	0.25	0.85	0.88	55.39	16.01	0.29	1.89	7.46
0.50	0.75	51.64	12.44	0.24	0.89	1.20	52.06	14.08	0.27	1.92	8.44
0.60	0.60	65.59	16.03	0.24	0.55	-0.18	69.32	24.92	0.36	2.12	7.27
0.60	0.70	63.08	14.99	0.24	0.58	-0.15	68.35	27.97	0.41	2.60	10.22
0.60	0.75	60.67	14.32	0.24	0.63	-0.05	66.01	28.12	0.43	2.83	11.98
0.70	0.70	66.22	14.96	0.23	0.25	-0.81	86.43	53.11	0.61	2.47	8.67
0.75	0.75	67.59	13.74	0.20	-0.15	-0.82	125.50	101.81	0.81	2.24	7.22

Taking the natural logarithm of the moment generating function gives an alternative generating function known as the cumulant generating function. Let $\kappa^{pg}{}_{A}(t)=\ln[m^{pg}{}_{A}(t)]$ represent the cumulant generating function for the number of points in a game. This relationship can be inverted to give $m^{pg}{}_{A}(t) = e^{\kappa pgA(t)}$.

The moment generating function, $m^{ps}{}_{A}(t)$, can be written as:

$m^{ps}{}_{A}(t)=f^{gs}{}_{A}(6)e^{6\kappa pgAB(t)}+f^{gs}{}_{A}(7)e^{7\kappa pgAB(t)}+f^{gs}{}_{A}(8)e^{8\kappa pgAB(t)}+f^{gs}{}_{A}(9)e^{9\kappa pgAB(t)}+ f^{gs}{}_{A}(10)e^{10\kappa pgAB(t)}+ N^{gs}{}_{A}(5,5)e^{12\kappa pgAB(t)}\,[1-N^{gs}{}_{A}(1,1)] / [1-N^{gs}{}_{A}(1,1)e^{2\kappa pgAB(t)}]$, when $p_A = 1 - p_B$

where: $\kappa^{pg}{}_{AB}(t) = [\kappa^{pg}{}_{A}(t)+ \kappa^{pg}{}_{B}(t)]/2$ is the average (in this case equal) of two cumulant generating functions.

This can be expressed as:

$$m^{ps}{}_{A}(t) = m^{gs}{}_{A}(\kappa^{pg}{}_{AB}(t)) \quad (1)$$

Similarly, the following result is established for $m^{psT}{}_{A}(t)$, when $p_A = 1 - p_B$:

$$m^{psT}{}_{A}(t) = m^{gsT}{}_{A}(\kappa^{pg}{}_{AB}(t)) + N^{gsT}{}_{A}(6,6)e^{12\kappa pgAB(t)}(e^{\kappa pgTA(t)} - e^{\kappa pgAB(t)}) \quad (2)$$

Notice the last term does not vanish due to the difference in the scoring system for a tiebreaker game compared with a regular game. Equations 1 and 2 can be used to obtain approximate results for the parameters of distributions for the number of points in a set, when p_A is not equal to $1 - p_B$.

The number of points in a match

From this point an advantage match is considered as a match where the first four sets played are tiebreaker sets and the fifth set is an advantage set.

The moment generating functions for the number of points in an advantage and tiebreaker match, $m^{pm}(t)$ and $m^{pmT}(t)$, when $p_A = 1 - p_B$ become:

$m^{pmT}(t) = m^{sm}(\kappa^{psT}_{AB}(t))$

$m^{pm}(t) = m^{sm}(\kappa^{psT}_{AB}(t)) + N^{sm}(2,2)\, e^{4\kappa psTAB(t)}\, (e^{\kappa psAB(t)} - e^{\kappa psTAB(t)})$

where: $\kappa^{psT}_{AB}(t) = [\kappa^{psT}_{A}(t) + \kappa^{psT}_{B}(t)] / 2$ and $\kappa^{ps}_{AB}(t) = [\kappa^{ps}_{A}(t) + \kappa^{ps}_{B}(t)] / 2$

The following approximation results can be established for the number of points in a match, similar to the approximation results established for the number of points in a set:

$m^{pmT}(t) \approx m^{sm}(\kappa^{psT}_{AB}(t))$ for all values of p_A and p_B.

$m^{pm}(t) \approx m^{sm}(\kappa^{psT}_{AB}(t)) + N^{sm}(2,2)\, e^{4\kappa psTAB(t)}\, (e^{\kappa psAB(t)} - e^{\kappa psTAB(t)})$ for all values of p_A and p_B.

Approximation results for distributions of points in a match, could also be established for tennis doubles by using the above results established for singles. The probability of a team winning a point on serve is estimated by the averages of the two players in the team.

When $p_A = 1\text{-}p_B$, the distribution of number of points played each set if player A serves first in the set, is equal to the number of points played each set if player B serves first in the set. This leads to the following result:

The number of points played each set in a match are independent, if $p_A = 1 - p_B$.

Suppose $Z = X + Y$, where X and Y are independent, then it is well known that $m_Z(t) = E[e^{Zt}] = E[e^{Xt}]E[e^{Yt}] = m_X(t)m_Y(t)$. By taking logarithms it follows that $\kappa_Z(t) = \kappa_X(t) + \kappa_Y(t)$.

An extension of this property of cumulants is given by the following theory (Brown, 1977) and can be applied to points in a tiebreaker match when the number of points played each set in a match are independent. When the independence assumption fails to hold the theory remains approximately correct according to the approximation result established for points in a tiebreaker match.

Theorem

If $Z = X_1 + X_2 + \ldots\ldots + X_N$ where X_i are *i.i.d.* then $\kappa_Z(t) = \kappa_N(\kappa_X(t))$
Taking the derivatives of the result and setting $t = 0$ gives the following useful results in terms of cumulants:

$k^{(1)}_Z = k^{(1)}_N k^{(1)}_X$
$k^{(2)}_Z = k^{(2)}_N [k^{(1)}_X]^2 + k^{(1)}_N k^{(2)}_X$
$k^{(3)}_Z = 3k^{(1)}_X k^{(2)}_N k^{(2)}_X + [k^{(1)}_X]^3 k^{(3)}_N + k^{(1)}_N k^{(3)}_X$
$k^{(4)}_Z = 3k^{(2)}_N [k^{(2)}_X]^2 + 6[k^{(1)}_X]^2 k^{(2)}_X k^{(3)}_N + 4k^{(1)}_X k^{(2)}_N k^{(3)}_X + [k^{(1)}_X]^4 k^{(4)}_N + k^{(1)}_N k^{(4)}_X$

For example the mean number of points in a tiebreaker match, M^{pmT}, with the associated variance, V^{pmT}, can be calculated from the cumulant generating function as:

$$M^{pmT} = M^{psT} M^{sm}$$

$$V^{pmT} = V^{sm}(M^{psT})^2 + M^{sm} V^{psT}$$

where:
M^{psT} represents the mean number of points in a tiebreaker set
M^{sm} represents the mean number of sets in a tiebreaker match
V^{psT} represents the variance of the number of points in a tiebreaker set
V^{sm} represents the variance of the number of sets in a tiebreaker match

Let M^{pm} , U^{pm} , C^{pm} , S^{pm} and K^{pm} represent the mean, standard deviation, and coefficients of variation, skewness and kurtosis for the number of points in an advantage match. Let M^{pmT} , U^{pmT} , C^{pmT} , S^{pmT} and K^{pmT} represent the mean, standard deviation, and coefficients of variation, skewness and kurtosis for the number of points in a tiebreaker match. Tables 3 and 4 represent the exact parameters of the distributions for an advantage and tiebreaker match for different values of p_A and p_B. The results agree with Pollard (1983) for a best-of-three sets tiebreaker match. It shows that the mean, standard deviation, coefficients of variation, skewness and kurtosis of the number of points played are greater for an advantage match, compared to a tiebreaker match. Also included in the tables are the probabilities of the match lasting for at least n points, represented by $P(n)$ for an advantage match and $Q(n)$ for a tiebreaker match. These probabilities were calculated using the NP-expansion technique (Pesonen, 1987). Notice that when p_A and p_B become "large", the probability of playing at least 400 points in an advantage match is considerably greater than for a tiebreaker match. This is some justification as to why an advantage match can seemingly never end with two strong servers.

Table 5 represents the exact parameters of distributions for a tiebreaker and an advantage match using 50-40 games, along with the probability of a match going beyond 300 points. For an extreme case, when $p_A = p_B = 0.75$, the probability of an advantage match going beyond 300 points is 0.06. In comparison to Tables 3 and 4, the probability of an advantage or tiebreaker match going beyond 300 points is 0.38. This shows that replacing standard 'deuce' games with 50-40 games, substantially decreases the likelihood of long matches occurring.

It is often the case that by shortening the length of matches, decreases the probability of winning for the better player. However this is not necessarily the case as shown by replacing standard 'deuce' games with 50-40 games. Table 6 represents the probabilities of winning under four different scoring systems, for different values of p_A and p_B.

Notice when p_A=0.75 and p_B = 0.70, the probability of player A (the stronger player) winning using 50-40 games is greater than using standard 'deuce' games.

Table 3: The parameters of the distributions of points in an advantage match for different values of p_A and p_B

p_A	p_B	M^{pm}	U^{pm}	C^{pm}	S^{pm}	K^{pm}	$P(300)$	$P(350)$	$P(400)$
0.50	0.50	272.27	62.58	0.23	0.13	-0.54	0.33	0.12	0.02
0.55	0.55	272.10	62.55	0.23	0.15	-0.51	0.33	0.12	0.02
0.60	0.60	271.96	62.85	0.23	0.21	-0.37	0.33	0.12	0.02
0.65	0.65	273.48	65.13	0.24	0.40	0.14	0.32	0.12	0.03
0.70	0.70	280.72	74.16	0.26	0.92	1.96	0.34	0.15	0.06
0.75	0.75	300.52	103.49	0.34	1.89	6.23	0.38	0.22	0.12

Table 4: The parameters of the distributions of points in a tiebreaker match for different values of p_A and p_B

p_A	p_B	M^{pmT}	U^{pmT}	C^{pmT}	S^{pmT}	K^{pmT}	$Q(300)$	$Q(350)$	$Q(400)$
0.50	0.50	271.56	61.40	0.23	0.06	-0.67	0.33	0.11	0.02
0.55	0.55	271.25	61.12	0.23	0.06	-0.67	0.33	0.11	0.02
0.60	0.60	270.56	60.42	0.22	0.06	-0.69	0.32	0.10	0.01
0.65	0.65	270.52	59.77	0.22	0.05	-0.73	0.32	0.10	0.01
0.70	0.70	273.14	59.64	0.22	0.02	-0.79	0.34	0.11	0.01
0.75	0.75	278.81	59.54	0.21	-0.04	-0.88	0.38	0.13	0.02

Table 5: The parameters of the distributions of points in a tiebreaker and advantage match using 50-40 games for different values of p_A and p_B

p_A	p_B	M^{pmT}	U^{pmT}	C^{pmT}	S^{pmT}	K^{pmT}	$Q(300)$	M^{pm}	U^{pm}	C^{pm}	S^{pm}	K^{pm}	$P(300)$
0.50	0.50	198.43	44.01	0.22	0.05	-0.71	0.01	198.94	44.93	0.23	0.14	-0.54	0.02
0.55	0.55	199.71	44.39	0.22	0.05	-0.71	0.02	200.09	45.08	0.23	0.11	-0.58	0.02
0.60	0.60	201.93	44.89	0.22	0.05	-0.71	0.02	202.31	45.58	0.23	0.11	-0.58	0.02
0.65	0.65	205.18	45.52	0.22	0.05	-0.71	0.02	205.71	46.47	0.23	0.14	-0.53	0.03
0.70	0.70	209.79	46.40	0.22	0.06	-0.71	0.03	210.71	48.10	0.23	0.21	-0.38	0.04
0.75	0.75	216.64	47.81	0.22	0.06	-0.73	0.05	218.71	51.70	0.24	0.39	0.12	0.06

Table 6: The probabilities of winning a tennis match under different scoring systems

		standard games		50-40 games	
p_A	p_B	p^m	p^{mT}	p^m	p^{mT}
0.51	0.50	0.567	0.567	0.554	0.554
0.55	0.50	0.800	0.799	0.754	0.754
0.60	0.50	0.952	0.951	0.918	0.917
0.61	0.60	0.565	0.564	0.557	0.557
0.65	0.60	0.789	0.785	0.764	0.763
0.70	0.60	0.941	0.938	0.927	0.926
0.71	0.70	0.560	0.558	0.559	0.559
0.75	0.70	0.772	0.760	0.775	0.772

CONCLUSION

The mathematical methods of generating functions have been used to calculate the parameters of distributions of the number of points in a tennis match. The results show that the likelihood of long matches can be substantially reduced by using the tiebreak game in the fifth set, or more effectively by using the 50-40 game throughout the match.

We used the number of points played in a match as a measure of its length. This measure is related to the time duration of the match and avoids the complications of delays between points, at change of serve, at change of end, injury time and weather delays. Further work could involve calculating the time duration of a match from the results presented in Subsection 3.4. This could then be used to calculate the probabilities of the match going beyond a given amount of time. This would provide commentators and tournament officials with very useful information on when the match is going to finish.

REFERENCES

Barnett, T. and Clarke, S.R. (2005) Combining player statistics to predict outcomes of tennis matches. *IMA Journal of Management Mathematics*. **16**(2), 113-120.

Brown, A. (1977) Cumulants of convolution-mixed distributions. *Astin Bulletin*. 9(1,2), 59-63. Croucher, J.S. (1986) The conditional probability of winning games of tennis. *Research Quarterly for Exercise and Sport*. **57**(1), 23-26.

Pesonen, E. (1975) NP-technique as a tool in decision making. *Astin Bulletin*. **8**(3), 359-363.

Pollard, G.H. (1983) An analysis of classical and tie-breaker tennis. *Australian Journal of Statistics*. **25**(3), 496-505.

Pollard, G.H. and Noble, K. (2002) The characteristics of some new scoring systems in tennis. *In Proceedings of the 6M&CS*. G. Cohen ed., 221-226.

Pollard, G.H. and Noble, K. (2004) The benefits of a new game scoring system in tennis the 50-40 game. *In Proceedings of the 7M&CS*. R.H. Morton and S. Ganesalingam eds., 262-265.

Stuart, A. and Ord, J.K. (1987) *Kendall's Advanced Theory of Statistics*. Charles Griffin & Company Limited. London, **1**(5).

A COMPARISON OF COMPUTER AND HUMAN PREDICTIONS OF AFL

Clarke, S.[1] and Clarke, R.[2]

[1] Faculty of Life and Social Sciences, Swinburne University of Technology, Hawthorn, Vic., Australia

[2] 9 Wiarando Court, Doncaster East, Vic., Australia

ABSTRACT

The Swinburne Computer has been predicting match winners, margins, odds, and final ladder order since 1981. Although its predictions are published early in the week, before team selections are made, this paper shows the computer's performance compares favourably with that of media experts and the public. This suggests that far from making predictive use of extra information, humans do not individually process correctly the publicly available information. Combining their forecasts may produce better predictions. Evidence suggests the computer's predictions can be used to exploit market inefficiencies, not only early in the week when prices are first set, but also late in the week when prices have adjusted following team selections.

KEY WORDS

sports, Australian rules football, forecasting, betting

INTRODUCTION

All sporting events are preceded by predictions of the outcome, almost invariably made by human experts. Since the experts are close to the sport, have their ear to the ground, are privy to information not available to the general public, we may believe their predictions are somehow more accurate. But rarely is the performance of these (so called) experts critically examined. In AFL, a whole army of experts give detailed predictions of the winner and margins of each match. In addition newspapers publish bookmaker lines and odds for each match. These presumably contain the public's forecast (or at least the bookmaker's perception of the public forecast).

The Swinburne Computer tips (Clarke, 1988) have been published in the media for over 20 years. Clarke (1993) discusses the program's methods, which have been unchanged since 1986 when it was rewritten using a more complicated error function. The program uses as input data only the order of the matches, the names of the competing teams and venue, and the final scores of previous matches. It does not use other information which many followers would believe is critical to selecting winners - the actual players selected, the time of the match (day or night), the elapsed time since a team's last match, the weather conditions, etc. The program predicts not only the winner of an upcoming match, but the expected margin in points, the chance of each team winning, and via a simulation of the remainder of the year, the chance of each team finishing the home and away series in any position.

Despite the paucity of input data, indications are that the program has performed better than the average expert tipster. Clarke (1992) gives a table showing that in 1991 only 2 the 22 expert tipsters from The Age and the Herald/Sun had a greater percentage correct than the computer, and the computer's percentage correct was 3.4% higher than the average of the experts. In terms of predicted margins, only one of the 21 Herald/Sun experts and celebrity tipsters was closer than the computer, and the computer's margin of error was 1.5 points per match closer than the average of the experts.

In 2002 the program was marketed via Ozmium Ltd to punters (see www.smartgambler.com.au). Over the intervening years, information has been collected in an ad hoc manner on the computer's and various humans' performance. This paper compares the performance of the program to human tipsters in picking winners, margins and odds of winning, and investigates the profit made by using various betting strategies. We demonstrate that the computer in general performs better than humans, and that inefficiencies do exist in the AFL betting market.

COMPARISON WITH EXPERTS IN SELECTING WINNERS AND MARGINS.

In 2002 the computer correctly predicted 126 winners, including the finals and with draws counting as wins. This is about 68% correct, which is close to its long term average. The average margin of error in the predicted margin of victory of 28.8 points is again close to its long term average. This compares with an average of 121 winners and average error of 30.8 points for the 14 experts in The Herald/Sun and 16 experts in The Age. On a combined table of 30 experts and two computers (Swinburne and The Age), the Swinburne computer comes fifth in terms of winners. There appears to be no difference in the ability of The Herald Sun and The Age tipsters to pick winners – both Herald Sun and Age tipsters averaged just fewer than 121 winners, and the average rank of the Age was 16.9 compared with 16.7 for The Sun. In terms of margins, the Swinburne computer would come in second on the table. There did appear to be a difference in the ability of The Sun and The Age tipsters to pick margins – the average rank of the Age was 13.0 compared with 21.6 for The Sun. While Pearce topped the table in both categories, clearly some tipsters performed well at selecting winners, but poorly at margins. For example, both Timms and Horan were in the top 7 in the first category, but the bottom 8 in the second. It remains to be seen whether these trends are consistent from year to year.

It is interesting to compare how some standard methods would have performed. For example, picking the home side to win by 0 points would give 123 winners and have placed you twelfth on the combined table of winners but 29th in margins. Picking the favourite of The Age tipsters (ie the team selected to win by more than half the tipsters) would do the same. But calculating the average home margin of The Age tipsters and selecting the home side if this is positive would have resulted in 127 winners and an average margin of error of 28.8 points. This puts you equal third on the winner's table, and equal second on the margin table. This seems to indicate that combining the expert's tips along with their strength of belief as contained in their margin prediction can produce better results than combining their predicted winners.

However, care should be taken in judging performance on a single year. This applies particularly to the number of winners, which can be quite variable from year to year. Tips have been collected and collated for the 17 tipsters (16 humans and 1 computer) in

The Age for each year from 2000-2002. With the Swinburne computer this made 18 tipsters. The Swinburne computer came 2nd, 5th and 2nd in each year's winners tables, and 1st, 1st and 2nd in the margins tables. Table 1 shows the combined table for the 3 years for the 12 tipsters who predicted each year, along with the Swinburne computer, The Age computer and the records of some other possible strategies. The computer comes top in both number of winners and total margin of error. The 12 experts averaged 363.3 winners with a standard deviation of 8.2, and averaged an error in the margin of 32.6 with a standard deviation of 0.7. This placed the computer 2.0 and 2.7 standard deviations better than their average. Clearly in the long run the Swinburne computer is better than human tipsters in both winners and margins.

Table 1: Comparison of Swinburne computer and The Age experts: winners and margin tipping totals 2000-2002.

Tipster	**Number of winners**	**Tipster**	**Average Error in Margin**
Swin Computer	380	Swin Computer	30.7
Ryan	380	Aver Age Margin	31.0
Aver Age Margin	375	Pearce	31.9
Pearce	374	Walls	31.9
Age computer	370	Connolly	32.0
Age favourite	369	Ryan	32.0
Baum	368	Niall	32.4
Walls	367	Lyon	32.4
Connolly	364	Baum	32.4
Reilly	364	Age computer	32.7
Lyon	362	Reilly	32.7
Niall	360	Johnson	32.8
Johnson	359	Brereton	32.9
Brereton	356	McClure	33.0
McClure	356	Wilson	34.4
Wilson	350	Age favourite	35.2
Home side	327	Home side	35.5

In 2003 the computer had an ordinary season with winners, coming equal eighth when compared with 18 Herald Sun experts, but still topped the margin table. Its 120 winners was slightly better than the 118.8 averaged by the experts, but its average error of 27.4 was more then 2 better than the average of the experts.

In 2004, with 126 winners at a strike rate of 68.1% at the end of the Finals the computer would have been placed fourth when compared with the Herald Sun 18 experts, three behind the winner. The computer had 10 winners more than their major writer Sheahan, and nearly 6 winners more than the average of the Herald Sun experts (120.1 winners). Again the computer outperformed all the experts in margin prediction. With an average error of 29 points the computer averaged nearly 2 pts a game closer than the best expert, and 3.5 pts per game better than the average of the experts.

COMPARISON OF COMPUTER ODDS WITH BETTING MARKETS

The computer gives a probability of victory for each team. While media experts do not regularly give this, comparisons can be made with humans via tipping competitions and published bookmaker odds.

Monash University run three tipping competitions, one of which requires entering the probability of each team winning. The score is based on the logarithm of the probability you give to the team that actually wins. This is a scientifically based score, which statisticians call the likelihood, and computer scientists the information content. The results are obtainable from www.csse.monash.edu.au/footy . Note that not all of the entrants in this competition are human – it is to be expected that in a competition run by a computer science department many may be driven by computer algorithms.

In 2004 Rod entered the Swinburne computer tips probabilities each week in the Monash probabilistic tipping site under the pseudonym Drtipper. We missed putting in one tip on a Friday, which cost us about half a point. Nevertheless at the end of the home and away season Drtipper finished top of the 914 tippers that started the year. We dropped to third after the finals, but this is probably due to end play by the competitors who choose unrealistic probabilities at the end of the competition to maximize their chance of reaching the top. Outperforming humans in estimating probabilities suggests the predicted odds can be used to exploit betting markets.

Betting for profit depends on getting value – betting on teams for which the price paid by the bookmaker is better than justified by the true probability of the team winning. The overlay is the expected profit you would make on an individual bet if the computer estimate of the probability is correct. It is obtained by multiplying the computer's estimate of winning by the dividend or price. Thus a 23% chance of winning a bet paying $5 (4 to 1) gives 23% x $5 = $1.15 for a 15% overlay. Now of course the computer is not always correct, so it is usual to leave a margin for error. Betting on small overlays results in lots of bets at what should be reasonable odds; large overlays results in fewer bets at better odds.

Kelly (1956) optimises the percentage growth rate in your pool, with the proportion of pool bet depending on the overlay and chance of winning. Because your bet grows with the size of the pool, the Kelly system is extremely volatile, and carries a large risk that your pool will disappear. To reduce this risk, many punters prefer to use a half Kelly, one third Kelly or some other fraction, or else calculate the percentage of a constant pool (I call this constant Kelly). Kelly also assumes that you know the result of one bet before another is placed. In AFL betting we may have to bet on several matches simultaneously. It is difficult to replicate testing of Kelly systems, as the final total

depends on the order of the bets. Again, proportions based on a constant pool eliminate this.

Various analyses performed in the past have suggested it is possible to exploit market inefficiencies using the computer predictions. In 2001, research student Jonathan Lowe obtained the past odds on each team winning and simulated various betting strategies based on the computer's tips for 3 years from 1998-mid 2001. These showed a substantial profit for some strategies. Yelas (2003) also analysed several years data. An analysis by one of Ozmium's AFL subscribers Anderson (2002) showed certain Kelly based betting strategies produced profits for 2002. Because of the problems with betting order, we could not replicate his calculations. Our own analysis for 2002 used four simple strategies – Full Kelly, Half Kelly, Kelly on a fixed bank of $100, and a constant bet of $20. Each strategy was evaluated for betting on any match showing an overlay greater than 0, 2%, 4%, …20%. In general the full Kelly made losses, and the other three made gains. All the systems showed the bank less than the starting point at some stage of the season. For all systems, overlays from 0 to 10% seemed the most consistently profitable. From this analysis it appeared the Kelly system on a fixed bank produced the most consistent results of these three strategies in 2002, with the initial bank of $100 accumulating to at least $245 for all overlays of 10% or less.

The computer's performance in 2004 is analysed in more detail in Clarke and Clarke (2004). Bookmaker's odds were collected for each round. These were generally those offered by TAB Sportsbet as published each Monday in The Herald Sun, and globalsportsbet odds as published each Friday. There were odd occasions when odds were not available and others from the Web were used. Betting $10 early in the week on the computer's predicted winner for each match would have resulted in 185 bets, $1,850 invested for winnings of $1803.20. This represents a return on investment (ROI) of -2.5%, or a loss of 2.5%. Note this is never recommended, but is still a good result given the bookies margin of about 8% for the TAB Sportsbet odds. Betting $10 on every team for which the computer predictions showed an overlay would have resulted in 116 bets, $1,160 invested for winnings of $1,331.80. This represents a ROI of 14.8%. Depending on overlay safety margins chosen, profits ranged from 22% to 58% betting Head to Head at Sportsbook odds as published in Monday's Herald Sun, 19% to 28% betting Head to Head at Friday's odds, and from 11% to 95% betting on margins at Friday's odds. We give here a couple of relevant tables.

Head to head betting

Table 2 shows the number of bets, number of winning bets, amount bet, amount won, profit and return on investment for the strategy of betting a fixed amount of $10 on overlays of 0%,2%,4%,…20%. All show a positive return, with ROI generally increasing as the overlay increases. Note this is an increasing percentage of a smaller amount, as the number of bets decreases as the required overlay increases.

Table 2: Head to head betting early in the week with TAB Sportsbet 2004:
Returns from a constant bet of $10 on all teams with given overlay

Overlay	Number of bets	Number of winning bets	Amount bet $	Amount won $	Profit $	ROI
0%	116	80	1,160.00	1,331.80	171.80	14.8%
2%	96	66	960.00	1,112.80	152.80	15.9%
4%	87	59	870.00	1,006.20	136.20	15.7%
6%	72	48	720.00	858.70	138.70	19.3%
8%	52	40	520.00	727.20	207.20	39.8%
10%	37	27	370.00	487.90	117.90	31.9%
12%	26	21	260.00	398.70	138.70	53.3%
14%	18	13	180.00	277.50	97.50	54.2%
16%	13	10	130.00	220.50	90.50	69.6%
18%	12	9	120.00	206.50	86.50	72.1%
20%	10	7	100.00	173.50	73.50	73.5%

While not recommended due to the volatility, Table 3 give the results of using a Kelly Percentage of a variable pool starting with $100. All overlays under 12% result in final pools over 10 times the original. While for 2004 there appears to be little downside, with the pool rarely going below the starting pool, this is not recommended as a long term strategy. It might be fun or exciting to try each year with a small discretionary starting pool, but might often result in losing all the starting pool.

Table 3: Head to head betting early in the week with TAB Sportsbet 2004:
Returns from betting the Kelly percentage of a variable pool with a $100 start

Overlay	Amount bet $	Amount won $	ROI	Final Pool $	Min Pool $	Max Pool $
0%	14,790.74	16,764.12	13.3%	2,073.38	87.08	3,076.44
2%	13,531.98	15,351.61	13.4%	1,919.63	87.08	2,884.76
4%	12,241.38	13,906.79	13.6%	1,765.41	87.08	2,653.00
6%	9,610.94	11,153.14	16.0%	1,642.20	87.08	2,467.86
8%	12,926.32	15,401.69	19.1%	2,575.38	100.00	3,870.20
10%	5,424.45	6,740.57	24.3%	1,416.12	100.00	2,128.11
12%	3,772.33	4,736.18	25.6%	1,063.84	100.00	1,598.71
14%	1,078.85	1,383.24	28.2%	404.39	100.00	607.71
16%	914.86	1,268.08	38.6%	453.22	100.00	681.08
18%	739.35	1,024.28	38.5%	384.93	100.00	578.46
20%	454.20	623.69	37.3%	269.49	100.00	404.98

Odds were also given in the Herald Sun on Friday from Globalsportsbet. These tended to have a lower bookmaker's margin of about 4%, but would be expected to better approximate public opinion and allow for team selections. Table 4 showS the results using these odds.

Table 4: Head to head betting late in the week with Global Sportsbet:
Returns from a constant bet of $10 on all teams with given overlay

Overlay	Number of bets	Number of winning bets	Amount bet $	Amount won $	Profit $	ROI
0%	142	92	1,420.00	1,803.20	383.20	27.0%
2%	118	72	1,180.00	1,445.40	265.40	22.5%
4%	100	59	1,000.00	1,241.40	241.40	24.1%
6%	87	52	870.00	1,120.90	250.90	28.8%
8%	72	39	720.00	884.90	164.90	22.9%
10%	65	35	650.00	827.30	177.30	27.3%
12%	53	30	530.00	716.80	186.80	35.2%
14%	40	23	400.00	587.60	187.60	46.9%
16%	33	17	330.00	446.50	116.50	35.3%
18%	32	16	320.00	427.00	107.00	33.4%
20%	26	12	260.00	353.20	93.20	35.8%

Because of the lower bookmaker's margins, more bets are made using Friday's odds. The returns on investment are generally higher for small overlays and lower for the higher overlays. An analysis of the individual bets would be needed to discover the process at work here, but clearly the market is still inefficient even after team selections.

Margin Betting

Since the computer selects margins relatively better than winners, one might expect betting on margins to be a better proposition than head to head. Lines, which allow betting at the margin for a return of $1.95, were given in the Herald Sun each Friday from Globalsportsbet. These have a low bookies margin of only 2.5%, and given the computer's superiority in predicting margins should be a prime opportunity. The lines were obtained from The Herald Sun each Friday for each week except for round 3 and one match in Round 13. For these missing values, an approximate line based on the Monday TAB sportsbet odds was used.

For line betting, the computer gives an expected margin, so either side of this margin would be expected to occur with 50% chance. A safety margin can be built in by requiring the line to differ from the computer's prediction by at least a given figure. However betting on the line with no safety margin in 2004 would have resulted in 105 winning bets out of 184, for a return on investment of 11.3%. Details for this and other safety margins are shown in Table 5.

Table 5: Results of fixed bet size $10 on the line at globalsportsbook $1.95 odds

Safety margin	Number of bets	Number of winning bets	Amount bet $	Amount won $	Profit $	ROI
0	184	105	1,840	2,048	208	11.3%
1	163	95	1,630	1,853	223	13.7%
2	148	89	1,480	1,736	256	17.3%
4	120	77	1,200	1,502	302	25.1%
6	92	60	920	1,170	250	27.2%
8	72	48	720	936	216	30.0%
10	53	35	530	683	153	28.8%
12	41	27	410	527	117	28.4%
14	30	20	300	390	90	30.0%
16	19	12	190	234	44	23.2%
18	13	10	130	195	65	50.0%
20	11	8	110	156	46	41.8%
22	9	6	90	117	27	30.0%
25	3	3	30	59	29	95.0%

While we have not analysed profits using Kelly type bet sizing for line betting, for those interested in dabbling in the full Kelly with an increasing pool, this sort of betting may have the advantage that all bets would have a greater than 50% chance of success. This should reduce the chance of a bank busting losing sequence.

CONCLUSION

There is a body of evidence that the computer performs better in the long run than the general newspaper expert in selecting winners and margins. While the occasional expert outperforms the computer in winners, none matches the accuracy of the computer's margin. Evidence from several sources also suggests that it is possible to use the computer's odds to develop profitable betting strategies. This suggests the computer's predictions of a team's chances are better than both the experts and the general public.

The superiority of the computer over humans is surprising given the humans have much more supposedly relevant information than the computer (particular player personnel, training form, injuries, time from last match etc). This implies the humans are not processing that extra information correctly. Bailey and Clarke (2005) have shown that the computer can incorporate this extra information to further improve the computer's predictions, particularly in the area of odds.

REFERENCES

Bailey, M and Clarke, S. (2005) Deriving a profit from Australian rules football: a statistical approach. In *Seventh Conference on Mathematics and Computers in Sport* H. Morton (Ed). Palmerston Nth, NZ.

Clarke, S. (1988). Tinhead the tipster. *OR Insight* **1**(1), 18-20.

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

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

Clarke, S. and Clarke, R. (2004). AFL Computer tipping - Sum up of 2004 season. Private report. (available from the author)

Anderson, G (2002) 2002 AFL football results analysis. Private communication.

Kelly, J. (1956). A New Interpretation of Information Rate. *Bell System Technical Journal* **35**, 917-926.

Yelas, S. (2003). Analysis of the Swinburne Computer performance in identifying inefficiencies in the AFL betting market. MSc thesis. School of Mathematical Sciences, Swinburne University of Technology.

REAL-TIME PREDICTIVE 3D TRACKING OF CRICKET BALLS FOR KNOWLEDGE OF RESULTS AND STRATEGY TRAINING

Gao, H., **Green, R.**

University of Canterbury, Dept. of Computer Science & Software Engineering, NZ.

ABSTRACT

A substantial amount of cricket training is done in 'nets' because they control the travel of the ball and make it possible to practice without spending a lot of time retrieving cricket balls. Baseball also uses 'batting cages' for this purpose. One key problem with batting in a net is the difficulty in working out if the shot selected would result in runs scored. In motor learning terminology this is called Knowledge of Results (KR). Since the aim of batting is to score runs, nets have a low KR quotient and so low feedback into training. This research introduces SeeUPlay® using cameras mounted at a height of 3.5 metres above the batter that tracks the cricket ball trajectory and speed as it is bowled and angle and acceleration as it is played to generate substantial information about the performance of the batter and the bowler. In this research, a State-based "Observation, Analysis and Prediction" (SOAP) target tracking algorithm is proposed. This algorithm is able to tracking objects with discontinuous trajectories and dynamically changing motion properties. The experiment results show that this system detects up to 93.5% of moving objects in noisy environment with a tracking accuracy of up to 97.42%.

KEYWORDS

computer vision, cricket, knowledge of results, tracking.

INTRODUCTION

Computer vision techniques are used in sport training applications for detecting and tracking trajectories of players (Needham, 2001) and balls (Ren et al., 2004; D'Orazio et al., 2002; Wedge et al., 2004). The captured information (e.g. trajectories) are recorded and analysed to support human coaches to improve the performance of players.

For example, Kovacic and Pers have developed a computer vision based handball tracking application (Kovacic and Pers, 2000). In this system, two stationary cameras are mounted directly above the court so that all players can be observed. To segment players from the captured images, an image of an empty court (without players or balls) is modelled as the background, which contains static components of the observation only. The players are detected by differencing the background image and the latest image captured during the match because the latest image contains both static components and dynamic components (players and a ball). Both temporal and colour matching approaches are used to track players. The trajectories of players are recorded according to the tracking results. This data is used to analyse the performance of each player.

In addition, Wedge presented a football tracking system in (Wedge et al., 2004). In this system, the trajectory of a ball moving over a cluttered background is found in panning and zooming camera video sequences. The video sources used in this system were the digital broadcast videos of Australian Football League matches. In this system, two consecutive greyscale images are aligned and one is subtracted from the other to search for motion regions. These regions are usually triggered by the motion of a football. Its position will be treated as the start point of a trajectory and is feed into a Kalman filter (Welch and Bishop, 2004) to predict the location and appearance of this target in the succeeding frame.

Although computer vision based computer aided sports training (CAST) systems are well researched, these systems usually expect the tracked targets have continuous trajectories with stable motion properties. In this paper, we propose a state based target tracking algorithm (the SOAP algorithm) that can track targets with discontinuous trajectories and dynamic changing motion properties. In addition, a cricket ball tracking system that implements this algorithm is introduced. This application is tested in both indoor and outdoor environments and shows good tracking results.

TARGET TRACKING

Relative Work

Tracking is "the problem of generating an inference about the motion of an object given a sequence of images" (Lipton, 1998).

The position of an object in the succeeding frame can be estimated as in the area that surrounds the object's current position (Funk, 2003). Kovacic and Pers (2000) used this approach for tacking handball players. Although this approach is simple and fast in calculation speed, it has one significant drawback: low prediction accuracy. This problem is caused by ignoring the target's motion properties.

In contract to the above approach, some tracking algorithms estimate the appearance of the target by both its last observed location and its motion properties (such as velocity and acceleration). The prediction accuracy is improved because more information (e.g. the priori knowledge of the motion properties) is involved in the estimation.

Kalman filter (Welch and Bishop, 2004) is a common algorithm to estimate the position of a moving object with the priori knowledge of its motion properties. It is a recursive mathematical algorithm that infers the variable of interest based on observed variables in the past (the measurement data) and prior knowledge of the system (Mayeck, 1979). Kalman filter contains two stages in each iteration step; they are the prediction stage and the measurement stage. The prediction stage, also referred to as the time update stage, is responsible for projecting forward in time the current state and error covariance estimators to obtain an estimate of the next time step. The measurement stage, on the other hand, incorporates a new measurement into the previous estimation (from the prediction stage) to obtain an improved posterior estimate (Welch and Bishop, 2004). In other words, the prediction stage predicts the target state in the next frame and the measurement stage corrects the estimation by observed information in the succeeding frame.

Although a Kalman filter improves the prediction accuracy of a tracking, it may fail when the target's motion properties are changed dynamically. For example, when a Kalman filter is applied to sport applications, it may fail when the target is hit by a player since there is a discontinuity in the ball's motion properties and the new trajectory is unable to be estimated.

SOAP Algorithm

We present a state based target tracking algorithm, the SOAP algorithm, to overcome the drawbacks of existing object tracking systems. Powered by a state based logic, the SOAP algorithm is able to robustly track objects with rapidly changing motion properties and discontinuous trajectories.

As the SOAP algorithm relies on priori knowledge of the target object's motion properties, we use an example of cricket ball tracking system to introduce this algorithm. This cricket ball tracking system is discussed in the "Application" section.

SOAP algorithm is a state based target tracking algorithm. A finite state machine (FSM) is used in the SOAP algorithm to determine its behaviours. There are six states in the SOAP algorithm designed for tracking a cricket ball. The six states are:

- S1: Inward cricket ball detection state.
- S2: Inward cricket ball tracking state.
- S3: Outward cricket ball detection state.
- S4: Outward cricket ball tracking state.
- S5: Idle state.
- S6: Stop state.

In addition to the six states, there are eleven input symbols in the cricket tracking SOAP algorithm. These input symbols are summarised in Table 1 and are illustrated in Figure 1. The number of input symbols may different when apply the SOAP algorithm in other domains, for example football or tennis.

Table 1: The cricket ball tracking SOAP algorithm has eleven input symbols.

No.	Symbol	Comment
1	[a]	An inward cricket ball candidate is detected.
2	[b]	The inward cricket ball target is elected among candidates.
3	[c]	The inward cricket ball target is tracked for another frame.
4	[d]	The inward cricket ball target is merged into the batter's influence area.
5	[e]	The target is close to image boundaries
6	[f]	An outward cricket ball candidate is detected.
7	[g]	The outward cricket ball target is elected among candidates.
8	[h]	The idle state timer set off.
9	[i]	The cricket ball target is not able to be found.
10	[j]	The SOAP algorithm is stopped
11	[k]	The SOAP algorithm is restarted

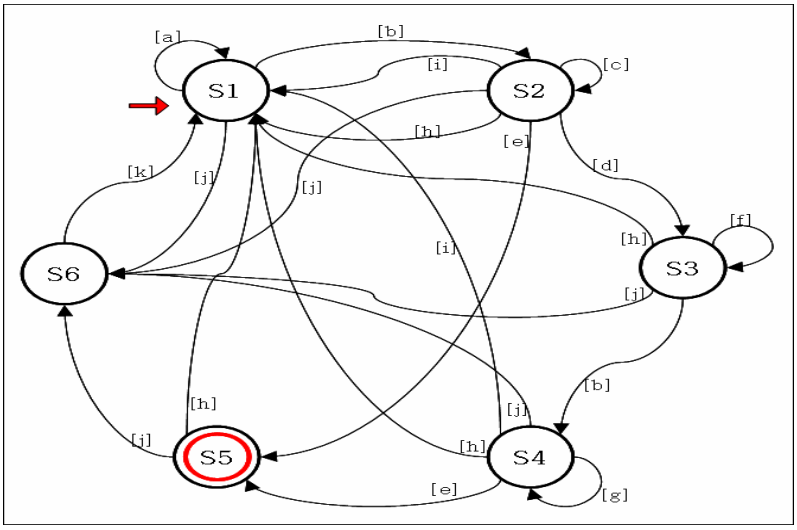


Figure 1: The state transition diagram of SOAP algorithm. State S1 is the initial state. State S5 is the acceptance state.

As discussed above, the FSM forms the core framework of the SOAP algorithm. In each tracking state (S1 to S4), a trilogy of "Observation", "Analysis" and "Prediction" is implemented to elect the cricket ball from its candidates and track the ball as long as it is moving and visible. The "Observation", "Analysis" and "Prediction" steps are discussed below.

Observation:
The observation step achieves the following goals:

1. Seek for the appearance of a cricket ball in the predicted areas.
2. If the cricket ball is not found, the observation step will adjust image segmentation threshold value and notify the prediction step to redo the prediction (with looser constraints).
3. If the cricket ball is still not found even after threshold value and prediction adjustments, the observation will abort the tracking process for this target.

Analysis:
The analysis step is to analyse the target's trajectory information, which was collected by the "Observation" step. The analysis step has two main tasks:

1. In both detection states (S1 and S3), the analysis step assigns a score for each observed candidate (objects that look like cricket balls). By thresholding the scores of all candidates, the real cricket ball target can be selected. Once the target has been chosen, the analysis step will change the SOAP algorithm to a tracking state (S2 or S4).
2. Analyse the trajectory of a cricket ball target. If a target is going to merge into the batter's influence area, it will report an inward trajectory captured event and change the state of the SOAP algorithm to the outward ball detection state (S3). Otherwise, if a target is going to fly out of image boarders, it will report an outward trajectory captured event and change the state of the SOAP algorithm to the idle state (S5).

Prediction:
The goal of the prediction step is to estimate the target's appearance in the successive frame based on the priori knowledge of its motion features. It completed the following tasks:

1. In each cricket ball detection state (S1 or S3), the prediction step defines a searching area based on the priori knowledge of cricket ball's motion properties (Figure 2 and 3). Any objects in this area will be detected by the observation step.
2. In each cricket ball tracking state (S2 or S4), it estimates the appearance of a cricket ball object by its past motion features. An estimation error value (shown in Figure 4 with the radius of the yellow circle surround the estimated location) is also given as a tolerance to the prediction error. This tolerance error value may be adjusted if the observation step requires it.

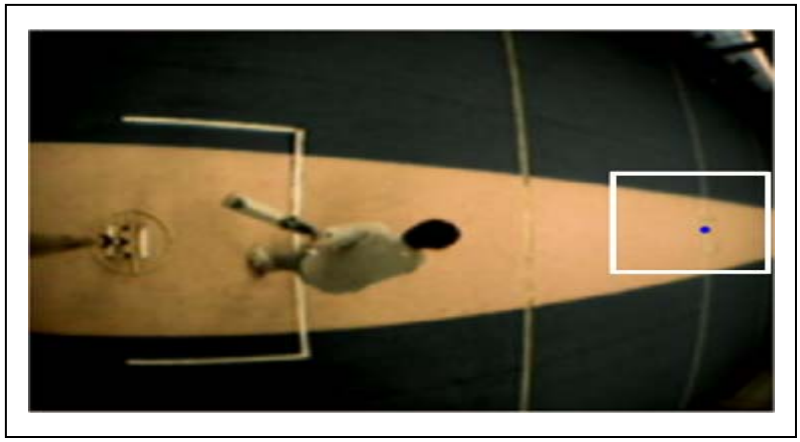


Figure 2: The predefined searching area in the inward cricket ball detection state (S1).

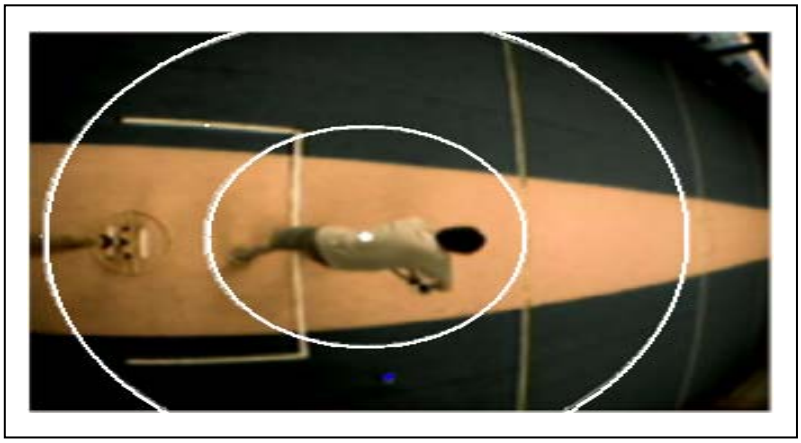


Figure 3: The predefined searching area in the outward cricket ball detection state (S3).

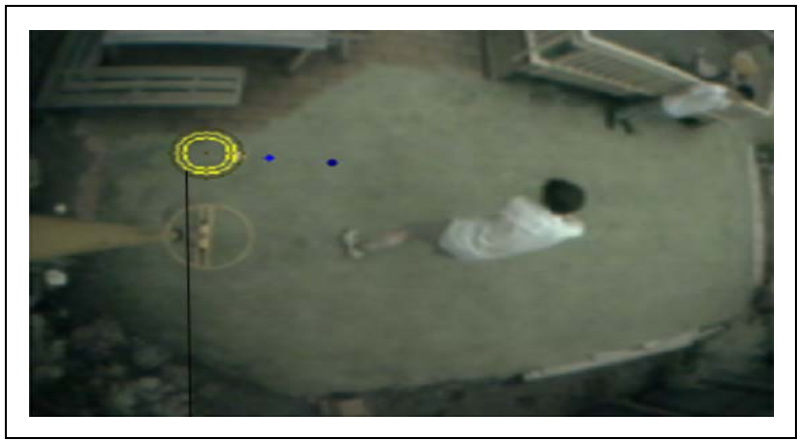


Figure 4: The tolerance estimation error.

As shown in the figure, the first circle (the internal one) indicates the initial tolerance error value. The two larger circles indicate the adjusted tolerance values.

APPLICATION

By utilising the SOAP algorithm discussed above, a cricket ball tracking application, SeeUPlay® Cricket, has been developed. It captures the cricket balls' trajectories with a camera mounted at a height of 3.5 meters above the batter. The SeeUPlay® Cricket application also analyses the trajectories information and calculates the cricket ball's bowling and pitching speed and the direction of cricket balls.

The user interface (UI) of the SeeUPlay® Cricket application is shown in Figure 5. As illustrated, the UI of this application contains three main parts:

- The wagon wheel (WW) part is located at the left half of the window. The wagon wheels are plotted in this area. The direction and length of a wagon wheel is related to the ball's tracked trajectory. Different balls are plotted in different colours. The red colour wagon wheel always represents the latest ball.

- The live video is rendered at the right top part of the window. This video image is updated at approximately 30 frames in a second. The ball's bowling information is also displayed on top of the live video image. For example, the length (the bounce position of an inward cricket ball) is indicated as a yellow dot on the image.

- The text based information part is displayed at the bottom right of the UI window. This text based information includes: total overs, current overs and balls, total runs, runs per over and average bowling speed. The latest bowling speed is shown in the blue round button on the top of the middle column between the wagon wheel and the live video window.

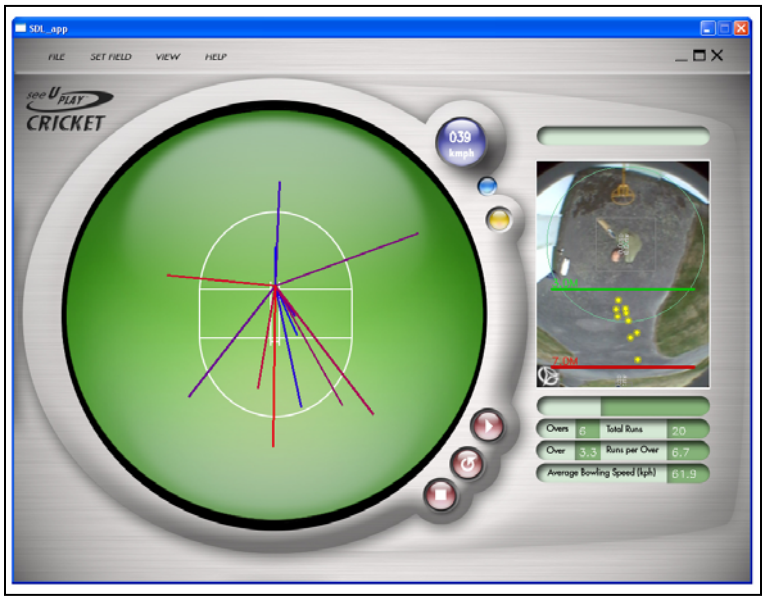


Figure 5: The Graphical User Interface (GUI) of the SeeUPlay® Cricket application.

Apart from the three main components of the UI, some control buttons are located at the bottom of the middle column. These buttons are: live play button, replay button and stop button.

The SeeUPlay® Cricket application is designed for two purposes. Firstly, for professional players, this application provides real time training support. As a substantial amount of cricket training is done in 'nets', which are used to control the travel of the ball and make it possible to practice without spending a lot of time retrieving cricket balls from around a field. One key problem with batting in a net is the difficulty in working out if the shot selected would result in runs scored. In motor learning terminology this is called Knowledge of Results (KR). This problem is solved by the SeeUPlay® Cricket system because it is able to calculate the cricket ball's trajectory parameters in real time and provide live feedbacks.

Secondly, the SeeUPlay® Cricket application will encourage young people to both be outside playing cricket, and also inside in front of a computer. With the current child obesity and fitness concerns being blamed on television and computer games, this application will provide an ideal model for future play: a compromise between the indoor and the outdoor, between what is "cool" for young people, and what is good for their health.

EXPERIMENTS

Two experiments have been done to evaluate the detection rate and tracking accuracy of the SeeUPlay® Cricket application. These experiments are done in both indoor environments (Figure 3) and outdoor environments (Figure 4).

Target Detection Test:
The target detection test is to evaluate the percentage of cricket balls that can be detected. In this test, the successful detection rates are calculated for both inward and outward targets. Manual observation is involved in this test to classify successful detections and failed detections.

Table 2 illustrates the results obtained from the target detection test. In this table, the numbers of balls that have been observed by the human observer are listed in the "Balls Observed" column. The "Total detected" column lists the number of balls that the test system has detected. The "False Negatives" column lists the number of balls observed by human subjects but ignored by the test system. The "False Positives" column lists the number of detected balls that are actually caused by camera noise or moving objects other than a cricket ball.

Table 2: The result of target detection experiment

Environment	Balls Observed	Total detected	Correct	False Positives	False Negatives
Indoor	62	62	58	4	4
Outdoor	56	55	51	4	5
Total	118	117	109	8	9

In the indoor environment, 93.5% of the cricket balls (58 out of 62) were detected by the SeeUPlay® Cricket system successfully. The rest of balls (6.5 percent) are missed. Apart from the total of 58 balls, the test system also reports four appearances of cricket balls. However, these four balls cannot be confirmed by the human subject. These four balls are classified as false positive cases. Thus the rate of false positives in an indoor environment is 6.5%.

In an outdoor environment, 91.1% of balls have been detected successfully. Besides the missed balls, another four ghost balls (false positives) have been classified by the human subject. The false positive rate is 7.3%.

Tracking Accuracy Test:
The tracking accuracy test is carried out to evaluate the accuracy of the SeeUPlay® Cricket application. The accuracy of the tracking results is defined as the distance between two positions: the position that the tracking system localized and the position that is localized by a human subject. The tracking error is defined as the distances between the manually specified position and the position tracked by the test system. The manual measurements may introduce positioning error as well. However, the errors introduced by human factor are ignored in this experiment.

The tracking accuracy test has been done with 86 tracked trajectories (in both indoor environment and outdoor environment). The accuracy of 620 balls observations along the 86 tracked trajectories has been calculated.

Table 3 illustrates the statistical results in this experiment. The mean distance error in this experiment is 1.37 pixels and 97.42% of tracking errors are less than or equal to three pixels in distance.

Table 3: The results of tracking accuracy test.

Function name	Subjects	Sample Size	Results
mean	distance	620	1.37 pixel
max	distance	620	15.52 pixels
standard deviation	distance	620	1.09

CONCLUSION AND FUTURE WORK

In this paper, we present the SOAP algorithm that utilises a FSM framework to solve the trajectory discontinuity problem. Since the ball's trajectory in sports usually contains discontinuous segments as a consequence of external forces (e.g. player bats a ball), the SOAP algorithm is especially useful in tracking balls in sports.

A cricket ball tracking application, SeeUPlay® Cricket, is also presented in this paper. This system implements the SOAP algorithm to track cricket balls and plots their trajectories as wagon wheels. This system also analyses cricket balls' trajectories to provide KR information for users to benefit both professional and domestic users.

A set of experiments have been done over the SeeUPlay® Cricket application in both indoor and outdoor environments. Up to 93.5% of cricket balls have been tracked successfully and up to 97.42% of the tracked balls are accurate.

In the future, this research could be improved in two directions. Firstly, as the framework of the SOAP algorithm, a FSM, has to be established manually with priori knowledge of the target's motion properties, future research can be done to create a method to establish FSM automatically. Secondly, the cricket ball tracking application analyses a ball's trajectory to plot its speed, direction and bounce location. Other trajectory parameters, such as spin angle and acceleration can also be calculated through the ball's tracking results. Future research can be done to develop other analysis techniques over the ball's tracked trajectory to provide more useful descriptions of a ball for coaches and players.

REFERENCES

Cai, Q. and Mitiche, A. and Aggarwal, J.K. (1995) Tracking human motion in an indoor environment. *Image Processing Proceedings., International Conference.*215-218.

D'Orazio, T. and Ancona, N. and Cicirelli, G. and Nitti, M. (2002) A ball detection algorithm for real soccer image sequences. *Pattern Recognition Proceedings*. 210-213.

Forsyth, D. and Ponce, J. (2003) *Computer Vision: a Modern Approach*. Prentice Hall, Upper Saddle River, N.J. London. 373-397.

Funk, N. (2003) *A Study of the Kalman Filter applied to Visual Tracking*. Project Report CMPUT 652, University of Alberta.

Gao, H. and Green R. (2004) State-based ball detection and tracking for cricket training. In *Proceedings of Image and Vision Computing New Zealand*, Akaroa, New Zealand. 423-428.

Iannizzotto, G. and Villari, M. and Vita, L. (2001) Hand tracking for human-computer interaction with graylevel visualglove: Turning Back to the Simple Way. *ACM International Conference Proceeding Series, Proceedings of the workshop on Perceptive user interfaces, Orlando, Florida*. 1-7.

Kovacic. S. and Pers. J. (2000) Computer vision system for tracking players in sports games. *First Int'l Workshop on Image and Signal Processing and Analysis*. 201-204.

Lipton, A. and Fujiyoshi, H. and Patil, R. (1998) Moving target classification and tracking from real-time video. *In Proceeding of 4th IEEE Workshop on Application of Computer Vision*. 8-14.

Mayeck, P. S. (1979) *Stochastic Models, Estimation, and Control*. Academic Press; New York. 1-16.

Milan S. and Vaclav H. and Roger B. (1998) *Image Processing, Analysis, and Machine Vision*. Brooks and Cole Publishing.

Needham, J. C. (2001) *Tracking Multiple Sports Players Through Occlusion Congestion and Scale*. Technique Report, University of Leeds. 1-15.

Ren, J. and Orwell, J. and Jones, G.A. and Xu, M. (2004) A general framework for 3D soccer ball estimations and tracking. *IEEE International Conference on Image Processing*. 1-8.

Tian, Y.L. and Hampapur A. (2005) Robust salient motion detection with complex background for real-time video surveillance. *WACV-Motion.* 30-35.

Toyama, K. and Krumm, J. and Brumitt, B. and Meyers, B. (1999) Wallflower: principles and practice of background maintenance. *The Proceedings of the Seventh IEEE International Conference.* 255 – 261.

Wedge, D. and Huynh, D. and Koves, P. (2004) Tracking footballs through clutter in broadcast digital videos. *In Proceedings of Image and Vision Computing New Zealand, Akaroa, New Zealand*. 155-160.

Welch, G. and Bishop, G. (2004) *An Introduction To the Kalman Filter.* Technique Report, University of North Carolina. 1-24.

PREDICTING THE MATCH OUTCOME IN ONE DAY INTERNATIONAL CRICKET MATCHES, WHILE THE GAME IS IN PROGRESS

Bailey M. [1] and Clarke S.[2]

[1] Department of Epidemiology & Preventive Medicine, Monash University, Australia

[2]Swinburne University of Technology, Melbourne, Australia.

Millions of dollars are wagered on the outcome of one day international (ODI) cricket matches, with a large percentage of bets occurring after the game has commenced. Using match information gathered from all 2200 ODI matches played prior to January 2005, a range of variables that could independently explain statistically significant proportions of variation associated with the predicted run totals and match outcomes were created. Such variables include home ground advantage, past performances, match experience, performance at the specific venue, performance against the specific opposition, experience at the specific venue and current form. Using a multiple linear regression model, prediction variables were numerically weighted according to statistical significance and used to predict the match outcome. With the use of the Duckworth-Lewis method to determine resources remaining, at the end of each completed over, the predicted run total of the batting team could be updated to provide a more accurate prediction of the match outcome. By applying this prediction approach to a holdout sample of matches, the efficiency of the "in the run" wagering market could be assessed. Preliminary results suggest that the market is prone to overreact to events occurring throughout the course of the match, thus creating brief inefficiencies in the wagering market.

KEY WORDS

linear regression, live prediction, market efficiency, BetFair

INTRODUCTION

The first official one day international (ODI) match was played in 1971 between Australia and England at the Melbourne Cricket Ground. Whilst ODI cricket has developed over the past 35 years (2300 matches), the general principles have remained the same. Both sides bat once for a limited time (maximum 50 overs) with the aim in the first innings to score as many runs as possible, and in the second innings to score more than the target set in the first innings. The high scoring nature of ODI matches ensures that team totals and differences between scores can be well approximated by a normal distribution. As shown by (Bailey, 2005), this facilitates the use of multiple linear regression to predict a margin of victory (MOV) prior to the commencement of the match. Using a similar approach, a multiple linear regression is also used to predict the number of runs scored by the team batting first. With the use of (Duckworth and Lewis, 1999) approach of converting resources available into runs, as each over is bowled, the current total and the predicted total for the remaining overs are combined to produce an updated predicted total for the batting team. The difference between the pre-match predicted total and the updated

predicted total provides a measure of how the batting team is performing through the course of their inning. This difference is then used to provide an updated prediction for the MOV.

METHOD

In ODI cricket the aim of the team batting first is to score as many runs as possible in the allotted time (usually 50 six ball overs). If the first team scores more runs than the second team, the MOV can readily be expressed in terms of runs difference between the two teams. The aim of the side batting second is to score more runs than the first team. Because the game is deemed to be finished if the team batting second achieves their target, the MOV is usually expressed in terms of resources (wickets and balls) remaining, rather than runs. In order to develop a predictive process for match outcomes, a consistent measure of the MOV is required. This can be achieved by following the work of Duckworth and Lewis to convert resources available into runs.

Table 1: Percentage of resources available for overs remaining and wickets lost

Overs remaining	**Wickets lost**									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
50	100.0	93.4	85.1	74.9	62.7	49.0	34.9	22.0	11.9	4.7
40	89.3	84.2	77.8	69.6	59.5	47.6	34.6	22.0	11.9	4.7
30	75.1	71.8	67.3	61.6	54.1	44.7	33.6	21.8	11.9	4.7
25	66.5	63.9	60.5	56.0	50	42.2	32.6	21.6	11.9	4.7
20	56.6	54.8	52.4	49.1	44.6	38.6	30.8	21.2	11.9	4.7
15	45.2	44.1	42.6	40.5	37.6	33.5	27.8	20.2	11.8	4.7
10	32.1	31.6	30.8	29.8	28.3	26.1	22.8	17.9	11.4	4.7
5	17.2	17.0	16.8	16.5	16.1	15.4	14.3	12.5	9.4	4.6
1	3.6	3.6	3.6	3.6	3.6	3.5	3.5	3.4	3.2	2.5

Frank Duckworth and Tony Lewis invented a now well-known system for resetting targets in ODI matches that were shortened due to rain. Although this system has undergone several refinements in recent years, the general way in which the Duckworth-Lewis (D-L) method is calculated has not changed, with wickets and balls remaining expressed as resources available and converted to runs. Table 1 shows an abbreviated version of the remained resources (R) for wickets lost and balls remaining. A complete tables and detailed account of the derivation of this table is given by Duckworth and Lewis (1999).

Whilst the D-L approach was specifically designed to improve 'fairness' in interrupted one-day matches, (de Silva et al., 2001) found that when used to quantify the MOV, the D-L approach sometimes overestimated the available resources when the second team to bat won easily, and underestimated the available resources when the second team to bat only just won. By minimizing the Cramer-von Mises statistic for the differences between actual and predicted runs, de Silva derived a formula to reduce bias by modifying the remaining resources. This is given by

$$R_{mod} = (1.183 - 0.006R)R \quad (1)$$

where R_{mod} = modified resources and R=resources given using D-L (see Table 1).

When an ODI match is won by the team batting first, the MOV is readily determined by the difference in runs scored. When the match is won by the team batting second, the MOV can be found by multiplying the first innings run total by the corresponding modified percentage of resources remaining as given by (1). *By referencing the MOV so that a 'home' win has a positive value and an 'away' win has a negative value, it can be seen from Figure 1, that the underlying distribution for MOV can be well approximated by a Normal distribution.*

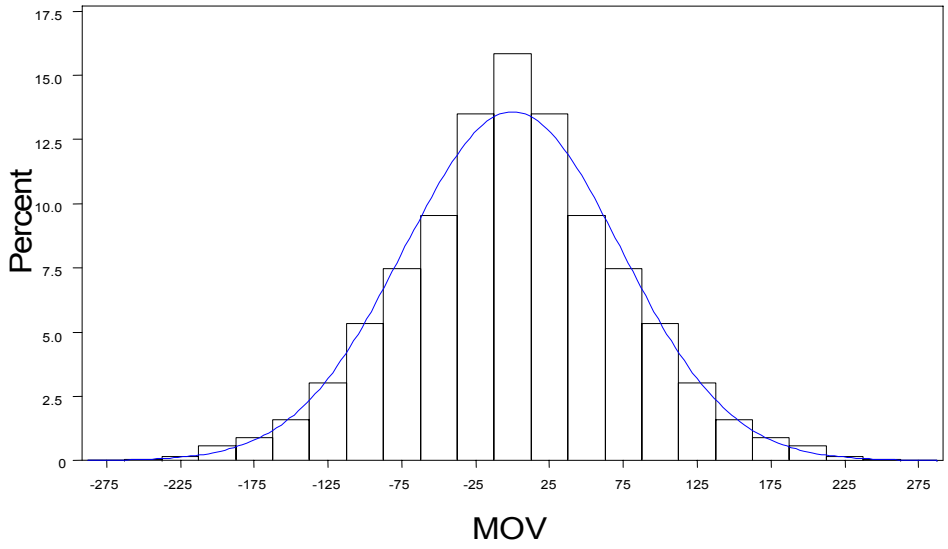


Figure 1: Histogram of MOV referenced against the home team

Statistical Analysis

All analysis was performed using SAS version 8.2.[1] Multiple linear regression models were constructed using a stepwise selection procedure and validated a backward elimination procedure. To increase the robustness of the prediction models a reduced level of statistical significance was incorporated with all variables achieving a level of significance below p=0.005. Comparisons between continuously normally distributed variables were made using student t-tests.

Prediction models for MOV

Using match and player information from 1800 ODIs played prior to Jan 2002, (Bailey, 2005) combined measures of recent form, experience, overall quality and home ground

[1] SAS Institute Inc., Cary, NC, USA

advantage (HA), to produce a prediction model that was successfully used to identify inefficiencies the betting market for ODI matches. Using 2200 matches played prior to January 2005 an updated version of this model was created and compared to the original.

Prediction variables of experience, quality and form were derived by developing separate measures for both teams and then subtracting the away team values from the home team values. This effectively references the final result in term of the home team. Indicator variables were created to identify matches played at a neutral venue and matches where the two competing teams were clearly from different class structures (established nation versus developing nation).

From Table 2 it can be seen that the results of ODI matches are becoming more predictable, with the updated model explaining 3.5% more of the variation in ODI outcomes (R-square: 23.4% vs. 19.6% $p<0.0001$).

Table 2: Multivariate models for MOV constructed with 1800 & 2200 ODI matches.

Variable	Bailey model (n=1800)			Updated Model (n=2200)		
	Estimate	**P-value**	**Partial R^2**	**Estimate**	**P-value**	**Partial R^2**
Intercept / HA	13.4±1.9	<0.0001		13.9±1.8	<0.0001	
Average Ever	0.6±0.1	<0.0001	17.3%	0.6±0.06	<0.0001	20.7%
Class*	-29.6±6.7	<0.0001	1.2%	-25.1±5.9	<0.0001	1.0%
Experience	0.2±0.1	0.002	0.4%	0.2±0.07	0.0003	0.4%
Ave. last 10	0.1±0.04	0.003	0.4%	0.2±0.04	<0.0001	0.7%
Neutral Venue	-8.6±3.2	0.007	0.3%	-8.2±3.2	0.005	0.3%
Total R^2			**19.6%**			**23.1%**

* When a developing cricket nation played host to an established cricket nation

Because the MOV in the regression model is nominally structured in favour of the home team, the intercept term in the regression equation reflects HA. It can be seen from Table 2 that HA is equivalent to about 14 runs and is highly statistically significant ($p<0.0001$). Because one third of all ODI have been played at neutral venues, a binomial indicator variable was created to negate the HA for these games. As the regression process requires a 'Home' and 'Away' team, when playing at neutral venue, the team with the most experience at the venue was assigned to be the 'Home' team. If all matches played at neutral venues were devoid of HA then the binomial variable for a neutral venue would be the exact negative of the intercept term. This was not the quite the case, with the neutral variable equivalent to about eight runs, suggesting a HA in neutral matches equivalent to about six runs. This six run difference could be thought of as a surrogate marker for the difference in familiarity between the competing teams.

The difference in quality, as measured by the difference in averages between the two teams for all past matches, was by far the strongest predictor, explaining 20.7% of the variation in

the updated model. The best measure of current form was the difference in averages for the past 10 matches, whilst the difference in overall experience (games played) between the home and away team was also statistically significant. Whilst no statistically significant difference could be found in parameter estimates, the difference in class declined (29.6 runs vs. 25.1 runs) as developing nations gain more experience. Similarly, the effect of HA rose slightly (13.4 runs vs. 13.9 runs) with more data, while the effect of a neutral venue was slightly lower (8.6 runs vs. 8.2 runs). Not surprisingly, all variables in the model achieved a higher level of statistical significant when additional data was used.

Prediction model for team totals

From Figure 2 it be clearly seen that the total of the team batting first can be well approximated by a normal distribution (mean=229.7±1.2). When the score of the team batting first was shortened due to rain, (about 13% of matches), the DL method was once again incorporated to determine a projected total.

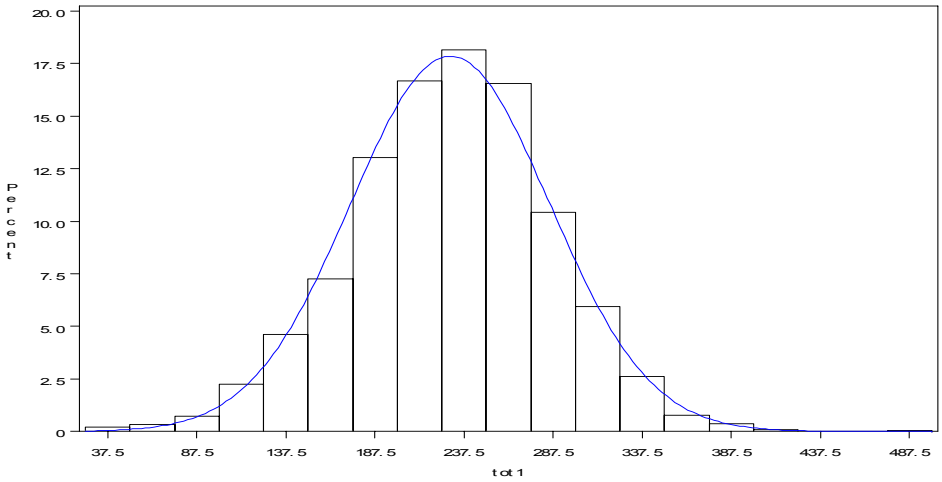


Figure 2. Histogram of first inning scores

Using past averages and exponential smoothing, prediction variables relating to past performance were created. Using a multiple linear regression, a six variable model was constructed.

Table 3: Multivariate model predicting the total of the team batting first

Variable	Estimate	P-value	Partial R^2
Ave. MOV against opposition	0.13±0.04	<0.0001	9.7%
Exp. Smooth past totals 1st inning batting team	0.25±0.04	<0.0001	3.6%
Ave. total conceded in 1st inning by bowling team	0.53±0.06	<0.0001	2.6%
Home Country	15.3±2.3	<0.0001	1.6%
Ave. MOV ever	0.31±0.05	<0.0001	1.1%
Exp. Smooth past totals 1st inning at venue	-8.6±3.2	0.0004	0.5%
Total R^2			**19.1 %**

Interestingly, when using a stepwise selection procedure, the strongest predictor of the total scored by the team batting first was in fact the average of the past MOV between the two teams. The next strongest predictors in the model were derived from the past first innings scores achieved by the batting team as well as scores conceded by the bowling team. HA was the next predictor of importance, with a team playing in it home country scoring an additional 15 runs. A second surrogate marker for the quality of the batting team was given by the average past MOV for the batting team. The final variable that was found to be highly statistically significant ($p=0.0004$) was derived from all past first innings played at the venue. This helped account for pitch conditions and venue size.

Whilst over 23% of the variation in MOV could be explained by the multivariate model, the total of the team batting first was not as predictable, with an R-square statistic of 19.1%.

Using a holdout sample of 100 completed matches played in the year 2005, the regression model successfully predicted the winning team 71% of the time and had and Absolute Average Error (AAE) between the predicted and actual margin of 55.8±4.1 runs. These results compare favourably against the original prediction model of (Bailey, 2005), who accurately identified the winning team 69.6% of the time, and had an AAE of 54.6±0.9 runs for a sample of 336 matches played between 2002 and 2004.

Using the same holdout sample of 100 matches, the AAE for the difference between the predicted and actual totals of the team batting first was 42.5±3.2 runs. By referencing the MOV in terms of the team batting first rather than the home team, a predicted total for the team batting second can be given by

$$\textit{Predicted Total2} = \textit{(Predicted Total1)} + \textit{(Predicted MOVordered)} \quad (2)$$

From the chosen holdout sample of 100 matches, the AAE for the difference between the predicted and actual totals of the team batting second was 47.1±4.0 runs.

RESULTS

With the use of the D-L method to convert available resources into runs, at the completion of each over, an updated total for the team batting first is calculated by combining the actual total with the predicted total for the remainder of the innings.

$$\textit{Updated Total} = \textit{(existing score)} + \textit{(projected total for remaining resources)} \quad (3)$$

Using complete over by over information for the 100 match holdout sample, it can be seen from Figure 3 that the accuracy with which the total of the batting team can be predicted, progressively improves throughout the course of the innings, with first innings totals significantly more accurate that those of the second innings.

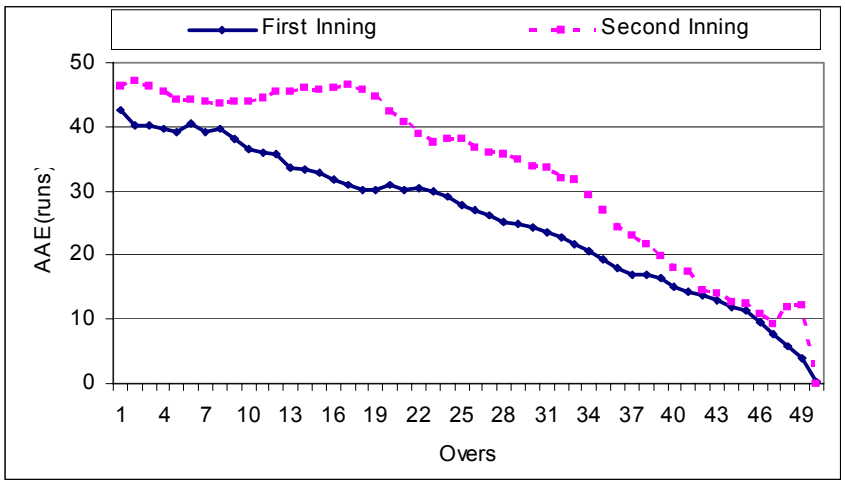


Figure 3: AAE for difference between predicted and actual total

By subtracting the pre-match predicted total from the updated prediction of the total, a performance indicator can be derived for whether each batting team is performing above or below expectation.

Performance indicator = (updated total) – (pre-match predicted total) (4)

With the use the performance indicator, an updated prediction for the MOV can then be readily obtained

Updated MOV = (Pre-match MOV) + (Performance indicators) (5)

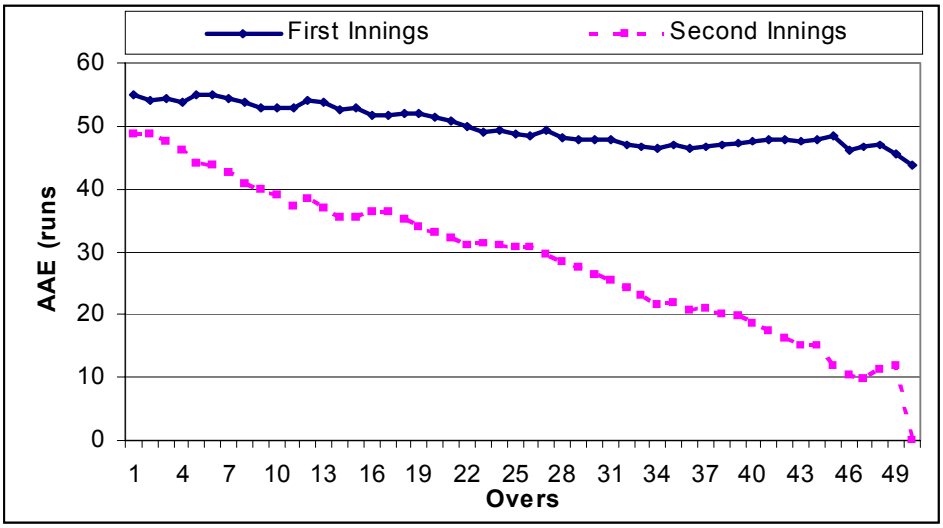


Figure 4: AAE for the difference between the predicted and actual MOV

From Figure 4 it can be seen that during the course of the first innings, the AAE for the difference between the predicted and actual MOV reduces by about 10 runs. In the second innings the reduction in AAE is much greater as the game draws nearer to its conclusion.

As shown by (Bailey, 2005), by dividing the predicted MOV by its standard error and comparing with a standard Normal distribution, the approximate probability that either side will win the match can be readily calculated.

Example

On December 7 2005, Australia played New Zealand in a day/night match at Westpac Stadium in Wellington. After winning the toss and electing to bat Australia proceeded to score a very respectable total of 322. The betting exchange Betfair fielded a betting market for this match, with just over $1,000,000 AUD of matched bets occurring before the start of the game. As betting on this match remains open for the duration of the game, by the completion of the Australian innings, just over $4,000,000 AUD of matched bets had been placed. Figure 5 shows both the volume of bets placed and the price matched. From Figure 5 it can be seen that the opening price for Australia was about $1.38, with the price dropping to $1.30 after Australia won the toss. After losing 3 early wickets, the price drifted out to $1.70, but as Australia rallied, the price continued to drop and by the completion on the 50th over, the best price available for Australia to win was $1.08.



Figure 5: Betfair volume and price for Australia vs. New Zealand ODI 2302 (pre match until end over 50).

Using prediction models for the team total and MOV, the predicted probability for Australia to win was calculated both before and during the match, and compared with the market price offered by Betfair[2]. Where the predicted probability can be seen to exceed the market probability, the 'in play' market can be thought to be inefficient. From Figure 6 it can be seen that while Australia was batting, the predicted probability for Australia to win was consistently below the market probability, with only one inefficiency occurring throughout the course of the Australian innings.

[2] Market probabilities included 5% for commission

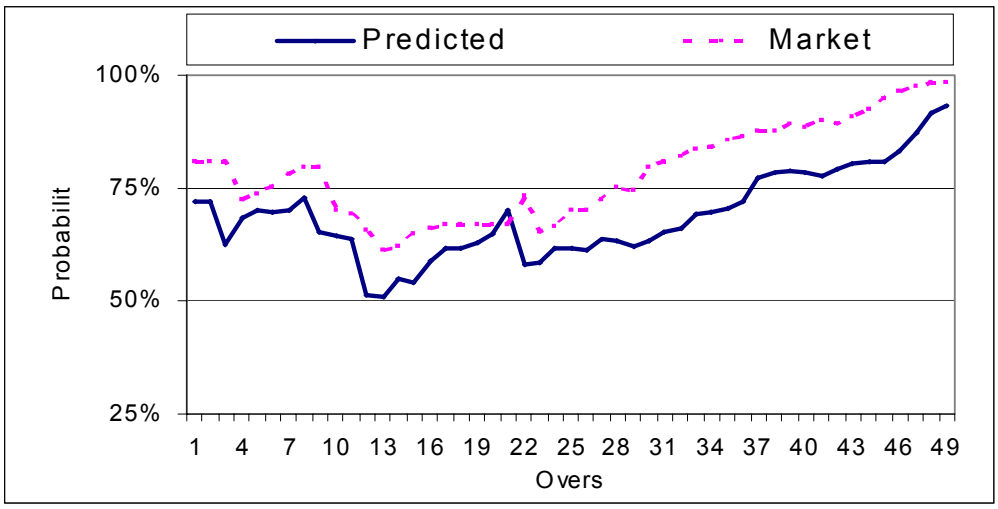


Figure 6: Predicted probability and market price for Australia to win against New Zealand ODI 2302 (pre match until end over 50)

Chasing 323 runs to win the match, New Zealand started slowly. With some big hitting towards the end of the innings, the black caps clawed their way into contention and started the final over as favourites, only requiring six runs to win. Unfortunately, two wickets falling in the final over gave victory to Australia by 2 runs. Figure 7 shows that several inefficiencies were present in the betting market with the predicted probability of success often exceeding the market price. By the completion of the 100th over, more than $9,000,000 AUD had been wagered on the outcome of the match.

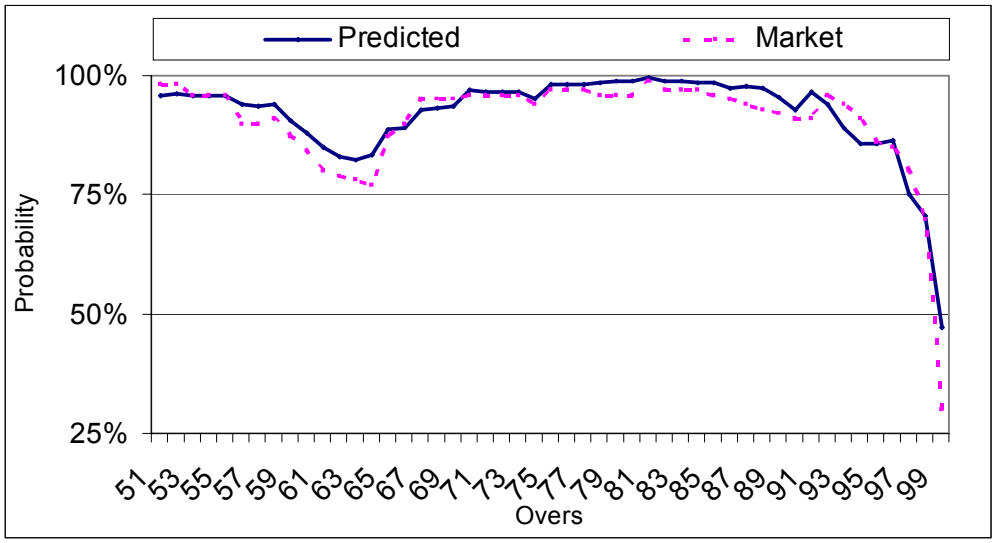


Figure 7: Predicted probability and market price for Australia to win against New Zealand ODI 2302 while New Zealand batted (overs 51-100)

DISCUSSION

In July 2005 the International Cricket Council (ICC) announced a new set of rules to be applicable to ODI matches. An increase in fielding restrictions and the introduction of a substitute player (super-sub), significantly increased the total achieved by the team batting first by more than 20 runs. (252.7±8.0 vs. 229.7±1.2 p=0.002). As these changes occurred within the holdout sample of the data used, it is unsure how these modifications would impact upon the prediction process.

Whilst the price and volume of bets traded are available through Betfair (see Figure 5), this information is not time coded by over, ensuring that if the efficiency of the market is to be accurately determined, information must be gather manually at the completion of each over. This would undoubtedly prove time consuming should a definitive appraisal of the market inefficiency be required.

In Australia, federal laws prevent Australian citizens from placing bets over the internet after a sporting event has commenced. Paradoxically, Australian citizen can place bets 'in the run' provided the bets are placed over the phone. This inconvenience causes a greater delay between observing an inefficient price and actually placing a bet.

CONCLUSION

Multiple linear regression provides a useful way to assign the winning probabilities to the competing teams in ODI matches. With the use of D-L approach, this process can be readily modified to produce 'in the run' predictions. Whilst a definitive analysis of the efficiency of the betting market is yet to be conducted, preliminary evidence suggest punters may be prone to over or under estimate the true probability of the competing teams as the game progresses.

REFERENCES

Bailey, M., 2005, *Predicting Sporting Outcomes: A Statistical Approach*: PhD thesis, Swinburne, Melbourne, 212 p.

de Silva, B., G. Pond, and T. Swartz, 2001, *Estimation Of The Magnitude Of Victory In One-Day Cricket*: Australian & New Zealand Journal of Statistics, v. 43, p. 259-268.

Duckworth, F., and T. Lewis, 1999, *Your Comprehensive Guide To The Duckworth/Lewis Method For Resetting Targets In One-Day Cricket,* University of the West of England.

COUNTBACK METHODS AND AUXILIARY SCORING SYSTEMS IN TENNIS

Pollard, G. and **Noble, K.**

Faculty of Information Sciences and Engineering, University of Canberra, Australia

ABSTRACT

In tennis the tiebreak game is used if the games-score reaches six games all. As its name implies, it is used to 'break-the-tie'. It is not uncommon for the luckier rather than the better player to win the tiebreak game. Indeed, some people think that it seems a little unfair when it happens that the player who has played the better tennis overall up to six games all, loses the tiebreak game (and the set!). In many tournaments using computers with data recorded after each point, it would be a simple matter to use a countback rule based on all of the information in the set up to six games all, in order to 'break-the-tie'. It would also be simple even if only manual records of the score were kept by the umpire. Several countback methods are considered in this paper. They are compared with the tiebreak game for their effectiveness at 'breaking-the-tie', and the best countback method identified. The notion of auxiliary (or parallel) scoring systems is also developed. These auxiliary scoring systems make use of only part of the total information up to six games all, and can be easily used in practice when there is no umpire keeping (at least) a written record of the score. A few auxiliary scoring systems are considered and the most effective is identified.

KEY WORDS

countback methods in tennis, effectiveness of the tiebreak game

INTRODUCTION

In a tiebreak set of tennis singles, the two players alternately play service games until one player wins 6-0, 6-1, 6-2, 6-3, 6-4 or 7-5, or until the score reaches 6-6, when the tiebreak game is played to determine the winner. For the situation in which 6-6 has been reached, as many as (say) 70 points would typically have been played, so there is considerable information available on each player at that stage. This information could be used in an attempt to determine who should win. By comparison, the tiebreak game (when played) has relatively limited information in it as it lasts on average only about 12 points. For this reason, it is not uncommon for the luckier player rather than the better player to win the tiebreak game.

The question arises as to whether the use of a countback method at 6-6 is more (or less) likely to correctly identify the better player. Also, which countback method is ‘best’? These are the purposes of this study. Indeed, if a countback method turned out to be more effective than the tiebreak game at correctly identifying the better player, it would seem a little unusual to play an additional 12 points on average in order to achieve an outcome which was worse (at least probabilistically or ‘in the long run’) than one that could be achieved by playing no points at all.

It would seem that there are two main reasons for playing a tiebreak game. One is to provide entertainment for the spectators (it is seen by many spectators as exciting, although often it is seen as a lottery by the players), and the other is to determine the winner. In a large number of matches, there are no spectators at all. For example, many non-professional matches, local competition matches and social matches have no spectators most of the time. In such cases it would appear that the sole purpose of playing the tiebreak game is to determine the winner.

In this paper we consider 13 individual countback methods, and see how they perform relative to each other. These countback methods are:

C1: The proportion of points won on service by each player (maximum)
C2: The total number of points won in the set by each player (maximum)
C3: The total number of points won as receiver (maximum)
C4: The total number of points played as server (minimum)
C5: The total number of points won as server (minimum)
C6: The total number of break points played as receiver (maximum)
C7: The number of games with break points as receiver (maximum)
C8: The number of times 30-30 or deuce points are played as receiver (maximum)
C9: The number of games with 30-30 or deuce as receiver (maximum)
C10: The number of deuce points played as receiver (maximum)
C11: The number of games with deuce point(s) as receiver (maximum)
C12: The number of games won to love or 15 as server (maximum), and
C13: The number of games won to love, 15 or 30 as server (maximum).

The proportion of points won as *receiver* by each player was also initially considered as a countback method, but the authors quickly realised that this method is equivalent to method C1 above. It is noted that criteria C4 and C5 are measures of how easily a player wins his/her serve, and so a minimum value is selected as the indicator of the better player. Two general methods of combining the best few countback methods are then considered, with a view to producing an improved composite method and seeing whether it is really much better than simpler ones.

In many professional matches (even some with a small number of spectators) the result of every point is entered into a computer, whereas in many matches at the 'local' level there is either an umpire who keeps a manual record of the match, or alternatively the players themselves keep a (verbal) record of the score. These three scenarios of umpiring lead to three different amounts of information being available for countback purposes at 6-6. It is clear that more sophisticated countback methods are possible when 'total' information is available on a computer which can be used to process that information at the press of a button. (And the winner is…!). This is not to say that manual methods and methods based on the players keeping their own records may not work reasonably well.

Records that are not kept at present but are simple enough that they can be kept by the players themselves as the match proceeds have been termed 'auxiliary scoring systems'. These scoring systems can be used if needed for countback purposes at 6-6. Auxiliary

scoring systems may not be as good as computer-based countback methods, but might still be quite reasonable. This is investigated in the study.

METHODS

In most tennis matches the probability of winning a point is greater when serving than it is when receiving. For many singles matches this probability or p-value is close to 0.6, and for many doubles matches it is close to 0.7.

Tennis matches have been categorized into various groups (Pollard and Noble, 2004). For example, 'Group 60' has p-values near 0.6 and includes 'stronger professional women and average professional men on slower to average surfaces'. It would also include many average to higher level 'weekend' matches. 'Group 65' has p-values near 0.65 and includes 'stronger professional men on average to fast surfaces, and men's doubles on slow to average surfaces'. 'Group 70' has p-values near 0.70 and includes 'strong-serving professional men on fast grass surfaces, and men's doubles on average to fast surfaces'. This study focuses on 'Group 60', 'Group 65' and 'Group 70' matches.

Firstly we consider 'Group 60' matches with (pa,pb) = (0.6+0.04,0.6-0.04) = (0.64,0.56) where player A is the better player and has a probability of 0.64 of winning a point on service, and player B has a probability of 0.56 of winning a point on service. Many sets of tennis were simulated with (pa,pb) = (0.64,0.56) until 1,000,000 sets in which 6-6 had been reached, were obtained. (Samples of 1,000,000 typically give statistics that are accurate to 3 significant digits. Further, using a sample size that is a power of 10 leads to easy interpretation of the tables). For each of these sets in which 6-6 had been reached, various countback methods were applied to see whether the set would be allocated to player A or player B. For example, using countback method 2 (see row 2 of Table 1) in which the set is awarded to the player who won the greater number of points, player A was awarded the set on countback 532,907 times, player B was awarded the set on countback 319,124 times, and there was a draw on 147,969 occasions. The results for 12 other countback methods are also given in Table 1. It can be seen that the countback methods with the highest ratios of the number of times the set was awarded to A divided by the number of times it was awarded to player B were the four methods C1, C2, C3 and C12.

It can be seen that C12 has the highest ratio (1.716) of these four methods, but it has the disadvantage of having the highest number of draws (248,544). Also, C1 has the least number of draws (22,379), but it also has the second lowest ratio (1.636). Focusing on just these four criteria C1, C2, C3 and C12, consideration was given as to whether combinations of these criteria might perform better than any one of them by itself, and two broad methods of combining them were considered.

Firstly, one criterion is applied and the set awarded to player A or player B, or if a draw resulted using that criterion, a second criterion is applied and the set awarded to player A or player B, or if a draw resulted using that criterion, a third criterion is applied,…. For example, this sequential application of different countback criteria is denoted by C1>C3>C12>C2 when the sequential order is C1, C3, C12 and finally C2.

Table 1: Countback Results for (pa,pb) = (0.64,0.56) for individual countback criteria (cf. Frequency for the tiebreak game = 627,456)

	A wins	B wins	Draws, D	Ratio(A/B)	Freq(A+D/2)
C1	606,806	370,815	22,379	1.636	617,995.5
C2	532,907	319,124	147,969	1.670	606,891.5
C3	568,193	352,798	79,009	1.611	607,697.5
C4	569,009	383,446	47,545	1.484	592,781.5
C5	509,289	381,543	109,168	1.335	563,873
C6	484,512	357,632	157,856	1.355	563,440
C7	381,728	259,022	359,250	1.474	561,353
C8	535,729	369,672	94,599	1.449	583,028.5
C9	462,408	313,452	224,140	1.475	574,478
C10	512,728	361,326	125,946	1.419	575,701
C11	442,776	302,643	254,581	1.463	570,066.5
C12	474,792	276,664	248,544	1.716	599,064
C13	429,028	277,319	293,653	1.547	575,854.5

Secondly we considered combinations of criterion in the following way. Considering (say) the three criteria Cx, Cy and Cz, a counter, SUM, is initially set to zero, +1 is added to SUM if player A is awarded the match under Cx, -1 is added to SUM if player B is awarded the match under Cx, and 0 is added if Cx indicated a draw; +1 is added to SUM if Cy indicated A, -1 is added if Cy indicated B, and 0 if Cy indicated draw; identically for Cz. Finally, if SUM is greater than zero the set is awarded to player A, and if it is less than zero it is awarded to player B, and if it is zero, we have a draw. This composite criterion is denoted by C(x,y,z).

The results of various combinations of C1, C2, C3 and C12 are given in Table 2. The last column in this table is obtained by adding half the number of draws to the number of 'A wins'. Thus, this column represents the number of times player A would win (on average) if draws were allocated at random to each player. This is a useful column for comparing the various countback methods, and for comparing them with the tiebreak game. It can be seen that when C1 was used in sequence with C2, C3 or C12, it was always better to start the sequence with C1. Note that C12 resolved quite a lot of draws after C1 had been applied in the first instance. Typically, all of the non-sequential combinations C(x,y), C(x,y,z) and C(1,2,3,12) produced too many draws, as might have been expected. There was one exception however. Surprisingly, the combination C(1,2,3) performed better than the criterion C1, although not by much. (A simulation of 10,000,000 scores at 6-6 was carried out to verify that this observation was not just a random or chance occurrence). These observations suggested starting with C(1,2,3) as the first criterion, and considering only sequencing from then on. The results for the best two composite criterion along these lines, C(1,2,3)>C3>C12 and C(1,2,3)>C12>C3 are given in Table 2, and it can be seen that they are virtually identical in performance. Thus, for 'Group 60' players, either of these complex, composite countback systems seems to be about as good as is possible. Thus, in

the situation in which the outcome of every point is entered into a computer, the use of either of these two countback rules would be quite feasible.

Table 2: Countback Results for (pa,pb) = (0.64,0.56) for individual and more complex countback criteria (cf. Frequency for the tiebreak game = 627,456)

	A wins	B wins	Draws, D	Ratio(A/B)	Freq(A+D/2)
C1>C2	608,637	372,611	18,752	1.633	618,013
C2>C1	604,709	376,539	18,752	1.606	614,085
C1>C3	610,019	373,705	16,276	1.632	618,157
C3>C1	603,566	380,158	16,276	1.588	611,704
C1>C12	612,582	376,293	11,125	1.628	618,144.5
C12>C1	608,375	380,500	11,125	1.599	613,937.5
C1	606,806	370,815	22,379	1.636	617,995.5
C(1,2,3)	607,324	370,587	22,089	1.639	618,368.5
C(1,2,3)>C3>C12	614,236	377,172	8,592	1.629	618,532
C(1,2,3)>C12>C3	614,222	377,186	8,592	1.628	618,518
C12>C11	569,201	358,489	72,310	1.588	605,356
C11>C12	553,017	374,673	72,310	1.476	589,172
C12	474,792	276,664	248,544	1.716	599,064
C11	442,776	302,643	254,581	1.463	570,066.5

The above two countback rules are not of any use when a computer is not available. A 'best' sequential combination of C2, C3 and C12 could easily be determined, and could probably be managed by a competent and alert umpire using manual record keeping. It can be seen that C1, C2 and C3 are not suitable criteria for use by the players themselves, whereas C12 and C11, for example, are suitable. Thus, a sequential combination of C12 and C11 (which had a moderate performance as a separate criterion) may be able to be handled by just the players themselves. (Recording/updating the relevant information at each change of ends would assist the players in this task.) The performances of the auxiliary scoring systems C11, C12, C11>C12 and C12>C11 are given in Table 2. Firstly, it is noted that C12 is better than C11 as an auxiliary scoring system, since C12 has the larger ratio and a smaller number of draws. Also, criteria C12>C11 is a better auxiliary scoring system than C11>C12 as it has a larger ratio. More generally, it is noted here that all of the conclusions in the above paragraphs for 'Group 60' matches also applied for this group of matches when the simulations were carried out with the smaller player differences pa-pb=0.04 and pa-pb=0.02.

We now compare these various countback methods and auxiliary scoring systems with the use of the tiebreak game at 6-6. The probability that the tiebreak game correctly identifies the better player when pa=0.64 and pb=0.56 is 0.627456 (an exact result obtained by forward recurrence methods), and so one would expect player A to win on 627,456 occasions out of 1,000,000 sets in which the tiebreak game was played. Thus, it can be seen that, when pa and pb are 'near 0.6', the countback methods are not as good as the tiebreak game at identifying the better player, but they are perhaps somewhat better than one might have expected. The performances of the 'better-performing' countback, composite-countback and auxiliary scoring systems identified above are now investigated for 'Group

70’ matches. Table 3 is a replica of the lower part of Table 2 for the case in which pa=0.74 and pb=0.66. Given that the probability player A wins a tiebreak game is now 0.638188 when pa=0.74 and pb=0.66, it can be seen from Table 3 that, for ‘Group 70’ matches, all of these countback methods except C11 work better than actually playing the tiebreak game! Even some of the auxiliary scoring systems work better at identifying the better player. As above, these conclusions also applied for ‘Group 70’ matches when the simulations were carried out with the smaller player differences given above.

Table 3: Countback Results for (pa,pb) = (0.74,0.66) for individual and more complex countback criteria (cf. Frequency for the tiebreak game = 638,188)

	A wins	B wins	Draws, D	Ratio(A/B)	Freq(A+D/2)
C1	673,872	298,027	28,101	2.261	687,922.5
C(1,2,3)	673,875	298,027	28,098	2.261	687,924
C(1,2,3)>C3>C12	681,404	305,735	12,861	2.229	687,834.5
C(1,2,3)>C12>C3	681,436	305,703	12,861	2.229	687,866.5
C12>C11	630,895	291,444	77,661	2.165	669,725.5
C11>C12	619,658	302,681	77,661	2.047	658,488.5
C12	550,928	234,332	214,740	2.351	658,298
C11	480,805	226,457	292,738	2.123	627,174

Table 4 gives the corresponding results for the ‘Group 65’ matches for the case in which pa=0.69 and pb=0.61. The relevant probability for player A winning a tiebreak game is now 0.631622, so it can be seen that the various countback systems and even the auxiliary scoring system, C12>C11, perform better than the tiebreak game for ‘Group 65’ matches. These results also applied for the smaller differences.

Table 4: Countback Results for (pa,pb) = (0.69,0.61) for individual and more complex countback criteria (cf. Frequency for the tiebreak game = 631,622)

	A wins	B wins	Draws, D	Ratio(A/B)	Freq(A+D/2)
C1	641,214	335,745	23,041	1.91	652,734.5
C(1,2,3)	641,310	335,744	22,946	1.91	652,783
C(1,2,3)>C3>C12	647,711	342,230	10,059	1.893	652,740.5
C(1,2,3)>C12>C3	647,734	342,207	10,059	1.893	652,763.5
C12>C11	601,648	325,847	72,505	1.846	637,900.5
C11>C12	588,639	338,856	72,505	1.737	624,891.5
C12	514,362	257,982	227,656	1.994	628,190
C11	467,736	265,492	266,772	1.762	601,122

The point at which the best composite countback method is about equal to the tiebreak method is when pa and pb averaged about 0.616. Below this value the tiebreak game performs increasingly well relative to the various countback methods, until it performs at its best relatively when we reach ‘Group 50’ matches. As an example of this best relative scenario, when pa=0.54 and pb=0.46, the probability player A wins the tiebreak game is 0.624390 whilst the total frequency (corresponding to those in Tables 2, 3 and 4) is only

574,954 for C(1,2,3)>C12>C3 and 539,772 for the auxiliary scoring system C12>C11. It is noted here that we explored the notion of extending the definition of a draw for the basic criteria (for example, extending the draw definition for C2 as a point difference between the players of -1, 0, +1 rather than a difference of 0) with the view to increasing the ratio for each criterion. The ratios (for these modified criterion C2, C3,…) certainly did increase, but so did, of course, the number of draws. There appeared to be no particular benefit in extending draws prior to combining or sequencing the (draw-modified) criteria into composite ones. This technique of extending the draws in order to achieve higher ratios could be revisited in other ways (maybe successfully) but this has not yet been investigated.

RESULTS

By simulating 1,000,000 sets that reached 6-6, a range of different countback procedures have been analysed to see which procedures are the best. For matches in which pa and pb average 0.6 or more, several countback methods work quite well. Indeed, provided this average is about 0.62 or more, the best computer-based countback method is more likely to correctly identify the better player than is a tiebreak game. Countback methods do not work so well when the average of pa and pb falls below 0.6. For example, for the worst case when pa and pb average only 0.5, the tiebreak game was shown to be quite a bit more effective than countback methods. Some simple quite easy-to-manage countback methods, called auxiliary scoring systems are only marginally worse than the best countback systems for the situations in which one would consider using a countback approach.

DISCUSSION

The approach taken in this paper is to consider using a countback method for all matches, or to not use it at all and to use the tiebreak game as at present. An alternative approach to this could be taken. This approach involves using a countback method some of the time (namely when one player has a sufficient lead on the countback criterion), and to play the tiebreak game the remainder of the time. This approach would involve awarding the set to the player with a ‘large enough’ lead on the countback criterion, and to play the tiebreak game if the lead is not ‘large enough’. The amount required for a ‘large enough’ lead could be determined so that the relevant ratio is no less than that of the tiebreak game. Thus, for any tennis match, we would have a dynamic structure in which the probability of the better player winning would be at least as large as what it is for the tiebreak game. It is noted that this approach could be used for any countback method. Another way of looking at this is that the tiebreak game should always be played unless there is ‘large enough’ evidence to award the match to one of the players. This approach would appear to have considerable merit, and is presently being studied by the authors.

There would appear to be several practical advantages in using a countback procedure:

- For the situation in which two teams play simultaneously on two courts (and then swap around in some way), it is clear that the two courts are more likely to be better synchronized when tiebreak games are replaced by countbacks or draws.

- Matches in general will clearly have more predictable durations (smaller variance of duration) if tiebreak games are replaced by countback or draw (Pollard & Noble, 2003).
- It is noted that, under the present scoring system, players sometimes 'throw away' a point or two in certain circumstances. For example, players who are down 40-0 against strong servers have been observed to do this. This practice is not attractive for the spectators, and it is clear that in a countback situation it is likely to cease.

CONCLUSIONS

Countback methods at 6-6 in a set of tennis can be more effective at correctly identifying the better player than is the playing of a tiebreak game. This is particularly so when serving is a strong advantage in a match, and this occurs in many doubles and in many men's singles situations. Some countback methods work better than others, and not surprisingly the more complex ones work a little better. These complex methods require the recording of every point played up to 6-6, and sometimes a computer to work out the actual winner. Simpler countback methods, called auxiliary scoring systems, although not as good as the more complex ones, nevertheless perform quite well in the situations where they might be considered. It would appear that spectators like the excitement of a tiebreak game, even though the players themselves realize that the tiebreak game is a bit of a lottery, and that the better player quite often loses. Thus, it would seem that the most relevant occasions on which a countback might be used are when there are no spectators. This is probably most of the time for tennis played at the local level. The simplest auxiliary system is 'the number of service games won to 15 or better', possibly used in conjunction with 'the number of games in which a player reaches deuce on their opponent's serve'. These two auxiliary systems are actually reasonably effective compared with the most complex ones. The best complex countback system found makes use of four measures: the proportion of points won on service by each player, the total number of points won by each player, the total number of points won by each player when receiving, and the number of games won to 15 or better by each player on their service. This complex system is described in Section 2.

Acknowledgement
The authors would like to thank Mark Fowler for initial discussions that led to this research.

REFERENCES

Pollard, G. and Noble, K. (2003) Scoring to remove long matches, increase tournament fairness and reduce injuries. *Medicine and Science in Tennis* **8(3)**, 12-13.

Pollard, G. and Noble, K. (2004) The benefits of a new game scoring system, the 50-40 game. In R. Morton and S. Ganesalingam (Eds). *Proceedings of the Seventh Australasian Conference on Mathematics and Computers in Sport.* Massey University, New Zealand, 262-265.

DEVELOPING A MODEL THAT REFLECTS OUTCOMES OF TENNIS MATCHES

Barnett T., **Brown A.,** and Clarke S.
Faculty of Life and Social Sciences, Swinburne University, Melbourne, VIC, Australia

ABSTRACT

Many tennis models that occur in the literature assume the probability of winning a point on service is constant. We show this assumption is invalid by forecasting outcomes of tennis matches played at the 2003 Australian Open. A revised model is formulated to better reflect the data. The revised model improves predictions overall, particularly for the length of matches and can be used for index betting. Suggestions on further improvements to the predictions are discussed.

KEY WORDS

tennis, sport, Markov chain, index betting

INTRODUCTION

Many tennis models that occur in the literature assume the probability of winning a point on service is constant (Barnett and Clarke 2002, Carter and Crews 1974, Fischer 1980, Schutz 1970). On the other hand, there are works in the literature to show that the assumption of players winning points on serve being i.i.d. does not hold. Jackson (1993) and Jackson and Mosurski (1997) show that psychological momentum does exist in tennis, and set up a "success-breeds-success" model for sets in a match, and find that this model provides a much better fit to the data, compared to an independence of sets model. Klaassen and Magnus (2001) test whether points in tennis are i.i.d. They show that winning the previous point has a positive effect on winning the current point, and at important points it is more difficult for the server to win the point than at less important points.

In this paper, an i.i.d. Markov Chain model is used to predict outcomes of tennis matches. The predictions indicate that the i.i.d. assumption may not hold since there are fewer games and sets actually played than predicted. A revised Markov chain model is then formulated for sets in a match that allows for players that are ahead on sets, to increase their probability of winning the set, compared to their probabilities of winning the first set. This is then followed by a revised model for points in a match that has an additive effect on the probability of the server winning a point. The revised models better reflect the data and the latter model is most useful for predicting lengths of matches, as demonstrated through index betting.

MARKOV CHAIN MODEL

Modelling a game

A Markov chain model of a game for two players, A and B, is set up where the state of the game is the current point score (a, b), where both $a \geq 0$ and $b \geq 0$. With a constant probability p the state changes from (a, b) to (a + 1, b) and with probability 1 - p it changes from (a, b) to (a, b + 1). Therefore the probability P(a, b) that player A wins the game when the point score is (a, b), is given by:

$P(a, b) = p\, P(a + 1, b) + (1 - p)\, P(a, b + 1)$

where p is the probability of player A winning a point.

The boundary values are $P(a, b) = 1$ if $a = 4$, $b \leq 2$,

$P(a, b) = 0$ if $b = 4$, $a \leq 2$, $\qquad P(3, 3) = \frac{p^2}{p^2 + (1 - p)^2}$.

Similarly, the mean number of points M(a, b) remaining in the game at point score (a, b) is given by

$M(a, b) = 1 + p\, M(a + 1, b) + (1 - p)\, M(a, b + 1)$

The boundary values are $M(a, b) = 0$ if $b = 4$, $a \leq 2$ or $a = 4$ $b \leq 2$,

$M(3, 3) = \frac{2}{p^2 + (1 - p)^2}$.

Let N(a, b|g, h) be the probability of reaching a point score (a, b) in a game from point score (g, h) for player A. The forward recurrence formulas are

$N(a, b|g, h) = N(a - 1, b|g, h)$, for $a = 4$, $0 \leq b \leq 2$ or $b = 0$, $0 \leq a \leq 4$,
$N(a, b|g, h) = (1 - p)\, N(a, b - 1|g, h)$, for $b = 4$, $0 \leq a \leq 2$ or $a = 0$, $0 \leq b \leq 4$,
$N(a, b|g, h) = p\, N(a-1 , b|g, h) + (1 - p)\, N(a, b - 1|g, h)$ for $1 \leq a \leq 3$, $1 \leq b \leq 3$.

The boundary values are $N(a, b|g, h) = 1$ if $a = g$ and $b = h$.

Modelling a set

Let $P_A^{gs_T}(c, d)$ $\{P_A^{gs}(c, d)\}$ and $P_B^{gs_T}(c, d)$ $\{P_B^{gs}(c, d)\}$ represent the conditional probabilities of player A winning a tiebreaker {advantage} set from game score (c, d) for player A or player B serving respectively. The formulas below are for player A serving. Similar formulas apply for when player B is serving.

For a tiebreaker set

$P_A^{gs_T}(c, d) = p_A^g\ P_B^{gs_T}(c + 1, d) + (1 - p_A^g)\ P_B^{gs_T}(c, d + 1)$

The boundary values are $P_A^{gs_T}(c, d) = 1$ if $c = 6, 0 \leq d \leq 4$ or $c = 7, d = 5$,

$P_A^{gs_T}(c, d) = 1$ if $d = 6, 0 \leq c \leq 4$ or $c = 5, d = 7, P_A^{gs_T}(6, 6) = p_A^{g_T}$

where p_A^g and p_B^g represents the probability of player A and B winning a game on serve respectively and $p_A^{g_T}$ represents the probability of player A winning a tiebreaker game.

For an advantage set $P_A^{gs}(c, d) = p_A^g \; P_B^{gs}(c + 1, d) + (1 - p_A^g) \; P_B^{gs}(c, d + 1)$

The boundary values are $P_A^{gs}(c, d) = 1$ if $c = 6, 0 \leq d \leq 4$,

$P_A^{gs}(c, d) = 0$ if $d = 6, 0 \leq c \leq 4$, $\quad P_A^{gs}(5, 5) = \dfrac{p_A^g (1 - p_B^g)}{p_A^g (1 - p_B^g) + (1 - p_A^g) p_B^g}$

Recurrence formulas can be obtained for the mean number of games remaining in sets and the probability of reaching score lines in sets.

Table 1: Percentage of matches correctly predicted at the 2003 Australian Open

Round	Percentage correct(%)	No. of matches
1	78.1	64
2	62.5	32
3	68.8	16
4	75.0	8
5	75.0	4
6	50.0	2
7	100.0	1
Total	72.4	127

Modelling a match

Let $P^{sm}(e, f)$ represent the conditional probabilities of player A winning an advantage match from set score (e, f).

The recurrence formula is represented by

$P^{sm}(e, f) = p^{s_T} \; P^{sm}(e + 1, f) + (1 - p^{s_T}) \; P^{sm}(e, f + 1)$

The boundary values are

$P^{sm}(e, f) = 1$ if $e = 3, f \leq 2$, $P^{sm}(e, f) = 0$ if $f = 3, e \leq 2$, $P^{sm}(2, 2) = p^s$

where p^{s_T} represents the probability of player A winning a tiebreaker set, and p^s represents the probability of player A winning an advantage set.

Recurrence formulas can be obtained for the mean number of sets remaining in a match and the probability of reaching score lines in matches.

MATCH PREDICTIONS

2003 Australian Open Men's predictions

When two players, A and B, meet in a tournament, forecasting methods (Barnett and Clarke 2005) are used to obtain estimates for the probability of each player winning a point on serve. These two parameters (each player winning a point on serve) are then used as input probabilities in the Markov Chain model to obtain match outcomes. We will compare the accuracy of the predictions to the actual outcomes for the 2003 men's Australian Open.

For a match between two players, the player who has greater than a 50% chance of winning was the predicted winner. Table 1 represents the percentage of matches correctly predicted for each round and shows that overall 72.4% of the matches were correctly predicted. Based on the ATP tour rankings only 68.0% were correctly predicted.

If p_i represents the probability for the predicted player for the i^{th} match, then the proportion of matches (P) correctly predicted and the variance (V) of the proportion can be calculated by

$$P = \frac{\sum_i p_i}{n}$$

$$V = \frac{\sum_i p_i \, q_i}{n^2}$$

where $q_i = 1 - p_i$ and n = total number of matches played in the tournament.

Applying these equations gives values of P = 0.753 and V = 0.0013. The 95% confidence interval is represented by
$(0.753 - 1.96\sqrt{0.0013}, 0.753 + 1.96\sqrt{0.0013}) = (0.682, 0.824)$
which includes the value of 0.724.

Out of 127 scheduled matches for the 2003 Australian Open men's singles, only 118 were completed. For the other 9 matches, players had to withdraw prior to the match or retire injured during the match. Therefore, only the 118 completed matches were used for predicting the number of games and sets played. Table 2 gives the results. Overall, there were 487.7 fewer games played than predicted. This equates to 487.7/118 = 4.13 fewer games per match. Also, there were more 3 set matches played than predicted and fewer 5 set matches. This gives some indication that the i.i.d. model may need to be revised.

Table 2: Predicted and actual number of games and sets played
at the 2003 Australian Open men's singles

	Games played	3 sets	4 sets	5 sets
Prediction	4737.7	41.1	42.7	34.2
Actual	4250.0	50.0	42.0	26.0

Using the model for gambling

For head-to-head betting we will place a bet only when there is a positive overlay as represented by:

$$Overlay = [Our\ Probability \times Bookmakers\ Price] - 1$$

A method developed by Kelly, discussed in Haigh (1999), calculates the proportion of bankroll you should bet depending on your probability and the bookmaker's price and is represented below:

$$Proportion\ of\ bankroll\ to\ gamble = \frac{Overlay}{Bookmaker's\ Price - 1}$$

For example: Suppose player A was paying \$2.20 to win, and player B was paying \$1.65 to win. Suppose we predicted player B to win with probability 0.743. In this situation we would bet on B as given by a positive overlay [0.743 * 1.65] - 1 = 0.226

$$Proportion\ of\ bankroll\ to\ gamble = \frac{0.226}{1.65 - 1} = 0.348$$

Figure 1 represents how we would have performed by adopting a constant Kelly system (fixed bankroll) of \$100 for the head-to-head matches played at the 2003 Australian Open. It can be observed that we would have suffered a \$195 loss by our 72nd bet but still ended up with a \$45 profit. This recovery came from round 3 (bet number 75) onwards, where at that point we were down \$147. By updating the parameters after each round by simple exponential smoothing some important factors such as court surface, playing at a particular tournament, playing in a grand slam event and recent form would be included in the predictions.

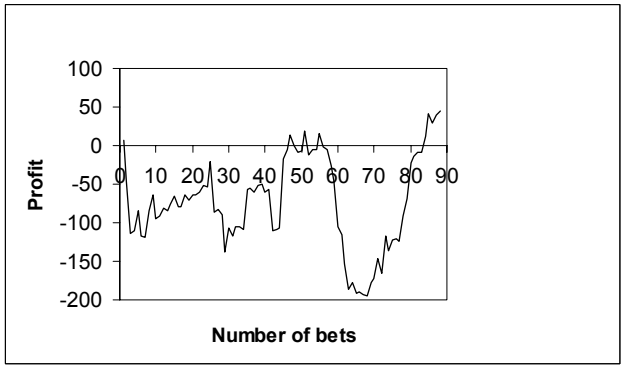


Figure 1: Profit obtained from betting on head-to-head matches played at the 2003 Australian Open

Jackson (1994) outlines the operation of index betting with some examples in tennis through binomial type models. The outcome of interest X is a random variable and for our situation is the number of games played in a tennis match. The betting firm offers an interval (a, b), known as the spread. The punter may choose to buy X at unit stake y, in which case he receives y $(X - b)$ if $X > b$ or sell X at unit stake z, in which case receives z $(a - X)$ if $X < a$.

We will place a bet only when our predicted number of games is greater than b or less

than *a*. For example if an index is (35, 37), we would sell if our prediction is less than 35 games or buy if our prediction is greater than 37 games. We will use a very simple betting system, and that is to trade 10 units each time the outcome is favourable. Figure 2 represents how we would have gone by using our allocated betting strategy, for a profit of $435. This was as high as $480 but as low as -$220. We also made $420 from one match alone being the El Aynaoui versus Roddick match where a total of 83 games were played. Without including this match we would have still made a profit of $60.

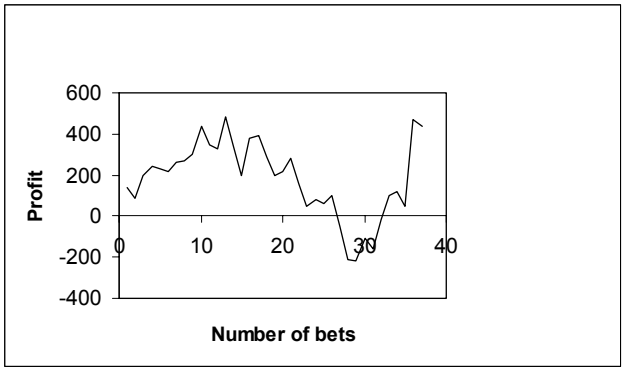


Figure 2: Profit obtained from index betting on matches played at the 2003 Australian Open

Unlike head-to-head betting, there does not appear to be any advantage by betting from later rounds. We can generate a profit from the start of the tournament. Perhaps the bookmakers are not as proficient in estimating the number of games played in a match as they are with the probabilities of winning the match. The bookmakers are always trying to balance their books where possible so that they gain a proportion of the amount gambled each match regardless of the outcome. This implies that the general public are unable to predict the number of games played in a match as well as probabilities of players winning. Figure 3 represents the results by subtracting an additional 4.13 games per match from our predictions. This gave a profit of $285, despite the fact that no money was bet on the El Aynaoui versus Roddick match, which made a $420 profit previously.

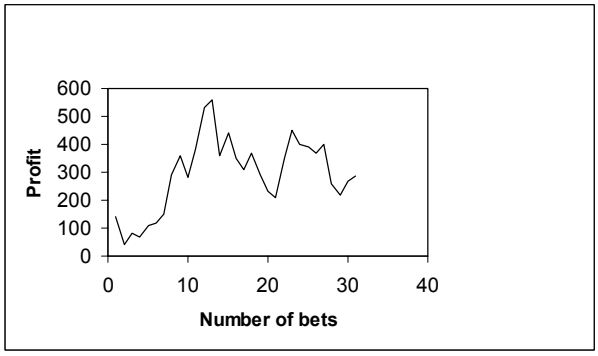


Figure 3: Profit obtained from index betting on matches played at the 2003 Australian Open by subtracting 4.13 games per match from our predictions

REVISED MARKOV CHAIN MODEL

Probabilities of reaching score lines within an advantage match

Suppose that a player is ahead on sets, then they can increase their probability of winning a set by α . The forward recursion formulas become

$N^{sm}(e, f|k, l) = p^{s_T} N^{sm}(e - 1, f|k, l)$, for $(e, f) = (1, 0)$

$N^{sm}(e, f|k, l) = (1 - p^{s_T}) N^{sm}(e, f - 1|k, l)$, for $(e, f) = (0, 1)$

$N^{sm}(e, f|k, l) = p^{s} N^{sm}(e - 1, f|k, l)$, for $(e, f) = (3, 2)$

$N^{sm}(e, f|k, l) = (1 - p^{s}) N^{sm}(e, f - 1|k, l)$, for $(e, f) = (2, 3)$

$N^{sm}(e, f|k, l) = (p^{s_T} + \alpha) N^{sm}(e - 1, f|k, l)$, for $(e, f) = (3, 0), (2, 0)$ and $(3, 1)$

$N^{sm}(e, f|k, l) = (1 - p^{s_T} + \alpha) N^{sm}(e, f - 1|k, l)$, for $(e, f) = (0, 3), (0, 2)$ and $(1, 3)$

$N^{sm}(e, f|k, l) = (p^{s_T} - \alpha) N^{sm}(e - 1, f|k, l) + (1 - p^{s_T}) N^{sm}(e, f - 1|k, l)$, for $(e, f) = (1, 2)$

$N^{sm}(e, f|k, l) = p^{s_T} N^{sm}(e - 1, f|k, l) + (1 - p^{s_T} - \alpha) N^{sm}(e, f - 1|k, l)$, for $(e, f) = (2, 1)$

$N^{sm}(e, f|k, l) = (p^{s_T} - \alpha) N^{sm}(e-1, f|k, l) + (1 - p^{s_T} - \alpha) N^{sm}(e, f - 1|k, l)$, for $(e, f) = (1, 1)$ and $(2, 2)$

where $0 \le p^{s_T} + \alpha \le 1$ and $0 \le p^{s} + \alpha \le 1$.

The boundary values are $N^{sm}(e, f|k, l) = 1$ if $e = k$ and $f = l$.

Table 3: Distribution of the number of sets in an advantage match when $\alpha = 0$ and 0.06

			$\alpha = 0$			$\alpha = 0.06$		
p_A	p_B	p^{s_T}	3 sets	4 sets	5 sets	3 sets	4 sets	5 sets
0.60	0.60	0.50	0.25	0.38	0.38	0.31	0.38	0.30
0.61	0.60	0.53	0.25	0.37	0.37	0.32	0.38	0.30
0.62	0.60	0.57	0.26	0.37	0.36	0.33	0.38	0.29
0.63	0.60	0.60	0.28	0.37	0.35	0.35	0.38	0.28
0.64	0.60	0.63	0.30	0.37	0.32	0.37	0.37	0.26
0.65	0.60	0.66	0.33	0.37	0.30	0.40	0.37	0.23
0.66	0.60	0.69	0.36	0.37	0.27	0.43	0.36	0.21
0.67	0.60	0.72	0.40	0.36	0.24	0.47	0.34	0.19
0.68	0.60	0.75	0.43	0.35	0.21	0.51	0.33	0.16
0.69	0.60	0.77	0.47	0.34	0.19	0.55	0.31	0.14
0.70	0.60	0.79	0.51	0.33	0.16	0.60	0.29	0.11

Table 3 represents the probabilities of playing 3, 4 and 5 set matches when $\alpha = 0$ and 0.06, for different values of p_A and p_B. The probability of playing 3 sets is greater when $\alpha = 0.06$ compared to $\alpha = 0$, for all p_A and p_B. The probability of playing 5 sets is greater when $\alpha = 0$ compared to $\alpha = 0.06$, for all p_A and p_B.

From our forecasting predictions, it was noticed that on average the proportion of 3 set matches played are about 7% more than the model predicted and the proportion of 5 set matches are about 7% less than the model predicted, based on the assumption that the probability of players winning a point on serve are i.i.d. Notice from Table 3, the probability of playing 4 sets is about the same for both values of $\alpha = 0$ and 0.06, and the

differences in probabilities for playing 3 sets is about 0.07 greater when α = 0.06 compared to α = 0, if $p^{s_T} \leq 0.75$. This is the reason α = 0.06 was chosen for the revised model.

Conditional probabilities of winning a match
The recurrence formulas are represented by:

$$P^{sm}(e, f) = p^{s_T}\ P^{sm}(e+1, f) + (1 - p^{s_T})\ P^{sm}(e, f+1), \text{ for } e = f$$
$$P^{sm}(e, f) = (p^{s_T} + \alpha)\ P^{sm}(e+1, f) + (1 - p^{s_T} - \alpha)\ P^{sm}(e, f+1), \text{ for } e > f$$
$$P^{sm}(e, f) = (p^{s_T} - \alpha)\ P^{sm}(e+1, f) + (1 - p^{s_T} + \alpha)\ P^{sm}(e, f+1), \text{ for } e < f$$

The boundary values are $P^{sm}(e, f) = 1$ if $e = 3, f \leq 2$, $P^{sm}(e, f) = 0$ if $f = 3, e \leq 2$, $P^{sm}(2, 2) = p^s$.

When $\alpha = 0$, the formulas reflect the Markov chain model presented earlier.

Table 4 represents the probabilities of player A winning an advantage match for α = 0 and 0.06, for different values of p_A and p_B. It can be observed that the probabilities remain essentially unaffected for all values of p_A and p_B by comparing the probabilities of winning the match when $\alpha = 0$ to $\alpha = 0.06$.

Table 4: Probabilities of player A winning an advantage match when $\alpha = 0$ and 0.06

p_A	p_B	p^{s_T}	p^s	$p^m : \alpha = 0$	$p^m : \alpha = 0.06$
0.60	0.60	0.50	0.50	0.500	0.500
0.61	0.60	0.53	0.54	0.565	0.564
0.62	0.60	0.57	0.57	0.627	0.627
0.63	0.60	0.60	0.61	0.686	0.685
0.64	0.60	0.63	0.64	0.740	0.739
0.65	0.60	0.66	0.67	0.789	0.787
0.66	0.60	0.69	0.71	0.831	0.829
0.67	0.60	0.72	0.74	0.867	0.865
0.68	0.60	0.75	0.76	0.897	0.895
0.69	0.60	0.77	0.79	0.921	0.920
0.70	0.60	0.79	0.81	0.941	0.939

Mean number of sets remaining in a match
The recurrence formulas are represented by

$$M^{sm}(e, f) = 1 + p^{s_T}\ M^{sm}(e+1, f) + (1 - p^{s_T})\ M^{sm}(e, f+1), \text{ for } e = f$$
$$M^{sm}(e, f) = 1 + (p^{s_T} + \alpha)\ M^{sm}(e+1, f) + (1 - p^{s_T} - \alpha)\ M^{sm}(e, f+1), \text{ for } e > f$$
$$M^{sm}(e, f) = 1 + (p^{s_T} - \alpha)\ M^{sm}(e+1, f) + (1 - p^{s_T} + \alpha)\ M^{sm}(e, f+1), \text{ for } e < f$$

The boundary values are $M^{sm}(e, f) = 0$ if $e = 3, f \leq 2$ or $f = 3, e \leq 2$, $M^{sm}(2, 2) = 1$.

Table 5 represents the mean number of sets played in an advantage match M^{sm} for $\alpha = 0$ and 0.06, for different values of p_A and p_B. The mean number of sets played when $\alpha = 0.06$ is less than that when $\alpha = 0$ for all p_A and p_B.

Table 5: Mean number of sets played in an advantage match when $\alpha = 0$ and 0.06

p_A	p_B	p^{sT}	$M^{sm}:\alpha = 0$	$M^{sm}:\alpha = 0.06$
0.60	0.60	0.50	4.13	3.99
0.61	0.60	0.53	4.12	3.98
0.62	0.60	0.57	4.10	3.96
0.63	0.60	0.60	4.06	3.93
0.64	0.60	0.63	4.02	3.89
0.65	0.60	0.66	3.97	3.83
0.66	0.60	0.69	3.91	3.78
0.67	0.60	0.72	3.85	3.71
0.68	0.60	0.75	3.78	3.65
0.69	0.60	0.77	3.71	3.58
0.69	0.60	0.79	3.65	3.52

Mean number of games in a tiebreaker match

Note at the outset that a tiebreaker match was chosen to simplify the analysis. The number of games in a tiebreaker set is 13 at most, and the number of sets is 5 at most. However we are faced with the problem of relating the momentum for winning a set to the momentum for winning a game, or even winning a point.

In simple approach to dealing with this problem is to assume the momentum factor m has a linear form:

$$m = c_p k_p + c_g k_g + c_s k_s$$

where c_p, c_g, c_s are coefficients for terms of points, games and sets respectively, and k_p, k_g, k_s are the leads on the scoreboard in terms of points, games and sets respectively.

We can make further simplifying assumptions for the coefficients c_p, c_g, c_s, such that $c_s = 6\ c_g$ and $c_g = 4\ c_p$. The factors 6 and 4 come from the stopping rules for sets and games respectively. Putting $c_p = c > 0$ we then have

$$m = c\ (k_p + 4\ k_g + 24\ k_s)$$

where $-3 \le k_p \le 3$ in a standard game, $-6 \le k_p \le 6$ in a tiebreaker game, $-5 \le k_g \le 5$ in a set, and $-2 \le k_s \le 2$ in a 5-set match. Combining all these inequalities we find that: $-71c \le m \le 71c$.

Now if we model momentum as an additive effect on the probability of the server winning a point we require $p' = p + m$, where $0 < p + m < 1$. This puts theoretical limits on the values of c that can be assumed. It is possible to avoid such difficulties by following Jackson and Mosurski (1997) and express the momentum effect of leading on the scoreboard in terms of log odds

$$\ln(\frac{p'}{q'}) = m \ln(\frac{p}{q})$$

This can be re-expressed as p' = $\frac{p}{p + q\, e^{-m}}$. It is easy to show that p < p' < 1 if the player is ahead, and 0 < p' < p if the player is behind. Whilst this theory appears nicer it is of no practical consequence, as the transformation is very close to linear in the main area of interest.

The linear model for the momentum factor enables us to calculate the probability of a player winning a point on serve in a tiebreaker game in a manner consistent with the probability of a point in a standard game. Furthermore wider consistency can be maintained with the probability of winning a game or a set. The model calculations can be carried out with each player having his own base probability of winning a point on serve, as well as his own individual momentum factor to allow for temperament or other personality factors. Table 6 represents the number of games played per match and the percentage of matches finishing in 3, 4 or 5 sets, for different values of c. The results indicate, that under this momentum model, values of $0.0004 \leq c \leq 0.0007$ give a better fit to the data, when compared to an independence model (c = 0). Nevertheless the linear model fails to fit the data on game difference and the pattern of sets simultaneously.

Table 6: The number of games played per match and the percentage of matches finishing in different score lines

	c	games	3 set	4 set	5 set
observed		36.0	42.4%	35.6%	22.0%
predictions	0.0000	40.1	34.8%	36.2%	29.0%
	0.0004	37.9	43.0%	34.5%	22.5%
	0.0007	36.2	50.5%	31.8%	17.7%

CONCLUSION

Using an Markov chain model to predict outcomes of tennis matches, where the probability of each player winning a point on serve is i.i.d. overestimates the number of games and sets played in a match. Revised Markov chain models are then formulated to give a better fit to the data. Magnus and Klaassen (2001) have tested for independence of points from 4 years of Wimbledon point-by-point data. Further research for testing for independence, could involve analyzing Australian Open point-by-point data. This could involve finding a suitable momentum factor for individual players. In the current paper we set up a revised model with the focus of estimating the number of games played in a match. This model could also be used for estimating the number of points played in a match, by choosing suitable values for the coefficients.

REFERENCES

Barnett T. and Clarke S.R., (2002), Using Microsoft Excel to model a tennis match, *In Proceedings of the 6M&CS, G. Cohen and T. Langtry eds.*, 63-68.

Barnett T. and Clarke S.R., (2005), Combining player statistics to predict outcomes of tennis matches, *IMA Journal of Management Mathematics*, **16**(2), 113-120.

Carter W.H. and Crews S.L., (1974), An Analysis of the Game of Tennis, *The American Statistician*, **28**(4), 130-134.

Fischer G., (1980), Exercise in probability and statistics, or the probability of winning at tennis, *Am. J. Phys.*, **48**(1), 14-19.

Haigh J., (1999), *Taking Chances: winning with probability*, New York: Oxford University Press.

Jackson D.A., (1993), Independent trials are a model for disaster, *Appl. Statist.*, **42**(1), 211-220.

Jackson D.A., (1994), Index betting on sports, *The Statistician*, **43**(2), 309-315.

Jackson D.A. and Mosurski K., (1997), Heavy Defeats in Tennis: Psychological Momentum or Random Effect?, *Chance*, **10**(2), 27-33.

Klaassen F.J.G.M. and Magnus J.R., (2001), Are points in tennis independent and identically distributed? Evidence from a dynamic binary panel data model, *Journal of the American Statistical Association*, **96,** 500-509.

Schutz R., (1970), A mathematical model for evaluating scoring systems with specific reference to tennis, *The Res. Quart.*, **4**1(4), 552-561.

FAIRER SERVICE EXCHANGE MECHANISMS FOR TENNIS WHEN SOME PSYCHOLOGICAL FACTORS EXIST

Pollard, G.[1] and **Barnett, T.**[2]

[1]School of Information Sciences and Engineering, University of Canberra, Australia

[2]Faculty of Life and Social Sciences, Swinburne University of Technology, Australia

ABSTRACT

In a tennis match it is not uncommon for games to 'go with service' (ie 1-0, 1-1, 2-1, 2-2, 3-2,...). When this occurs, the player who serves first is either ahead by one game, or the games' score is equal. Some commentators, players,...argue that the person who serves first has a psychological advantage in that his/her opponent is very often 'playing catch-up'. Assuming that such a (non-zero) psychological advantage of 'being ahead in the games' score' exists, the advantage of serving first in a set between two equal players, is determined. In the presence of such 'front-runner' psychological effects, alternative methods or rules for allocating service to the players are considered, and some are shown to be fairer than the present rule. A proposal consisting of two modifications to the present rules is put forward for consideration. One of these modifications is very easy to apply.

The reverse psychological effect to the above, the 'back-to-the-wall' effect, occurs when a player performs better when behind. The proposal is seen to be fairer than the present method for the cases in which both player A has either a positive or negative psychological effect and player B also has an equivalent positive or negative effect. Further, the application of the proposal to doubles is also considered and a modification for doubles suggested for consideration.

KEY WORDS

rules in tennis, psychological advantage, back-to–the-wall effect in tennis, cricket fairness

INTRODUCTION

Many people believe that the person who serves first in a set of tennis has an advantage. This is because games often 'go with service', so that the first server is quite often ahead on the games' score, giving that player a psychological advantage. In this paper the extent of this advantage is analysed, by considering two identical players.

A fair scoring system has the characteristic that in a match between two equal players, each player has a probability of 0.5 of winning. A scoring system that does not have this characteristic is unfair. Several methods for attempting to overcome the advantage noted above, are considered in Section 2 of this paper, and two methods that reduce the advantage are proposed. Consideration is then given to the case of the reverse psychological effect, the 'back-to-the-wall' effect. The performance of the two proposed methods is then analysed in the presence of both psychological effects. Finally, the case of doubles is analysed. With

four players on the court, there is scope for considering additional methods for overcoming the advantage of serving first. Minor changes to the above two methods are seen to apply.

Earlier studies have considered the advantage gained by lifting play in certain circumstances (Morris, 1977; Pollard, 2002). There is however very little reported empirical evidence of psychological advantages in tennis. A 'first game effect' in a match, namely that fewer breaks occur in the first game of the match, has been identified (Magnus and Klaasen, 1999). More recently, in men's singles grand slam tennis, the better player in a match has been shown to possess a 'back-to-the-wall' effect (Pollard, Cross and Meyer, 2006).

METHODS

(a) Singles

It is assumed that player A has a probability PA of winning a game on service when the games' scores are equal (ie 0-0, 1-1, 2-2, …), that it is PA+D when he/she is ahead in the games' score, and that it is PA-D when he/she is behind. Correspondingly, it is assumed that player B has a probability PB of winning a game on service when the games' scores are equal, that it is PB+D when he/she is ahead in the games' scores, and PB-D when behind. Thus, the psychological advantage of being ahead might be called the 'front-runner' effect, and is represented by D.

(b) A set of singles

Firstly we consider a tiebreak set of tennis between two equal players (PA=PB) with equal psychological factors, D. For simplicity, it is assumed throughout this paper that the two equal players have an equal chance of winning the tiebreak game if it is played (at 6-6). Assuming player A serves in the first game of the set, the probability player A wins the set can be evaluated using a branching diagram or using recurrence methods. For example, when PA=PB=0.6 and D=0.1, the probability that the games' score reaches 2-0, 1-1 and 0-2 is (0.6)(0.5) = 0.30, (0.6)(0.5) + (0.4)(0.3) = 0.42 and (0.4)(0.7) = 0.28 respectively. Further, the probability the games' score reaches 3-0, 2-1, 1-2 and 0-3 is (0.30)(0.7) = 0.210, (0.30)(0.3) + (0.42)(0.6) = 0.342, (0.42)(0.4) + (0.28)(0.5) = 0.308 and (0.28)(0.5) = 0.140 respectively. Continuing in this manner, and adding the probabilities that player A wins 6-0, 6-1, 6-2, 6-3, 6-4, 7-5 or 7-6, it follows that the probability player A wins a tiebreak set is equal to 0.5164 (see Table 1).

(c) A match of the present best-of-three tiebreak sets

We now consider a match of the best-of-three tiebreak sets between two such equal players. We note that if player A serves in the first game of a set and the set lasts an even number of games (ie the set score is 6-0, 6-2, 6-4 or 7-5) , then, under the present service exchange rules, player A also serves first in the next set (otherwise player B serves first). Thus, if player A has an advantage of serving first in the first set, he/she also has that advantage in the second set when an even number of games is played in the first set. The probability that player A wins a tiebreak set in an even number of games, and the probability he/she wins the set in an odd number of games are given in Table 1. Corresponding probabilities are also given for player B.

Table 1: The probabilities two equal players A and B win a tiebreak set in an even and odd number of games when there is a probabilistic advantage D in being ahead in scores

PA	0.5	0.55	0.6	0.65	0.7	0.75	0.8
PB	0.5	0.55	0.6	0.65	0.7	0.75	0.8
D	0.1	0.1	0.1	0.1	0.1	0.1	0.1
P(A wins even)	0.2770	0.2642	0.2499	0.2342	0.2175	0.2001	0.1826
P(A wins odd)	0.2230	0.2438	0.2665	0.2913	0.3187	0.3490	0.3832
P(A wins)	0.5000	0.5080	0.5164	0.5256	0.5362	0.5491	0.5658
P(B wins even)	0.2770	0.2882	0.2976	0.3052	0.3103	0.3122	0.3089
P(B wins odd)	0.2230	0.2038	0.1860	0.1693	0.1535	0.1387	0.1253
P(B wins)	0.5000	0.4920	0.4836	0.4744	0.4638	0.4509	0.4342

Table 2 lists ten mutually exclusive outcomes for a best-of-three sets match won by player A, given player A serves in the first game of the first set. Note that in Table 2 a prefix is used to denote the server in the first game of the set, a capital letter is used to denote the winner of the set, and odd/even classifies the number of games played in the set. The fourth column in Table 2, headed probability, is obtained by multiplying the probabilities of the events in columns 1 and 2, or columns 1, 2 and 3 for the case in which PA=PB=0.6 and D=0.1. These probabilities can in turn be obtained from Table 1. Note that to obtain probabilities for sets with player B serving first, we can just reverse the roles of player A and player B (as they are equal players). Thus, for example, the probability of the event bBeven equals the probability of the event aAeven.

Table 2: The probabilities of ten mutually exclusive outcomes for a best-of-three tiebreak sets match won by player A, when PA = PB = 0.6 and D = 0.1

First set	**Second set**	**Third set**	**Probability**
aAeven	aAeven or odd		0.1290
aAodd	bAeven or odd		0.1289
aAeven	aBeven	aAeven or odd	0.0384
aAeven	aBodd	bAeven or odd	0.0225
aAodd	bBeven	bAeven or odd	0.0322
aAodd	bBodd	aAeven or odd	0.0367
aBeven	aAeven	aAeven or odd	0.0384
aBeven	aAodd	bAeven or odd	0.0384
aBodd	bAeven	bAeven or odd	0.0268
aBodd	bAodd	aAeven or odd	0.0179
		TOTAL	0.5091

It can be seen from Table 2 that when PA=PB=0.6 and D=0.1, the probability that player A wins a best-of-three tiebreak sets match given he/she serves first in the match, is equal to 0.5091. Column (c) in Table 3 gives corresponding results for the other values of PA, PB

and D in Table 1. Thus, it is clear that player A gains a match advantage by serving first in the first set.

Table 3: The probability player A wins a best-of-three tiebreak sets match when there is a probabilistic advantage D in being ahead in the score and player A serves first in the match

PA	PB	D	(c) P(A wins)	(d) P(A wins)	(e) P(A wins)
0.5	0.5	0.1	0.5000	0.5000	0.5000
0.55	0.55	0.1	0.5045	0.5040	0.4996
0.6	0.6	0.1	0.5091	0.5082	0.4993
0.65	0.65	0.1	0.5140	0.5128	0.4991
0.7	0.7	0.1	0.5195	0.5182	0.4993
0.75	0.75	0.1	0.5259	0.5248	0.5002
0.8	0.8	0.1	0.5340	0.5335	0.5022

(d) An alternative best-of-three tiebreak sets system

We now consider the effect of modifying the service exchange mechanism 'across-sets'. The case in which service alternates at the beginning of each set is considered. It can be shown using values from Table 1 that, given player A serves first in the first set, player B first in the second set and player A first in the third set (if necessary), the probability that player A wins the match is equal to 0.5082 when PA=PB=0.6 and D=0.1. This is a slight improvement on the present situation analysed in (c) above. The corresponding probability values for other values of PA, PB and D are given in column (d) of Table 3. Also, it can be seen that if we modify this service exchange mechanism so that player B serves first in both the second and third sets (if necessary), player A's probability of winning the match is now 1-0.5082 when PA=PB=0.6 and D=0.1. Thus, this modification to the third set server leads to no overall difference in fairness on simply alternating service at the beginning of each set. Also, note that when a set lasts an even number of games under this system, the same person serves the last game in that set and the first game of the following set (if played). This should not be a problem as the players have a two minute rest between sets.

(e) An alternative 'across-sets' service exchange mechanism

We now consider a slight variation in the third set to the service exchange mechanism considered in (d) above. We suppose the server in the third set is determined as at present. That is, given player B served first in the second set, player A serves first in the third set if there is an odd number of games in the second set, and player B serves first if there is an even number of games. It can be shown using values from Table 1 that, with this variation to (d) above, the probability player A wins the match is equal to 0.4993 when PA=PB=0.6 and D=0.1. This represents a considerable improvement on the situation analysed in (d) above. The corresponding probability values for other values of PA and PB are given in column (e) of Table 3. (Another variation in the third set to this service exchange mechanism could be where the server in the third set is the player who won the most number of games in the first two sets. If this countback procedure leads to a tie in the number of games won, we use the present service exchange mechanism as in (c) above. This two stage countback mechanism in fact leads to a very small increase in fairness, but

this is not considered to be worthy of further discussion. Other countback methods have been considered in another context (Pollard and Noble, 2006).)

(f) A 'within-set' service exchange mechanism

A service exchange mechanism similar to that used in the tiebreak game is considered. Player A serves in the first game, player B serves in the next two games, player A serves in the following two games, …(ie A,B,B,A,A,B,B,A,A,B,B,A). The present stopping rules (6-0, 6-1, 6-2,…7-5) are used and the tiebreak game is played if the games' score reaches 6-6. Under this service exchange mechanism, assuming that the two equal players have an equal chance of winning the tiebreak game if played, the probability player A wins the set is equal to 0.5046 when PA=PB=0.6 and D=0.1. It can be seen by comparing Table 4 column 4 with Table 1 that this service exchange mechanism within a set considerably reduces the advantage that player A obtains by serving first. A major disadvantage of this mechanism is that player B is required to serve two games in a row on (up to) three occasions, and player A is required to do the same on (up to) two occasions. This mechanism is not considered to be of particular practical relevance. However, if this mechanism was used, change-of-ends might occur after an even number of games is played, so that, when a player serves two games in a row, they are from different ends of the court with a time-break between those service games.

Table 4: The probability player A wins a tiebreak set, given player A serves first and the service-game order is A,B,B,A,A,B,B,A,A,B,B,A (column 4) and when the service order is A,B,B,A;B,A,A,B;B,A,A,B (column 5)

PA	PB	D	Col4	Col5
0.5	0.5	0.1	0.5000	0.5000
0.55	0.55	0.1	0.5023	0.5017
0.6	0.6	0.1	0.5046	0.5032
0.65	0.65	0.1	0.5069	0.5046
0.7	0.7	0.1	0.5092	0.5056
0.75	0.75	0.1	0.5116	0.5062
0.8	0.8	0.1	0.5141	0.5060

(g) Another 'within-set' service exchange mechanism

A variation of the 'tiebreak-like' service exchange mechanism in the above paragraph is the following-A,B,B,A; B,A,A,B; B,A,A,B. Using the present stopping rules and this mechanism, it can be shown that the probability player A wins a set is equal to 0.5032 when PA=PB=0.6 and D=0.1 (see Table 4 column5). It can be seen from Table 4 that this mechanism gives a slight improvement on that in the previous paragraph. Players A and B would each have to serve two games in a row on (up to) two occasions, and change-of-ends could again be 'on-the-even'. This mechanism is also considered to be of little practical relevance.

(h) A third 'within-set' service exchange mechanism

A further 'within-set' mechanism is now considered. Suppose player B, the server in the second game of the set, is allowed to serve two games in a row on (up to) one occasion in

the set (whilst player A never serves two games in a row). The possibilities for the (maximum of) twelve games in a set (up to 6-6) are

(i) A,B;B,A,B,A,B,A,B,A,B,A
(ii) A,B,A,B;B,A,B,A,B,A,B,A
(iii) A,B,A,B,A,B;B,A,B,A,B,A
(iv) A,B,A,B,A,B,A,B;B,A,B,A
(v) A,B,A,B,A,B,A,B,A,B;B,A

For these five alternatives it can be shown that the probability player A wins the set when PA=PB=0.6 and D=0.1 is

(i) 0.4976
(ii) 0.5037
(iii) 0.5078
(iv) 0.5111 and
(v) 0.5140 (see Table 5)

Thus, the mathematics suggests that it would be in player B's interest to elect to serve the two games in a row early (rather than later) in the set. However, he/she might prefer to elect to do it later in the set when the games are more important, or alternatively just after having played an 'easy' service game. It would seem that such a system would increase the 'excitement' of the set. 'Change-of-ends' might again be 'on-the-even'.

The player who serves first in a match against an equal opponent has been shown to have an overall advantage in the situation in which each player has the same psychological advantage when ahead in games' score within the set. Several methods of decreasing this advantage have been considered, and two of them would seem appropriate for consideration. Firstly, if service alternates at the beginning of each set (except the final third or fifth set), the benefit a player receives from serving first in the match is reduced. Secondly, if the player who serves second in a set is allowed to serve on two consecutive occasions within that set, the benefit the player receives from serving first in the set is reduced.

Table 5: The probability player A wins a tiebreak set for each of the five service order cases (i) to (v) in Section 2(h)

PA	PB	D	(i)	(ii)	(iii)	(iv)	(v)
0.5	0.5	0.1	0.5000	0.5000	0.5000	0.5000	0.5000
0.55	0.55	0.1	0.4989	0.5018	0.5038	0.5054	0.5068
0.6	0.6	0.1	0.4976	0.5037	0.5078	0.5111	0.5140
0.65	0.65	0.1	0.4959	0.5059	0.5124	0.5174	0.5220
0.7	0.7	0.1	0.4935	0.5083	0.5178	0.5250	0.5313
0.75	0.75	0.1	0.4898	0.5109	0.5245	0.5345	0.5427
0.8	0.8	0.1	0.4840	0.5136	0.5330	0.5469	0.5576

(i) Another psychological factor

It has been argued by some players, commentators, spectators,…that some players possess a different psychological factor called the 'back-to-the-wall' effect. In this case the player is assumed to have a higher probability of winning a game when behind. We firstly consider

the case in which both players possess this factor. Thus, player A is assumed to have a probability PA of winning a game on service when the games' scores are equal, and that it is PA+D when he/she is behind and that it is PA-D when ahead. Correspondingly, it is assumed that player B has a probability PB of winning a game on service when the games' scores are equal, and that it is PB+D when he/she is behind and that it is PB-D when ahead. Thus, the psychological advantage of being behind is represented by D. Similarly to (b) above, and assuming player A serves first in the set, the probability player A wins a tiebreak set is 0.4806 (refer to Table 6) and the probability player A wins a best-of-three tiebreak sets match is 0.4896 (refer to Table 7 column (c)) when PA=PB=0.6 and D=0.1. Using the 'across-sets' service exchange mechanism described in (e) above, player A's probability of winning such a modified best-of-three tiebreak sets match is 0.5007 when PA=PB=0.6 and D=0.1 (refer to Table 7 column (d)). This represents a considerable improvement on the number immediately above.

Table 6: The probabilities two equal players A and B win a tiebreak set in an even and odd number of games when there is a probabilistic advantage D in being behind in scores

PA	0.5	0.55	0.6	0.65	0.7	0.75	0.8
PB	0.5	0.55	0.6	0.65	0.7	0.75	0.8
D	0.1	0.1	0.1	0.1	0.1	0.1	0.1
P(A wins even)	0.2712	0.2502	0.2269	0.2012	0.1726	0.1405	0.1035
P(A wins odd)	0.2288	0.2403	0.2537	0.2685	0.2845	0.3013	0.3184
P(A wins)	0.5000	0.4905	0.4806	0.4697	0.4571	0.4417	0.4219
P(B wins even)	0.2712	0.2901	0.3066	0.3206	0.3315	0.3382	0.3384
P(B wins odd)	0.2288	0.2194	0.2128	0.2097	0.2114	0.2200	0.2397
P(B wins)	0.5000	0.5095	0.5194	0.5303	0.5429	0.5583	0.5781

Table 7: The probability player A wins a best-of-three tiebreak sets match when there is a probabilistic advantage D in being behind in the games' score and player A serves first in the match.

PA	**PB**	**D**	**(c) P(A wins)**	**(d) P(A wins)**
0.5	0.5	0.1	0.5000	0.5000
0.55	0.55	0.1	0.4949	0.5004
0.6	0.6	0.1	0.4896	0.5007
0.65	0.65	0.1	0.4843	0.5008
0.7	0.7	0.1	0.4788	0.5006
0.75	0.75	0.1	0.4730	0.4997
0.8	0.8	0.1	0.4669	0.4974

The 'within-set' modifications considered in (f), (g) and (h) all decrease player B's probability of winning a set from (1-0.4806)=0.5194. The reason for this is that under these modifications player B is required to serve (on average) earlier in the match so he/she is less often behind (when his/her p-values are higher). The mechanisms in (f) and (g) were

considered to be of little practical relevance. With respect to the mechanism in (h), as player B's probability of winning the set is decreased for all cases (i) to (v), player B would presumably not elect to serve two games in a row as he/she would only decrease his/her probability of winning the set.

(j) The combination of the two psychological factors

We now assume that player A possesses a 'front-runner' factor D1, and player B possesses a 'back-to-the-wall' factor D2. Player A's probability of winning a service game is equal to PA when the games' scores are equal, PA+D1-D2 when he /she is ahead in games' scores, and PA-D1+D2 when he/she is behind. Correspondingly, player B's probabilities on service are PB when equal, PB+D1-D2 when B is ahead and PB-D1+D2 when behind. It can be seen that player A's probability of winning a game on service is always PA when D1=D2, and player B's is always PB when D1=D2. Thus, the present scoring system is fair for this situation, as are the two recommendations in (h) above.

(k) Doubles

The situations for doubles are very similar. As an example, we consider section (b) above for the case in which PA1=PB1=0.65 and PA2=PB2=0.55 (PA and PB both average 0.6), the only psychological factor being the 'front-runner' effect for every player and it is assumed to be D=0.1, and the teams' chances at the tiebreak game are assumed to be equal.

When the service order is A1, B1, A2, B2,… (the typical case in which each team uses their more effective server first) the probability team A wins the tiebreak set is 0.5189. When the service order is A2, B2, A1, B1,…the probability team A wins the set is 0.5139 (a fairer outcome than for the order in the previous sentence). This suggests a minor adjustment to the 'across-sets' modification in (e) above. Namely, if team A serves first in the first set (with service order A1, B1, A2, B2) and team B serves first in the second set (with service order B1, A1, B2, A2), then if team A serves first in the third set, the first two servers should be reversed (ie A2, B2, A1, B1,…), and if team B serves first in the third set, the order should be B2,A2,B1,A1.

Looking at the 'within-set' changes considered in section (h) (cases (i) and (ii) in particular), when the order is A1,B1;B2,A2,B1,A1,…, the probability team A wins the set is 0.4820, and when the order is A1,B1,A2,B2;B1,A1,B2,A2,…, the probability team A wins the set is 0.5054. This suggests that after four games have been played within a set, team B be allowed to play two service games in a row.

RESULTS

Given two equal singles players with an equal psychological advantage when ahead, the player who serves first is shown to have a probability of winning the set greater than 0.5. His/her probability of winning a best-of-three sets match is also greater than 0.5. Thus, given the existence of a psychological advantage when ahead, the present best-of-three tiebreak sets scoring system is unfair.

It has been shown in the previous section that the present scoring system can be made fairer by two methods. Firstly, alternating service at the beginning of each set (with the server in

the final third or fifth set being determined as at present) reduces the unfairness. Secondly, allowing the player who serves in the second game of a set to serve (only (up to) once) on two consecutive games within that set also reduces the unfairness.

The reverse psychological effect is when a player lifts his/her game when he/she is behind in games' score (the "back-to-the-wall" effect). The two methods above have also been shown to be applicable to the situation in which one player has the psychological advantage of being ahead or its reverse, whilst the other player also has this psychological advantage or its reverse.

Also for doubles, the above two methods were shown to decrease unfairness. Interestingly, the unfairness is further reduced by reversing the service order within each doubles pair for the final third or fifth set.

DISCUSSION

The problem that the person or team that serves first in a set of tennis, has an advantage, has been long recognised. Indeed, it is an intrinsic difficulty within the tennis scoring system, and in this paper it has been quantified. The solution presently used is to toss a coin, so that each player or team has an equal chance of getting the advantage of serving first. A better solution is to modify the scoring system so that the advantage of serving first is decreased or reduced to zero. Scoring systems in which this advantage is zero have been devised (Miles, 1984), but their structures are quite different to the present tennis scoring system. A change to such a structure would be regarded by many people as a major change, and hence would be unlikely to gain acceptance. In this paper, minor changes to just the service exchange mechanism within the scoring system have been considered and shown to decrease the advantage gained by the person or team that serves first.

The methods of this paper can be used to analyse the one-day and test versions of a series of (say) three or five cricket matches. At present there is a toss before each match within the series. Assuming there is a psychological advantage in batting first in a match, then it can be shown that it is better to toss only before the first match within the series, and then alternate the first team to start the batting after that. The team to bat first in the final (third or fifth) match could be determined by some countback procedure. More generally, it can be seen that the toss of a coin is often used to create fairness in a situation that is intrinsically unfair. The irony of the situation in cricket is that the use of the toss of a coin three or five times only makes the rules about the first team to bat in each match not as good as they can be. One toss is not only enough, but it is better when followed by an alternating structure.

CONCLUSIONS

For the situation in which players have a psychological advantage when ahead in games' score, the player who serves first in a set of tennis has been shown to have an advantage. This (set) advantage can be decreased by alternating service *at the beginning of each set* (with the exception of the final third or fifth set which would be determined under the

present rules). This change to the present service exchange mechanism would be very easy to implement, and it appears to have no real drawbacks.

The advantage of serving first in a set can be further reduced by allowing the player who serves in the second game of the set to serve on two consecutive games at some stage within that set. This might seem to be a little unusual at first, but it would appear to create some additional excitement in the set for the spectators. It would also create a strategic and additional dynamic element in the set for the players. However, it might be more difficult for such a change to gain acceptance. Also, these two changes in the service exchange mechanism have been shown to be applicable when either or both players have either this psychological effect (the 'front-runner' effect) or its reverse (the 'back-to- the-wall' effect).

Finally, the two changes have been shown to be applicable to doubles. The advantage of serving first in the match is further reduced in doubles by reversing the service order within each doubles pair for the final third or fifth set in a match.

REFERENCES

Magnus, J. and Klaasen, F. (1999) On the advantage of serving first in a set: Four years at Wimbledon. *The Statistician* **48**, 247-256.

Miles, R. (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.

Morris, C. (1977), The most important points in tennis. In S. Ladany and R. Machol (Eds). *Optimal Strategies in Sports*. North-Holland, Amsterdam, 131-140.

Pollard, G. (2002), The effect of a variation to the assumption that the probability of winning a point in tennis is constant. *Proceedings of the Sixth Australian Conference on Mathematics and Computers in Sport*. Bond University, Australia, 227-230.

Pollard, G., Cross, R. and Meyer, D.(2006), An analysis of ten years of the four Grand Slam Men's Singles data for lack of independence of set outcomes. *Proceedings of the Eighth Australasian Conference on Mathematics and Computers in Sport.* Coolangatta, Australia.

Pollard, G. and Noble, K. (2006) Countback methods and auxiliary scoring systems in tennis. *Proceedings of the Eighth Australasian Conference on Mathematics and Computers in Sport.* Coolangatta, Australia.

AFL FOOTBALL – HOW MUCH IS SKILL AND HOW MUCH IS CHANCE?

Forbes, D, Clarke, S. and Meyer, D.

School of Life Sciences, Swinburne University

ABSTRACT

It is widely accepted that luck has some influence on all sporting events, whether it be an injury at a crucial time, an unfortunate bounce of the ball or a controversial decision by an official. But just how much does the result of a match come down to bad or good luck i.e. the random events within a game? A Markov process model with 18 states has been developed to approximate AFL football using match statistics. The models have been applied to matches from the 2003 and 2004 AFL seasons and displayed very encouraging results with minimal errors. Using transition probabilities derived from these matches, simulated goal distributions and probabilities of victory have been calculated. Using these distributions and the standard errors for each team's match scores, the randomness that occurs in AFL matches is identified. Some of these matches have been investigated with the results presented in this paper.

KEY WORDS
Markov process, Poisson distribution, computer simulation, random events

INTRODUCTION

In almost all sports played worldwide, there is an element of luck associated with performance. The game of Australian Rules football is no different. The fact that the ball used is oval in shape and often does not bounce as expected probably exacerbates the importance of being on the right side of lady luck. Historically, the issue of chance in sport and the role it plays has not been a focal point of academic research. Investigation into chance playing a part in sporting outcomes has mostly been concentrated on soccer.

Early work on modelling sporting events using statistical distributions found that the best descriptor of goals scored in soccer was the negative binomial distribution (Moroney, 1956). The author expressed surprise that weather conditions and the strength of the competing teams did not exert as great an effect as is often supposed. Later work also found the negative binomial distribution to give the best fit due to it being generated by random or chance mechanisms that underlie the conclusion that soccer is a game dominated by chance. It was suggested that, due to this notion, a team who recognised this random element would be able to develop a successful style of play that harnessed the importance of chance on the game (Reep and Benjamin, 1968). A direct style of play was deemed to be the best method by which to harness the chance occurrences that appear in soccer. By putting the ball into goal scoring situations as often as possible, a team was more likely to take advantage of these opportunities when they arose. Recent work has traced this direct and successful style of play back to the Arsenal side of the 1930s, but does acknowledge the part Reep played in bringing it to the successful Norwegian side of the 1990s (Larson, 2001).

The negative binomial distribution was applied to a number of facets of different sports (Pollard, Benjamin and Reep, 1977). Events looked at included passing chains in soccer (and goal scoring), points scored in gridiron, runs in a baseball half-inning, goals scored in ice-hockey, strokes per rally in tennis and runs scored per partnership in cricket. It was found that the negative binomial produced a good fit where there was an occurrence of infrequent events in a team environment e.g. soccer goals. However, when individual performances were looked at, such as in the tennis or cricket examples, there was not a close fit, indicating that individual skill was more significant than chance in these instances.

These discoveries lend weight to chance playing a significant part in matches played in the Australian Football League (AFL), where the negative binomial has been found to be the best approximation for goal scoring in the competition, and the Poisson is the best approximation for individual team returns (Forbes, 2006). This paper uses a Markov process model to investigate matches from the 2005 AFL season, highlighting the importance of chance and making the most of it to win AFL games. Furthermore, simulated matches give a better understanding of the amount of variation and randomness that can be expected in a team's score for a match.

THE MODEL AND SIMULATION PROCESS

An 18-state Markov process model has been developed to approximate Australian Rules football (Forbes, 2006). This model builds on an earlier, cruder model that was made up of only seven states (Forbes and Clarke, 2004). The accuracy of this model in reflecting the events of a match is impressive. It gives rise to a number of applications that can be used by AFL clubs, as well as media outlets and sports bookmakers. Some of the applications of the model will be presented in the case studies that follow, investigating the role of chance and random events in AFL football.

A simulation program has been developed to utilise the model in a post match environment. The underlying idea of the program is to generate random numbers that determine which state the model will move into next. The process has been broken down into four parts to reflect more accurately an AFL match, which is divided into four quarters. The starting state for each quarter is state seven (centre bounce) and the simulation is run according to the number of transitions contained in the match divided by four. The simulation is run 10,000 times in general, but there is no restriction, save CPU processing time, to the number of times it can be run. To give an example of speed, on a notebook with 256MB of RAM, a season comprising 185 matches can have each match simulated 10,000 times in roughly 20 minutes.

Upon running a simulation of a match, or set of matches, there is a variety of information that can be gleaned from the process. Perhaps the most important and useful data are the projected score of each team for each simulated match. Goals are calculated for each team by counting the number of times either Team A or B had possession directly before a centre bounce, excluding the centre bounce at the start of each quarter. These occurrences are then multiplied by six and added to the number of

times the model entered the behind state for each team to come up with a match score, which is then used to ascertain the match outcome. The probability of victory for each team and the likelihood of a draw is then a simple calculation according to the expected outcome. Other information that can be extracted from a simulation is the proportion of time the 'match' is likely to spend in any one state and the number of occurrences of each state within a simulated match.

The simulation process was validated by the results obtained under different scenarios. Home advantage in the AFL competition between 1998 and 2003 amounted to 12.3 points on average. When the 2004 season was simulated using a transition matrix for each match, the average margin in favour of the home side amounted to a very similar 11.8 points. A matrix was derived for the 2005 season based on the 185 matches and used to simulate an 'average' match. The score line that resulted from the simulation was 99-90 in Team A's favour, whereas for the actual data from the 2005 season, the average score was 102-87, again suggesting a favourable comparison between the simulated and actual results. Further results from the simulation program will be presented in the following case studies.

CASE STUDY 1: SYDNEY V WEST COAST

The 2005 AFL Grand Final was a classic, not so much for the quality of football played, but more so due to the commitment of both teams and the knife edge atmosphere that was created by the closeness of the score line. In the end Sydney won the match by just four points, 58-54, and no doubt West Coast spent the summer licking their wounds and pondering upon just how close they came. Some analysis has been done on the match to investigate how much of a part randomness played in the final result and to determine whether this was a game that could have been won by the bounce of the ball.

The transition matrix for the match was derived using the 18-state zone model. The match contained 885 transitions, which amounts to 221 a quarter. Using the observed transition matrix and the number of transitions for a quarter, the match was simulated 10,000 times. From the simulation, expected scores for each team were obtained, along with probabilities of victory and these are contained in Table 1.

Table 1: Sydney v West Coast simulated output

Team	Mean	Std Dev	Probability of Victory
Sydney	58.4	16.9	0.53
West Coast	56.0	16.0	0.45

What is evident from Table 1 is that, with the transition probabilities as they were in the match, the West Coast were expected to score, on average, two more points than they did. They have a mean score of 56.0 points from the simulation results, yet in the Grand Final they scored 54 points. Do we then consider West Coast unlucky not to get the return on the scoreboard that was expected from the way the match was played? They had a 45% chance of winning the match and a 2% chance of the draw.

To continue on from this match, the distribution of goals for each team has been calculated from the simulated matches. It has been discovered in other research that the

distribution of goals for teams in the AFL is best approximated by the Poisson distribution (Forbes, 2006). The assumptions that must be met for the Poisson distribution to apply are that the probability of an event within a certain interval does not change over different intervals and that the probability of an event in one interval is independent of the probability of an event in any other non-overlapping interval. The first assumption indicates that club scoring rates in the AFL do not vary from week to week and therefore, are not affected by factors such as weather conditions, venue size, strength of the opposition or influential players missing. Although this is somewhat of a surprising result, the distributions presented below for both Sydney and the West Coast, highlight the variability associated with how many goals a team can be expected to score in an AFL match.

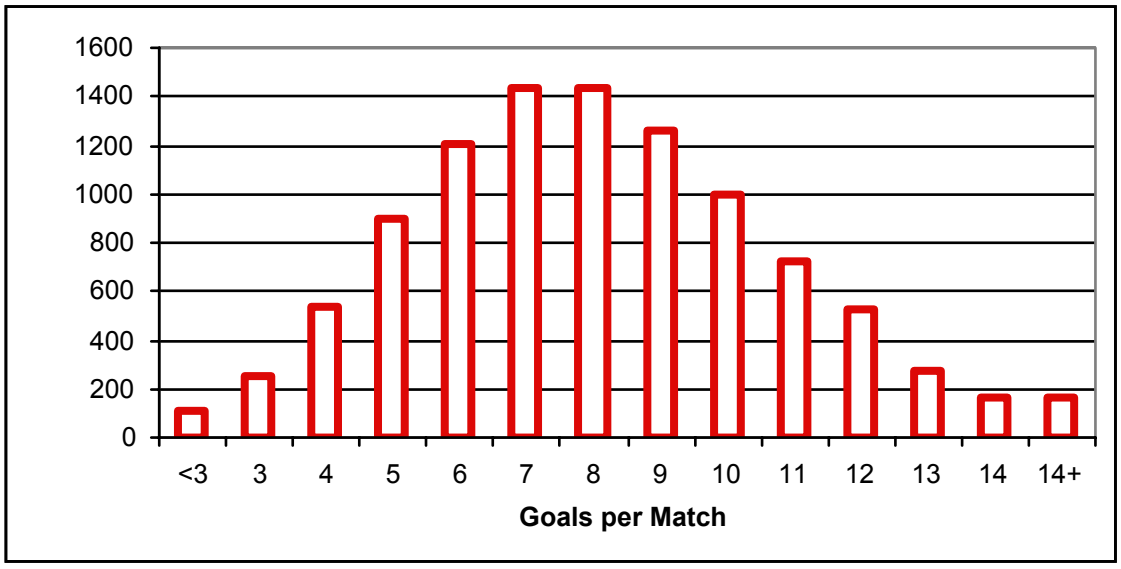


Figure 1: Sydney goal distribution from simulated Grand Final

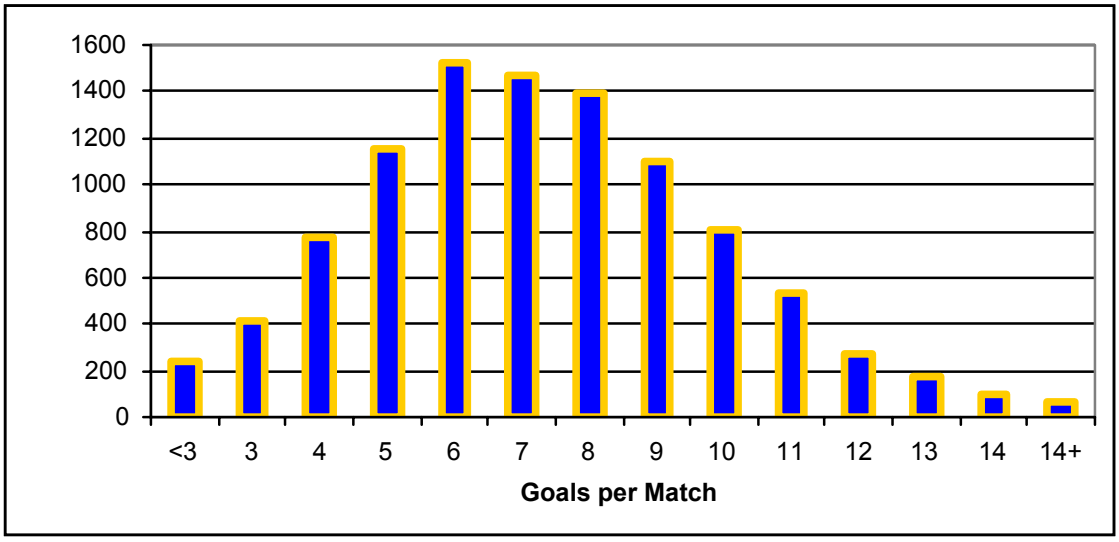


Figure 2: West Coast goal distribution from simulated Grand Final

The final analysis of the Grand Final involves what turned out to be the ultimate play of the match. West Coast surged into attack and West Coast players were queuing up to mark the ball when Sydney's Leo Barry came from the side for a spectacular contested

mark. Barry's mark was the kind of mark that a player might pull off a couple of times during a career and further indicates the importance of taking your chances when they arise. The ball could quite easily have been marked by a West Coast player, who would have had a relatively straight forward shot on goal to win a premiership.

To investigate the importance of Barry's mark, the last three plays of the game have been simulated starting from West Coast possession in the midfield, with the next transition resulting in an Inside 50 play for the West Coast. Using the matches from the simulation where these conditions were satisfied, with Sydney leading the match by four points at the time, the West Coast were found to have an 11% chance of winning the match. Even though this is still a relatively low chance of winning, compared to Sydney's chances, it indicates the importance of the play by Barry. Paul Roos, the coach of Sydney, would not like to endure those final few seconds again, knowing that he had an 11% chance of having the Premiership snatched away from him if Barry spills the mark.

CASE STUDY 2: KANGAROOS V RICHMOND

The 2005, round 12 match between the Kangaroos and Richmond, saw the Roos walk away with a comfortable 29 point victory, 109 - 80. On paper, the issue of chance does not seem to be a concern. A five goal victory is comprehensive in anyone's language, but upon drilling deeper into the match statistics, chance played a big part in such an easy win. The transition matrix for the match, when used in the simulation program produced an expected score line of 111 – 78, which is actually four points more than the observed margin, indicating that perhaps Richmond were 'lucky' to get as close as they did. Not surprisingly, the Kangaroos were an 85% chance of winning the match.

That aside, the goal kicking of both teams was the crucial difference, with the Kangaroos kicking 17 goals, 7 behinds and Richmond 10 goals, 20 behinds. Richmond had six more scoring shots than their opponents, yet they were soundly beaten by nearly five goals. The possession profiles of the teams were very similar and this is highlighted by how they entered their attacking zones. Table 2 contains the profile for each team when they put the ball into their attacking zone.

Table 2: Distribution of ball into attacking zones, Kangaroos v Richmond, round 12 2005

Team		BUBO	THIN	DISP	APOS	BPOS	BEHI	CEBO	Total
Kangaroos	Count	0	0	21	17	5	0	3	46
Kangaroos	%	0.0	0.0	45.7	37.0	10.9	0.0	6.5	100.0
Richmond	Count	0	0	22	10	20	3	0	55
Richmond	%	0.0	0.0	40.0	18.2	36.4	5.5	0.0	100.0

There is no clear indication from Table 2 as to why the Kangaroos were such easy victors. In their favour was the ability to kick long goals from outside 50m. Aside from this, the profiles are very similar. Deeper investigation has to be carried out on what happened when the ball was in the attacking zones to see why there was such a discrepancy on the scoreboard. Richmond did better than the Kangaroos at getting the ball from dispute in the Kangaroos attacking zone. In Richmond's attacking zone the teams broke even in this area. The lopsided scoreboard can only be put down to Richmond's inaccuracy both from inside and outside 50. Table 2 shows that the

Kangaroos kicked three goals from beyond 50m whilst Richmond could manage only three behinds. When the ball got into the attacking zones, the Kangaroos kicked a goal 40.0% of the time and a behind only 20.0% of the time, whereas, Richmond could manage a goal only 16.8% of the time and a behind 30.4% of the time. This is a clear case where the relationships in the game are relatively equal; however, poor conversion by one side has led to a comprehensive defeat.

To investigate the random effect of goal kicking accuracy, adjustments have been made to the transition matrix so that each side has a similar conversion rate of 50%. Using this adjusted matrix in the simulation program, the expected score line is now 111 – 104 in the Kangaroos favour, but they are now only a 59% chance of winning. This is a powerful example of the importance of accurate kicking in the AFL. Richmond, cruelled their chances of victory with poor conversion and this is a purely random occurrence as the week before they kicked 15 goals, 13 behinds and the week after 9 goals, 8 behinds.

CONCLUSION

Clearly, randomness is a big part of success in the AFL competition. With the aid of an 18 state Markov process model and a simulation program, two case studies have been presented that identify the importance of chance in AFL matches. The first game looked at was the 2005 grand final and the simulated goal distributions gave a clear indication of the variability one can expect from game to game in goals kicked. Furthermore, the expected score line was closer than the actual score, highlighting that perhaps the luck did not run with the West Coast. Finally, a once in a lifetime mark at the end of the game saved the premiership for Sydney, when a different outcome produced an 11% chance of victory for the opposition with only three plays left in the match.

The second case study highlighted the random effect of goal kicking in an AFL match; Richmond had more shots than their opposition, but got beaten by five goals due to their poor conversion. With an adjusted conversion rate of 50%, surpassed by Richmond in the week before and the week after the match, their chances of victory improved by over 25%. This is another example of the effect of chance in the competition and, for Richmond, the match against the Kangaroos was obviously one of those days.

REFERENCES

Forbes, D. (2006) Dynamic prediction of Australian Rules football using real time performance statistics. *PhD Thesis*, Swinburne University of Technology.

Forbes, D. and Clarke, S. R. (2004) A seven state Markov process for modeling Australian rules football. In H. Morton (Ed). *Seventh Conference on Mathematics and Computers in Sport.* Palmerston North: Massey University.

Larson, O. (2001) Charles Reep: A Major Influence on British and Norwegian football. *Soccer and Society* **2**(3)**,** 58-78.

Moroney, M. (1956) *Facts from Figures,* Middlesex, Penguin Books.

Pollard, R., Benjamin, B. and Reep, C. (1977) Sport and the negative binomial distribution. In S. Ladany and R. Machol (Eds). *Optimal Strategies in Sports*. Amsterdam, North Holland.

Reep, C. and Benjamin, B. (1968) Skill and chance in association football. *Journal of the Royal Statistical Society Series A,* **131,** 581-585.

A PROBABILITY BASED APPROACH FOR THE ALLOCATION OF PLAYER DRAFT SELECTIONS IN AUSTRALIAN RULES FOOTBALL

Bedford, A. and Schembri,A.

School of Mathematical and Geospatial Sciences, RMIT University, Bundoora, VIC, Australia

ABSTRACT

Australian Rules Football, governed by the Australian Football League (AFL) is the most popular winter sport played in Australia. Like North American team based leagues such as the NFL, NBA and NHL, the AFL uses a draft system for rookie players to join a team's list. The existing method of allocating draft selections in the AFL is simply based on the reverse order of each team's finishing position for that season, with teams winning less than or equal to 5 regular season matches obtaining an additional early round priority draft pick. Much criticism has been leveled at the existing system since it rewards losing teams and does not encourage poorly performing teams to win matches once their season is effectively over. We propose a probability-based system that allocates a score based on teams that win 'unimportant' matches (akin to Carl Morris' definition of importance). We base the calculation of 'unimportance' on the likelihood of a team making the final eight following each round of the season. We then investigate a variety of approaches based on the 'unimportance' measure to derive a score for 'unimportant' and unlikely wins. We explore derivatives of this system, compare past draft picks with those obtained under our system, and discuss the attractiveness of teams knowing the draft reward for winning each match in a season.

KEY WORDS

AFL, probability, draft, importance

INTRODUCTION

The AFL draft system has been designed to favour teams anchored to the bottom of the ladder. This enables those teams to improve their player lists and propel themselves up the ladder during future seasons by having a first choice of picking rookie players. Currently, the order of the AFL draft coincides with the inverse order of the ladder as it stands at the conclusion of the home and away season. This system allocates the first draft choice to the team that finished last (or sixteenth), the second draft choice to the team that finished fifteenth, and the sixteenth draft choice to the team that finished first. Subsequent rounds of the draft replicate the exact order of the first round (www.afl.com.au).

A highly contentious issue surrounding the AFL draft has been the allocation of priority picks. A priority pick is a draft choice provided to a team prior to the first round of the national draft. From 1997 to 2005, priority picks were provided to teams that won fewer than five games during the regular season (www.afl.com.au). In effect, if a team finished last with less than five wins, they received a priority pick in addition to the first choice in the national draft, thus enabling the club to receive the two best available

players. This system provided teams with poor win/loss records with little incentive to win games during the latter part of the season and actually provided clubs with an incentive not to win five games.

The draft systems of other major world sports are somewhat comparable to that of the AFL. Major League Baseball (MLB) and the American National Football League (NFL) both allocate draft choices using inversed final season standings (Grier & Tollison, 1994; Spurr, 2000). However, neither competition provides priority picks and thus does not provide additional reward for winning only a handful of games.

Several studies have assessed the incentive effects of draft systems and their impact on team performance. Taylor and Trogdon (2002) assessed the performance of teams following initiatives by the NBA to reduce team incentives to win or lose games. After controlling for venue and the quality of each team, these authors found that when draft choices were decided on inverse rankings (1983-84 season), non-playoff teams were 2.5 times more likely to lose games than teams likely to feature in the playoffs. However, when the NBA modified the draft system and gave all teams an equal probability of obtaining the first draft choice (1984-85 season), non-playoff teams were as likely to win as play-off bound teams. Finally, when the lottery system became weighted during the 1989-1990 season, non-playoff teams were 2.2 times more likely to lose when compared to teams qualifying for the playoffs (Taylor & Trogdon, 2002). These findings demonstrate the profound impact of providing incentives for teams to lose games and head towards the bottom of the ladder. Furthermore, it highlights that an incentive based system such as the process employed by the AFL increases the likelihood that teams will lose additional matches during the latter part of the season since they are unlikely to feature in the finals.

In this paper we use a model that allocates a *Draft Point Reward* (*DPR*) to each team when they win a match. This reward varies in value from 0 to 1 depending upon the Unimportance of the match. The cumulative sum of *DPR*, known as the *DScore*, is used to determine the final draft picks at the conclusion of the regular season.

We will begin our work by defining the criteria our model should meet. We then outline the methods employed, and consider the operational aspects of the model.

METHODS

Our system is based on Carl Morris' famous work on the most important points in tennis (Morris, 1977). He defined the importance of a point as the difference between two conditional probabilities: the probability a server wins the game given that he wins the next point, minus the probability a server wins the game given that he loses the next point. Here we are not considering points in a game, but rather matches in a season, and it is the Unimportant matches that appeal to us. We calculated Unimportance so that it is independent of the opposing team.

Criteria

In devising the system of selection for the AFL national draft, we designed our model based on the following:

1. Teams with a reduced probability of making the finals are rewarded incrementally higher for winning matches of high Unimportance
2. Teams that have qualified for the finals are ineligible for any reward.

3. *DPR* is restricted to a 16-week period commencing from the end of Round 6.
4. *DPR* is higher for teams unlikely to win, and is further enhanced by the Unimportance of a match in terms of making the finals.
5. No *DPR* is given in defeat, so teams must win to obtain a reward.
6. A priority system is in place to protect teams that have continuous runs of losses, but it is not implemented at the expense of rewarding victory.

The way in which the number of matches 'needed to win' is calculated is based upon the minimum number of wins needed by a team in the remainder of the season based solely upon making the final 8 (*F*8). Obviously this is not precisely known until the end of the season; however a reasonable estimation can be made.

Probabilistic model

The heart of our model is based upon reworking Morris' equation to suit our purpose of determining how Unimportant a match is to a team's finals aspirations. There are a number of things that we need to evaluate first, such as what measures are required in our assessment of what makes a match Important, and, in turn, Unimportant. A regular AFL season constitutes 22 matches and we need to consider the probability of a team making the finals based upon the number of matches won at round *r*. There are a number of features in our probabilistic model that were used to determine how much a team was rewarded for winning a match. The process is as follows:

1. Determine the minimum number of wins (*Par Wins*) required for team *i* to make the final 8 after round *r*.
2. Check if team *i* at round *r* has already made the final 8 or cannot make the final 8. If neither of these events are true, we determine the probability of team *i* making the finals at the completion of round *r*.
3. Calculate the Unimportance of match *r* +1 for team *i* using the above results.
4. Allocate the *Draft Points Reward* (*DPR*) based on the above measures.

Determination of projected wins to make the final 8

There are two possible approaches to determining the number of wins required to make the final 8 at round *r* for team *i*. We could either use the final season's required wins and impose that retrospectively on the completed season, or use a projected requirement during the season and keep this result even at the end of the season. For example, in season 2004, the eighth placed team won 12 of 22 matches to make the finals. Ultimately, differing results make it difficult to predict this result during the season. However, the attraction of our model is that teams must know the rewards of winning their next match prior to the game as an incentive to win. They should also be confident this reward does not change post game. So we used a projected final 8 wins, or *Par Wins*, during the season and maintain these values to seasons end, despite minor variations in predictions.

We formally define the number of wins required after round *r* for team *i* as $Par\ wins_i(r)$; and the total number of wins for team *i* at the completion of round *r* as $TW_i(r)$. Using the 8th ranked team at any round *r* as the ideal *Par* proportion in determining the wins required to make the finals, we obtain

$$Par\ wins_i(r) = \max\left(\left\lfloor \frac{\left(TW_{8^{th}\ \text{ranked team}}(r)\right)\times 22}{r} \right\rfloor - TW_i(r), 0\right) \quad (1)$$

This does, on occasion, return a result that is not possible. For example, a team with 4

wins at the completion of round 7, and sitting in 8th place, yields a *Par Wins* of 8.57. Therefore we round to the nearest 0.5, using 0.25 and 0.75 as the round off points. In this example, we round to 8.5, and this is interpreted as team *i* requiring 8.5 wins (minimum) from the remaining 15 games to make the finals.

Determination of the probability of making the final 8

At the heart of the second stage of the process is the binomial distribution. A number of other methods were considered, such as simulating the remainder of the season using success probabilities for each team using $p = 0.5$, or varying p; also averaging the number of wins of all teams and forward multiplying to determine the number of wins needed to make the final 8. Ultimately, it was both simplicity and a reduction of variability that settled our choice. We define the probability of team *i* at the completion of round *r* making the final 8 as $\Pr_i(F8 \mid r)$. Using $B(x; n, p)$ (the cumulative binomial distribution function with x = number of successes, n = number of trials and p = probability of success), we have

$$\Pr_i(F8 \mid r) = 1_{\{ParWins_i(r)=0\}} + 1_{\{(ParWins_i(r)>0)\cap(ParWins_i(r)\le 22-r)\}}\left[1 - B\left(\text{Par wins}_i(r) - 1; 22 - r, p_i\right)\right] \quad (2)$$

where $1_{\{a\}}$ is the indicator function taking value 1 if condition *a* is true and 0 if false.

Notably, one must consider the value of p_i. We have chosen to look at two methods, the first, and predominant choice in our results, is the classic coin toss model $(p_i = 0.5)$. The second method uses the winning ratio $\left(p_i = \frac{\text{TW}_i}{r}\right)$. One could be tempted to use successful prediction probabilities such as those determined by Stefani and Clarke (1992); or the simpler winning ratio. However, we wanted the system to be as simple as possible, and the introduction of a nested probability model may complicate this idea. A brief treatment of this is given in the discussion section.

The unimportance of a match

We define the Importance for team *i* at the end of round *r*, or $I_i(r)$, as:

$$I_i(r) = \Pr_i(Make\,F8 \mid Win\ match\ r+1) - \Pr_i(Make\,F8 \mid Lose\ match\ r+1) \quad (3)$$

Now we unpack the two components of Importance:

$$\Pr_i(Make\,F8 \mid Win\ match\ r+1) = 1_{\{ParWins_i(r)=0\}} + 1_{\{(ParWins_i(r)>0)\cap(ParWins_i(r)\le 22-r)\}}\left[1 - B\left(ParWins_i(r) - 2; 22 - (r+1), p\right)\right] \quad (4)$$

and

$$\Pr_i(Make\,F8 \mid Lose\ match\ r+1) = 1_{\{ParWins_i(r)=0\}} + 1_{\{(ParWins_i(r)>0)\cap(ParWins_i(r)\le 22-r)\}}\left[1 - B\left(ParWins_i(r) - 1; 22 - (r+1), p\right)\right] \quad (5)$$

By using the binomial cumulative density function to model the probability of making the finals based on winning or losing the next match, we can, in turn, calculate the Unimportance of a match. Through some neat cancellation of terms we obtained a simple result for the Unimportance:

$$U_i(r) = 1 - I_i(r)$$
$$= 1 + \Pr_i(Make\,F8 \mid Lose\,match\ r+1) - \Pr_i(Make\,F8 \mid Win\,match\ r+1)$$
$$= 1 + [1 - B(x;n,p)] - [1 - B(x-1;n,p)]$$
$$= 1 + \left[1 - \sum_{k=0}^{x} b(k;n,p)\right] - \left[1 - \sum_{k=0}^{x-1} b(k;n,p)\right]$$
$$= 1 + [1 - (b(0;n,p) + b(1;n,p) + \cdots + b(x;n,p))] - [1 - (b(0;n,p) + b(1;n,p) + \cdots + b(x-1;n,p))]$$
$$= 1 - b(x;n,p)$$

So,

$$U_i(r) = 1 - b(Par\,Wins_i; 22 - (r+1), p) \qquad .(6)$$

Noting $b(x;n,p)$ (the discrete binomial distribution function with x = number of successes, n = number of trials and p = probability of success) in the final result, Unimportance is simple to evaluate, relying on a discrete rather than continuous result, and given the values of *Par Wins*, can be easily computed using a scientific calculator.

Allocation of Draft Point Reward (DPR)

The allocation of *DPR* is simply the Unimportance probability multiplied by the probability of not making the final 8 at round *r*. In this way, the Unimportance is tempered by the likelihood of making the final 8. Teams that cannot make the final 8 receive the highest weight possible (1), that is, the full Unimportance probability, as long as they win the match. The allocation of *DPR* for team *i* at round *r* is given by the following:

$$DPR_i(r) = 1_{\{i\,wins\,match\,r\}} \cdot 1_{\{r>6\}} \cdot U_i(r) \cdot (1 - \Pr_i(F8 \mid r)) \qquad (7)$$

where $1_{\{a\}}$ is the indicator function taking value 1 if condition *a* is true and 0 if false. The Draft Score, or *DScore*, for team *i* at round *r* is simply the sum of the *DPR*:

$$DScore_i(r) = \sum_{k=7}^{r} DPR_i(k), r \in \{7,..,22\} \qquad (8)$$

Using the DScore for the national pre-season draft

The use of the *DScore* towards draft selections encompasses some parts of the AFL's latest policy on priority picks. For our *DScore* system, the teams are ranked 1 through 16, with the highest *DScore* attracting pick 1, and the lowest pick 16. This ordering remains for the subsequent iterations of the draft with one exception. Current AFL policy dictates that a team that wins less than or equal to 4 matches in a season receives a priority pick in the second round of the draft. As a method of protecting teams that may never win another match after round 6, we employ a similar priority pick system, whereby a team that wins less than or equal to 5 matches in a season receives a priority pick at the start of the second round of the draft. This is a little more generous than the AFL system, however the bottom side will not necessarily end up with the first draft pick under the *DScore* model.

RESULTS

We begin by examining how the system operated for 2005 in finer detail. We then cover some interesting scenarios, and investigate the implications of the model.

The 2005 season

For season 2005, a number of teams remained in contention for the final 8 right through to the last round. The final round saw five teams competing for three finals places. One win separated 6th through 10th at seasons end. Notably, half a win separated last (16th) from 14th and all three bottom sides received a reward from the AFL for winning less than or equal to 5 matches. Table 1 outlines the final results of three draft systems; first the variable success *DScore* model, then the 50-50 *DScore* model, and finally the AFL system. Note that there is little variation when using a team's win ratio to determine $\Pr_i(F8 \mid r)$ instead of the simpler $p_i = 0.5$, and henceforth we will only consider the equal success probability model.

Table 1 : Draft pick comparison for *DScore* and actual draft system for the 2005 AFL season.

Team	*DScore,* $p_i = \frac{\mathrm{TW}_i}{r}$	*DScore,* $p_i = 0.5$	AFL Draft System
Carlton	9	9	4
Collingwood	7	7	5
Hawthorn	5	3	6
Essendon	2	2	7
Richmond	11	11	8
Brisbane	4	4	9
Fremantle	6	6	10
Western Bulldogs	1	1	11
Port Adelaide	3	5	12
Melbourne	14	13	13
Geelong	15	14	14
Kangaroos	13	15	15
St Kilda	10	10	16
Sydney	8	8	17
West Coast	16	16	18
Adelaide	12	12	19
Priority	17	17	1
Priority	18	18	2
Priority	19	19	3

Carlton, Collingwood and Hawthorn received priority picks 1, 2 and 3 respectively under both the AFL model and our model, although ours comes into effect in round 2 of the draft. Variations in the 2005 season round-by-round results are given in Figure 1.

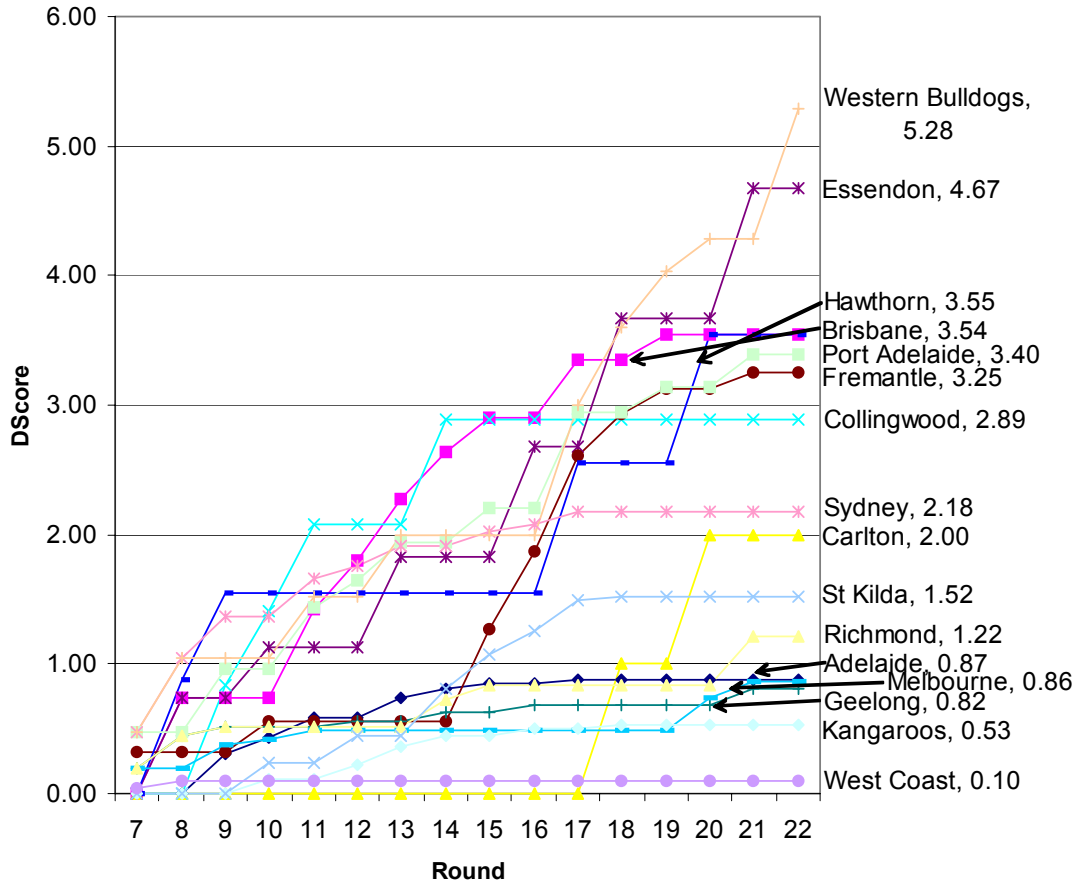


Figure 1: *DScore* by Round by teams for 2005

Variation of the *DScore* throughout the season is evident; with the number 1 pick changing teams 11 times during the season - twice in the last three rounds. Also, picks 3 to 7 provided extremely close results in the final round, given that if Collingwood had won its last match against the Western Bulldogs they could have secured pick 3 (instead of 7) and cost the Western Bulldogs first pick. So a win to Collingwood under the *DScore* model would see a rise to pick 3, however a win under the AFL model would have seen a drop to pick 5.

An evaluation of the incentive of the DScore model

Ideally the *DScore* model should evidence high *DPR* continuously for low placed teams, given they win. Table 3 outlines the number of teams in contention for the number one pick in the last round, and three rounds before the end of the home and away season under the *DScore* model. The first overall draft pick changed teams in the last round during seasons 2001, 2005, and in the last 3 rounds during seasons 1997, 1999, 2001, 2003, 2004, 2005. There was a blowout in the *DScore* in 2000 and thus, the race for the top draft pick was over by round 20. However, six teams fought it out for picks 2 to 7. Of course, these matches were not played with the *DScore* incentive and therefore imposing it retrospectively is hypothetical.

Table 3 : Number of teams in contention for the number one draft pick going into the final round of seasons 1997 to 2005

Season	Number of teams in contention for Pick 1, Round 22	Teams in contention for Pick 1, Round 19
2005	2	5
2004	2	3
2003	3	5
2002	2	5
2001	3	4
2000	1	5
1999	2	5
1998	1	6
1997	4	6

Aside from the 1998 and 2000 seasons, the race for the number one draft pick would have remained alive and well prior to the final round.

Importance

Of interest to us was when the maximum value of importance occurs for each team. We then sorted the teams by final ladder position (*FLP*) and calculated the mean and standard deviation of the round, as given in Figure 2.

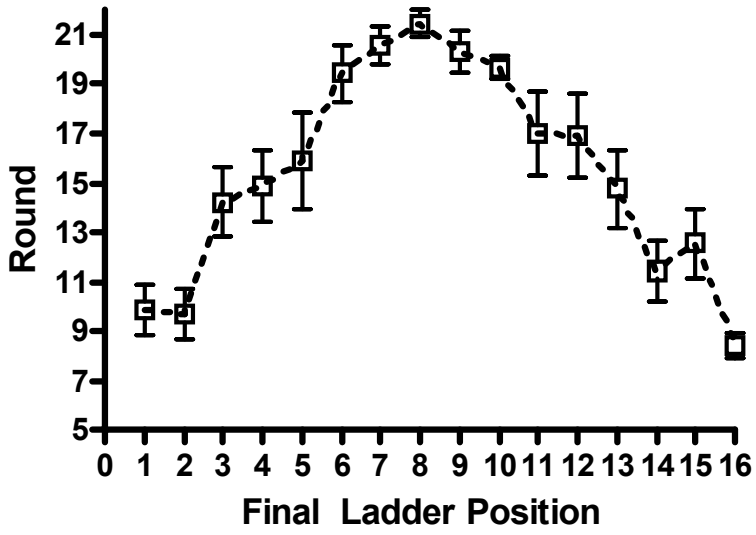


Figure 2: Error Bars of Maximum Round of Importance by Final Ladder Position (1997-05)

As shown in Figure 2, the teams finishing in the top 2 and bottom 3 have their most important games generally in the early rounds of the season (note that we have only considered round 6 onwards). All other teams heading towards the middle of the ladder have maximal important matches later in the season. As one would expect, the 8th *FLP* has the maximal importance match in the last three rounds.

DISCUSSION

It is somewhat difficult to measure the effect of our model on past results as we are implementing our method retrospectively. As a consequence, where players would end up under our model would be different to reality and therefore team success may change. Even so, the findings are still an eye-opener, and indeed motivate poorer teams toward success. As was shown in the results section, for the final round of 1998, the 1st and 2nd draft pick had been decided. However, 11 teams could still be playing in expectation of a change in their draft pick with a victory. The 'ideal' advocate of our system was the final round of 2003. Geelong played St Kilda, and it could be argued they were playing for 'nothing', sitting 10th and 13th on the ladder - no finals place or priority pick at stake. Under our *DScore* system the winner of that match would take 1st pick and the loser potentially 3rd. The match played on Saturday had Geelong prevail by 19 points, snatching 1st pick. Remarkably, the result was not yet settled, with the Sunday encounter between Hawthorn (9th) and Richmond (10th), (again two sides with nothing to play for), pivotal in the *DScore* outcome. Hawthorn won by 4 points, winning their fourth game in a row, snatching the number 1 pick on the last game of the home and away season!

Alternatives

A criticism that may be leveled at the *DScore* system is that teams which continually lose are never rewarded. A possible way of assisting teams that consistently lose may be to reward a 'gallant' defeat. Calculating an expected and actual margin, then smoothing the difference, is an approach used in other areas of sport analysis, such as tennis as in Bedford and Clarke (2000). They used their model to predict and improve upon ATP ratings in tennis based on margin of victory rather than win or loss. Once again, a team may play so poorly as to never get within the expected margin, and the same problem arises. We believe the priority criteria is a reasonable approach to combat this, and we can only hope teams would 'try harder' to win to obtain better draft picks, and in turn, enhance their future chances, rather than 'lie down' and be rewarded for defeat.

A point of interest raised in the methods section was the possible inclusion of a team's relative skill into the system, either using probabilities such as those pioneered by Stefani and Clarke (1992), or more arbitrary measures such as a win ratio to weight the *DPR*. The use of an 'opponent' weight would see some rather unattractive scenarios. Specifically, the use of a probability based multiplier on the *DPR* introduces only occasional need for lowly placed teams to win, as they need only defeat one successful team and reap a high *DPR*, thereby obtaining a high draft pick. This is a clear disincentive as the *DScore* system is designed to encourage teams to win every game possible.

CONCLUSION

In this paper, we have developed a unique system for player allocation in the AFL draft using probabilistic principles designed to encourage success. Whilst the AFL system was not designed to encourage teams to lose, it does reward teams that only win a small amount of games. Our model, known as the *DScore* model, uniquely encourages teams to strive for victory with a high draft pick as the prize, especially when the game (and their season) is - in terms of the finals - Unimportant. Utilizing this principle of unimportance, we cited exciting and motivating cases whereby otherwise 'meaningless'

encounters become a battle for high draft picks. The *DScore* model may also have a broad appeal, with potential outcomes easily publishable in daily newspapers and on the internet, with the relevant draft permutations providing a motivator not only for the club, but for the supporters alike.

REFERENCES

Bedford, A. and Clarke, S. (2000) A comparison of the ATP ratings with a smoothing method for match prediction. In: *Proceedings of the Fifth Australian conference on Mathematics and Computers in Sport*. Eds: Cohen G. and Langtry T. Sydney: University of Technology Sydney. 43-51.

Grier, K. and Tollison R. (1994) The rookie draft and competitive balance: The case of professional football. *Journal of Economic Behavior and Organization* **25**, 293-298.

Morris, C. (1977) The most important points in tennis. In: *Optimal strategies in sport*. Eds: Ladany, S.P and Machol, R.E. Amsterdam: North-Holland Publishing Company. 131-140.

Spurr, S. (2000) The baseball draft: A study of the ability to find talent. *Journal of Sports Economics* **1**, 66-85.

Stefani, R. and Clarke, S. (1992) Predictions and home advantage for Australian rules football. *Journal of Applied Statistics* **19**, 251-261.

Taylor, B.and Trogdon, J. (2002) Losing to win: Tournament incentives in the National Basketball Association. *Journal of Labor Economics* **20**, 23-41.

A BROWNIAN MOTION MODEL FOR THE PROGRESS OF AUSTRALIAN RULES FOOTBALL SCORES

Glasson, S.

Department of Mathematical and Geospatial Sciences, RMIT University, Melbourne, VIC, Australia

ABSTRACT

In financial markets, the price movements of shares are often modelled as following Brownian motion. The score difference in a sporting contest can often be seen to move in much the same way. This paper shows how Australian Rules football in particular satisfies the assumptions needed to implement a Brownian motion model. Such a model can be used to estimate a team's chance of winning a match that is already underway. It is shown how one can estimate the probability of a team winning at any time during the match using only the bookmaker's line, current score and time of the match remaining as the inputs. As an example, the 2005 AFL Grand Final is discussed illustrating the random walk and how the estimated probability of winning varies over time. Some applications to sports betting are presented, including actual results from two seasons of betting using probabilities generated from the model.

KEYWORDS

Australian Rules football, Brownian motion.

INTRODUCTION

During an Australian Rules football match, the chance of a team winning depends primarily on the score difference and the time remaining in the match. The purpose of this paper is to present a simple method for estimating the in-game chance of a team winning a football match.

Australian Rules football is the dominant spectator sport in the southern Australian states, and a typical Australian Football League (AFL) match draws over 30,000 spectators. A team score in two ways, either by scoring a goal-worth six points, or a behind-worth one point. In the 1480 matches from 1998 to the end of 2005, 28.0 goals and 23.8 behinds were scored on average per match with an average absolute winning margin of 34.4 points. For a more detailed description of the sport see Stefani and Clarke (1992).

The progress of the lead during a match can be charted as shown in Figure 1. The line moves up and down somewhat randomly with the vertical jumps corresponding to scoring events. Similar patterns have been observed in various phenomena, such as the movement

of the price of a stock, or the motion of a speck of dust across a room. This type of process is known as a random walk.

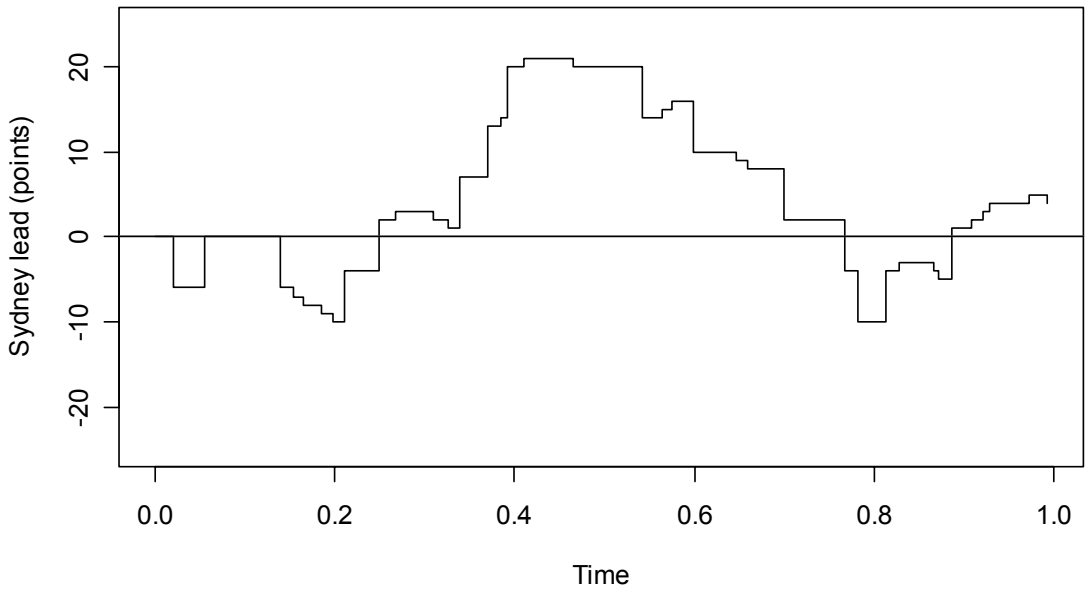


Figure 1: The score progress of and Australian Football League match, in this case, the 2005 grand final between Sydney and the West Coast Eagles

If the number of steps in the random walk is sufficiently large, guided by the central limit theorem, a random walk approaches Brownian motion. Brownian motion is a continuous process, and is easier to analyse than the more general random walk. It is described by two parameters; the variance, and drift. The ubiquitous Black-Scholes equation in finance (Black and Scholes, 1973) is perhaps the most famous application of Brownian motion for predictive purposes.

Stern (1994) first considered a Brownian motion model to provide in-game estimates of the chance of winning baseball and basketball games. Even for baseball, a relatively low scoring sport, the model provided promising results. This paper extends Stern's work to AFL matches, and also incorporates any prior beliefs of difference in skill between the teams. AFL is particularly suited to this type of continuous analysis because it is high scoring, particularly when compared to baseball.

METHODS

First, transform the time elapsed during the match, where $t \in (0,1)$ is the fraction of the match has been completed. Quarter breaks naturally occur at t=0.25, 0.5, and 0.75. Scores

are most commonly reported at these quarter breaks, although this model has the ability to make predictions continuously over the entire interval.

Let $X(t)$ represent the score difference between two teams at time t. The sign of $X(t)$ indicates who is leading the match. If $X(t)>0$, then Team 1 is leading. Team 2 is winning if $X(t)<0$. Finally, if $X(t)=0$, the scores are level. Let's now assume that $X(t)$ can be modelled as a Brownian motion process with drift, μ, and variance, σ^2 per unit time. μ can be thought of as the difference of skill between the teams measured in points over an entire match. If, for example, our estimate of μ is 10 points, then on average the home team will win by 10 points. The variance, σ^2, is a measure of the uncertainty, or "luck", involved in the football game.

Brownian motion is closely related to the Normal distribution, hence $X(t)$ can be described as

$$X(t) \sim N(\mu t, \sigma^2 t) \tag{1}$$

Therefore, at the start of the match, the estimated probability of Team 1 winning the match, given skill difference, μ and standard deviation, σ is

$$\Pr(X(1) > 0) = \Phi\left(\frac{\mu}{\sigma}\right), \tag{2}$$

where Φ is the CDF of the standard normal distribution. This expression is closely related to the ordered probit model proposed by Brailsford et al. (1995) for analysing Footybet scores (where gamblers try to pick which point margin the scores will fall).

Once the match is underway, we can now use the Brownian motion model to estimate Team 1's chance of winning. Here,

$$\Pr(X(1) > 0) = \Phi\left(\frac{l + (1-t)\mu}{\sigma\sqrt{(1-t)}}\right) \tag{3}$$

where l is the current lead in points of Team 1. The probability of Team 2 winning is simply the complement of (3).

For reasons of simplicity, draws have been ignored so far, but can be incorporated into the model as follows. Here the continuity correction proposed by Stern (1994) is used, where it is assumed that the observed difference in scores is $X(t)$ rounded to the nearest integer. Therefore the estimate of the probability of a draw is given by

$$\Pr(-0.5 < X(1) < 0.5) = \Phi\left(\frac{l + 0.5 + (1-t)\mu}{\sigma\sqrt{(1-t)}}\right) - \Phi\left(\frac{l - 0.5 + (1-t)\mu}{\sigma\sqrt{(1-t)}}\right). \tag{4}$$

Bookmakers offer other bets based on the final margin of the match. For example, the probability that Team 2 wins by more than 39.5 points can be calculated by

$$\Pr(X(1) < -39.5) = 1 - \Phi\left(\frac{39.5 + \mu}{\sigma}\right). \quad (5)$$

Similar calculations can be made to estimate any other offered point margin bets.

RESULTS

Prior to using the model, we need to estimate the parameters μ and σ for each match. Stern (1994) uses a global home ground advantage for all teams as an estimate of μ. A more accurate estimate can found be by using a ratings algorithm such as that described by Stefani and Clarke (1992). Another method, the one used here, is to use the lines provided by bookmakers for each match as an estimate of μ.

The standard deviation, σ, can also be estimated by utilising information provided by bookmakers. Bookmakers express the chance of a team winning in two ways, odds and lines. The line gives an estimate of μ, and the odds provide $\Pr(X(1)>0)$. By rearranging (3), one can estimate σ for a given line and odds, (provided that $\mu \neq 0$). Generally for AFL matches, $\hat{\sigma} = 38$ provides an adequate fit between the bookmakers' lines and odds. It is possible that σ may vary between matches due to weather conditions, but for the purposes of this paper it is assumed to be constant since varying σ by a few points either way does not appear make large differences in predictions.

The Brownian motion model requires some assumptions for it to be effective. To check these assumptions, first define the error for the bookie's line for the entire match as

$$\varepsilon = X(1) - \mu. \quad (6)$$

Likewise, for the first quarter, the error is

$$\varepsilon_{q1} = X(0.25) - \frac{\mu}{4}, \quad (7)$$

and the second quarter

$$\varepsilon_{q2} = X(0.5) - X(0.25) - \frac{\mu}{4}. \quad (8)$$

ε_{q3} and ε_{q4} are calculated similarly.

For our model to be valid, the errors should not be significantly correlated, their standard deviations should be equal, and they should be normally distributed. These assumptions ensure that μ and σ remain stationary over the course of the match, and that the normality assumption of the Brownian motion model holds. For testing purposes, a database of 1480 matches from 1998 through to the end of the 2005 season was used. Recorded data for each match includes the bookies odds and line prior to the start of the match, the final score, and the score at each quarter break.

Figure 2 shows a histogram and density plot of ε for all 1480 matches. The striking part of this chart is the almost perfectly bell-shaped curve of the distribution. A Shapiro-Wilk test

confirms the validity of the normality assumption (n=1480, W = 0.9988, p = 0.4167). Despite the discrete nature of the scoring (1 and 6 point increments), the many scoring shots within a match ensure that ε approaches normality. Likewise, for each individual quarter, the errors about ε/4 appear normal (None are significantly non-normal at the 0.05 level).

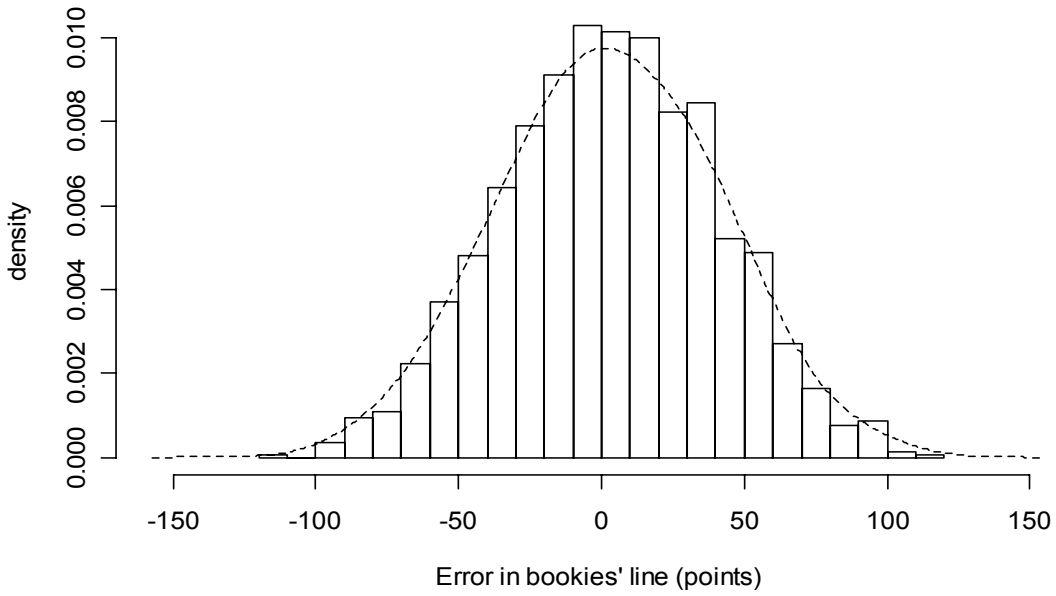


Figure 2: Histogram and density plot of the error in the bookies' line, (ε). The bell-shaped curve highlights the normality of the errors.

Table 1 contains the mean and standard deviations for each of the ε's. If ε is unbiased then the errors for each quarter should be equal to zero. Note that they are all positive here. This is not an artefact of our model but indicates there is some bias in the bookies' line. The most likely reason for the bias is because the bookmakers underestimate the size of the home ground advantage.

Table 1: Mean and standard deviation for the errors around the bookies line for all quarters.

Variable	Mean (std err)	Standard deviation
ε_{q1}	0.97 (0.44)	16.75
ε_{q2}	0.45 (0.45)	17.20
ε_{q3}	1.07 (0.46)	17.81
ε_{q4}	0.48 (0.46)	17.66
ε	2.97 (0.98)	37.50

Another key assumption of the model is that variance remains stationary over the course of the match, $\left(\text{that is, } \text{Var}\left(\varepsilon_{q1}\right) = \text{Var}\left(\varepsilon_{q2}\right) = \text{Var}\left(\varepsilon_{q3}\right) = \text{Var}\left(\varepsilon_{q4}\right)\right)$. We can test whether at least

one pair of variances is different through Bartlett's test for homogeneity of variance. For these data, Bartlett's test (Bartlett's K^2 = 6.8775, df = 3, p = 0.0759) shows that assuming equal variances is not unreasonable here.

Commentators often describe one end as the "scoring end", usually due to a strong wind blowing in that direction. The strength of this effect has never been quantified. If this was a general effect there should be positive correlation between ε_{q1}, and ε_{q3}, and also between ε_{q2}, and ε_{q4}. The Brownian motion model assumes there are no obvious correlations between any of $\varepsilon_{q1}, \varepsilon_{q2}, \varepsilon_{q3}$, and ε_{q4}. Table 2 presents the correlation matrix with associated p-values. The correlations are reasonably small. At the 1% level of significance, only the correlation between the third and fourth quarters is significant (positively). One possible explanation for this correlation is that teams give up when they are behind by a large margin. There is also some evidence of correlation between other quarters, although more detailed research needs to be undertaken to conclude whether the scoring end effect exists. Although the correlation between quarters does give some doubt as to the adequacy of the Brownian motion model, it appears the model still fits well when used as a predictive model.

Table 2: Correlation matrix of the errors for each quarter. P-values are in brackets. Under the Brownian motion model, the correlations are assumed to be negligible.

	ε_{q1}		ε_{q2}		ε_{q3}		ε_{q4}
ε_{q1}	1						
ε_{q2}	-0.0056	(0.829)	1				
ε_{q3}	0.0591	(0.023)	0.0503	(0.053)	1		
ε_{q4}	0.0456	(0.080)	0.0669	(0.010)	0.1122	(<0.001)	1

Finally, to assess the adequacy of the model, we can compare the standard deviation of ε, with the value of σ used to fit the model. From Table 1, sd (ε) = 37.5 compared to σ=38. The closeness of these values indicates the model is representative of the data.

DISCUSSION

Having taken the steps to fit a predictive model in the previous section, implementing the model to produce meaningful predictions comes is the next step. Take, for instance, a hypothetical match where Team 1 are 20 point favourites (μ=20). Figure 3 summarises predictions for Pr(X(1)) as a set of contour lines, where each contour line represents a particular lead. So if scores are level we use the lead=0 contour line. At the start of the match Team1 has approximately 70% chance of winning, but if they are still level at half time the chance of winning is only 65%. As each team scores, we jump to a different

contour line corresponding to the lead at that time. Note that as *t* approaches 1, even a small lead means victory becomes assured.

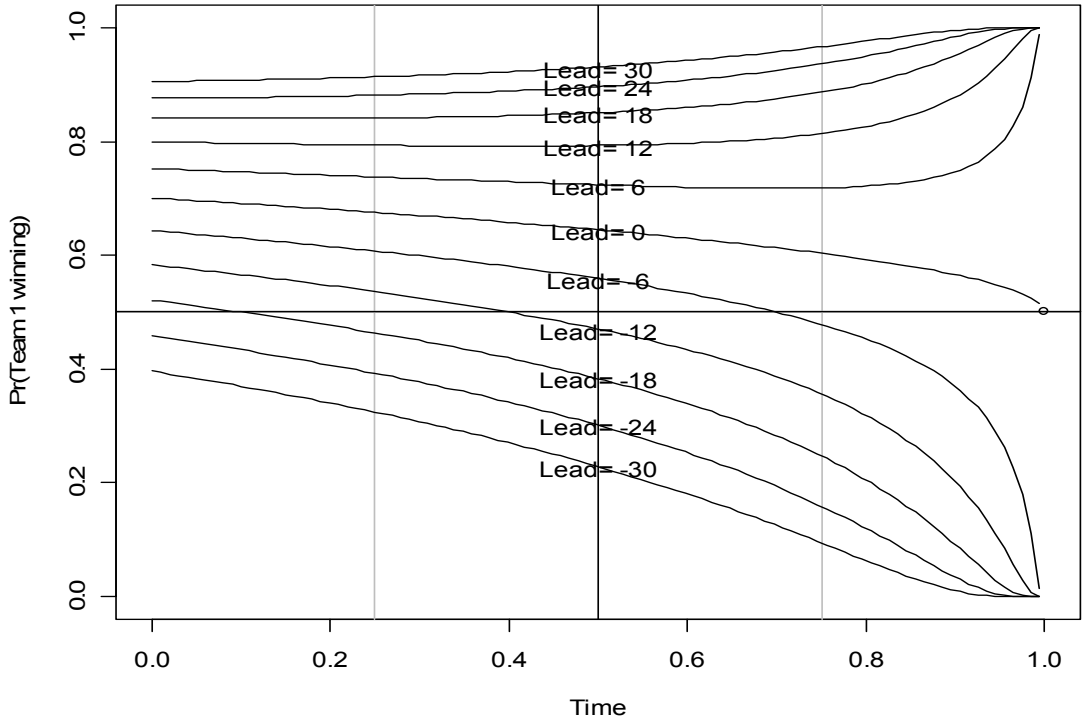


Figure 3: A plot to estimate a team's chance of winning at any time during the match. In this example $\mu = 20$ points. Each contour line represents a particular lead.

Figure 4 shows how the model can be applied in real-time to an actual match, in this case the 2005 grand final between Sydney and West Coast. According to the bookies this was an even contest so we take $\mu = 0$. The two charts illustrate a number of features of the Brownian motion model. Early in the match, each score causes a small difference in the estimated outcome, but as time progresses each score becomes more and more important, especially as the match is close. A simple program or Excel spreadsheet can be made to provide these estimates while the match is in progress.

One downfall of this model is that for close matches such as this grand final, the model doesn't take into account the discrete nature of scoring late in the match. For the last few minutes of the match the model predicts Sydney to be almost certain winners even though most observers of the match would conclude that the match was still in the balance. As it turns out, Sydney did in fact win the match, despite several last ditch attacking moves by West Coast.

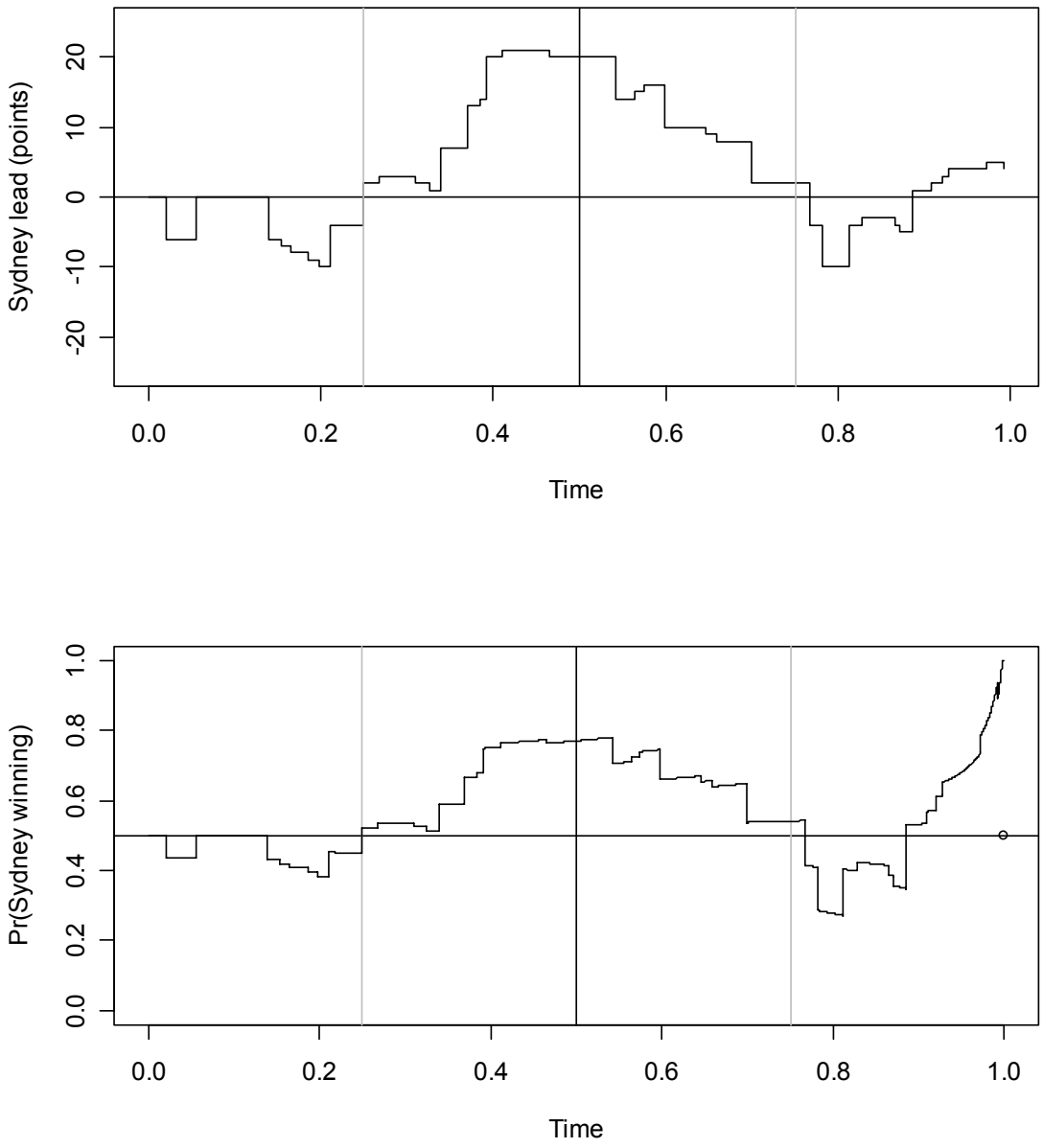


Figure 4: The 2005 grand final revisited. The top panel shows the progress of the score, whereas the bottom panel provides estimates of Sydney winning the match as time progresses $(\mu = 0, \sigma = 38)$.

Predicting the outcome of the match in progress is of more use than just to satisfy curiosity. Many Australian bookmakers offer "in-the-run" betting, where they offer odds of a team winning while the match is in progress. If a model, such as the one outlined in this paper, can make more accurate predictions than the bookies, an astute bettor can profit by betting

on teams which the model estimates have a greater chance of winning than is reflected in the odds.

Figure 5 shows the bankroll growth for 206 actual bets made over the course of the 1999 and 2000 AFL seasons at a major Australian bookmaker. The average return on turnover was 16% at average odds of $2.35. These bets are part of a larger sample of bets made at other bookmakers, with similar results to those reported here. Estimates were made using the Brownian motion model and bets were sized roughly according to the Kelly criterion (see Kelly (1956) and Epstein(1977)) to maximise the long-run growth rate.

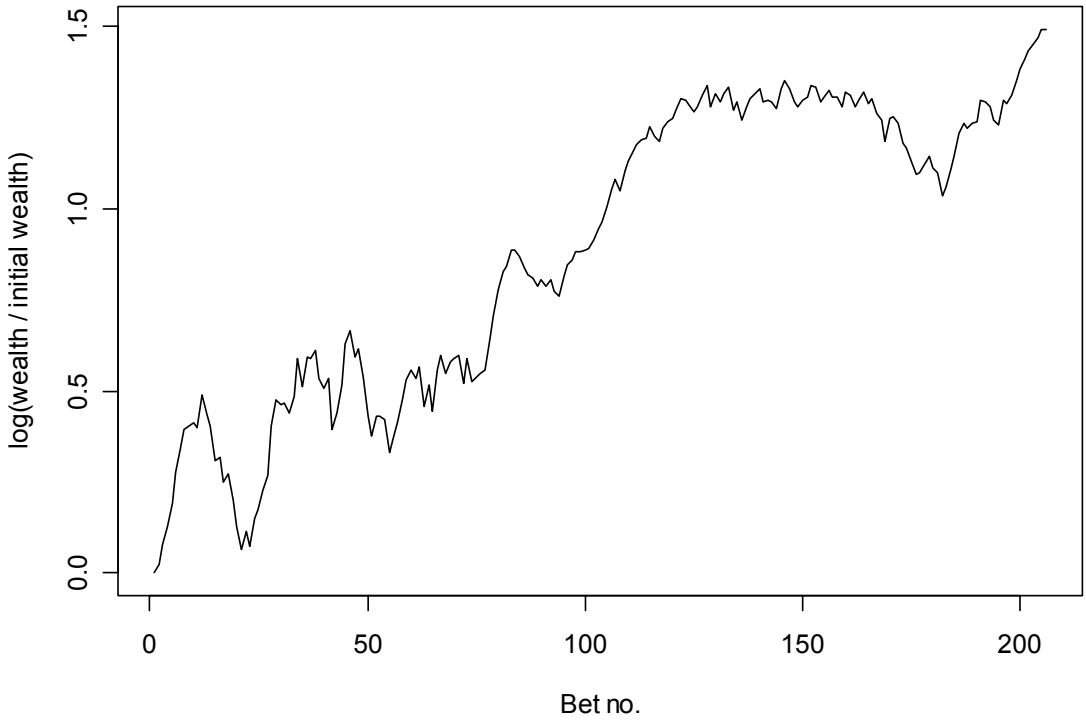


Figure 5: Bankroll growth from betting "in-the-run" at a major Australian bookmaker over the 1999 and 2000 seasons using the Brownian motion model.

The vast majority of advantageous in-the-run bets were on the pre-game underdogs leading during the match. It appears that the general public and the bookies who set the odds overestimate the chances of pre-game favourites winning when they are trailing during the match. The success of a mathematical model in this situation can be partly attributed to the complexity of analysing the in-the-run betting markets. Benter (1994) notes that advantages are more likely to be gained by mathematical models in more complicated betting products. The complexity in setting odds for in-the-run markets is high and the bookmaker has severe time constraints when choosing the odds during a match with a rapidly changing score. A

model, such as the one described in this paper, takes away much of the need for human judgement.

CONCLUSIONS

Sports commentators often make comments about how likely a team is to win while the match is in progress. This paper introduces the Brownian motion model as a simple way of quantifying these predictions. What was once a "miracle comeback" could now be a "comeback from having less than 5% chance of winning at half time". The Brownian motion model is not necessarily the most accurate way to make predictions during a game, but it outperforms estimates given by bookmakers. One other possible method for predicting in-game probabilities is to apply a probit regression to each quarter using relevant predictors. While this method may end up being more powerful, it lacks the elegance of the Brownian motion model and requires more information to fit.

REFERENCES

Benter, W. (1994) Computer based horse racing handicapping and wagering systems: a report. In: *Efficiency of Racetrack Betting Markets*. Eds: Hausch, D., Lo, S. and Ziemba, W. San Diego: Academic Press. 183-198

Black, F., and Scholes, M. (1973) The pricing of options and corporate liabilities. *Journal of the Political Economy* **81**, 637-654.

Brailsford T. J., Easton, S. A., Gray, P. K., Gray, S.F. (1995) The efficiency of Australian football betting markets. *Australian Journal of Management* **20**, 167-195.

Epstein, R. A. (1977) *The Theory of Gambling and Statistical Logic*. New York: Academic Press.

Kelly Jr, J. (1956) A new interpretation of information rate. *Bell System Technical Journal* **35**, 917-926

Stefani, R. T. and Clarke, S. R. (1992) Predictions and home advantage for Australian Rules Football. *The Journal of Applied Statistics* **19**(2), 251-261.

Stern, H. (1994) A Brownian motion model for the progress of sports scores. *Journal of the American Statistical Association* **89**, 1128-1134.

POSSESSION VERSUS POSITION: STRATEGIC EVALUATION IN AFL

O'Shaughnessy, D.M.

Champion Data, Melbourne, Australia

ABSTRACT

In sports like Australian Rules football and soccer, teams must battle to achieve possession of the ball in sufficient space to make optimal use of it. Ultimately the teams need to score, and to do that the ball must be brought into the area in front of goal – the place where the defence usually concentrates on shutting down space and opportunity time. Coaches would like to quantify the trade-offs between contested play in good positions and uncontested play in less promising positions, in order to inform their decision-making about where to put their players, and when to gamble on sending the ball to a contest rather than simply maintain possession.

To evaluate football strategies, Champion Data has collected the on-ground locations of all 350,000 possessions and stoppages in the past two seasons of AFL (2004, 2005). By following each chain of play through to the next score, we can now reliably estimate the scoreboard "equity" of possessing the ball at any location, and measure the effect of having sufficient time to dispose of it effectively. As expected, winning the ball under physical pressure (through a "hard ball get") is far more difficult to convert into a score than winning it via a mark. We also analyse some equity gradients to show how getting the ball 20 metres closer to goal is much more important in certain areas of the ground than in others. We conclude by looking at the choices faced by players in possession wanting to maximise their likelihood of success.

KEY WORDS

notational analysis, Australian Rules football, tactical coaching

INTRODUCTION

Australian Rules Football (informally known as "AFL" after the Australian Football League) is played with an oval ball on an oval field at high speed, leading to it sometimes being called "What Rules?" by the unschooled observer. Compared to more structured football codes such as American football or rugby league where a "phase of play" always starts in a simply-defined formation, the free-flowing nature of Australian football creates extra dimensions for analysis. This paper describes the qualitative framework for evaluating the phases of AFL and presents empirical interpretation of data from the 2004 and 2005 seasons.

AFL coaches are clamouring for this sort of analysis to inform their strategies and training procedures. They know that being in possession of the ball is important, but this research can show exactly how much it's worth on the scoreboard to take a contested mark, compared with someone from the opposition grabbing the loose ball spilled from the pack. They also know that position is important. They must create opportunities in

positions near goal, but their players often have to choose whether to aim at a riskier proposition close to the goalmouth or maintain possession in a worse position. Dynamic programming based on empirically derived parameters can answer this dilemma.

Dynamic programming was first applied to AFL (Clarke and Norman, 1998) to answer the question of whether players should concede a point on the scoreboard in order to gain clean possession afterwards. A new thesis (Forbes, 2006) based on Champion Data's statistics uses a Markov model approach to map out the probabilities of transitions between AFL's phases to predict scoreboard results.

American football, where position is effectively one-dimensional and there are only four phases – the "downs" – has been analysed using dynamic programming in a famous paper (Romer, 2002), and a rating system (Schatz, 2005) called DVOA (Defence-adjusted Value Over Average) evaluates actions with respect to a model of scoreboard value similar to the one created in this paper. The fast-flowing and open sport of ice hockey has recently been modelled using a "semi-Markov" approach (Thomas, 2006).

The modelling undertaken here is largely exploratory – this is a mass of new data which requires further detailed research.

METHODS

Match Equity and Field Equity

Various authors have employed a plethora of terms to describe the expected value of actions on sporting fields. Studeman (2004) describes the repeated reïnvention and relabelling of "Win Probability Added" in baseball. Bennett (2005) has a good simple description of how to value an action that alters the probability of winning the match.

The terminology we use in this paper is derived from the theory of backgammon (Keith, 1996), a game in which the players compete to win points, the first to n points winning the match. We assume teams of equal strength, although much of the reasoning below is still valid for uneven teams. ***Match Equity*** is the probability of the team to win the match from this moment, or more specifically:

$$E_M(m,t,\boldsymbol{x},\varphi) = p_{win} + \frac{1}{2} p_{draw} \qquad (1)$$
$$0 \le E_M \le 1$$

The Match Equities of each team in the contest sum to one. A team is always aiming to increase its Match Equity until it reaches one – certain victory. I.e., it is looking for actions which maximise ΔE_M, or at the very least have $\Delta E_M \geq 0$. As noted in Equation 1, Match Equity is a function of four parameters:

- the score margin, m
- the time remaining in the match, t
- the position on the field, $\boldsymbol{x}$
- the possession state or phase of play, φ

AFL typically has about 50 scores in a match of 80 live minutes. We define s_{typ} as the typical score of a game (in AFL's case, the goal worth 6 points is dominant), and t_{typ} as

the typical time between scores (approximately 100 seconds in AFL). We can roughly decouple the first two parameters from the others by noting that if we discard any knowledge of $\boldsymbol{x}$ or φ, we can build a satisfactory model of winning probability based only on the time remaining t and changes to the margin m. The phase and location information can be treated as a perturbation of the match-winning probability model E_M.

To model the net potential value on the scoreboard of the current state of play, we introduce ***Field Equity***:

$$E_F(\boldsymbol{x},\varphi) = \sum_i (p_{i,team} s_i - p_{i,opp} s_i)$$
$$-\max(s_i) \le E_F \le \max(s_i) \tag{2}$$

where

- s_i is the value of the i^{th} type of score
- $p_{i,q}$ is the probability of the next score being of type i by team q

The Field Equities of each team in the contest always sum to zero. The Field Equity fluctuates as play progresses until either team scores, at which team it precipitates an actual change to the margin m and E_F is reset to zero. AFL has two different restart phases, one being a centre bounce after a goal (where obviously each team has equal chances and $E_F = 0$), the other being a kick-in from the goalmouth after a behind. Remarkably, empirical evidence suggests that the average team has zero residual equity in the behind restart phase (see Table 1 in RESULTS).

Changes to Match Equity, Decoupled

$$\Delta E_M \approx \Pi(m,t) \cdot \Delta E_F(\boldsymbol{x},\varphi)$$
$$\Pi = \frac{\partial E_M}{\partial m} \tag{3}$$

The "Pressure Factor" multiplier Π is the impact an instantaneous change to the margin would have on the match-winning chances of the teams. Empirically, kicking the first goal in an evenly-matched contest increases E_M from 0.50 to about 0.56. The decoupling transfers the potential held in the field position into improved match-winning probability. It allows us to assume that a team that increases E_F to +2 soon after the start of a game increases its match-winning probability to about 0.52, but if only a quarter of the match is left and $m = 0$, ΔE_F of +2 could imply ΔE_M of +0.04, from 0.50 to 0.54. A detailed formula for Π is beyond the scope of this paper. Henceforth the term "equity" (E) will refer to Field Equity and we will assume the time remaining is effectively unlimited.

The decoupling assumption only breaks down when both t and m are of the order of t_{typ} and s_{typ} respectively – i.e., when the game goes down to the wire, the added quantum of a major score could be the difference between a win ($E_M = 1$) and a loss ($E_M = 0$), and the time left on the clock must be considered.

Data Collection

Champion Data has been logging qualitative AFL statistics by computer since 1996. All statistics are classified live by a caller at the venue, connected by phone to a reviewer watching a monitor, and a data entry operator. Traditionally, AFL statisticians had only

captured the numbers of kicks, marks, handballs, and scores for each player. The system introduced in 1996 imposes a structure on the flow of play, so that every disposal or use of the ball must be preceded by a "possession".

We need to be able to say which player is in possession, in which circumstances he got the ball, where he was on the field, how much time he had to think once he got it, a rough idea of what his options were, which option he chose, and whether he successfully executed his choice. Each of these events has to be put in context, with respect to what happened before and after the ball was in his control. The data capture software executes a model of the sport, which only allows certain events to take place in certain circumstances. Every statistic is time-coded, and since 2004 all possessions are given a position on the field by an independent operator whose sole responsibility is to pinpoint the location of the ball on a map of the field for each of these 1000 data points per match.

Testing has shown that the quantity of statistics for each player is logged at better than 99% accuracy, time is accurate to within about five seconds, and position to within approximately 5-10 metres.

AFL Phases of Play

Possession of the football has been qualitatively stratified to become the descriptive framework of AFL's Phases. Phases of Play with a team in possession include:

- Mark. The player has caught the ball from a kick and according to the rules is entitled to consider his options without being tackled.
- Handball Receive. The player has received a handball from a teammate, uncontested.
- Loose Ball Get. The ball has indiscriminately spilled loose and a player has been in the right place to pick it up.
- Hard Ball Get. The player has taken usable possession of the football while under direct physical pressure from an opponent.

Play can also be in an active neutral phase, after a smother of the ball or a similar random collision. There are also passive neutral phases where the umpire holds the ball, before launching it back into play. Lastly there are a couple of set-play phases such as a kick-in after a behind.

For the purposes of this paper we will consider five Phases of Play, which experience and analysis have shown cover most important facets of AFL:

- "Set" (approximately 35% of possession is granted this way). A player has taken a mark or received a free kick, or has been given another set-play role. He has an optimal amount of time to consider options and make the right choice. We will ignore kick-ins from goal in this paper.
- "Directed" (approx 38%). The ball was directed into the player's possession by a teammate, either via a handball, a kick to the player's advantage without achieving a mark, or a knock-on or hit-out intended for the player. Generally the player has space to run onto the ball and some time to make a good decision.
- "Loose" (approx 17%) - Player won a virtually random ball via a loose ball get, and while he is not yet under physical pressure there is little time to evaluate the situation.

- "Hard" (approx 10%) - Player won the ball under direct physical pressure and often must take the quickest option available to avoid being caught with the ball.
- "Umpire" - Umpire has the ball and restarts play with equal chances for both teams.

We have ignored quasi-possession states like knock-on, hit-out and kick off the ground for this paper. A full description of Phase of Play would also include extra dimensions such as: how fast the ball travelled to where it is (catching the defence napping, for instance); who is currently on the field (is it the best 18 players available?); what formation the team is playing (flooding the backline to reduce the odds of uncontested ball near the opposition's goal).

Assumptions

AFL is regularly played at a dozen different venues, each with slight variations from the ideal oval shape and various lengths and widths. The shortest ground is the SCG at 148.5 metres, meaning that the 50m-wide centre square touches the 50m arcs at each end of the ground. At Subiaco in Perth, on the other hand, there are 175.6 metres between the goal-lines and therefore 12.8 metres of territory between the top of the arc and the centre-square. When plotting locations, it is important to note that some areas of the ground simply don't exist at some grounds, and that the wings are much wider at the SCG (length:width ratio of 1.09:1) than Geelong (1.47:1).

The positional capture software assumes that every ground is a perfect ellipse, and only the lengths of the axes vary, so the operators can accurately pinpoint play. For analysis, we use the MCG (160 × 138 metres) as the standard ground and transform the other venues into this shape to utilise their data. This transformation preserves fixed areas of the ground such as the centre-square, boundary and the corridor leading to goal, while distorting distances and angles in other regions. We will always show teams attacking the goal to the right of the page. Contour maps have been generated using ComponentOne Chart3D v8. Other diagrams have been designed by the author.

An implicit assumption in the equity model is that the expected value of the next score is a good measure of the current phase of play, no matter how many minutes in the future that score may be. This has advantages over a Markov Model in that we do not assume that future states are exactly classifiable, instead there may be subtle repercussions of actions which are evident further down the track and should not be washed away by repeated normalising. Coaches value the players who can see three or more moves ahead, and don't just look for an easy option in front of them. The disadvantage of the equity approach is that the further we go from the source phase, the less relevance it has to the developing play, as more randomness floods in. Standard error measurements are quite high because of the number of data points ignored.

Method of Calculating Estimated Equity

For each data point, the value of the next score has been noted. This could be +6 (a goal for this team), +1 (a behind for this team), -1 (a behind for the opposition), or -6 (a goal for the opposition). Data points are excluded from analysis if there is no further scoring in the quarter. An example appears in Tables 1 and 2 at the start of the RESULTS section below. It has been assumed that left/right and north/south biases are inconsequential, so the standard ground has been folded down the spine and data points from each half are analysed together.

We have used two different positional filters in this paper. The contour graphs are generated using a six metre square grid. All points within a six metre radius of the vertex are taken into account in the calculation, meaning that each point appears in roughly three map points – this is an attempt at smoothing, knowing the natural sampling error in the data. Parts of the map with insufficient data (fewer than ten points in the disc or an equity standard error of greater than 0.5) are shown blank. Where we want to measure true statistical deviations and start to develop a model, the zones must not overlap. The semi-ellipse (remembering that the ground has been folded along its spine) is divided into 200 zones of equal area. First the length-wise (X) axis is divided into 25 sections to segment the ellipse into 25 equal areas. Then seven curves are drawn equidistant from each other, between the spine and the boundary to cut each strip into eight zones.

Error figures presented are two standard errors (95% confidence) except where noted.

RESULTS

Table 1 has a simple example of how to estimate the scoreboard value of two well-defined phases: after a goal, and after a behind. This is Summarised in Table 2.

Table 1: How to Calculate Estimated Equity

	Team Scored a Goal, Centre Bounce follows	Team Scored a Behind, Opposition will Kick-In
Event Count	31236	26160
Discard (no further score)	2376	2064
Team Goal (+6)	8340	7045
Team Behind (+1)	6823	5803
Opposition Goal (-6)	7435	6065
Opposition Behind (-1)	6262	5183
Sum of Next Scores	5991	6500
Sum of Squares	580985	482946
Equity Mean Estimate	0.208	0.270
Standard Error in Mean	0.026	0.029

Table 2: Residual Equity in Restart Phase as measured in seasons 2000-2005

After scoring a goal	Measured E_F	+0.21(5)
After scoring a behind	Measured E_F	+0.27(6)

The measured residual equity in real matches is non-zero since the stronger team is more likely to be the scorer of both the previous and the next score. It appears irrelevant

whether the restart is via the umpire in the centre of the oval, or via a player kick-in at the end of the ground (95% confidence interval for the difference: [-0.02, +0.14]).

Equity Maps

The value of taking a mark and having a set shot at goal directly in front can be seen in this map, with an expected value of more than four points extending all the way out to about 40 metres from goal. A free kick within 25 metres makes the goal a virtual certainty. The tight bunching of contour lines from 40 to 60 metres out along the spine shows the natural limit of an AFL footballer's kick, being about 50-55 metres. To get within one kick of goal, and have the time to execute it, is extremely valuable.

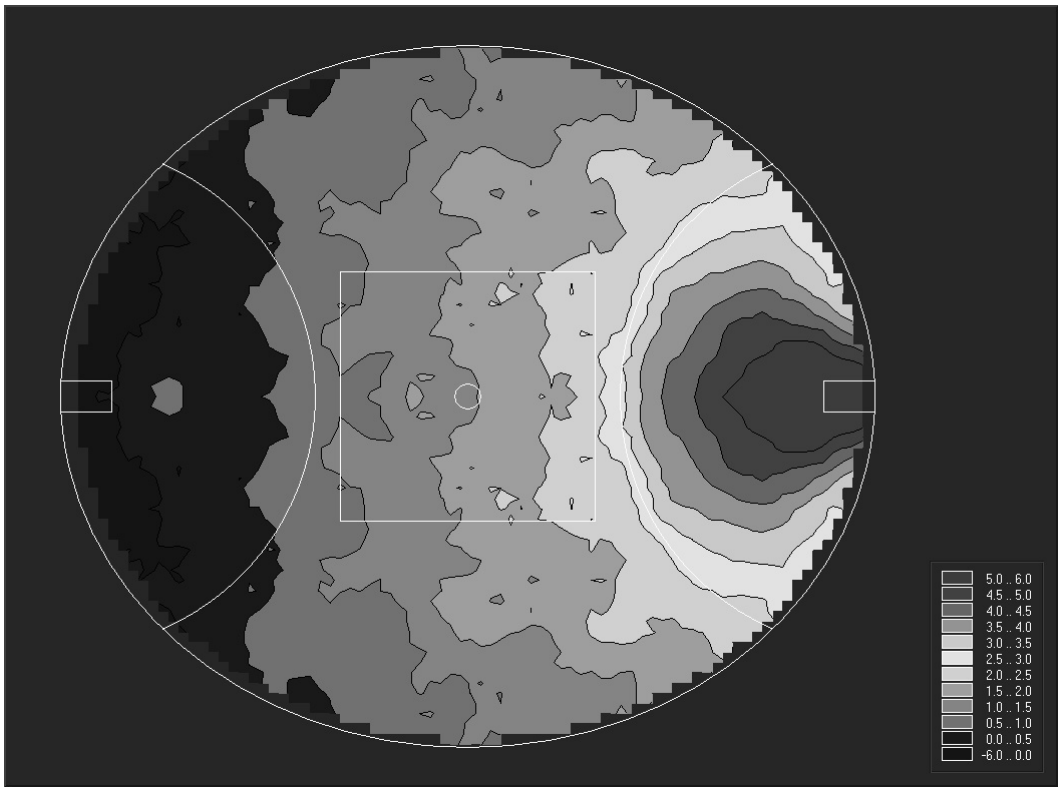


Figure 1: "Set" Phase Contour Map

"Directed" (Figure 2) is the second-best phase for a footballer to receive the ball in. Usually he has received a handball in some space and should be able to execute his preferred option. But often he will have to take critical time to swivel as the defence closes in, and it's only within ten metres of goal that the maximum six points can almost be assumed. The gradient we saw at 40-60 metres in Figure 1 is completely missing here, showing the greater difficulty of a snap shot on the run – the attacker wants to be within 30 metres.

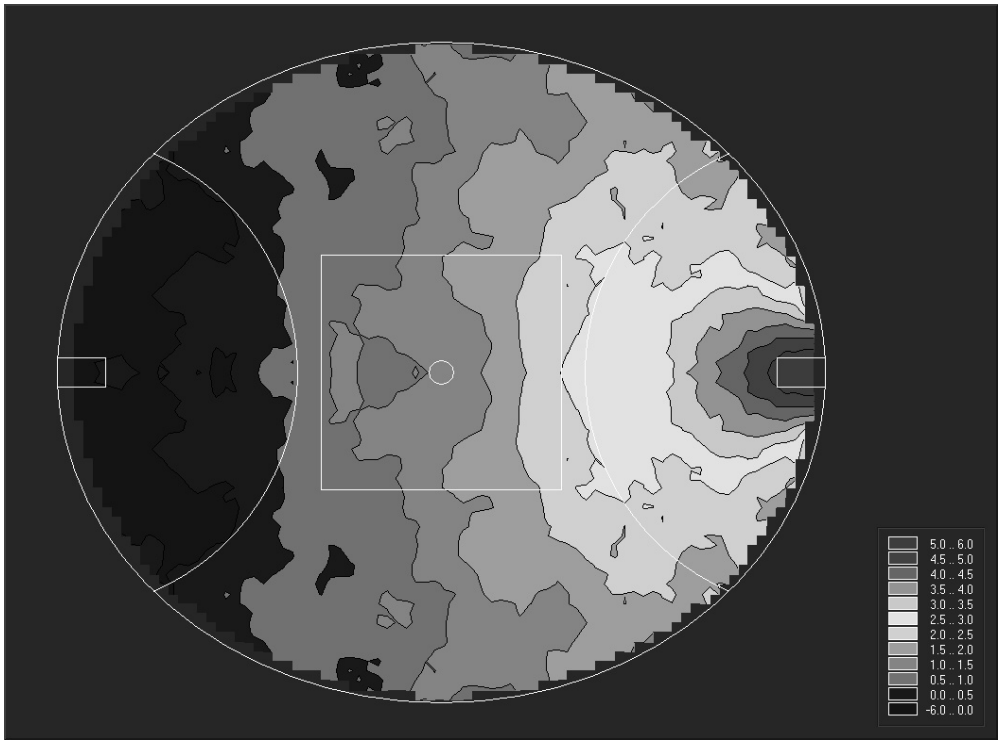


Figure 2: “Directed” Phase Contour Map

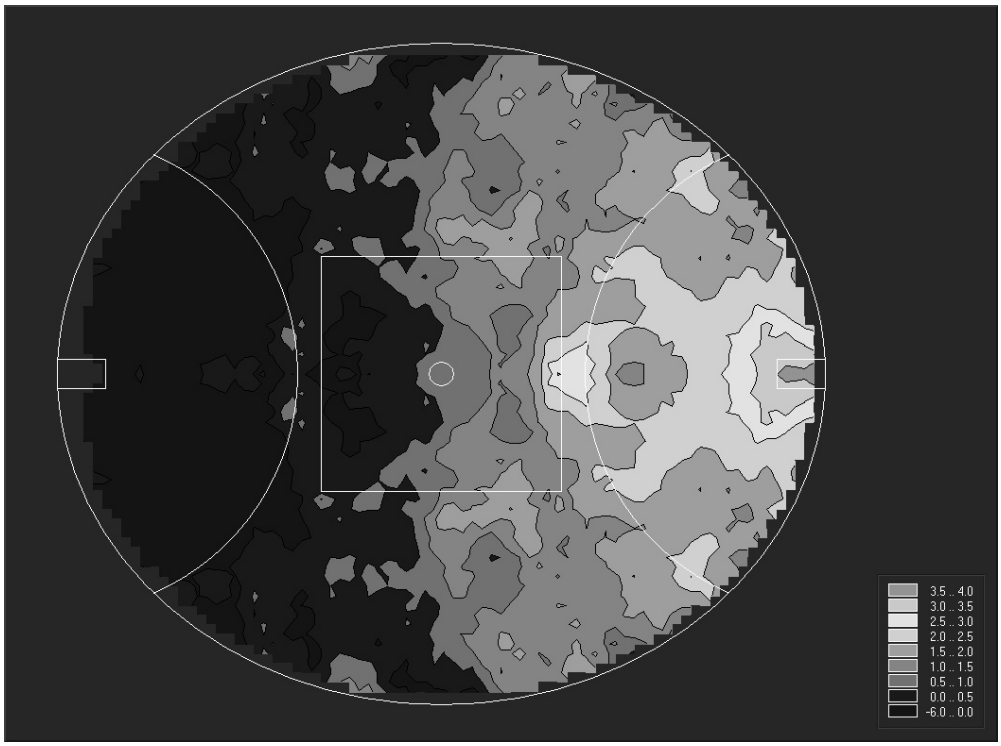


Figure 3: “Hard” Phase Contour Map

An utterly different picture (Figure 3) awaits the player who faces the extreme pressure of a hard ball get. Even within ten metres of goal the expected scoreboard outcome is just 3.5 points. Equity is below zero for the entire defensive zone, but interestingly there is a peak at the top of the forward arc, indicating that perhaps this is one place on the ground where he has two reasonable areas either side of him to shoot out a handball and

find a teammate who suddenly has options within range of goal. This circumstance often happens after the centre bounce when a quick kick lands at the congested top of the arc with the opposition still rushing the centre square.

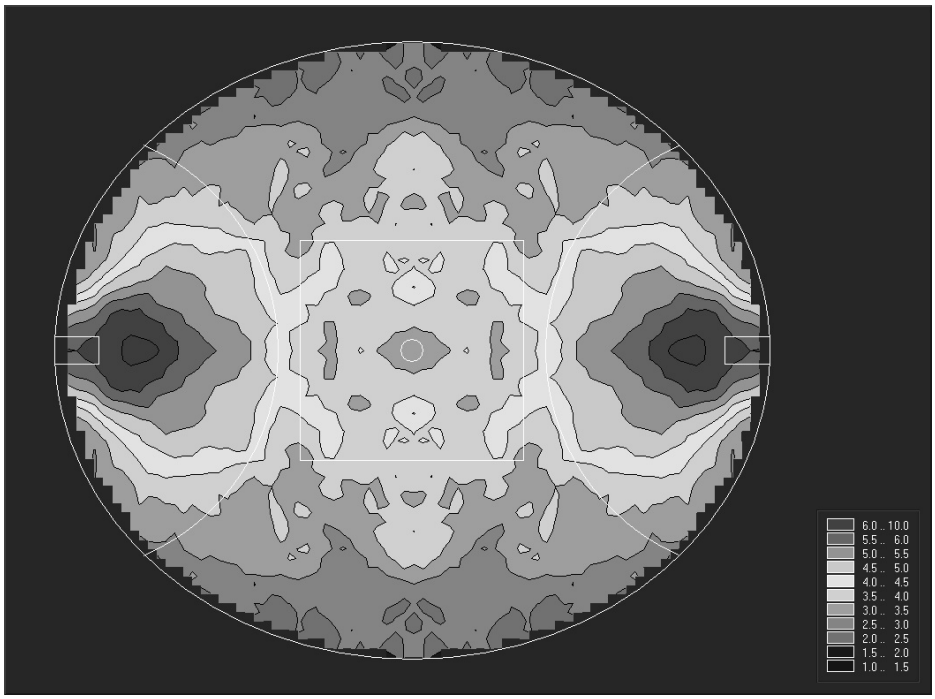


Figure 4: Advantage of "Set" over Opposition "Set"

This shows clearly the "hot spot" favoured by AFL coaches. There is volatility of more than six points in contesting a mark or winning a free kick twenty metres out directly in front of goal, rather than letting your opponent have the same.

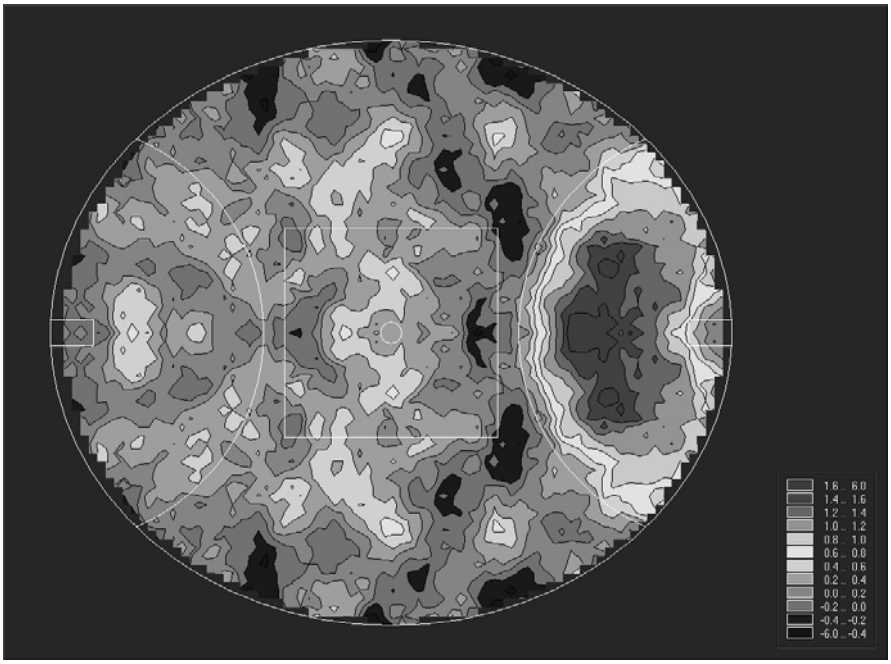


Figure 5: Advantage of "Set" over "Directed"

Also of interest is how cool the wings are – an equity swing of less than 2.5 points for taking a mark over his opponent, as neither player can directly make use of the extra time. A kick from a set shot near the boundary will often travel straight down the boundary to a settled pack, which is very low in volatility.

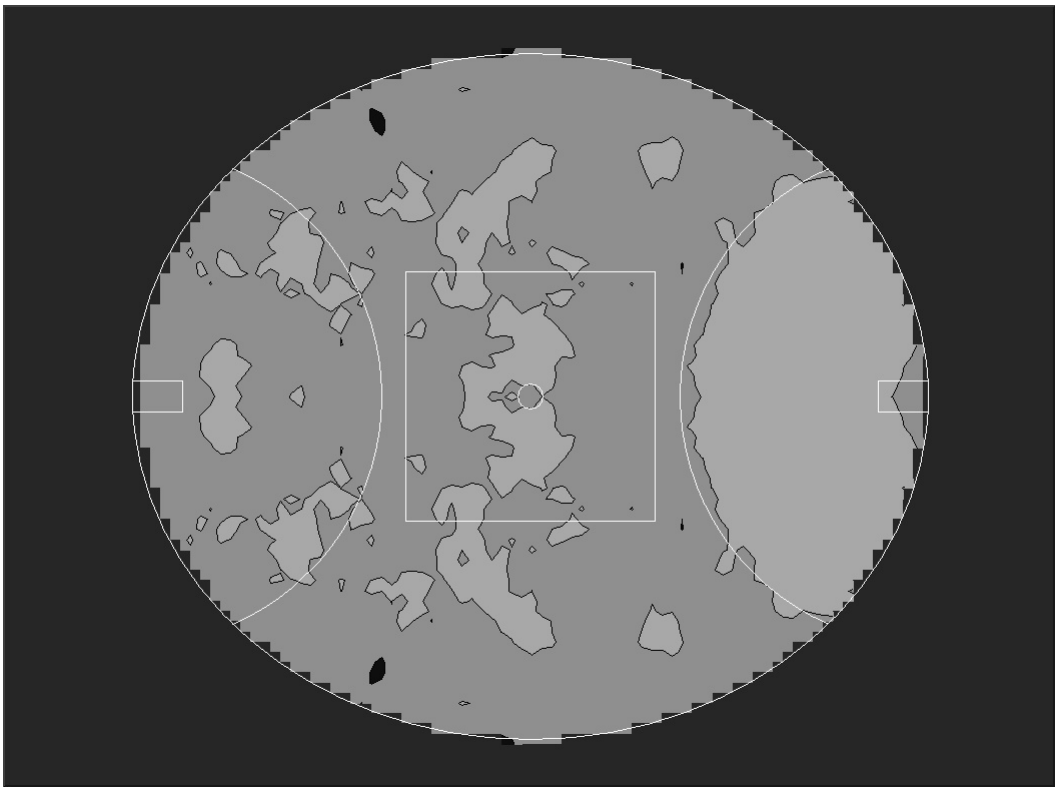


Figure 6: Significance Test: "Set" versus "Directed"

Calculated as an average over the ground, there is only a boost of 0.3 points to be gained by taking a mark instead of receiving a handball. In modern football uncontested marks across the half-back-line are cheap, with the opponent barely interested in forcing the man to go back and take the set shot. But the advantage is wholly concentrated in the forward-50 arc, with an extra 1.5 points available on the scoreboard for having a set shot rather than a running shot at goal between 25 and 45 metres out. The light areas on Figure 6 show the regions where it is significantly better, at the two-sigma level, for a player to take a mark rather than gather it uncontested.

Average Phase Equity

The mean net value of each of the phases was calculated by averaging over the 200 zones on the field. This works as a "standard candle" to investigate deviations by teams or in certain situations.

A Player's Choices: What Happens Next

Imagine a player who has just taken a mark 70 metres out from goal, on about a 40°-45° angle. It's unlikely he can score himself, and he faces an unenviable choice between bombing it long in hope of improved field position without turning the ball over, or

picking out a nearby teammate to do the dirty work for him. This scenario – within six metres – has played out 822 times over the seasons 2004-2005. On average, a team in this position can expect to convert to about two points on the scoreboard (2.06(14)).

Table 3: Equity of Possession Phases, Averaged Over Field

Phase	Mean Equity	Comments
Set	+1.61(2)	Being awarded a free kick in the centre circle is worth more than 1½ points on the scoreboard
Directed	+1.32(2)	About half-way between Loose and Set, this Phase tends to show up the good decision makers
Loose	+1.11(2)	Even if the options aren't great, it's still worth more than two points on the scoreboard to be in the right place instead of his opponent
Hard	+0.80(3)	Half the value of a set shot, compared to a 50/50 Phase

It's immediately obvious from Figure 7 below that if the player passes short and keeps it near the boundary, he almost always finds a teammate. Even more encouragingly, the team scores from there virtually every time. On the other hand, directing the ball long into the central corridor seems to be about a 50/50 proposition to hold onto the ball. Is it worth the risk? And should he play on, relinquishing the set shot time to gain some ground by running? Figure 7 shows the results of the 195 marks at the MCG from this position.

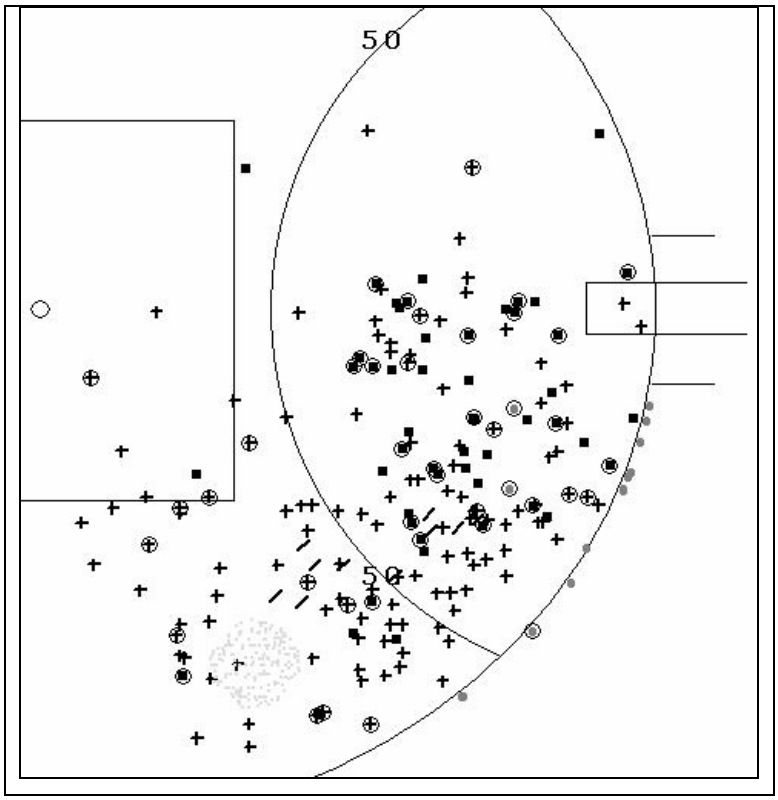


Figure 7: Mark on the MCG Half-Forward Flank

Showing all venues made the picture too crowded. The grey speckle in the lower left is the collection of points where a player marked. The plus signs (+) show where he managed to get the ball to a teammate, while the dark squares are immediate turnovers. The grey circles indicate the ball went into the umpire's control. A ring around the marker means that the next score was to the opposition – no ring indicates a score for the marking player's team. The nine diagonal slashes are the rare occasions that the player managed to run to this point and scored for himself.

The results are inconclusive, but they do highlight the dilemma. By choosing to handball, his team keeps the ball 97% of the time. With a short kick (gaining less than 35 metres or not moving closer to goal), the retention rate is 78%, but just 48% with a long kick. And yet the improved position gained from the long kick is worth the risk: a slightly higher equity as more of the scores are goals.

Table 4: Choices from a Mark (All Venues), 70 metres out on a 40-45 degree angle

Choice:	Handball	Kick 35m+	Kick <35m	Play On	Wait
N (%)	116 (14%)	382 (47%)	319 (39%)	347 (42%)	475 (58%)
Teammate	97%	48%	78%	73%	62%
Umpire	1%	14%	5%	8%	9%
Turnover	3%	37%	17%	19%	29%
Next Score	78%	77%	80%	81%	76%
Equity	2.08(35)	2.15(21)	1.96(21)	2.21(20)	1.95(18)

(All errors are one standard error)

Players who waited rather than played on saw the defence coalesce around their options and could only find a teammate 62% of the time, for a slightly lower equity result. It should be noted that some of these would not have had the choice to play on, due to a close opponent or falling to ground after the mark.

DISCUSSION

Players and fans understand the scoreboard. Telling them that giving the won ball to the opponent at this point on the field is effectively taking three points off the scoreboard is a strong message, and should foster a new way of thinking about the game.

It has long been noticed that defenders have higher "kicking effectiveness" percentages, a measure of how often they find a teammate as a percentage of total kicks. The pictures in this paper make it obvious why – there is little pressure on them, and a wealth of options to hit. There is an implicit "funnel" in many team sports due to the location of the goals – trying to kick into the neck of the funnel at centre-half-forward is very risky, but as seen by the equity gradient also very rewarding if the team has strong marking forwards in the corridor. It is much easier to advance along the gentle equity gradient in the back half of the ground, the funnel gaping open as teammates have more space to run to. The next step is to identify clubs' equity signatures, and find out where they are breaking down compared to the league standard. Where do they mostly direct the ball? Sydney are known to hug the boundary, but can this tactic be exploited?

CONCLUSION

This is just a first look at a huge body of data which is ready for exploitation by AFL researchers. Even these preliminary results are informing AFL coaches about the risks and rewards associated with some patterns of play. Future directions include looking at the effect of speed of play on equity – how much of an advantage is it to be able to advance the ball quickly? Or should the players switch play across the ground to exploit open space? A semi-Markov approach as advocated for ice hockey (Thomas, 2006) could also be useful, to reduce the number of data points needed for conclusive evidence of strategic advantage. Following the lead of baseball, an application to player ratings would be a significant opportunity. Identifying which players consistently increase equity for their team is a major goal.

Acknowledgements

The author would like to thank the dedicated Champion Data data-capture staff who methodically logged the 350,000 data points used in this analysis.

REFERENCES

Bennett, J. (2005) World Series Player Game Percentages. *Statistics in Sports section of the American Statistical Association*. Available from URL: http://www.amstat.org/sections/sis/pgp/

Clarke, S.R. and Norman, J.M. (1998) When to rush a behind in Australian Rules football: a Dynamic Programming approach. *Journal of the Operational Research Society* **49**(5): 530-536. Available from URL: http://www.swin.edu.au/lss/staff/staff_bios/s_clarke/Chapter8.html

Forbes, D.G. (2006) *Dynamic prediction of Australian Rules football using real time performance statistics*. Ph.D. Thesis submitted to Swinburne University.

Keith, T. (1996) Backgammon Glossary. *Backgammon Galore*. Available from URL: http://www.bkgm.com/glossary.html#equity

Romer, D. (2002) It's Fourth Down and what does the Bellman Equation say? A Dynamic-Programming Analysis of Football Strategy. *National Bureau of Economic Research Working Paper 9024*. Available from URL: http://www.nber.org/papers/w9024

Schatz, A. (2005) Methods to our Madness: DVOA Explained. *Football Outsiders*. Available from URL: http://www.footballoutsiders.com/methods.php

Studeman, D. (2004) The One About Win Probability. *Hardball Times*. Available from URL: http://www.hardballtimes.com/main/article/the-one-about-win-probability/

Thomas, A.C. (2006) The Impact of Puck Possession and Location on Ice Hockey Strategy. *Journal of Quantitative Analysis in Sports* **2(1)**, Article 6. Available from URL: http://www.bepress.com/jqas/vol2/iss1/6

AN ANALYSIS OF TEN YEARS OF THE FOUR GRAND SLAM MEN'S SINGLES DATA FOR LACK OF INDEPENDENCE OF SET OUTCOMES

Pollard, G.[1], Cross, R.[2] and **Meyer, D.**[3]

[1]Faculty of Information Sciences and Engineering, University of Canberra, Australia

[2]Physics Department, Faculty of Science, University of Sydney, Australia

[3]Faculty of Life and Social Sciences, Swinburn University of Technology, Australia

ABSTRACT

The objective of this paper is to use data from the highest level in men's tennis to assess whether there is any evidence to reject the hypothesis that the two players in a match have a constant probability of winning each set in the match. The data consists of all 4883 matches of grand slam men's singles over a 10 year period from 1995 to 2004. Each match is categorised by its sequence of win (W) or loss (L) (in set 1, set 2, set 3,...) to the eventual winner. Thus, there are several categories of matches from WWW to LLWWW. The methodology involves fitting several probabilistic models to the frequencies of the above ten categories. One four-set category is observed to occur significantly more often than the other two. Correspondingly, a couple of the five-set categories occur more frequently than the others. This pattern is consistent when the data is split into two five-year subsets. The data provides significant statistical evidence that the probability of winning a set within a match varies from set to set. The data supports the conclusion that, at the highest level of men's singles tennis, the better player (not necessarily the winner) lifts his play in certain situations at least some of the time.

KEY WORDS

data analysis, independence in tennis, constant probabilities, psychological development

INTRODUCTION

Several authors have carried out probabilistic analyses of tennis (Carter and Crews, 1974; Miles, 1984). A common assumption is that player A has a constant probability PA of winning a point on his/her service and that player B also has a constant probability PB of winning a point on service. Under this assumption and the assumption that points are independent, it can be shown that the better player does not always win and that each player has a constant probability of winning each set, no matter who serves first in the set (Pollard, 1983). Player A is the better player if PA is greater than PB.

There is little published research on testing whether players do have constant probabilities on service, that points (and hence games and sets) are independent and identically distributed (iid). A 'first game effect' in a match, namely that fewer breaks occur in the first game of the match, has been identified (Magnus and Klaassen, 1999). However, it would

appear that any non-iid effects such as the 'hot-hand effect' (in which winning a point, game or set increases ones chances of winning the next point, game or set) and the opposite effect, the 'back-to-the-wall effect', are small when analyzing large data sets (Klaassen and Magnus, 2001).

Many players believe, and commentators often state, that the winner of a set of tennis is not infrequently determined by merely a couple of points within that set. Given that a set lasts about (say) 60 points on average, and the couple of critical points can occur almost anywhere in the set, it would appear to be difficult to use statistical methods to identify a couple of non-iid points amongst approximately 60 other iid points. It would be like 'searching for a needle in a haystack'.

In this paper we focus on sets rather than points. If sets are not iid, it follows that points and games cannot be strictly iid, even if only a very small percentage of points contribute to the non-iid nature of the data. The data consists of ten years (1995 to 2004) of the four major annual tournaments for men's singles. These tournaments are the Australian Open, the French Open, Wimbledon and the US Open, are known as the Grand Slam tournaments, and are played on different types of surfaces. Using W to represent a set won by the eventual winner of the match and L to represent a set lost by the eventual winner, there are several possible match categories from WWW to LLWWW. Each of the 4883 singles matches for this period were classified into the relevant categories, and the frequencies of the categories were analysed to check for lack of independence of set outcomes.

METHODS

Assuming without loss of generality that player A is the better player, the results of a best-of-five sets singles match can be recorded as WWW, WWLW, WLWW, LWWW, WWLLW, WLWLW, WLLWW, LWWLW, LWLWW, LLWWW, and LLL, LLWL, LWLL, WLLL, LLWWL, LWLWL, LWWLL, WLLWL, WLWLL and WWLLL where W represents a set won by player A, and L represents a set lost by player A. When we do not know who the better player is, a win in three sets for example (WWW or LLL above) is simply a win WWW to the winner of the match (not necessarily player A). Thus, when we do not know who the better player is, the above twenty outcomes reduce to the ten mutually exclusive outcomes WWW, WWLW, WLWW, LWWW, WWLLW, WLWLW, WLLWW, LWWLW, LWLWW and LLWWW where W represents a set won by the eventual winner of the match and L represents a set lost by the eventual winner.

The data consisted of ten years of men's singles grand slam results. There were 4883 matches in total, and spurious data such as matches where one player 'retired' (presumably injured) before the match was finished were omitted. The number of matches in each of the above categories was:

WWW	2330;	WWLW	503;	WLWW	487;	LWWW	609;
WWLLW	151;	WLWLW	135;	WLLWW	186;	LWWLW	138;
LWLWW	156;	LLWWW	188						

The first model fitted involved a constant probability, p, of player A (the notionally or theoretically better player) winning each set. A short and simple search using a spreadsheet showed that the value of p which minimized Chi-Squared was 0.769, and the results are given in Table 1. For example, the expected value for the row WWW in Table 1 is 4883*(0.769*0.769*0.769 + 0.231*0.231*0.231) = 2280.77, allowing for both a win and a loss by the theoretically better player.

Table 1: Match outcomes when the probability player A wins each set is 0.769

Category	Observed	Expected	Obs-Exp	Chi-Squared
WWW	2330	2280.77	49.23	1.06
WWLW	503	559.24	-56.24	5.66
WLWW	487	559.24	-72.24	9.33
LWWW	609	559.24	49.76	4.43
WWLLW	151	154.09	-3.09	0.06
WLWLW	135	154.09	-19.09	2.36
WLLWW	186	154.09	31.91	6.61
LWWLW	138	154.09	-16.09	1.68
LWLWW	156	154.09	1.91	0.02
LLWWW	188	154.09	33.91	7.46
Total	4883	4883	0.00	38.68

The value of Chi-Squared was 38.68 with 8 degrees of freedom, so the fit is a poor one. This is not surprising as a constant p-value for all matches is clearly unrealistic. It can be seen from the Obs-Exp column in Table 1 that there was a greater number of three sets and five sets results observed than was expected under this model. Also, for the four sets matches, this model underestimated the number of LWWW matches, and overestimated the other two categories. Similarly, for the five sets matches, the model underestimated the number of WLLWW and LLWWW matches.

In order to attempt to overcome the shortage of three and five sets matches expected under the above model, it was decided to model the data using two p values, one greater than 0.769 and the other less than it, and combine the results. The value greater than 0.769 would increase the proportion of three set matches, and the value less than 0.769 would increase the proportion of five set matches. Thus, for simplicity, the data was modeled as consisting of 2 types of matches-'close' matches (with p less than 0.769) and 'not-so-close' matches (with p greater than 0.769).

Half the matches were assumed to be 'close', and half 'not-so-close'. Symmetric values about 0.769, p1 and p2, were considered, and the two p values which minimized Chi-Squared were identified. These two values were p1 = 0.705 and p2 = 0.833. The results for this model are given in Table 2. For example, the expected value for row 4 (LWWW) of Table 2 is given by 4883(0.5*((1-p1)*p1*p1*p1 + p1*(1-p1)*(1-p1)*(1-p1)) + 0.5*((1-p2)*p2*p2*p2 + p2*(1-p2)*(1-p2)*(1-p2))) = 541.71, allowing for both a win and a loss by the theoretically better player.

Table 2: Set outcomes when player A has in half the matches a probability of 0.705 of winning a set and in half the matches a probability of 0.833 of winning a set.

Category	Observed	Expected	Obs-Exp	Chi-Squared
WWW	2330	2340.77	-10.77	0.05
WWLW	503	541.71	-38.71	2.77
WLWW	487	541.71	-54.71	5.53
LWWW	609	541.71	67.29	8.36
WWLLW	151	152.85	-1.85	0.02
WLWLW	135	152.85	-17.85	2.08
WLLWW	186	152.85	33.15	7.19
LWWLW	138	152.85	-14.85	1.44
LWLWW	156	152.85	3.15	0.06
LLWWW	188	152.85	35.15	8.08
Total	4883	4883	0.00	35.59

The value of Chi-Squared for this model was 35.59 with 7 degrees of freedom, so the fit is again a poor one. Whilst this is a better fit with respect to the proportion of three and five set matches, the number of LWWW matches is still underestimated under this model, as is the number of WLLWW and LLWWW matches.

(It is noted here as an aside that if we remove the restriction that exactly half of the matches have a p-value of p1 and half of them have the value p2 whilst keeping p1 = 0.705 and p2 = 0.833, a slightly smaller value of chi-squared can be obtained. The lowest Chi-Squared value obtained was 34.35 with 6 degrees of freedom when the proportion of matches with p1 = 0.705 was 0.53, and the proportion of matches with p2 = 0.833 was 0.47. Thus, for this model (and indeed for the others considered in this paper), modifying the proportion of 'close' and 'not-so-close' matches had negligible effect on the Chi-Squared values. For this reason, no further reports on this modification are given in this paper.)

It can be seen from Table 2 that, under this model, the expected number of matches in each of the 3 four sets categories are equal. Correspondingly, the expected number of matches in each of the 6 five sets categories are also equal. It is clear that this characteristic remains true even if we fitted more (or even many many more!) than just two p values to the data. Further, it follows that if the p-value is constant for each set within each match (but possibly different for each of the 4883 matches) the expected number of matches in each of the 3 four set categories would be equal, and that the expected number of matches in each of the 6 five set categories would also be equal. It is possible to fit the best-fitting model to this data such that the 3 four set categories have equal expected values and the 6 five set categories also have equal expected values. Note that this is simply a data fitting exercise, and that there is no assumed underlying p-value(s) such as in the above analyses. When this is done, the expected values for the three, four and five set categories are 2322.2, 534.0, and 159.8 respectively, and the Chi-Squared value is 33.17 with 7 degrees of freedom. Again the fit is not a good one and **we conclude that the p-values for each set (within each match) are not constant.**

It can be seen from Table 2 that the (Obs-Exp) value was positive for the categories LWWW, WLLWW, LWLWW and LLWWW. These categories represent situations in which the winner (typically, but not always, the better player, player A) was behind (in sets) at some stage in the match. Thus, the data suggests that the better player might 'try harder' or 'lift his game' in situations in which he is behind. In order to address this 'trying harder when behind' effect, it was assumed that player A lifted his probability of winning a set by D1 when he was behind in the set score. A closer look at the data also suggests that player A might be 'on-a-roll' when he has just won a set and as a consequence lifts his probability of winning the next set. In order to address this 'on-a-roll' effect, it was assumed that player A lifted his probability of winning a set by D1 when he won the previous set. The categories WLLWW and LLWWW noted above represent situations in which the winner (probably more often player A) lost two sets in a row. These are situations in which player A has a real need to make an extra special effort to lift his game. Thus, it was further assumed that player A lifted his probability of winning a set by an amount D2 (anticipated to be somewhat bigger than D1) for the remainder of the match immediately after having lost two sets in a row (there are 3 such match categories). It is for reasons of simplicity that the parsimonious model with only two lifted levels was tested.

Given that p1 and p2 are increased by D1 or D2 in certain situations, it seemed appropriate, in order to get a reasonable overall fit, to lower both their 'starting' values (ie, those for set1) from those in Table 2. Given this, the notion of symmetric p-values about 0.769 also seemed irrelevant. The values of p1 and p2, D1 and D2 which minimized Chi-Squared were p1 = 0.704 and p2 = 0.798, D1 = 0.035 and D2 = 0.110, and the results are given in Table 3. For example, the expected value for the number of LLWWW matches is 4883(0.5*((1-p1)*(1-p1-D1)*(p1+D2)*(p1+D2)*(p1+D2) + p1*(p1+D1)*(1-p1-D1)*(1-p1)*(1-p1-D2)) + 0.5*((1-p2)*(1-p2-D1)*(p2+D2)*(p2+D2)*(p2+D2) + p2*(p2+D1)*(1-p2-D1)*(1-p2)*(1-p2-D2))) = 186.68.

Table 3: Match outcomes when p1 and p2 are 0.704 and 0.798 respectively, D1 = 0.035 and D2 = 0.110.

Category	Observed	Expected	Obs-Exp	Chi-Squared
WWW	2330	2333.26	-3.26	0.00
WWLW	503	485.11	17.89	0.66
WLWW	487	497.52	-10.52	0.22
LWWW	609	607.47	1.53	0.00
WWLLW	151	159.08	-8.08	0.41
WLWLW	135	128.06	5.94	0.27
WLLWW	186	184.21	1.79	0.02
LWWLW	138	143.72	-5.72	0.23
LWLWW	156	156.89	-0.89	0.01
LLWWW	188	186.68	1.32	0.01
Total	4883	4883	0.00	1.83

The value of Chi-Squared was 1.83 with 5 degrees of freedom, so the fit is a good one indicating that the model fits the data well.

In order to carry out a simple check on the model, it was decided to break the data into two time periods (1995-1999 and 2000-2004), and check for consistency across the periods. The above parameter values or estimates for p1, p2, D1 and D2 based on the full 10 year period 1995 to 2004 were used 'as estimates' for the period 1995 to 1999 (2448 matches) and for the period 2000-2004 (2435 matches). The fits were surprisingly good, with Chi-Squared values of 1.81 and 3.00 respectively. (It is clearly quite likely that lower values of Chi-Squared could be obtained by fitting p1, p2, D1 and D2 values specific to each period, but there is little point in doing this.

There appeared to be no evidence in the data that the weaker player could lift his game in situations where it would have been useful for him to do so.

Table 4: Observed match outcomes for the periods 1995-1999 and 2000-2004.

category	1995-1999	2000-2004
WWW	1189	1141
WWLW	242	261
WLWW	240	247
LWWW	304	305
WWLLW	76	75
WLWLW	66	69
WLLWW	98	88
LWWLW	67	71
LWLWW	76	80
LLWWW	90	98
Total	2448	2435

RESULTS

The 4883 completed men's singles matches at grand slam tournaments for the period 1995-2004 have been analysed to test the hypothesis that the probability of winning a set within a match is constant. This hypothesis was rejected.

A model which fits the data well has been found. It is a model in which the better player lifts his probability of winning a set in certain situations. These situations are
(i)when he is behind in the set score, needs to lift his game, and lifts his probability of winning the next set by (on average) 0.035,
(ii)when he has just won a set, is 'on-a-run'', and lifts his probability of winning the next set by (on average) 0.035, and
(iii)when he has just lost two sets in a row, desperately needs to lift his game, and lifts his probability of winning each remaining set by (on average) 0.110.

DISCUSSION

The results of this study are quite encouraging for the better player, but perhaps somewhat discouraging for the weaker player. The findings indicate that the weaker player needs to be 'on his guard' for a change in fortunes when the match is 'going well' for him.

The results of the analysis in this paper show that often the better player can increase his probability of winning a set by quite a substantial amount when it is really necessary to do so in order to reduce his probability of losing the match. A set can often be won rather than lost by winning just one, two, or a few particular important points (Morris, 1977). Thus, it would appear from the analysis in this paper that the better player is more able to lift his play on particularly important points than is the weaker player.

Further studies might include whether women's matches (although only best-of-three sets) have comparable characteristics or whether there are gender differences in this regard.
It would appear that the methodology used in this paper has a range of sporting applications, particularly for the often occurring situation in which the better player or team does not always win a match, or the 'best' player or team does not always win a series of matches. Another area of application might be assessment in which the 'best' student (or persons being assessed) does not always come first.

CONCLUSIONS

The conclusion is that matches turn around in favour of the better player significantly more often than would be expected under the usual randomness/independence assumptions of probability. As each point is a 'zero-sum' situation for the two players, it is not strictly possible to tell from just the statistical records whether this 'turn-around' characteristic is because the better player lifts his play or because the weaker player lowers his play. Nevertheless, it is useful for both players to know of the existence of this phenomenon as any player (except the best player in the world) should sometimes be the better player and sometimes the weaker on the court. The better player can take advantage of it, and the weaker player needs to guard against it.

REFERENCES

Carter, W. and Crews, S. (1974) An analysis of the game of tennis. *The American Statistician.* **28(4)**, 130-134.

Klaassen, F. and Magnus, J. (2001) Are points in tennis independent and identically distributed? Evidence from a dynamic binary panel data model. *Journal of the American Statistical Association*. **96**, 500-509.

Magnus, J. and Klaassen, F. (1999) On the advantage of serving first in a tennis set: Four years at Wimbledon. *The Statistician*. **48**, 247-256.

Miles, R. (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(1)**, 93-108.

Morris, C. (1977) The most important points in tennis. In S.Ladany and R. Machol (Eds). *Optimal strategies in sports* (Volume 5 in Studies in Management Science and Systems). Amsterdam: North-Holland, 131-140.

Pollard, G. (1983) An analysis of classical and tie-break tennis. *Australian Journal of Statistics*. **25(3)**, 496-505.

MAXIMISING HEIGHT, DISTANCE OR ROTATION FROM REAL-TIME ANALYSIS VISUALISATION OF TAKE-OFF ANGLES AND SPEED

Green, R

Department of Computer Science, University of Canterbury, New Zealand

ABSTRACT

Studies to optimise take off angles for height or distance have usually involved either a time-consuming invasive approach of placing markers on the body in a laboratory setting or using even less efficient manual frame-by-frame joint angle calculations with one of the many sport science video analysis software tools available. This research introduces a computer-vision based, marker-free, real-time biomechanical analysis approach to optimise take-off angles based on speed, base of support and dynamically calculated joint angles and mass of body segments. The goal of a jump is usually for height, distance or rotation with consequent dependencies on speed and phase of joint angles, centre of mass COM) and base of support. First and second derivatives of joint angles and body part COMs are derived from a Continuous Human Movement Recognition (CHMR) system for kinematical and what-if calculations. Motion is automatically segmented using hierarchical Hidden Markov Models and 3D tracking is further stabilized by estimating the joint angles for the next frame using a forward smoothing Particle filter. The results from a study of jumps, leaps and summersaults supporting regular knowledge of results feedback during training sessions indicate that this approach is useful for optimising the height, distance or rotation of skills.

KEY WORDS

gymnastics, jumping, three-dimensional kinematics, computer vision

INTRODUCTION

Sport skills are tracked[1] and biomechanically analysed by either requiring athletes to wear joint markers/identifiers (an approach with has the disadvantage of significant set up time) or manually marking up video frame-by-frame. Such complex and time consuming approaches to tracking and analysis is an impediment to daily use by coaches and has barely changed since it was developed in the 1970s. Using a less invasive approach free of markers, computer vision research into tracking and recognizing full-body human motion has so far been mainly limited to gait or frontal posing (Moeslund and Granum, 2001). Various approaches for tracking the whole body have been proposed in the image processing literature using a variety of 2D and 3D body models. However cylindrical, quadratic and ellipsoidal (Drummond and Cipolla, 2001. Kakadiaris and Metaxas, 1996. Pentland and Horowitz, 1991. Wren et al., 1997) body models of previous studies do not contour accurately to the body, thus decreasing tracking stability. To overcome this problem, in this research 3D clone-body-model regions are sized and texture mapped from each body part by extracting features during

[1] Commercially available trackers are listed at www.hitl.washington.edu/scivw/tracker-faq.html

the initialisation phase (Cham and Rehg, 1999). This clone-body-model has a number of advantages over previous body models:

- It allows for a larger variation of somatotype (from ectomorph to endomorph), gender (cylindrical trunks do not allow for breasts or pregnancy) and age (from baby to adult).
- Exact sizing of clone-body-parts enables greater accuracy in tracking edges, rather than the nearest best fit of a cylinder.
- Texture mapping of clone-body-parts increases region tracking and orientation accuracy over the many other models which assume a uniform colour for each body part.
- Region patterns, such as the ear, elbow and knee patterns, assist in accurately fixing orientation of clone-body-parts.

Neither joint markers nor manual frame-by-frame mark-up provide volume and 3D centre-of-mass (COM) estimates of a 3D body model – invaluable for 3D biomechanical analysis. In this study, joint angle velocities, together with the size and mass of body segments enabled more accurate optimisation of take-off angles supporting the goal of a jump whether for height, distance or rotation with consequent dependencies on phase of joint angles and base of support.

CLONE-BODY-MODEL

The clone-body-model proposed in this paper consists of a set of clone-body-parts, connected by joints, similar to the representations proposed by Badler (Badler et al., 1993). Clone-body-parts include the head, clavicle, trunk, upper arms, forearms, hands, thighs, calves and feet. Degrees of freedom are modeled for gross full body motion. Degrees of freedom supporting finer resolution movements are not yet modeled, including the radioulnar (forearm rotation), interphalangeal (toe), metacarpophalangeal (finger) and carpometacarpal (thumb) joint motions.

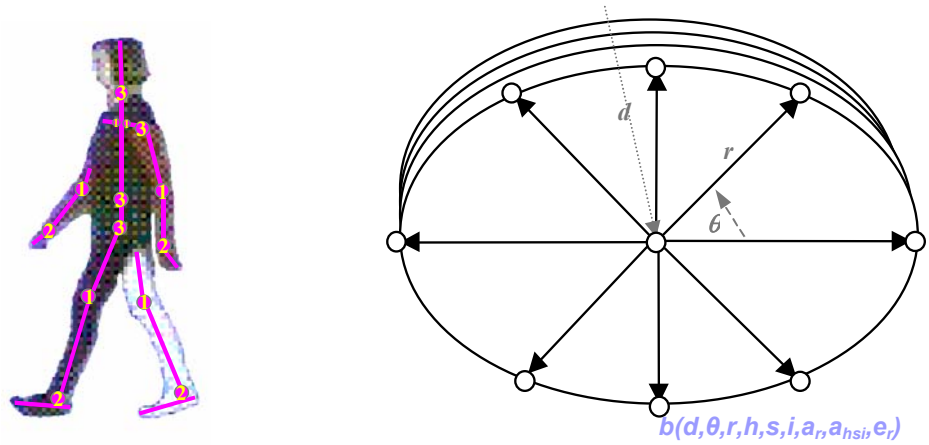


Figure 1: Clone-body-model consisting of clone-body-parts which have a cylindrical coordinate system of surface points *b()* and up to three DOF for each joint linking the clone-body-parts. Each surface point is a vector *b* with cylindrical coordinates (d,θ,r), colour (h,s,i), accuracy of radius (a_r), accuracy of colour (a_{hsi}), elasticity of radius (e_r).

Each clone-body-part consists of a rigid spine with pixels radiating out (Figure 1). Each pixel represents a point on the surface of a clone-body-part. Associated with each pixel is: radius or thickness of the clone-body-part at that point; colour as in hue, saturation and intensity; accuracy of the colour and radius; and the elasticity inherent in the body part at that point. Although each point on a clone-body-part is defined by cylindrical coordinates, the radius varies in a cross section to exactly follow the contour of the body as shown in Figure 2.

Automated initialisation assumes only one person is walking upright in front of a static background initially with gait being a known movement model. Anthropometric data (Pheasant, 1996) is used as a Gaussian prior for initializing the clone-body-part proportions with left-right symmetry of the body used as a stabilizing guide from 50^{th} percentile proportions. Such constraints on the relative size of clone-body-parts and on limits and neutral positions of joints help to stabilize initializations. Initially a low accuracy is set for each clone-body-part with the accuracy increasing as structure from motion resolves the relative proportions. For example, a low colour and high radius accuracy is initially set for pixels near the edge of a clone-body-part, high colour and low radius accuracy for other near side pixels and a low colour and low radius accuracy is set for far side pixels. The ongoing temporal resolution following self occlusions enables increasing radius and colour accuracy. Breathing, muscle flexion and other normal variations of body part radius are accounted for by the radius elasticity parameter.



Figure 2: Clone-body-model example rotating through 360 degrees.

KINEMATIC MODEL

The kinematical model tracking the position and orientation of a person relative to the camera entails projecting 3D clone-body-model parts onto a 2D image using three chained homogeneous transformation matrices as illustrated in Figure 3.

$$p\,(x,\, b\,) = I_i\,(x,\, C_i\,(x,\, B_i\,(x,\, b\,))) \qquad (1)$$

where x is a parameter vector calculated for optimum alignment of the projected model with the image, B is the Body frame of reference transformation, C is the Camera frame of reference transformation, I is the Image frame of reference transformation, b is a body-part surface point, p is a pixel in 2D frame of video (Rehg and Kanade, 1995).

Joint angles are used to track the location and orientation of each body part, with the range of joint angles being constrained by limiting the DOF associated with each joint. A simple motion model of constant angular velocity for joint angles is used in the kinematical model. Each DOF is constrained by anatomical joint-angle limits, body-part inter-penetration avoidance and joint-angle equilibrium positions modelled with Gaussian stabilizers around their equilibria. To stabilize tracking, the joint angles are

predicted for the next frame. The calculation of joint angles, for the next frame, is cast as an estimation problem which is solved using a Particle filter (Condensation algorithm).

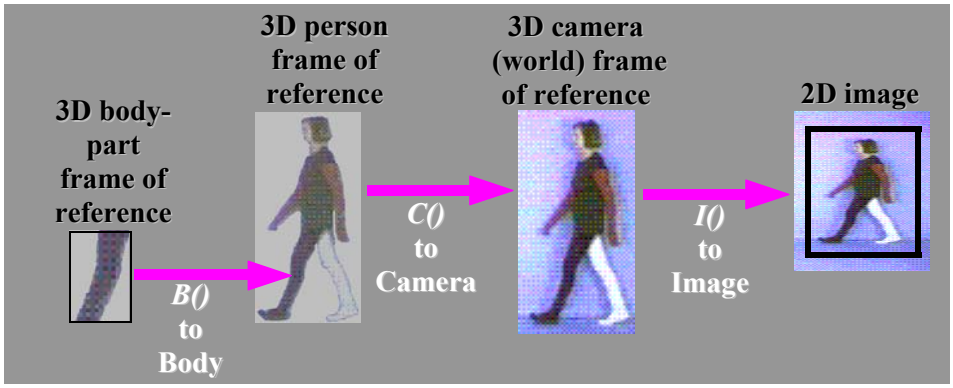


Figure 3: Three homogeneous transformation functions *B()*, *C()*, *I()* project a point from a clone-body-part onto a pixel in the 2D image.

PARTICLE FILTER

The Particle Filter was developed to address the problem of tracking contour outlines through heavy image clutter (Isard and Blake, 1996, 1998). The filter's output at a given time-step, rather than being a single estimate of position and covariance as in a Kalman filter, is an approximation of an entire probability distribution of likely joint angles. This allows the filter to maintain multiple hypotheses and thus be robust to distracting clutter.

With about 32 DOFs for joint angles to be determined for each frame, there is the potential for exponential complexity when evaluating such a high dimensional search space. MacCormick (MacCormick, 2000) proposed Partitioned Sampling and Sullivan (Sullivan, 1999) proposed Layered Sampling to reduce the search space by partitioning it for more efficient particle filtering. Although Annealed Particle Filtering (Deutscher et al., 2000) is an even more general and robust solution, it struggles with efficiency which Deutscher (Deutscher, 2001) improves with Partitioned Annealed Particle Filtering.

The Particle Filter is a considerably simpler algorithm than the Kalman Filter. Moreover despite its use of random sampling, which is often thought to be computationally inefficient, the Particle Filter can run in real-time. This is because tracking over time maintains relatively tight distributions for shape at successive time steps and particularly so given the availability of accurate learned models of shape and motion from the human-movement-recognition (CHMR) system. Here, the particle filter has:

- 3 probability distributions in problem specification:
  1. Prior density $p(x)$ for the state x
     - joint angles x in previous frame
  2. Process density $p(x_t|x_{t-1})$
     - kinematical and clone-body-models (x_{t-1}: previous frame, x_t: next frame)
  3. Observation density $p(z|x)$
     - image z in previous frame
- one probability distribution in solution specification:

1. State Density $p(x_t|Z_t)$
   - where x_t is the joint angles in next frame Z_t

1. Prior density: Sample s'_t from the prior density $p(x_{t-1}|z_{t-1})$ where x_{t-1}=joint angles in previous frame, z_{t-1}. The sample set are possible alternate values for joint angles. When tracking through background clutter or occlusion, a joint angle may have N alternate possible values (samples) $\boldsymbol{s}$ with respective weights $\boldsymbol{w}$, where *prior density*,

$$p(x) \approx S_{t-1} = \{(s^{(n)}, w^{(n)}), n=1..N\} = \text{a sample set}$$

(S_{t-1} is the sample set for the previous frame, $w^{(n)}$ is the n^{th} weight of the n^{th} sample $s^{(n)}$)
For the next frame, a new sample is selected, $s'_t = s_{t-1}$ by finding the smallest i for which $c^{(i)} \geq r$, where $c^{(i)} = \sum_t w^{(i)}$ and $\boldsymbol{r}$ is a random number $\{0,1\}$.

2. Process density: Predict s_t from the process density $p(x_t|x_{t-1}= s'_t)$. Joint angles are predicted for the next frame using the kinematic model, body model & error minimisation. A joint angle, $s_t^{(n)}$ in the next frame is predicted by sampling from the *process density,* $p(x_t|x_{t-1} = s_t'^{(n)})$ which encompasses the kinematic model, clone-body-model and cost function minimisation. In this prediction step both edge and region information is used. The edge information is used to directly match the image gradients with the expected model edge gradients. The region information is also used to directly match the values of pixels in the image with those of the clone-body-model's 3D colour texture map. The prediction step involves minimizing the cost functions (measurement likelihood density):

edge error E_e using edge information:

$$E_e(S_t) = \frac{1}{2n_e v_e} \sum_{x,y} (|\nabla i_t(x,y)| - m_t(x,y,S_t))^2 + 0.5(S-S_t)^T C_t^{-1}(S-S_t) \rightarrow \min S_t \qquad (2)$$

region error E_r using region information:

$$E_r(S_t) = \frac{1}{2n_r v_r} \sum_{j=1}^{n_r} (i_t[p_j(S_t)] - i_{t-1}[p_j(S_{t-1})])^2 + E_e(S_t) \rightarrow \min S_t \qquad (3)$$

where i_t represents the image at time t, m_t the model gradients at time t, n_e is the number of edge values summed, v_e is the edge variance, n_r is the number of region values summed, v_r is the region variance, p_j is the image pixel coordinate of the jth surface point on a clone-body-part.

3. Observation density: Measure and weigh the new position in terms of the observation density, $p(z_t|x_t)$. Weights $w_t = p(z_t|x_t = s_t)$ are estimated and then weights $\sum_n w^{(n)} = 1$ are normalized. The new position in terms of the observation density, $p(z_t|x_t)$ is then measured and weighed with *forward smoothing*:

- Smooth weights w_t over $1..t$, for n trajectories
- Replace each sample set with its n trajectories $\{(s_t, w_t)\}$ for $1..t$
- Re-weight all $w^{(n)}$ over $1..t$

Trajectories tend to merge within 10 frames

- O(N_t) storage prunes down to O(N)

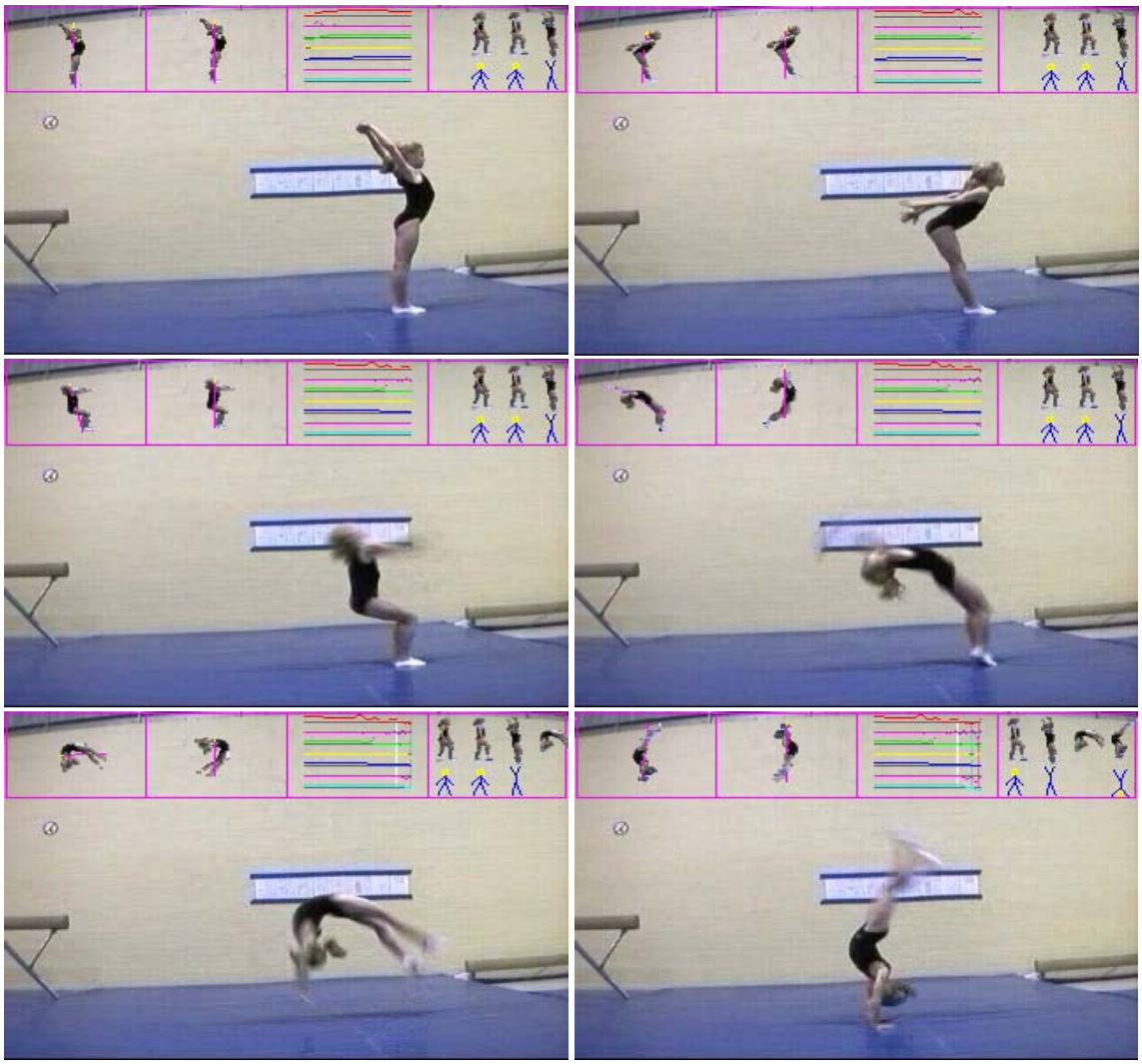


Figure 4 : Tracking jumping into a flic-flac with four overlaid information tiles.
Tile 1: principle axis through the body; Tile 2: body frame of reference (normalised to the vertical); Tile 3: motion vector trace (subset displayed); Tile 4: recognised skills.

In this research, feedback from the CHMR system utilizes the large training set of skills to achieve an even larger reduction of the search space. In practice, human movement is found to be highly efficient, with minimal DOFs rotating at any one time. The equilibrium positions and physical limits of each DOF further stabilize and minimize the dimensional space. With so few DOFs to track at any one time, a minimal number of particles are required, significantly raising the efficiency of the tracking process. Such highly constrained movement results in a sparse domain of motion projected by each motion vector.

Because the temporal variation of related joints and other parameters also contains information that helps the recognition process infer skill boundaries, the system computes and appends the temporal derivatives and second derivatives of these features to form the final motion vector. Hence the motion vector includes joint angles (32 DOF),

body location and orientation (6 DOF), centre of mass (3 DOF), principle axis (2 DOF) all with first and second derivatives.

PERFORMANCE

Hundreds of jumps and leaps were tracked and classified using a 2GHz, 640MB RAM Pentium IV platform processing 24 bit colour within the Microsoft DirectX 9, Intel OpenCV environment under Windows XP. The video sequences were captured with a Logitech USB 2.0 camera at 30 fps, 320 by 240 pixel resolution. Each person jumped in front of a stationary camera with a static background and static lighting conditions with minimal shadows. Only one person was in frame at any one time. Tracking began when the whole body was visible which enabled initialisation of the clone-body-model.

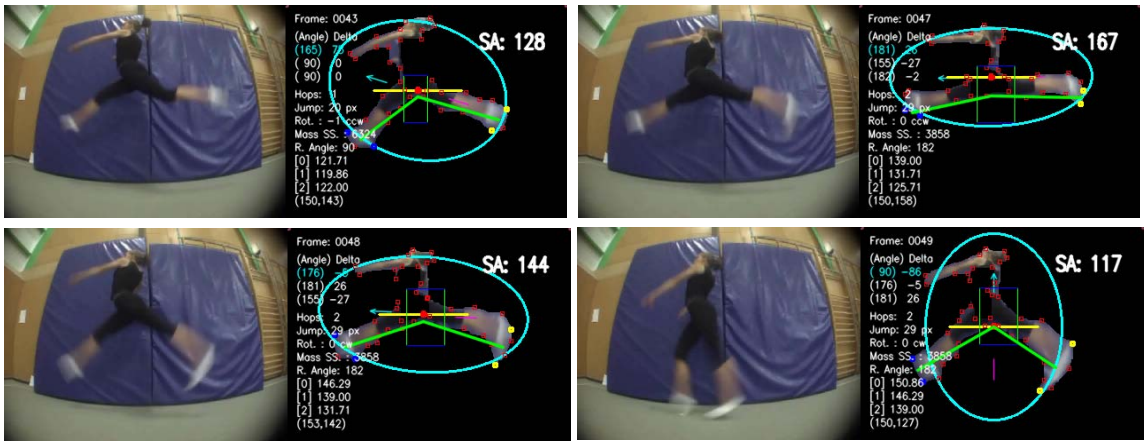


Figure 5: Tracking height, angle of splits, centre of mass and principal axis through a split jump.

The skill error rate quantifies CHMR system performance by expressing, as a percentage, the ratio of the number and magnitude of joint angle tracking errors to the number of joint angles in the reference set. Depending on the skill, CHMR system skill error rates can vary by an order of magnitude. The CHMR system results are based on a set of a total of 240 jump patterns, from straight jumps and split leaps (Figure 5) to jumping backward into flic-flacs (Figure 4). These were successfully tracked and evaluated with their respective biomechanical components quantified where a skill error rate of only 3.8% was achieved.

Motion blurring lasted about 10 frames on average with the effect of perturbing joint angles within the blur envelope. Given a reasonably accurate angular velocity, it was possible to sufficiently de-blur the image. There was minimal motion blur arising from rotation about the longitudinal axis during a twisting salto due to a low surface velocity tangential to this axis from minimal radius with limbs held close to a straight body shape. This can be seen in Figure 6 where the arms exhibit no blurring from twisting rotation, contrasted with motion blurred legs due to a higher tangential velocity of the salto rotation.

The CHMR system also failed for loose clothing. Even with smoothing, joint angles surrounded by baggy clothes permutated through unexpected angles within an envelope sufficiently large as to invalidate the tracking and evaluation.



Figure 6: Minimal motion blur arising from rotation about the longitudinal axis during a double twisting salto.

CONCLUSIONS

The 3.8% error rate attained in this research is not yet evaluating a natural world environment nor is this a real-time system with up to seconds to process each frame. The CHMR system did achieve 96.2% accuracy for the reference test set of skills. Although this 96.2% recognition rate was not as high as the 99.2% accuracy Starner and Pentland (Starner and Pentland, 1996) achieved, a larger test sample of skills were evaluated in this paper.

To progress towards the goal of lower error rates, the following improvements seem most important:

- Expand the clone-body-model to include a complete hand-model for enabling even more subtle movement domains such as finger signing and to better stabilize the hand position during tracking.
- Use a multi-camera or multi-modal vision system such as infra-red and visual spectrum combinations to better disambiguate the body parts in 3D and track the body in 3D.
- More accurately calibrate all movement skills with multiple subjects performing all skills on an accurate commercial tracking system recording multiple camera angles to improve on depth of field ambiguities. Such calibration would also remedy the qualitative nature of tracking results from computer vision research in general.
- Enhance tracking granularity using cameras with higher resolution, frame rate and lux sensitivity.

The results suggest that this approach has the potential to assist coaches and athletes optimise jump based skills during regular sessions by automatically displaying and logging biomechanical parameters of specific skills involving jumping and leaping.

REFERENCES

Badler, N. I., Phillips, C. B., Webber, B. L. (1993) *Simulating humans*, Oxford University Press, New York, NY.

Cham, T. J., Rehg, J. M. (1999) A multiple hypothesis approach to figure tracking, *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, **1**: 239–245.

Deutscher, J., Blake. A., Reid, I. (2000) Articulated body motion capture by annealed particle filtering, *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, **2**: 1144-1149.

Deutscher, J., Davison, A., Reid, (2001) A. Automatic partitioning of high dimensional search spaces associated with articulated body motion capture, *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, **2**: 669-676.

Drummond, T., Cipolla, R. (2001) Real-time tracking of highly articulated structures in the presence of noisy measurements, *IEEE International Conference on Computer Vision*, ICCV, **2**: 315–320.

Isard, M. A., Blake, A. (1996) Visual tracking by stochastic propagation of conditional density, *4th European Conference on Computer Vision*, 343-356, Cambridge, England.

Isard, M. A., Blake, A. (1998) A mixed-state Condensation tracker with automatic model switching, *6th International Conference on Computer Vision*, 107-112.

Kakadiaris, I., Metaxas, D. (1996) Model-based estimation of 3D human motion with occlusion based on active multi-viewpoint selection, *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, 81–87.

MacCormick, J., Isard, M. (2000) Partitioned sampling, articulated objects and interface-quality hand tracking, *European Conference on Computer Vision*, **2**: 3-19.

Moeslund, T. B., Granum, E. (2001) A survey of computer vision-based human motion capture, *Computer Vision and Image Understanding* **18**: 231–268.

Pentland, A., Horowitz, B. (1991) Recovery of nonrigid motion and structure, *IEEE Transactions on PAMI*, **13**: 730-742.

Pheasant, S. (1996) Bodyspace. *Anthropometry, Ergonomics and the Design of Work*, Taylor & Francis.

Rehg, J. M., Kanade, T. (1995) Model-based tracking of self-occluding articulated objects, *Fifth International Conference on Computer Vision*, 612-617.

Sullivan, J., Blake, A., Isard, M., MacCormick, J. (1999) Object localization by bayessian correlation, *7th International Conference on Computer Vision*, **2**: 1068-1075.

Wren, C., Azarbayejani, A., Darrell, T., Pentland, A. (1997) "Pfinder: Real-time tracking of the human body", *IEEE Transactions on PAMI*, **19**(7): 780-785.

AN ANALYSIS OF GOAL-KICKING ACCURACY IN AUSTRALIAN RULES FOOTBALL

Bedford[1], A and **Schembri[1], A**

School of Mathematical and Geospatial Sciences, RMIT University, Bundoora, VIC, Australia

ABSTRACT

Australian Rules Football, governed by the Australian Football League (AFL), is the most popular winter sport played in Australia. One criticism directed towards players of the game is that in over 100 years of football, accuracy of kicking for goal is the only aspect not to have improved. The truth of this claim is somewhat difficult to ascertain. However, using data from the 2005 AFL season, we investigate goal-kicking accuracy at a variety of levels to determine any differences in goal-kicking success. We chart accuracy at both a team and player level across various distances, angles and types of play. Unlike overseas codes of football, AFL was not played in a fully enclosed arena until the completion of Docklands Stadium in Melbourne in 2000. This investigation also determines if Docklands stadium has had any possible influence on accuracy.

KEY WORDS

goal kicking, accuracy, AFL

INTRODUCTION

Australian Rules Football is arguably Australia's most popular winter sport. It is a fast game played over four quarters of roughly 30 minutes in length, where players at the elite level need to be multi skilled to succeed. Some of its features include jumping and high marking (catching the ball), the use of handball, being able to bounce the ball on the run, and being able to kick over both short and long distances to either a player or for the goals. One common criticism directed towards players of the game is that in over 100 years of football, kicking for goal accurately is the only aspect not to have improved. However, this claim appears unfounded, since the ratio of goals to behinds has improved steadily since the game's inception. At its worst, the ratio of goals scored (6 points) to every behind (1 point) was 0.61 in 1900. This goal to behind ratio has exceeded one consistently since 1982, whereas it remained below 1.0 from 1897 until 73 years later. The highest ratio was recorded in 2000, with a ratio of 1.23. There may well be many factors that have contributed to this improvement, including the move away from suburban grounds, an improvement in player skill, the introduction of a national competition, or the use of an indoor stadium.

In this paper, goal-kicking accuracy will be examined at a variety of levels to ascertain any possible differences in goal-kicking success. Using data from the 2005 AFL season, we chart the accuracy at both a team and player level matched with various distances, angles and types of play. An additional aim of this investigation is to determine if the

recent addition of an indoor stadium has had any influence on the goal-kicking accuracy of teams in relation to other stadium and team combinations.

Unlike its North American equivalent, AFL football has only enjoyed the use of an indoor stadium since season's commencement in 2000. The stadium is located at Docklands in Melbourne and has the current commercial name of the Telstra Dome. Playing sports indoors removes many possible environmental variations experienced when playing sport outdoors. Although few studies have examined the impact of playing AFL football at Docklands Stadium, research has indicated that sporting teams based in indoor stadiums demonstrate a greater home field advantage when compared to those based in outdoor stadiums. To illustrate this notion, Zeller and Jurkovac (1989) found that playing home games in an indoor venue generated a 3-4% increase in home field advantage for American baseball teams, a figure that translated to three additional wins each season.

Several studies have emphasized the importance of shooting accuracy in sports such as basketball, water polo, and soccer. Onwuegbuzie (2000) incorporated a multiple regression model to examine the greatest predictors of winning a game of NBA basketball, and reported that field goal conversion was the single greatest predictor of success. Recent studies have also examined goal scoring accuracy during penalty shootouts in sports such as soccer and water polo (McGarry & Franks, 2000; Smith, 2004).

Research has suggested that during a game of water polo, the accuracy of each team's penalty shots does not vary as a function of the closeness of the game, the quarter the shot is taken, or the importance of the game in the context of the remainder of the season (Smith, 2004). In contrast, penalty shots in soccer demonstrate widespread variation with regard to accuracy. McGarry and Franks (2000) used data from the 1982-1998 FIFA World Cups and the 1996 European Championships and reported that penalty shots during free play were more accurate than those during a penalty shoot out. These authors contend that this may result from specialist penalty takers repeatedly shooting penalties during free play. In effect, these findings may have implications for AFL football since specialist goal kickers are often present within each team.

We will begin our investigation by discussing the intricacies of scoring in AFL football, the data collected in the present study, and the statistical methods used to analyze the data.

METHODS

Scoring in AFL football

There are several ways a team can score in AFL football. Kicking a goal earns a maximum score of six points. This is achieved by kicking the football between two center goal posts without the ball being touched by any other player. Goal kicking requires both accuracy and ability to keep the ball from other players reach. Figure 1 displays a typical forward line in AFL football and the cell distribution that was incorporated in the current analysis.

Definition of accuracy

We define accuracy by considering the total number of scoring plays by the player, team or teams of interest. We define the probability of an accurate shot for event i as

$p_i = \Pr(Accurate_i) = \frac{\sum goals\ scored_i}{\sum scoring\ shots_i}$. We have modeled accuracy using the binomial distribution, whereby each shot from certain positions under certain conditions is considered independent with either a goal (accurate) or a point (inaccurate) occurring. It is notable that missed shots (those that did not make the distance or were kicked out of bounds) were not considered due to this data being unavailable, as was consideration of rare cases (typically close games nearing the end of the match) where it may be advantageous to deliberately miss a goal to maintain a forward field position.

Data

AFL data is easy to source thanks to the endeavors of both dedicated fans of the game and the internet. A number of sources were used to obtain the necessary data for analysis in this paper. AFL tables (http://stats.rleague.com/afl/afl_index.html), All the stats (http://www.allthestats.com/) and Footywire (http://www.footywire.com/fw/web/ft_index) provide the public with readily accessible match data which we used to cross reference for our final sample of play by play data for season 2005, and match by match data from 1995 to seasons end 2005. We conducted our analysis on a number of variables. For every scoring play in season 2005, we obtained data on each scoring play - including the game time of a score, the player kicking for goal, the angle and distance from goal, and the type of play. We eliminated rushed behinds (when an opposition concedes a point) as this has no direct bearing on the scoring accuracy of a team or player.

Analysis methods

As indicated in the results section, we look at accuracy at a multinomial level – consisting of individual binomial success probabilities from a variety of distances from goal under differing types of play. When we considered two sets of events *i*, we simply used a two sample *z*-test on the success probabilities to determine significance. In order to determine significant differences between more than two events we used both Ryan's method (Ryan 1960) and a contingency table analysis reported as χ^2 (Pazer & Swanson, 1972). For post-hoc analysis, we used both Ryan's method and a Tukey-type multiple comparison test (Zar, 1999). Both procedures are comparable to the renowned Tukey's method for Wholly Significant Difference (WSD). For the Tukey-type multiple comparison test, the square root of each proportion was transformed to its arcsine so the resultant data will have an underlying distribution that is nearly normal. Then Tukey's test is performed on the data. Ryan's method differs slightly from Tukey's method in that it produces better results for smaller samples where the normal approximation is invalid. In this way we end up with results analogous to that obtained using Tukey's WSD.

RESULTS

The cost of inaccuracy since 1995

If we consider the 2035 matches played from 1995 to 2005, the team with the best scoring accuracy won 64.7% of matches played. However, this statistic fails to consider if the losing team could have won if they were more accurate. Remarkably, 10.7% of losing teams had a greater number of scoring shots than their opposition and could have won the match if they were more accurate.

Figure 1. Proportion of losses where scoring shots outnumbered the opponent 1995-2005

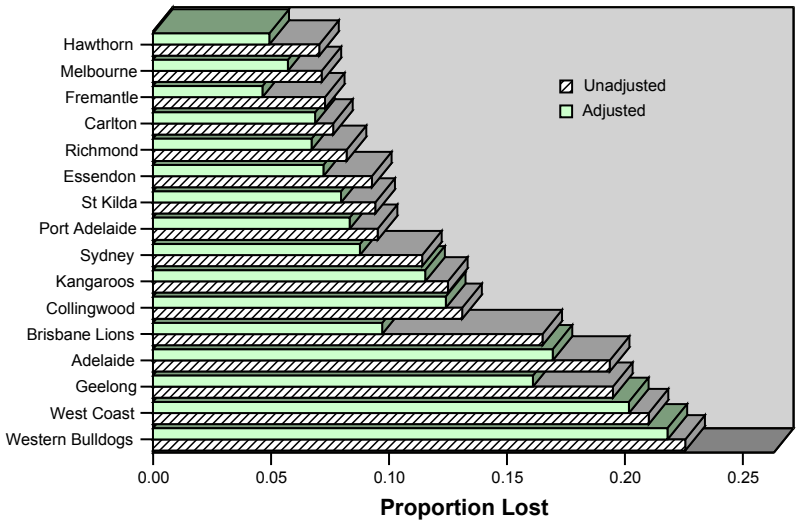


Figure 1 looks deeper at this cost, considering the proportion of lost matches that could have been won at a team level if their goal scoring were more accurate than their opposition. The unadjusted values include rushed behinds, whereas the adjusted removes all rushed behinds. The AFL team with the worst 'inaccuracy losses' record was the Western Bulldogs, suffering 30 defeats that could have been victories (an unadjusted 22.6% of losses). In both seasons 2002 and 2004, the Bulldogs lost 4 matches that could have been victories if they were more accurate – a huge amount considering there are only 22 regular season matches (in 2002 they missed the finals by 1.5 wins).

Accuracy for season 2005

Data from the 2005 AFL season were analyzed to determine those areas that generate scoring opportunities in AFL football. The following tables incorporate all home and away, and finals matches played during the 2005 season.

Table 1: Goals scored across varying distances and angles as a percentage of all goals scored.

Distance / Angle	Less 10m	10m-30m	30m-40m	40m-50m	Over 50m	Total
In Front	6.0	22.6	12.3	14.5	5.9	61.3
45 Degrees	1.7	11.0	8.9	10.1	3.8	35.5
On Boundary	0.4	1.8	0.6	0.4	0.0	3.2
Total	8.1	35.4	21.8	25.0	9.7	100.0

As shown in Table 1, the majority of goals were scored directly in front of goal as opposed to on a 45° angle or from the boundary line. Behinds scored as a function of distance and angle were also assessed and the results are displayed in Table 2.

Table 2: Behinds scored across varying distances and angles as a percentage of all behinds scored.

Distance / Angle	Less 10m	10m-30m	30m-40m	40m-50m	Over 50m	Total
In Front	0.3	8.5	8.9	11.4	7.9	37.0
45 Degrees	0.2	7.7	11.7	11.4	4.6	35.6
On Boundary	0.1	2.4	1.5	0.7	0.1	4.8
Total	0.6	18.6	22.1	23.5	12.6	100.0*

*Note. 22.6% of behinds were rushed

Unlike the distribution of goals in Table 1, behinds are evenly distributed when players kick from either in front of goal or at a 45° angle. However, the proportion of behinds scored directly in front of goal may be a reflection of the high number of shots taken from this position.

Comparison of set shots and general play accuracy.

To further examine the nature of goal kicking in AFL football, shots at goal were separated into set shots and general play categories. A set shot incorporates a kick for goal following a mark or free kick, whilst general play includes taking a snap shot for goal or kicking whilst on the run. Significance tests comparing accuracy in sets shots and general play were conducted using two sample *z*-tests. These analyses revealed that when kicking from directly in front, set shots were significantly more accurate than shots in general play at a distance of 10 to 30 meters ($p < .001$), or 30 to 40 meters ($p < .001$). Similar results were obtained when kicking for goal on a 45° angle. In particular, there was a significant increase in goal-kicking accuracy when kicking a set shot from a distance of 10 to 30 meters ($p < .001$), or 30 to 40 meters ($p < .02$). In contrast, when kicking for goal from greater than 40 meters, kicking whilst in general play was equally effective as kicking a set shot. However, the likelihood of scoring a goal from further than 40 meters remains moderate at best, regardless of the type of kick. Such a conclusion is verified in Table 3 which displays the overall accuracy of AFL clubs across all venues during 2005.

Table 3: Goal-kicking accuracy (%) across varying angles and distances from goal.

	Set Shot			General Play		
Distance from Goal	In Front	45˚ Angle	Boundary Line	In Front	45˚ Angle	Boundary Line
0-10	100.0 (90)	100.0 (27)	83.3 (12)	94.4 (213)	88.7 (62)	80.0 (10)
10-30	86.0 (664)	73.4 (331)	59.1 (66)	67.6 (772)	56.7 (513)	41.2 (119)
30-40	71.2 (510)	51.2 (516)	40.0 (55)	52.5 (446)	43.4 (385)	21.2 (33)
40-50	62.8 (744)	52.2 (644)	48.4 (31)	56.2 (418)	50.0 (304)	26.7 (15)
50+	51.3 (277)	53.6 (183)	50.0 (4)	44.0 (325)	45.9 (185)	0.0 (2)

Note. *n* in parenthesis

Clearly, when kicking from a distance less than 50 meters, the probability of scoring a goal from directly in front is considerably higher than an equivalent shot on a 45° angle or from the boundary line. Using Ryan's method, significance tests were conducted for each multiple comparison to determine whether goal-kicking accuracy varied across the distance and angle of the shot at goal, refer to Tables 4 and 5 respectively.

As shown in Tables 3, 4 and 5, goal-kicking accuracy declined significantly with each increasing distance interval for set shots taken in front of goal. Furthermore, when kicking a set shot on a 45° angle, AFL players were significantly more accurate when kicking between 0 and 30 meters, when compared with all other distances. During general play, kicking from in front of goal between 0 and 30 meters was significantly more accurate than kicking in excess of 30 meters. Additionally, kicking in general play between 0 and 30 meters was significantly more accurate than any other distance interval when on a 45° angle. There were no significant differences in goal-kicking

accuracy across the various distances when having a set shot or kicking in general play from the boundary line.

Table 4: Multiple comparisons of goal-kicking accuracy across varying distances from goal.

Type of Play	Angle	χ^2 *p-value*	Extreme Groups	ω	*p*
Set Shot	In Front	< .001	(0-10m, 50m +)	-0.334	< .001
	45° Angle	< .001	(0-10m, 50m +)	-0.334	< .001
			(30-40m, 50m +)	0.64	ns
	Boundary	ns	(0-10m, 50m +)	0.013	ns
General Play	In Front	< .001	(0-10m, 50m +)	-0.383	< .001
			(30-40m, 40-50m)	0.041	ns
	45° Angle	< .007	(0-10m, 50m +)	-0.261	< .001
			(30-40m, 50m +)	0.165	ns
	Boundary	.007	(0-10m, 50m +)	0.255	ns

Note. *n* in parenthesis

In regard to the angle from goal, kicking a set shot from the boundary line was significantly less accurate than kicking from in front or on a 45° angle when less than 30 meters out from goal. Furthermore, when kicking from 30 to 40 meters from goal, having a set shot from in front was significantly more accurate than from a 45° angle or from the boundary line. Finally, kicking from in front was significantly more accurate than kicking from a 45° angle or from the boundary line when kicking in general play between 10 and 40 meters from goal.

Goal-kicking accuracy across home and away venues.

Table 6 displays a comparison of goal-kicking accuracy across all AFL teams when playing at home and away venues. Two-sample *z*-tests revealed that when kicking a set shot from outside the 50 meter arc, there was a significant increase in the probability of kicking a goal when playing at home in comparison to an away venue, $p < .031$. However, no other significant differences emerged between goal kicking accuracy at home and away venues.

Goal-kicking accuracy across specialist and non-specialist goal-kickers.

Comparable to a striker in soccer, or a three point specialist in basketball, primary goal kickers exist in the majority of AFL teams. What remains unclear however is whether these forwards are superior at kicking for goal when compared to their teammates. Two sample *z*-tests revealed that leading goal kickers were significantly more accurate than all other AFL players when converting set shots from either 30 to 40 meters in front of goal ($p = .032$), or 10 to 30 meters on a 45° angle ($p = .004$). Set shots in front of goal were also bordering significance for the following distances: 10 to 30 meters ($p = .06$), 40 to 50 meters ($p = .055$), and in excess of 50 meters ($p = .056$). However, these distances may have lacked significance due to the small sample of specialist goal kickers. Surprisingly, there were no significant differences in shots taken during general play between specialist goal kickers and all other AFL players.

Table 5: Multiple comparisons of goal-kicking accuracy across varying angles from goal.

Type of Play	Distance	χ^2 *p-value*	Extreme Groups	ω	*p*
Set Shot	0m to 10m	< .001	In Front, 45°, Boundary	-0.066	< .001
			45°, Boundary	-0.004	.044
	10m to 30m	< .001	In Front, 45°, Boundary	-0.134	< .001
			In Front, 45°	-0.071	< .001
			45°, Boundary	-0.013	.029
	30m to 40m	< .001	In Front, 45°, Boundary	-0.128	< .001
			In Front, 45°	-0.135	< .001
			45°, Boundary	0.039	ns
			In Front, Boundary	-0.171	< .001
	40m to 50m	< .001	In Front, 45°, Boundary	0.095	ns
	50m +	ns	In Front, 45°, Boundary	0.631	ns
General Play	0m to 10m	ns	In Front, 45°, Boundary	0.079	ns
	10m to 30m	< .001	In Front, 45°, Boundary	-0.138	< .001
			In Front, 45°	-0.050	< .001
			45°, Boundary	-0.048	.004
	30m to 40m	< .001	In Front, 45°, Boundary	-0.075	.003
			In Front, 45°	-0.017	.013
			45°, Boundary	-0.031	.020
	40m to 50m	.030	In Front, 45°, Boundary	0.050	ns
	50m +	ns	In Front, 45°, Boundary	0.473	ns

Note. *n* in parenthesis

Table 6: Goal-scoring accuracy (%) across home and away venues.

		10-30		30-40		40-50		50+	
		In Front	45˚ Angle	In Front	45˚ Angle	In Front	45˚ Angle	In Front	45˚ Angle
Set Shot	Home	86.3 (342)	74.3 (183)	74.6 (252)	51.9 (243)	61.4 (373)	52.6 (321)	57.3 (150)	59.0 (105)
	Away	85.7 (322)	72.6 (146)	67.8 (258)	50.6 (269)	63.7 (364)	51.7 (317)	44.4 (124)	45.5 (77)
General Play	Home	69.5 (413)	55.9 (272)	50.6 (243)	46.6 (221)	56.5 (216)	53.9 (154)	46.5 (142)	48.6 (109)
	Away	65.5 (359)	57.7 (241)	54.7 (203)	39.0 (164)	55.9 (202)	46.0 (150)	42.1 (183)	42.1 (76)

Note. *n* in parenthesis

Quarter by quarter comparison.

Another aspect of AFL football that may impact goal-kicking accuracy is the duration of a game. AFL football is played over four 30 minute quarters, totaling 120 minutes of intense sport which demands peak levels of endurance and sustainability. Using both Ryan's method and the χ^2 approach, significance tests were conducted for each multiple

comparison. The four quarters differed significantly for a set shot taken 10 to 30 meters out when directly in front of goal ($Ryan's\ p = .028; \chi^2\ p = .017$). Follow-up analysis using both Ryan's method and Tukey's approach revealed that players from this position were more accurate during Quarter 1 when compared to Quarters 2 and 3. Additionally, goal-kicking accuracy was superior during Quarter 1 in comparison to Quarters 3 and 4 when kicking a set shot in excess of 50 meters from goal.

Close versus non-close game comparison.

A comparison of goal-kicking accuracy during the fourth quarter of close versus non-close games was also examined. Close games were defined as those with a points margin between the teams of less than or equal to 12 points at the time the kick was taken. Two-sample *z*-tests revealed no significant differences in goal-kicking accuracy, indicating that AFL footballers remain as accurate during pressurized and regular games.

A look at Docklands Stadium

To determine the impact that playing at an indoor stadium had on goal-kicking accuracy, the number of games lost due to inaccuracy at each stadium between the 2000 and 2005 seasons were plotted and the results are displayed in Table 7.

Table 7: Losses due to inaccuracy and overall accuracy across AFL venues during seasons 2000 to 2005.

Stadium	Losses due to inaccuracy		Losses due to Inaccuracy (%)^	Total Losses	Overall Accuracy (%)	First Played
	Unadjusted	Adjusted^				
Docklands	13	12	4.7	275	56.4	2000
Subiaco	9	7	7.0	129	54.2	1987
Princes Park	4	3	9.1	44	54.5	1897
Football Park	15	15	10.6	142	53.7	1991
MCG	30	23	10.8	279	54.8	1897
GABBA	9	6	11.5	78	53.1	1981
SCG*	9	5	13.6	66	52.9	1979
Kardinia	7	7	15.2	46	52.4	1941
York Park	3	1	20.0	15	53.0	2001
Stadium Aust.	4	4	28.6	14	53.8	2002
Manuka	5	4	35.7	14	53.5	1998

^Rushed behinds removed

*4 matches played in 80 years prior to 1979

As shown in Table 7, Docklands Stadium hosted the least number of games that were potentially lost as a result of inaccurate kicking and provides preliminary evidence that playing in an indoor stadium impacts on goal-kicking accuracy. Furthermore, a comparison of the two main Victorian stadiums (Docklands and MCG) showed statistically significant improvements in the goal kicking accuracy levels of games played at Docklands ($z = -2.06$, $p = .040$).

To assess whether playing in an indoor stadium increases goal-kicking accuracy for the home team, the accuracy of teams based at Docklands stadium was compared to those based at alternate venues, shown in Table 11. Using both Ryan's method and the χ^2 approach, no significant differences were found in goal-kicking accuracy for Docklands stadium tenants, other Victorian clubs, and non-Victorian clubs.

DISCUSSION

Table 1 revealed that 90.3% of goals scored during 2005 were kicked between 0 and 50 meters from goal, demonstrating the multidimensional nature of AFL footballers. In effect, forwards are capable of kicking goals from varying distances and angles which is analogous to the skills of professional basketball players, albeit on a much larger scale.

The benefits of directing the ball through the center corridor were also demonstrated throughout this paper. In fact, the probability of scoring a goal when kicking from directly in front was significantly greater than an equivalent shot on a 45° angle or from the boundary line; provided the player was within 50 meters from goal. These findings highlight the benefits of implementing a team strategy that encourages players to direct the ball through the middle of the ground rather than running the ball along the wing and boundary line. In addition, goal-kicking accuracy declined significantly when the distance from goal was increased which provides further evidence for the popular notion to kick the ball 'long and direct' into the forward line.

A surprising result was that goal-kicking accuracy did not differ significantly across home and away venues aside from set shots taken from outside the 50 meter arc. These findings are consistent with those of Varca (1980) who found no significant differences in field goal conversion rates for American college basketball teams playing either at home or away. Unlike basketball however, there is considerable variation in the playing field dimensions used in AFL football and thus inaccuracy should increase due to a lack of ground familiarity. In effect, this finding demonstrates the high adaptability of AFL footballers and provides further support for the multidimensional nature of their kicking skills.

Analogous to McGarry and Franks (2000) study of penalty shootouts in international soccer, the results presented in Table 7 indicate that set shots taken from directly in front of goal are better taken by specialist goal-kickers. There was little variation in goal-kicking success across the four quarters during a game. However, an interesting finding was that set shots in front of goal from greater than 50 meters were significantly more accurate during the first quarter than an equivalent kick in either the third or fourth quarters. This may be a reflection of player fatigue during the latter stages of a game and demonstrates that where possible; kicking for goal from outside 50 meters should be implemented during the first half.

To further examine the effects of player fatigue and heightened pressure on goal-kicking, an analysis of accuracy during the final quarter of close games was employed. However, there was only minor variation in goal-kicking accuracy between close and non-close games.

When assessing Docklands Stadium in comparison to all other AFL venues, accuracy levels were highest at the indoor stadium. Furthermore, despite it being a new stadium, the number of matches lost by teams when they could have won if more accurate than

their opponents was the lowest of all stadiums at 4.7%, or 13 of 275 matches. In stark contrast (see Table 10) other new outdoor stadiums have seen a high percentage of matches potentially lost as a result of inaccuracy. The findings in Table 11 also demonstrate that indoor stadiums benefit both home and away teams in regard to goal-scoring capabilities and fail to disadvantage the away team. However, this analysis merely examined goal-kicking accuracy and does not question the presence of home ground advantages such as ground familiarity or crowd factors which have been demonstrated in past research (Stefani & Clarke, 1992; Varca, 1980).

CONCLUSION

In this paper, we have investigated how goal-kicking accuracy in Australian Rules football varies as a function of distance, angle of shot and type of play. Notably, the type of player and venue also impact on goal kicking accuracy. The depth of inaccuracy cannot be underestimated, with potentially 10.7% of losses since 1995 attributed to errant goal kicking. We analysed the goal kicking accuracy of players at a multinomial level and found positional, play type, player and stadium significances. Extending this descriptive analysis, we intend to apply these findings to a tactical analysis whereby we investigate what the optimal path to goal should be based upon a series of field and type of play conditions. Other work could involve a more detailed look at complete misses for target and investigating the effect of rushing behinds on final match outcomes. In addition, a comprehensive examination of the interaction effects between the distance and angle from goal is necessary since these appear to be the primary determinants of the overall probability of scoring a goal.

REFERENCES

McGarry, T. and Franks, I. (2000) On winning the penalty shoot-out in soccer. *Journal of Sports Sciences* **18**, 401-409.

Onwuegbuzie, A. (2000) Factors associated with success among NBA teams. *The Sport Journal* **3** (2).

Pazer, H. and Swanson, L. (1972) *Modern methods for statistical analysis*. Intext educational publishers, Scranton, Pennsylvania.

Ryan, T. (1960) Significance tests for multiple comparison of proportions, variances, and other statistics. *Psychological Bulletin* **57**, 318-328.

Smith, H. (2004) Penalty shot importance, success and game context in international water polo. *Journal of Science Medicine and Sport* **7**, 221-225.

Stefani, R. and Clarke, S. (1992) Predictions and home advantage for Australian rules football. *Journal of Applied Statistics* **19**, 251-261.

Varca, P. (1980) An analysis of home and away game performance of male college basketball teams. *Journal of Sport Psychology* **2**, 245-257.

Zar, J. (1999) *Biostatistical analysis*.4th edition. Prentice-Hall, New Jersey.

Zeller, R. and Jurkovac, T. (1989) Doming the stadium: The case for baseball. *Sport Place International* **3**, 34-38.

GOLF PUTTING MODELS

de Mestre, N.

Bond University, Gold Coast, Queensland, Australia

ABSTRACT

The rolling of a golf ball on a flat grass green forms a significant part of the game of golf. Since putting strokes count for nearly half of the number of strokes taken in a regular game of golf, it is considered worthwhile to analyse this putting phase. This has been difficult to achieve even on a flat green. The putting surface is partly elastic and the grass springs back during the passage of the ball. To investigate the matter more fully, a fast film was made of a golf ball rolling next to a large rule on a flat green. When a golf ball is hit by a golf putter it slips initially but then rolls purely. In the experiment, a golf ball was released from a known height down an inclined plane and smoothly run onto the green. Its position was then recorded at many time intervals until it came to rest. The experimental results are compared with current models, including a constant-retardation model and the quadratic finite elastic model, which works well for heavier balls.

KEY WORDS

golf, putting, experiments

INTRODUCTION

The putting phase is a significant section of a round of golf, since more than 40% of the strokes taken in a professional golfer's game are made on the greens. A golf ball is 0.0427m in diameter, and is almost spherical in shape except for a large number of concave dimples on its surface. Its mass is 0.046kg, but the ball is not uniform in density throughout, consisting of a dense inner core, a less dense outer core of wound rubber and a thin plastic cover containing the dimples. Therefore it is inappropriate to consider its moment of inertia (I) as $0.4ma^2$, where m represents its mass and a represents its radius.

When a golf ball is hit by a putter, it leaves the club with an initial linear and angular speed. Pure rolling only occurs when

$$\dot{x} = a\dot{\varnothing} \tag{1}$$

where $\dot{x}$ is the linear speed at any instant, and $\dot{\varnothing}$ is the ball's angular speed at the same instant. Therefore, it is almost assured that the ball will skid along initially (Thomas, 2002). To avoid this initial skidding, a club head in the shape of a horizontal circular cylinder was suggested by Daish (1972), but it has never been manufactured commercially. To obtain the correct impact point for no skidding immediately on impact, the diameter of the club head should be 7/3 times the diameter of the golf ball, or almost 0.1m (Figure 1).

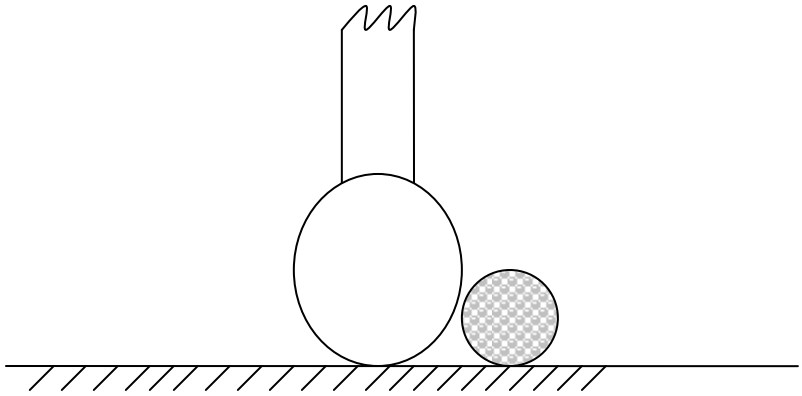


Figure 1: Daish's circular cylindrical putter

However, most professional golfers use a flat-faced putter with slight angle of loft. This causes the ball to skip over the top surface of the grass initially in almost imperceptible little hops. This makes it difficult to analyse a golf putt, even on a flat green.

An analysis of golf putting should also consider errors in the horizontal plane where by the ball is pushed or pulled away from its intended line of motion. Cochran and Stobbs (1968) reported on experiments in which the ball was deliberately hit off-centre towards the toe or heel of the clubface. The deviation was not enough to cause short putts to be missed, but for long putts the ball would not only miss the hole, but stop short. The proper hitting speed is always judged from an assumption of hitting the ball with the centre of the club face, and golfers should be aware that off-centre putting reduces the value of the initial linear speed of the ball.

Consequently, two modifications to putters were suggested by Cochran and Stobbs. The first was to mark on the top of the putting blade the point on the club face where the ball is to be contacted with the club. The second was to shift the weight distribution of the club head, so that there is more weight at both the toe and the heel. The club would then be less likely to twist in the grip during impact, making for greater consistency in direction and distance.

Further developments have occurred with the golf putter leading to the Zen Oracle Tour Blade putter which contains a vertical hole in the putter's head large enough to take a golf ball. This can be used for training, for aligning the putt, and it improves the stability of the head during the striking process (www.zenoracle.co.uk).

A great deal more can be said (and has already been said by others) about the physiological, anatomical and psychological aspects of putting techniques. Most of this has been adequately covered by Cochran and Stobbs.

The purpose of this paper is to consider the theories developed so far for a rolling ball on a flat green, and to compare their predictions with the many experiments that I and my co-researchers (John Scott and Maurice Brearley) have carried out since 1993. It proves to be useful to consider other balls besides golf balls, so that the difficulties associated with an analysis of golf putting can be pin-pointed.

CURRENT THEORIES

Any theory concerning the rolling of a non-deformable ball on a deformable surface must be based on the observation that the ball will be retarded until it comes to rest or, in the case of a successful golf putt, drops into the hole.

The initial speed of a putted golf ball depends on how hard the ball is hit, and this is governed by the distance from the hole that the putt is taken. For long putts, an accepted strategy is to aim for a pseudo-hole with the same centre as the real hole, but with a radius of approximately one metre. If successful, the final putt should then be less than one metre.

A putted golf ball will not roll purely in the early stages of a long putt across the green, but may do so in the latter stages, particularly during the last one metre of the putt. Therefore, two theories of motion are needed for these two different behaviours.

The final stage of the putt, when the ball is rolling purely, lends itself to a classical dynamics analysis. It has been shown (Scott and de Mestre, 1994) that the effect of air drag on a golf ball is negligible, except on very windy days.

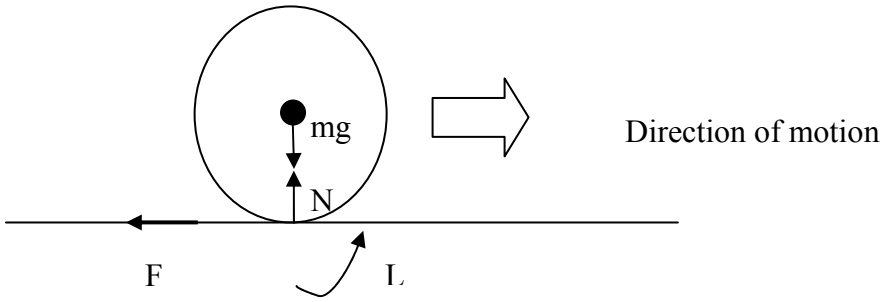


Figure 2: Forces and couple on a rolling sphere

If x denotes the distance that the ball has moved in the horizontal direction in time t, then the Principles of Linear and Angular Momentum yield

$$m\ddot{x} = -F \tag{2}$$

$$0 = N - mg \tag{3}$$

$$I\ddot{\varnothing} = Fa - L \tag{4}$$

where F and N are the horizontal and vertical components of the force exerted by the ground on the ball, L is the friction couple arising from the asymmetric foot print of the ball on the elastic deformable surface of the green, and a dot denotes differentiation with respect to time t. (Figure 2).

In the most basic model, F is assumed to be constant and equation (2) integrates to

$$x = -Ft^2/(2m) + Vt \tag{5}$$

where V is the initial linear speed of the putt at the start of the pure rolling phase. If T is the total time of this phase and X is the distance covered during this phase, then the retardation constant F/(2m) can be evaluated from

$$F/(2m) = (VT - X)/T^2 \tag{6}$$

The only difficulty with this pure rolling model is that the retardation has to be assumed to drop instantaneously to zero at the end of the putt. Note also that since the speed is zero at the end of the putt

$$F/m = V/T \tag{7}$$

Hence results (6) and (7) will only be consistent if

$$X = \frac{1}{2}VT \tag{8}$$

which is a simple check on whether the ball has rolled purely with constant retardation during the whole of any experiment.

When the ball is not rolling purely, Scott and de Mestre (1994) proposed a more complex model in which the frictional force F was assumed to be proportional to a power of $\dot{x}$, while the ball rolls purely yielding a solution

$$x = X\{1 - (1 - t/T)^{(VT/X)}\} \tag{9}$$

and therefore

$$\dot{x} = V\{1 - x/X\}^{1-X/VT} \tag{10}$$

The model is based on the condition that the ball should stay at rest when it finally comes to rest, but it over predicts the intermediate distances of experimental putts.

To take account of the elasticity of the grass on the green, finite elastic strain theory was employed by Brearley and de Mestre (2004). This pure-rolling model produced a frictional cum elastic force in the form

$$m\ddot{x} = -A - B\dot{x} + C\dot{x}^2 \tag{11}$$

with initial conditions $t = 0,\ x = 0,\ \dot{x} = V$ as before. The solution is

$$x = A_1 + A_2 t - A_3 \exp[1 + A_4 \exp(-A_5 T)] \tag{12}$$

where $A_i\,(i = 1\ to\ 5)$ depend on the constants A/m, B/m, C/m.

THE EXPERIMENTS

Since 1993 my colleagues and I have conducted rolling-ball experiments using golf balls, lawn bowls with the bias counteracted, bowling jacks, steel balls and billiard balls.

In the early experiments with golf balls, the ball was propelled by a putter over distances of less than 1 metre, and a fixed video camera (25 frames/second) recorded the position of the ball at specified times. The green used was a chipping green, and the ball bounced along in little hops indicating that pure rolling was not occurring (Scott and de Mestre, 1994).

To see if this occurred for all hard balls on a soft surface , a billiard ball was rolled down a ramp inclined at approximately 30° and ran across a carpet for a distance over 2 metres. The ramp enabled the velocity of the ball to be determined at the beginning of the carpet. A video camera was mounted on a trolley which ran on rails keeping up with the ball. The agreement between the constant retardation theory and the experiment was exceedingly close.

In 2001, experiments were conducted with a golf ball rolling down a ramp and across a green. The results were compared with the finite elastic strain model, since it was felt that the constant-retardation model could not apply to a golf ball. The time that the ball passed various markers was obtained using a multi-functional stop watch. The agreement between calculated and observed results was poor, presumably because a golf ball is too light to enforce the no-slip condition upon which the theory is based.

Experiments were next conducted on a bowling rink using respectively a lawn bowl (unbiased), a jack, and a steel ball which was the same size as the jack. Thus the steel ball and jack were smaller in size than a golf ball but much heavier. They were released down a ramp and ran for distances ranging from 5.45 to 18.33m. The agreement between finite elastic theory and experiments was very good for all these balls. However when similar experiments were conducted with a golf ball, on the same day over distances ranging from 2.9 to 6.75m, no such agreement could be obtained (Brearley and de Mestre, 2004). Again this indicated that the golf ball was hopping, but the others were rolling purely.

Therefore, we decided to obtain more detailed observations of the initial run of a golf ball on a green. In 2005 a hired Fastec Troubleshooter High Speed portable video camera was used with 640 x 480 resolution providing 125 frames/second (www.slowmotion.com.au). A lawn bowl and a golf ball where rolled down a ramp onto a practice putting green at Paradise Springs Golf Course, Robina. The lawn bowl had tape on its side and the golf ball had a great circle marked on it so that the change in angle could be recorded with time. All lawn bowl and short-distance golf experiments indicated that pure rolling was occurring. However the longer-distance golf

experiments showed that the golf ball was rising and falling by a few millimetres as it moved over the green, which indicated that long putts were not satisfying the pure-rolling conditions. This confirmed the observations made by Scott and de Mestre (1994). Quantitative measurements of these longer experiments were also restricted, as the ball came to rest well beyond the length of the 1-metre measuring rule. A representative summary of all these experiments is given in Table 1 below.

Table 1: Summary of typical experiments

Object	Surface	How Launched	Total T (s)	Total X (m)	Initial V (m/s)	Pure Rolling
Golf 1993	Grass	Putter	1.52	0.83	1.05	No
Golf 1993	Grass	Putter	1.52	0.78	1.0	No
Golf 1994	Grass	Putter	1.4	0.94	1.15	No
Golf 1994	Grass	Putter	0.96	0.47	0.75	No
Billiard 1999	Carpet	Ramp	3.92	2.41	1.2	Yes
Golf 2001	Grass	Ramp	4.66	6.78	3.58	No
Golf 2001	Grass	Ramp	3.07	2.89	2.08	No
Bowl 2002	Grass	Ramp	11.02	18.33	3.48	Yes
Bowl 2002	Grass	Ramp	8.31	9.91	2.62	Yes
Jack 2004	Grass	Ramp	8.49	14.81	3.53	Yes
Jack 2004	Grass	Ramp	7.88	12.22	3.14	Yes
Steel 2004	Grass	Ramp	6.68	11.53	3.51	Yes
Steel 2004	Grass	Ramp	5.19	6.63	2.66	Yes
Golf 2004	Grass	Ramp	4.66	6.75	3.58	No
Golf 2004	Grass	Ramp	3.23	4.50	2.72	No
Golf 2005	Grass	Ramp	1.96m	0.80	0.85	Yes
Golf 2005	Grass	Ramp	?	>1m	1.00	No

AN INTERESTING ANALYSIS

The $(t, x)-$values for all experiments were revisited, and plotted using Excel's Chart Wizard. Upon investigation of the trend lines, every experiment indicated a quadratic fit with correlation coefficients greater than 0.98.

For the lawn bowl, jack, steel ball and billiard ball, the coefficient of t agrees with the initial speed calculated at the bottom of the ramp, while the coefficient of t^2 is constant for repeated experiments with the same ball launched from different heights on the ramp on the same day. The term independent of t was very small in all these cases, indicating that the constant-retardation model seemed applicable for all the above balls, since they roll purely over nearly all of the distance moved. In addition, equation (8) is well satisfied for the billiard ball, jack and steel ball. So it appears that the constant-retardation solution (5) is just as effective for these balls as the finite elastic strain solution (12). Equation (8) is not so well satisfied by the lawn bowl, but this would be due to the significant effect of air drag because of the lawn bowl's much larger size.

However the golf ball experiments do not have these properties. Although they fit well a quadratic relation between x and t, the coefficient of t^2 is not the same for each experiment with the same ball and different initial speeds on the same day. This mirrors

the difficulty found by Brearley and de Mestre (2004) using the finite elastic strain model.

Nevertheless the high-resolution 2005 experiments do show that, for golf balls released down a ramp and travelling only a short distance (i.e. low initial speed at the bottom of the ramp), the motion of the ball is one of pure rolling. But as soon as the speed reaches a critical value at the bottom of the ramp, by release of the ball at a higher point on the ramp, the ball starts to bounce across the deformable green surface. This also happens for long putts with a golf club. Since no theory is yet available to describe accurately this hopping behaviour, it appears that the current best alternative is to assume a pseudo-constant-retardation model as an average of that which is really happening. This means that, with knowledge of an initial speed V and an average retardation constant $(F/2m)$, the distance of the putt can be estimated from equation (5). Perhaps this is what a golfer's brain does during his or her practice putts before a game anyway?

CONCLUSION

Putting by professional golfers usually involves two putts, a long one to arrive near the hole and a short one that hopefully drops the ball into the hole. A suitable model for the short putt (and the last metre of the long putt) is the constant-retardation model, neglecting the last few milliseconds when the acceleration drops sharply to zero. The experiments also indicate that the constant-retardation model is suitable for heavier balls which all appear to roll purely.

No suitable model has yet been determined for long golf putts, because the ball is light and not rolling purely, but hops along the top of the grassy surface. The more complicated problem of long putts on a sloping green cannot be analysed until a successful long-putt model on a flat green is resolved.

REFERENCES

Brearley, M. and de Mestre, N. (2004) Rolling of a rigid ball on a horizontal deformable surface. *ANZIAM Journal* **46**, 249-264.

Cochran, A. and Stobbs, J. (1968) *The Search for the Perfect Swing*. Lippincott, New York.

Daish, C.B. (1972) *The Physics of Ball Games*. English Universities Press, London, UK.

Scott, J. and de Mestre, N. (1994) Direct golf putting dynamics and strategies. In N. de Mestre (Ed). *Proceedings of 2nd Conference on Mathematics and Computers in Sport*, Bond University, Queensland. 135-146.

Thomas, Frank (2002) The physics of a putted ball. *Golf Digest* **July 1 issue**.

LOW COMPRESSION TENNIS BALLS AND SKILL DEVELOPMENT

Hammond, J.[1] and Smith, T.[2]

[1] School of Education, Southern Cross University, Lismore, NSW, Australia

[2] Department of Exercise Science & Sport Management, Southern Cross University, Lismore, NSW, Australia

ABSTRACT

Coaching aims to improve player performance and coaches have a number of coaching methods and strategies they use to enhance this process. If new methods and ideas can be determined to improve player performance they will change coaching practices and processes. This study investigated the effects of using low compression balls (LCBs) during coaching sessions with beginning tennis players. In order to assess the effectiveness of LCBs on skill learning the study employed a quasi-experimental design supported by qualitative and descriptive data. Beginner tennis players took part in coaching sessions, one group using the LCBs while the other group used standard tennis balls. Both groups were administered a skills at the beginning of a series of coaching sessions and again at the end. A statistical investigation of the difference between pre and post-test results was carried out to determine the effect of LCBs on skill learning. Additional qualitative data was obtained through interviews, video capture and the use of performance analysis of typical coaching sessions for each group. The skill test results indicated no difference in skill learning when comparing beginners using the LCBs to those using the standard balls. Coaches reported that the LCBs appeared to have a positive effect on technique development, including aspects of technique that are related to improving power of the shot. Additional benefits were that rallies went on longer and more opportunity for positive reinforcement. In order to provide a more conclusive answer to the effects of LCBs on skill learning and technique development recommendations for future research were established including a more controlled experimental environment and larger sample sizes across a longer period of time.

KEY WORDS

tennis, low compression balls, coaching

INTRODUCTION

One aim of a tennis coach is to improve player performance, and coaches will have a number of coaching methods and strategies they can employ to enhance this process. In addition, the knowledge base that underpins the coaching process is constantly changing due to research in coaching methodology and individual experiences (Fairweather, 1999). Such changes can take the form of alternative coaching styles and the use of new activities or equipment, amongst others. If new methods and ideas can be determined to improve player performance they can affect future coaching practice and enhance the coaching process. LCBs are typically used in modified versions of the game of tennis such as mini-

tennis (Cayer & Elderton, 2002; LTA, 2005). The balls used in these versions are variations of the standard ball that are softer, lighter and have lower bounce. The Lawn Tennis Association (LTA) also suggests that for very young players (4 - 8 years) the ball could be larger in order to make the game slower (LTA, 2005). While mini-tennis focuses on children the LTA suggest that beginners of all ages would benefit from playing the game with the slow moving balls, making skill learning easier.

Typically research involving tennis balls has used the standard type of ball or a standard ball that was modified by the researcher (Haake et al., 2003; Knudson, 1993; Mehta & Pallis, 2001). Recently the ITF (International Tennis Federation) have modified tennis ball specifications to include a faster (type 1) and slower, oversized (type 3) ball, to accompany the standard medium speed (type 2) ball which has the same ITF specifications that existed prior to 2000 (ITF, 2005). The new balls were developed in order to provide a greater degree of consistency to the game, the slowest ball to be used on fast courts the fastest ball on slow courts (ITF, 2005). Such changes have led to research on the effects of the new balls (particularly type 3). Metha & Pallis (2001) demonstrated that the larger cross-sectional area of a type 3 ball (approx. 6% bigger than type 1/2) increased drag on the ball, increasing the ball's flight time, which slows down the game. Research also suggests that the type 3 ball has the potential to change characteristics of game play, the type 3 ball having been shown to lead to less physiological strain and increased accuracy on a tennis skills test compared to the type 2 ball (Cooke & Davey, 2005). Cooke & Davey suggested that the improved accuracy of ground strokes with the type 3 ball may be beneficial to players with limited technical skills, such as those in the early stages of learning.

There are many dimensions to the coaching process, one important aspect being the development of skilled performance in players. As a result research has been conducted into the theory and practice of the coaching process aimed at improving skill learning, leading to recommendations for best coaching practice (Hodges & Franks, 2002; Schmidt & Lee, 1999). Challenges to the traditional style of coaching have led to the development of the game-based approach to coaching (Thorpe & Bunker, 1982). The success of this new method resulting in Tennis Coaches Australia (TCA) adopting the game-based approach to coaching as their preferred model (TCA, 2002). The use of non-traditional methods have also been investigated in younger players, where a series of pre-tennis activities using mini-tennis equipment led to improvements in fundamental motor skill acquisition in 5 year olds. This suggests that such improvements would ease the transition to learning specialist tennis skills (Quezada et al., 2000). In addition to modifications of coaching style, many researchers have investigated the use of new techniques or modified equipment on skill acquisition. Focusing on tennis, the types of techniques investigated include the effects of visualisation strategies and aids to performance as well as player reaction or movement time when playing with the larger type 3 ball (Andrew et al., 2003; Singer et al., 2001).

The purpose of this study was to investigate the effectiveness of LCBs on skill learning in beginners participating in an eight-week tennis coaching programme. Specifically, we investigated the effects of using low compression balls (LCBs) during regular coaching sessions on skill learning for beginning tennis players.

METHODS

Beginner tennis players took part in coaching sessions, one group using the LCBs while the other group used conventional (standard) balls. Both groups were administered a skills test at the beginning of a series of eight coaching sessions (pre-test) and again at the end (post-test). A statistical investigation of the difference between pre and post-test results was carried out to determine the effect of LCBs on skill learning. Additional qualitative data was obtained through interviews, video capture and the use of performance analysis software to analyse typical coaching sessions for each group. These multiple methods of data collection allowed for triangulation of data.

Fourteen boys and girls aged 5-11 years volunteered to participate in the study, they were members of a weekly beginners' class at a private tennis centre located in New South Wales. The participants were classified as beginners by the head coach and then self-selected into coaching groups, these groups were randomly assigned to an experimental or control condition. The experimental group used LCBs and the control group used standard balls. Information regarding demographics of each group is provided in Table 1. The beginners coaching programme was developed and overseen by the head coach, a level 2 accredited tennis coach with 26 years experience. The coaches responsible for delivering the sessions were all employees of the tennis centre. The coaching programme was consistent for both groups, in terms of strokes, drills and activities so that the coaching content and time was the same for both groups throughout the study.

Table 1: Age, gender and previous experience of participants by group

Group	n	Age (Mean±SD)	Gender	Prev.Experience (Mean±SD)
LCB	6	6.67 ±1.03 yrs	Male = 4 Female = 2	0.50 ±0.42 yrs
Standard	8	9.38 ±1.19 yrs	Male = 6 Female = 2	1.06 ±0.72 yrs

In order to describe the two types of ball used in the study, three LCBs and standard balls were randomly selected and their mass and size recorded. The LCB group used low compression balls that were softer, lighter and similar in size to the standard balls used in the study (Table 2).

Table 2: Mass and diameter (Mean ±SD) of LCB and standard ball

Tennis Ball	n	Mass (g)	Diameter (mm)
LCB	3	46.08 ±0.69	73 ±0
Standard	3	57.58 ±0.16	72 ±0

It is recommended that young children play tennis with a shorter, lighter racquet than adults (Cassell & McGrath, 1999; Harding, 1991). In this study children performed their skills test using a Pro Kennex Champ Ace Junior racquet suitable for their age (Cooper, 2005), all racquets being the same model. The participants underwent a traditional coaching programme, aimed primarily at introducing and developing forehand, backhand and serving skills. The sessions typically contained a warm up, drills, modified games and a cool down.

The sessions took place on an Astroturf court with regulation height nets. Each group attended one coaching session a week for eight weeks, sessions lasted for one hour.

A skill test was administered to each participant independently to establish performance levels prior to the study (pre test) and after the period of coaching (post test). The test was developed specifically for the study, in line with the coaching programme the three items tested were the forehand, backhand and serve. A review of the literature on existing skill tests and consideration of the participants' ability, learning context and time available for testing contributed to the development of the skill test, specific to the game of tennis. Skill test data was collected at the same venue as the coaching sessions. The pre-test occurring during week one and the post-test during week eight. Prior to the first test session age, gender, previous playing experience and hand dominance were recorded for each participant. Subjects were tested on the 3-item skills test. Verbal instruction and a demonstration were provided prior to testing each item, as well as indications of the scoring system. The same researcher administered all tests. For each participant, their score for each test trial on all three items (forehand, backhand and serve) were recorded and totalled (total test score). The score available for each trial ranged from 0-5, so a maximum total score of 90 could be achieved for the 3 test items across 6 trials each. For the skill test data, differences in total test scores between pre test and post test for both groups were analysed using the Kruskal-Wallis Test, in order to look for significant differences between groups. Skill test data was reported using group means and standard deviations. % difference calculations were used to compare a typical coaching session structure for each group.

A typical coaching session was videoed for both groups during week four of the coaching programme. The purpose of this was to obtain a record of a typical session and determine whether both groups spent similar amounts of time on each phase. A Macintosh OSX computer was linked to a digital 8 video camera so that the session content could be analysed using GameBreaker(Performance Analysis Software (GPAS) and excel. The GPAS was customised to allow the frequency and duration of each phase of the session to be logged in order to determine the typical session structure and amount of time spent on each phase of the session for both groups. Initially these events were logged on site while the sessions were in progress and edited as necessary during the post session. Table 3 lists the events logged to establish typical session content using the GPAS.

Table 3: Events used to analyse a typical coaching sessions for both groups.

Event	Explanation
INTRODUCTION & WARM-UP	Start to end of introductory / warm up activity(s)
DRILL/SKILL PRACTICE	Activity where groups practice whole or part tennis skills not including warm-up / cool-down activities
MODIFIED GAME	Game based / competition based activity involving whole or part tennis skills not including warm-up / cool-down activities of drill/skill practices
COOL-DOWN & DEBRIEF	Concluding game, activity, group address/debrief
COLLECTING BALLS	Whole group involved in collecting balls
GROUP INSTRUCTION	Whole group stopped for coach instruction/demo.
INDIVIDUAL INSTRUCTION	Individual/small groups stop for instruction/demo.

The coaches responsible for overseeing or coaching the weekly sessions were interviewed individually. Each coach was asked questions regarding their coaching experience, their approach to the coaching sessions and their perceptions on the effect of the LCBs on skill learning for players in the LCB group compared to those in the standard ball group. Each interview lasted approximately 30 minutes. The interviews transcriptions were analysed to identify categories of response made by interviewees. Individual categories that related to the responses for each question were derived and the substantive statements assigned into one of these (Gillham, 2000), allow for summaries of key points.

RESULTS AND DISCUSSION

The results of session content analysis for a typical session, indicate that both groups were receiving similar content and amount of time spent on each activity. Table 4 indicates the percentage time each group spent on each phase of the session and how long participants were involved in collecting balls or being instructed by the coach.

Table 4: Percentage of time spent on each phase of session & difference scores for groups

Event	% Time on each event (LCB Group)	% Time on each event (Standard Group)	% Difference
Intro & Warm-up	10.06	7.17	2.89
Drill/Skill Practice	27.9	32.58	-4.68
Modified Game	20.12	15.86	4.26
Cool-down/Debrief	7.06	6.54	0.52
Collecting Balls	9.27	5.35	3.92
Group Instruction	29.31	33.39	-4.08
Individual Instructn.	7.42	6.69	0.73

Table 4 shows the difference in time for each group is less than 5% for all events, the absolute difference between groups for each event ranging from 0.52% to 4.68%. A difference of less than 5% is generally considered acceptable when comparing the differences between two groups (Hughes & Franks, 2004). Data presented only represents one session for each group and cannot generalise to all sessions. However, these results do give an indication of the similarity of activities that both groups experienced throughout the study. In addition, the coaches strived to ensure that as many aspects of the session structure as possible was similar for both groups. A summary of the total skill test scores for pre and post tests and the difference scores (representing post-test score minus pre-test score) for each group is provided in Table 5.

Table 5: Summary of skill test scores (total score) for both groups.

Group	n	Pre Test Score (Mean ±SD)	Post Test Score (Mean ±SD)	Difference Score (Mean ±SD)
LCB	6	15.67 ±13.63	27.00 ±14.67	11.33 ±7.97
Standard	8	43.88 ±8.08	51.75 ±12.27	7.88 ±15.11

A positive difference score indicates improvement on the skills test between pre and post test, the larger the score the greater the improvement. The results show that the LCB group had a larger mean difference score (11.33 ±7.97) than the standard group (7.88 ±15.11). The difference in pre and post test (total) scores for each group were compared using the Kruskal-Wallis test, to determine whether the difference between the groups was significant and could be attributed to using the LCBs during the coaching sessions. The results of the statistical analysis (Table 6) indicate that the difference between groups was not significant ($p > 0.05$). A non-significant result suggests that using the LCBs during beginners coaching sessions over an eight week period does not significantly increase performance on a skills test when compared to a group of beginners coached using standard balls.

Table 6: Results of statistical analysis of skill test difference scores (post test - pre test)

χ^2 statistic	df	Asymp Sig.	Test Significant
0.038	1	0.846	No

Analysis of Figure 1, shows that the LCB group had a much lower mean score for the pre-test than the standard group (Table 5), suggesting that there were differences between the skill level of the two groups at the start of the study. At the end of the study the LCB group still had the lowest performance scores but had improved the most (although not significantly), such a trend follows the observations of Boyle and Ackerman (2003) where the lower performers initially remain the lowest performers at the end of a period of skill acquisition but have shown the biggest gain in improvement.

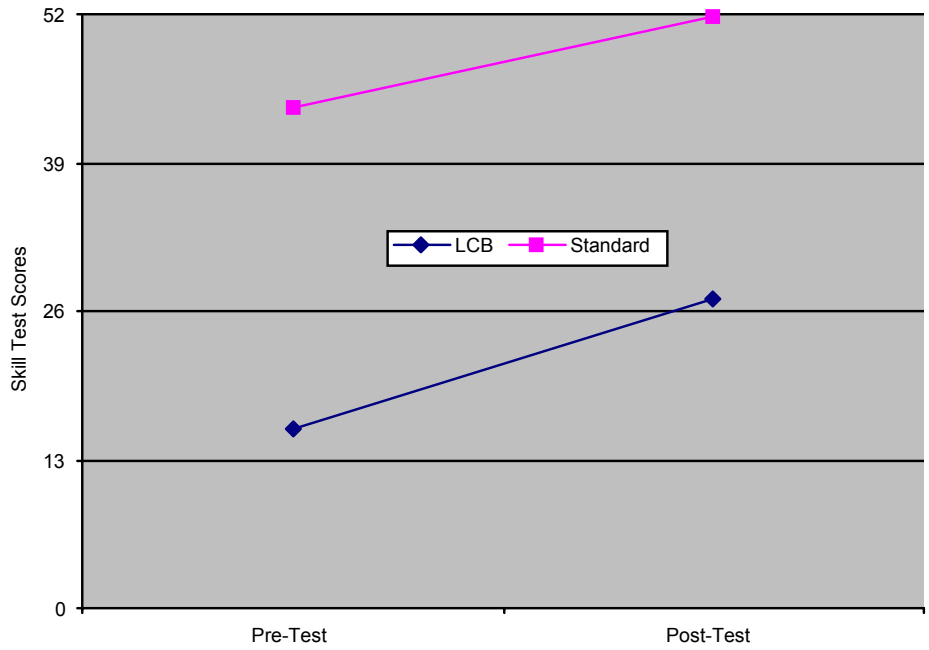


Figure 1: Mean skill test scores (total) for pre and post test for LCB and standard groups

The initial differences in skill level between the two groups could have been influenced by the differing characteristics of the participants in both groups, such as age and previous experience (Table 1). Age has been linked to the stage of motor development a child is in (Gallahue & Ozmun, 1997). The mean age of the LCB group (6.67 ±1.03 yrs) suggests that the majority of participants would still be developing their fundamental movement skills. The standard group, however, had a mean age of 9.38 ±1.19 years, indicating that the majority of participants in that group are more likely to have refined their fundamental movement patterns and progressed to a more advanced phase of motor development where they are more capable of developing the skills needed to play tennis (Gallahue & Ozmun).

It is suggested in the literature that children who attempt to learn specialist movement skills before they have developed the mature form of the fundamental movement skills necessary to perform the specialised movement, may be hindered in their progress (Gallahue & Ozmun, 1997). Therefore, due to the mean age of the LCB group it is possible they had not refined their fundamental movement skills prior to commencement of this study, which could result in less capability for the LCB group to learn specialised tennis skills such as the backhand stroke. In addition, the standard group was shown to have had more previous tennis coaching (1.06 ±0.72 yrs) than the LCB group (0.50 ±0.42 yrs). Although the difference between groups was not significant the mean difference scores for the LCB group were slightly higher than for the standard group (Table 5). Investigating the data by item it reveals that the LCB group showed the greatest improvement in the forehand stroke than any other, whereas the standard group showed the greatest improvement in the backhand score, both groups showed least improvement in the serve (Figure 2).

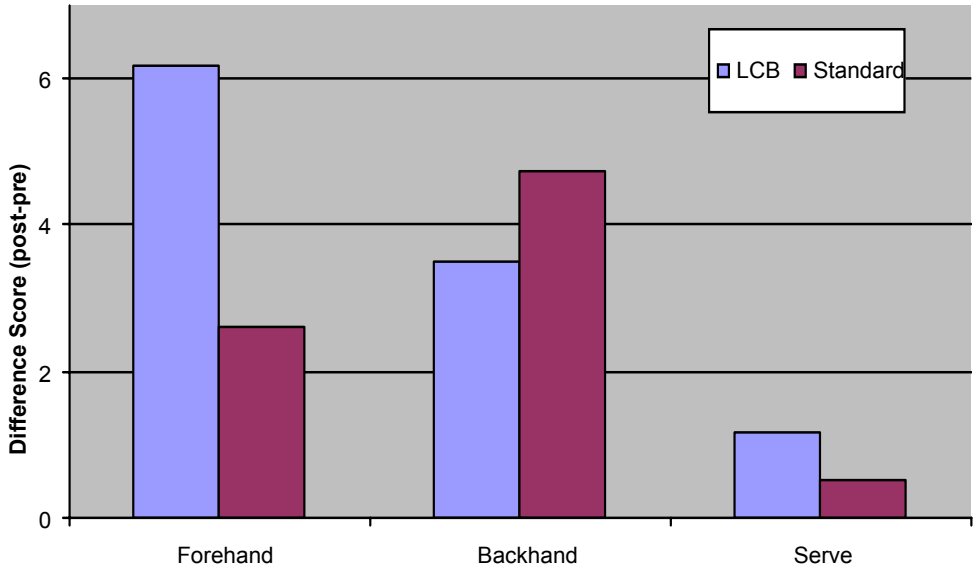


Figure 2: Mean skill test difference scores (pre - post test score) by item for each group

The focus of the interviews was the approach to the coaching sessions by the coaches and the observed effects of the LCBs on the learners' performance. There was agreement between coaches that they had followed a similar programme, which involved a combination of traditional approaches to coaching with a contemporary style, including fun participation games. Additionally activities provided in the TCA coaching manual and from Tennis Australia (TA) seminars were included. To ensure consistency the head coach monitored the sessions, with regular feedback to the coaches.

Regarding effects of the LCBs on learners' performance the coaches felt that the LCBs had the most positive effects on complete novices and the youngest players. Coaches suggested that players were aided by the lower ball bounce if they had no experience playing with the standard balls. For players with previous experience, using the LCBs were reported to have a negative effect on their attitude as these players considered using the LCBs as taking a step backwards. One coach reported that for accomplished beginners their confidence went down initially although this did not seem to be a problem after a couple of weeks. There was general agreement between the coaches that even for the beginners with previous experience the LCBs were good for overall development, especially development of technique. A similar point was observed for the less experienced beginners where one of the coaches indicated that it was easier to teach them the correct technique of hitting from low to high as the LCBs bounce closer to waist height.

The advantages of the LCBs in allowing technique development in children is that they can be taught the correct style, such as stepping in and hitting the ball as well as hitting from low to high with the ball remaining in court. One coach suggested that teaching players to step in and hit when using the standard balls often resulted in the ball being hit out of court as players at this level are not capable of applying the appropriate spin to keep the ball in court. The LCBs were reported as having a positive effect as they allowed correct technique to be taught yet allow rallies to continue and in turn provide greater positive reinforcement to the players. No major gender differences regarding the effect of the LCBs on learning were reported. Whilst interviews with coaches pointed to the benefits of using LCBs for technique development, there was no evidence from this study to suggest that these benefits will transfer to playing with the standard balls.

CONCLUSION

The skill test results indicate there is no difference in skill learning when comparing beginners coached using the LCBs to those coached using the standard balls. The non-significant differences between groups could have also been affected by differences in mean age and previous experience characteristics of the groups and the relatively short amount of practice time between the initial and final skills test. The coaches' reported that the LCBs appeared to have a positive effect on correct technique development in beginners including aspects of technique that are related to improving the power of the shot without the ball going out of court as much as when coaching with the standard ball. Additional benefits were that rallies went on longer providing more playing time and more opportunity for positive reinforcement. In order to provide a more conclusive answer to the effects of

LCBs on skill learning and technique development in beginners, recommendations for future research are suggested in the next section.

There were limitations imposed on this study that contributed to inconclusive results and these should be addressed in future research. Firstly, time and budget constraints determined the study had to be conducted within one school term, with the number of coaching sessions limited to 6x1 hour sessions between pre and post tests, limiting time for skill learning to take place. Future research would benefit from a longer period of coaching to ensure sufficient time for improvement, providing a more accurate assessment of LCB effect. Secondly, although participants were classified as attending the 'beginners' class, there were differing ages and levels of experience between the groups. Ages and playing experience should be standardised across control and experimental groups. Finally, results and observations from this study would suggest a more controlled, longitudinal study would enhance understanding of the effects of LCBs on beginners' skill learning.

Acknowledgements

The authors would like to acknowledge the assistance of the head coach, the coaching staff and the children who took part in the study. This research was funded by Tennis Australia.

REFERENCES

Andrew, D., Chow, J., Knudson, D. and Tillman, M. (2003) Effect of ball size on player reaction and racket acceleration during the tennis volley. *Journal of Science and Medicine in Sport* **6**, 102-112.

Boyle, M. and Ackerman, P. (2004) Individual differences in skill acquisition. In A. Williams and N. Hodges (Eds). *Skill Acquisition in Sport: Research, Theory and Practice.* Routledge, London, 84-92.

Cassell, E. and McGrath, A. (1999) *Lobbing Injury out of Tennis: A Review of the Literature*. Monash University Accident Research Centre. Victoria.

Cayer, L. and Elderton, W. (2002) Teaching techniques. *Development Coaching Manual*. TCA, Richmond.

Cooke, K. and Davey, P. (2005) Tennis ball diameter: the effect on performance and the concurrent physiological responses. *Journal of Sports Sciences* **23**, 31-39.

Cooper, J. (2005) *Buying the Right Length Tennis Racquet for a Junior Player*. Retrieved 02/05/2005, from http://tennis.about.com/od/forjuniors/a/raclengthjr_2.htm

Fairweather, M. (1999) Skill learning principles: Implications for coaching practice. In N. Cross and J. Lyle (Eds.). *The Coaching Process: Principles and Practice for Sport*. Butterworth-Heinemann, Oxford.

Gallahue, D. and Ozmun, J. (1997) *Understanding Motor Development: Infants, Children, Adolescents, Adults*. WCB/ McGraw-Hill, New York.

Gillham, B. (2000) *The Research Interview*. Continuum, London.

Haake, S., Carre, M. and Goodwill, S. (2003) The dynamic impact characteristics of tennis balls with tennis rackets. *Journal of Sports Sciences* **21**, 839-850.

Harding, W. (1991) Elbow pain in young tennis players. *The Physician and Sports Medicine* **19**,135-136.

Hodges, N. and Franks, I. (2002) Modelling coaching practice the role of instruction and demonstration. *Journal of Sports Sciences* **20**, 793-811.

Hughes, M. and Franks, I. (2004) *Notational Analysis of Sport: Systems for Better Coaching and Performance in Sport*. Routledge, London.

ITF. (2005) *ITF Approved Tennis Balls and Classified Court Surfaces*. ITF, London.

Knudson, D. (1993) Effect of string tension and impact location on ball rebound accuracy in static tennis impacts. *Journal of Applied Biomechanics* **9**, 143-148.

LTA. (2005) *Ariel Mini Tennis - Teachers and Coaches Information*. Retrieved 29/4/05, from http://www.mini-tennis.com/teachers_coaches.htm

Mehta, R. and Pallis, J. (2001) The aerodynamics of a tennis ball. *Sports Engineering* **4**, 177-189.

Quezada, S., Riquelme, N., Rodriguez, R. and Godoy, G. (2000) Mini-tennis. *Coaches Review* **1**, 21-24.

Schmidt, R. and Lee, T. (1999) *Motor Control and Learning: A Behavioral Emphasis*. Human Kinetics, Illinois.

Singer, R., Symons Downs, D., Bouchard, L., and de la Pena, D. (2001) The influences of a process verses an outcome orientation on tennis performance and knowledge. *Journal of Sport Behavior* **24**, 213-222.

TCA. (2002) *Development Coaching Manual*. Tennis Coaches Australia, Richmond.

Thorpe, R. and Bunker, D. (1982) From theory to practice: two examples of an understanding approach to the teaching of games. *Bulletin of Physical Education* **18**, 9-15.

THE APPLICATION OF AN EXPLORATORY FACTOR ANALYSIS TO INVESTIGATE THE INTER-RELATIONSHIPS AMONGST JOINT MOVEMENT DURING PERFORMANCE OF A FOOTBALL SKILL

Smith, T.[1], Gilleard, W. [1], **Hammond, J.**[2] and Brooks, L.[3]

[1] Department of Exercise Science & Sport Management, Southern Cross University, Lismore, NSW, Australia

[2] School of Education, Southern Cross University, Lismore, NSW, Australia

[3] Graduate Research College, Southern Cross University, Lismore, NSW, Australia

ABSTRACT

Many studies have investigated the kinematics of sports skills with the majority describing the kinematics of the technique or investigating significant kinematic variables that affect performance. Many sports skills are complex three-dimensional movements involving many joints. However, few studies have investigated the relationships between kinematic variables during performance of such skills. The aim of this study was to investigate the inter-relationships among three-dimensional kinematic variables during performance of a lofted instep soccer kick. A motion analysis system was used to collect kinematic data for 13 skilled amateur soccer players attempting a standardised lofted instep kick. Three-dimensional angular displacement patterns were reported for the thoracolumbar spine and right hip joints. Two-dimensional angular displacement data was reported for the right knee and ankle joints. An exploratory rather than confirmatory factor analysis was applied, as there is currently no established theory regarding the kinematics of a lofted instep kick. Factors were extracted using the Maximum Likelihood Solution and orthogonally rotated using Varimax with Kaiser normalisation. The inter-relationship among biomechanical variables within the seven extracted factors was analysed with each factor revealing previously unknown inter-relationships among variables for different aspects of the kick. The use of exploratory factor analysis has shown the complex three-dimensional kinematic inter-relationships for a lofted instep kick. An understanding of these relationships could prove useful to coaches when instructing, and in the development of coaching programmes related to the lofted instep kick.

KEY WORDS

soccer, kicking, three-dimensional kinematics

INTRODUCTION

The most widely studied skill in football is kicking (Lees and Nolan, 1998), with the majority of studies reporting on the two-dimensional (2D) and three-dimensional (3D) kinematics of the low or maximum velocity instep kick (Lees and Nolan, 2002, Lees et al., 2005, Isokawa and Lees, 1988, Barfield et al., 2002, Levanon and Dapena, 1998, Shan and Westerhoff, 2005). There are many types of kick used in a game of football,

including the lofted instep kick, the aim of which is to propel the ball high and over long distances. Few studies have analysed the 3D kinematics of a lofted instep kick (Browder et al., 1991, Prassas et al., 1990).

An understanding of the biomechanics of kicking can assist the coaching process (Lees, 2003). Coaching experience, combined with knowledge of a mechanical model of the desired performance, is regarded as necessary for a coach to correct performance (Elliott, 2001; Lees, 2002). More studies on the lofted instep kick are needed to provide detailed information on the kinematics of the skill and ensure that existing coaching literature is correct (Prassas et al., 1990). Anderson and Sidaway (1994) analysed the co-ordination of the low instep kick using timing variables and angle-angle plots. Few studies have used a factor analysis (or similar technique) to examine relationships between kinematic variables in kicking (Hodges et al., 2005).

The purpose of this study was to identify and interpret the inter-relationships amongst 3D kinematic variables for a lofted instep kick. As there is currently no established theory regarding the kinematics of a lofted instep kick, an exploratory rather than confirmatory factor analysis was applied to summarise the kinematic data.

METHODS

Thirteen male and female skilled amateur soccer players (23.9 ±6.1 yrs; 74.7 ±12.0 kg; 172.7 ±9.9 cm, previous experience 13.9 ±6.0 yrs), volunteered for the study. During data collection subjects were required to perform 20 trials of a right-foot lofted instep kick. They were required to take a two-step angled approach of 45 - 60° towards a stationary soccer ball and kick the ball over a 2m high net aiming for a target (which represented a kick of approximately 35m). The emphasis of the task was on height and distance not accuracy. Successful kicks were categorised according to distance, 15-27.6m, 27.7-34.9m and 35m+. Twelve retro reflective markers were used to define the thorax, pelvis, thigh, knee and foot, two markers were placed on the ball. Subjects were videoed using a four camera (50 Hz) motion analysis system. Up to three trials from each distance category were selected for further processing and analysis.

Video data of each kick from final toe-off of right foot preceding foot-ball impact to end of active follow-through, was digitised and processed using Peak Motus version 7.0. Spatial data was optimally filtered, the level chosen by the Jackson Knee Point Method and all angles calculated relative to a neutral standing posture. Post impact resultant ball velocity and pre impact resultant foot velocity were calculated manually from scaled co-ordinate data. 3D angular displacement patterns were reported for the thoracolumbar spine (relative motion between thorax and pelvis) and right hip joints. 2D angular displacement data was reported for the right knee and ankle joints. Range of motion (ROM) during follow-through for each joint was also calculated.

The kinematic variables for the thoracolumbar spine, right hip, knee and ankle joints chosen for inclusion in the exploratory factor analysis were: peak angular displacements between toe-off and foot-ball impact; angular displacements at toe-off, heel-strike and impact; and, ROM during follow-through. Additionally angular displacement data was time normalised between toe-off and foot-ball impact and timing of peak values were then reported as a percentage of total kick time. Timing of peaks were included in the analysis as were horizontal, vertical and resultant post-impact ball velocities and

resultant pre-impact foot velocity. Due to an insufficient number of 35m+ kicks in relation to the number of biomechanical variables, kicks of all distances were included in the exploratory factor analysis.

The factor analysis was carried out using SPSS version 11.0. Factors were extracted using the Maximum Likelihood Solution and orthogonally rotated using Varimax with Kaiser normalisation. Cattell's scree test (Stevens, 1996, Kim and Mueller, 1978) was used to determine the number of factors to be extracted. Examination of the scree plot indicated that no more than seven factors should be extracted. Factors were extracted from the rotated factor matrix by selecting variables with a factor loading of ≥ | 0.4 | for inclusion within that factor (Stevens, 1996, Hair et al., 1998). As a result a few variables were common to more than one factor. Timing of peak hip extension, abduction and external rotation, thoracolumbar spine extension, knee flexion, ankle plantar-flexion and ankle ROM during follow-through and foot velocity were poorly represented in the factor solution. As a result they were omitted from the interpretation of each factor.

RESULTS

Table 1 shows that the first seven factors obtained from the following factor analysis accounted for 67.6% of the variance. Factor one appears dominant, accounting for the largest amount variance (19.87%), subsequent factors account for decreasing amounts.

Table 1: Total variance explained by the first seven factors.

Factor	Initial Eigenvalues		
	Total	% of Variance	Cumulative % of Variance
1	11.523	19.868	19.868
2	6.304	10.868	30.736
3	5.679	9.791	40.527
4	4.804	8.284	48.810
5	4.115	7.095	55.906
6	3.539	6.101	62.007
7	3.254	5.610	67.618

Factor one (Table 2) was largely influenced by hip rotation and abduction variables. A decrease in hip joint internal rotation at impact and point of maximum internal rotation (heel-strike to impact) was associated with an increase in external hip joint rotation at toe-off, heel-strike and also peak motion (toe-off to heel-strike). Hip abduction angles at point of maximum (between heel-strike and impact) and impact increased along with the hip joint external rotation variables and hip extension at toe-off. The combination of increased hip abduction, external rotation and extension at toe-off indicated that the more hip extension at toe-off and external rotation in the earlier part of the kick (toe-off to heel-strike), the more hip abduction between point of maximum and impact. Decreased impact angles for hip internal rotation, and increased hip abduction and thoracolumbar spine rotation (thorax to right, pelvis to left) angular displacements, are seen to relate to increased knee flexion, at impact. Thus, hip motion prior to and at impact is associated with knee flexion at impact. This association may be result of compensatory movement by the knee to ensure appropriate foot placement at impact.

Table 2: Summary of inter-relationships for kinematic variables in factor one.

Column 1 Variables increasing / decreasing together	**Column 2** Variables increasing / decreasing together & opposite to those in column 1
Peak hip external transverse rotation	Hip internal rotation @ impact
Peak hip abduction	Hip peak internal rotation
Hip transverse rotation ROM during follow-through	
Hip abduction @ impact	
Thoracolumbar spine transverse rotation ROM during follow-through	
Thoracolumbar spine flexion/extension ROM during follow-through	
Hip external transverse rotation @ heel-strike	
Thoracolumbar spine abd/adduction ROM during follow-through	
Hip external transverse rotation at toe-off	
Hip flexion / extension ROM during follow-through	
Hip extension @ toe-off	
* Thoracolumbar spine transverse rotation (thorax to R, pelvis to L) @ impact	
* Hip abd / adduction ROM during follow-through	
* Knee flexion @ impact	

* Variables loading more strongly on other factors

Factor two (Table 3) indicates that increased knee flexion of the kicking limb at heel-strike is associated with an increase in peak knee flexion (heel-strike to impact) slightly later in the kick. Hip joint abduction at toe-off and heel-strike increases in line with the knee flexion variables suggesting that increased hip abduction in the earlier stages of the kick is related to increased knee flexion later.

Table 3: Summary of inter-relationships for kinematic variables in factor two.

Column 1 Variables increasing / decreasing together	**Column 2** Variables increasing / decreasing together & opposite to those in column 1
Peak knee flexion	Ankle plantar flexion @ toe-off
Knee flexion @ heel-strike	Min ankle plantar flexion
* Hip abd / adduction ROM in follow - through	Max ankle plantar flexion
Hip abduction @ heel-strike	Time of peak thoracolumbar spine transverse rotation: thorax to L, pelvis to R
Hip abduction @ toe-off	Time of min ankle plantar flexion
	Ankle plantar flexion @ heel-strike
	Time of peak hip internal transverse rotation
	Ankle plantar flexion @ impact

* Variables loading more strongly on other factors

Increases in hip joint abduction early on in the kick (at toe-off and heel-strike) also relate to increases in hip abduction/adduction ROM in follow-through. Increases in these hip abduction and knee flexion variables are associated with a decrease in plantar flexion of the ankle throughout the entire kick. A decrease in plantar flexion of the ankle at toe-off is associated with decreased plantar flexion values at heel-strike and impact as well as smaller maximum/minimum values.

Factor three (Table 4) indicates that an increase in thoracolumbar spine adduction (thorax up on right, pelvis down on right) at toe-off is associated with increased thoracolumbar spine adduction throughout the whole kick, and vice versa. In opposition to increases in thoracolumbar spine adduction, hip external rotation at heel-strike and hip abduction at impact decrease. The decreased external hip rotation at heel-strike may be associated with movements of the thorax and pelvic segments that increase thoracolumbar spine adduction in earlier parts of the kick. Also, as the side-to-side tilt of the pelvis influences the magnitude of both thoracolumbar spine and hip abduction/adduction a relationship between the movements of these joints is perhaps logical. The specific variables included in factor three suggest that only hip abduction/adduction variable related to spine adduction was hip abduction at impact.

Table 4: Summary of inter-relationships for kinematic variables in factor three.

Column 1 Variables increasing / decreasing together	**Column 2** Variables increasing / decreasing together & opposite to those in column 1
Thoracolumbar spine adduction (thorax up on R, pelvis down on R) @ toe-off Min thoracolumbar spine adduction (thorax up on R, pelvis down on R) Thoracolumbar spine adduction (thorax up on R, pelvis down on R) @ heel-strike Max thoracolumbar spine adduction (thorax up on R, pelvis down on R) Thoracolumbar spine adduction (thorax up on R, pelvis down on R) @ impact	* Hip external transverse rotation @ heel-strike * Hip abduction @ impact

* Variables loading more strongly on other factors

Thoracolumbar spine extension angles dominate factor four (Table 5), where a decrease/increase in spine extension at toe-off is associated with a change at point of maximum extension (toe-off to heel-strike) and at heel-strike.

Table 5: Summary of inter-relationships for kinematic variables in factor four.

Column 1 Variables increasing / decreasing together	**Column 2** Variables increasing / decreasing together & opposite to those in column 1
Hip extension @ heel-strike * Peak hip extension	Peak thoracolumbar spine extension Thoracolumbar spine extension @ toe-off Thoracolumbar spine extension @ heel-strike Knee flexion / extension ROM in follow-through

* Variables loading more strongly on other factors

The inter-relationships with the remaining variables suggest, that a decrease in thoracolumbar spine extension at toe-off, point of maximum, and at heel-strike is associated with an increase in hip extension at point of maximum (between toe-off and heel-strike) and at heel-strike.

Increases in ball velocities were associated with decreased peak hip extension, hip extension at heel-strike and peak hip external rotation for factor five (Table 6). As peak hip extension and external rotation occur between toe-off and heel-strike, these associations suggest a decreased external rotation in the earlier part of the kick immediately followed by a shorter backswing of the kicking leg as a result of decreased peak hip extension are related to increases in ball velocities. A decrease in external hip rotation suggests a reduced rotation away from the intended flight of the ball in the early stages of the kick is related to increases in ball velocities and reductions in hip extension but external hip rotation variable only accounted for 17.7% of variance on this factor.

Table 6: Summary of inter-relationships for kinematic variables in factor five.

Column 1 Variables increasing / decreasing together	**Column 2** Variables increasing / decreasing together & opposite to those in column 1
Resultant ball velocity	Peak hip extension
Horizontal ball velocity	* Hip extension @ heel-strike
Vertical ball velocity	* Peak hip external rotation

* Variables loading more strongly on other factors

Variables relating to orientation of the thoracolumbar spine during the initial part of the kick dominate factor six (Table 7). The inter-relationships indicate that increased thoracolumbar spine transverse rotation (thorax to left, pelvis to right) at toe-off is associated with increases at peak (toe-off to heel-strike) and heel-strike. Increases in spine rotation in the first part of the kick are also seen to relate to increased hip abduction/adduction ROM in follow-through. Time of minimum thoracolumbar spine adduction (thorax up on R, pelvis down on R) occurs just after toe-off and decreases in association with increases in spine rotation variables. The greater the spine rotation at toe-off the closer to toe-off minimum spine adduction occurs.

Table 7: Summary of inter-relationships for kinematic variables in factor six.

Column 1 Variables increasing / decreasing together	**Column 2** Variables increasing / decreasing together & opposite to those in column 1
Peak thoracolumbar spine transverse rotation (thorax to L, pelvis to R)	* Time of min thoracolumbar spine adduction (thorax up on R, pelvis down on R)
Thoracolumbar spine transverse rotation (thorax to L, pelvis to R) @ toe-off	
Thoracolumbar spine transverse rotation (thorax to L, pelvis to R) @ heel-strike	
Hip abd / adduction ROM in follow-through	

* Variables loading more strongly on other factors

For factor seven (Table 8), the inter-relationships indicates that the greater the knee flexion at toe-off the greater the knee flexion at impact (or vice versa), suggesting that a player who requires more or less knee flexion at impact may also be instructed to increase or decrease knee flexion (as appropriate) at toe-off. However, a decrease in knee flexion at toe-off indicates the players are taking a longer final stride prior to kicking leading to a relative increase in kick time allowing more time to swing the kicking leg backwards and to extend the knee at impact. An increase in knee flexion angles was associated with a delay in the time of maximum thoracolumbar spine adduction (thoracolumbar spine remained adducted the entire kick). The further the pelvic segment was orientated down to the right (thorax up on right, pelvis down on right) decreased the distance between the pelvis and the ground, and more knee flexion may have been required to clear the foot prior to impact.

Table 8: Summary of inter-relationships for kinematic variables in factor seven.

Column 1 Variables increasing / decreasing together	**Column 2** Variables increasing / decreasing together & opposite to those in column 1
Knee flexion @ toe-off	Hip flexion @ impact
Knee flexion @ impact	Time min thoracolumbar spine adduction (thorax up on R, pelvis down on R)
Time max thoracolumbar spine adduction (thorax up on R, pelvis down on R)	Thoracolumbar spine transverse rotation (thorax to R, pelvis to L) @ impact
* Knee flexion / extension ROM in follow-through	Thoracolumbar spine flexion @ impact

* Variables loading more strongly on other factors

Smaller hip joint flexion and thoracolumbar spine flexion and transverse rotation values at impact corresponded to a kick with increased knee flexion at impact, indicating that the body will be in a more upright and more forward position. The time of minimum thoracolumbar spine adduction was inversely related to time of maximum thoracolumbar spine adduction and knee flexion.

DISCUSSION

Kicking is a complex three-dimensional movement and exploratory factor analysis has proved an effective technique for describing and summarising the inter-relationships between the spine and hip, knee and ankle joints of the kicking limb for a lofted instep kick. Interpretation of the seven factors has provided a practical insight into the complexities of the inter-relationships apparent in lofted instep kicking.

Combined with knowledge of ideal characteristics of lofted instep kick performance, the identification of specific associations between similar or different joints in varying planes of motion is of potential benefit to a coach when attempting to improve a player's technique. Interpretation of the factors has allowed the identification of similar variables that increase (or decrease) in association with each other for every measurement throughout the entire kick. If a coach requires a performer to increase or decrease the magnitude of a specific movement at some later point in the kick, knowing these associations indicates whether increases or decreases in the same type of

movement earlier in the kick is likely to contribute to the desired response. Knowledge of these associations will remove the need for the coach to make assumptions regarding the inter-relationships amongst movement patterns and stop erroneous feedback being provided in an attempt to correct the motion.

Likewise, knowledge of positive and negative associations among variables for different joints and / or planes of motion, for similar and opposing phases of the kick, also identified in the preceding analysis, are potentially useful to a coach. These inter-relationships indicate which other movements are likely to be affected if one particular aspect of the kicking action is altered. Having identified a critical aspect of kicking movement to alter, knowledge of other associated movement characteristics provides a coach and performer with the choice of more than one variable to focus on altering. As suggested in the interpretations of some factors, associations between different joints could be due to the influence of a common segment to both, such as the pelvic segment common to thoracolumbar spine and hip joints and thigh common to hip and knee joints, regardless of whether the motion was in the same plane. Further investigation is required to determine the mechanisms of these inter-relationships.

Although factor analysis is a non-dependent statistical process, the interpretation of factor five indicated that decreases in peak hip extension just after toe-off and at heel-strike (suggesting a shorter backswing of the kicking leg) are associated with increased ball velocities. Lees and Nolan (2002) reported increased hip (thigh-trunk) ROM for instep kicks under speed compared to accuracy conditions and Lees et al. (2005) found increases in hip (thigh-trunk) ROM to correlate positively with ball velocity in low maximum velocity instep kicking. The contradictions between the literature and the interpretation of factor five could be due to the differing aims of a lofted instep kick compared to a maximal velocity instep kick. In addition, the hip extension variables discussed did not load very high on factor five, peak hip extension and extension at heel-strike accounting for 23.8% and 20.3% of the variance respectively, suggesting they are of limited importance to the interpretation overall. Further investigation of these inter-relationships is warranted for different types of kick and with differing aims, such as maximal distance, speed or accuracy, to understand the associations between hip extension and ball velocity in lofted instep kicking.

Similarly further analysis is recommended to explore the positive inter-relationship between knee flexion at toe-off and impact partially describing factor seven. An increase in knee flexion at toe-off indicates the players are taking a shorter final stride prior to kicking therefore, it is possible that a relatively shorter kick time will result in with less time to extend the knee in preparation for impact resulting in an increased knee flexion at impact. Isokawa and Lees (1988) suggested there might be two types of kicking patterns for a one-step instep kick. The first involving a large backswing and longer kicking time, the second a small backswing with the lower limb moved forward sharply by knee extension and shorter kicking time. The inter-relationships amongst variables in factor seven suggests two types of kicking action, although further investigation is needed on of expert technique to determine this definitively.

In combination with the existing coaching literature and developments in defining an ‘ideal’ kinematic model of a lofted instep kick, the inter-relationships among variables identified using factor analysis may be used to aid the development of coaching programmes and coaching points. Such knowledge proving particularly useful if the

kinematic variables of interest are difficult to observe or control by the performer, the developed coaching points could then be based on other variables that inter-relate with those deemed critical to performance.

CONCLUSION

The application of an exploratory factor analysis to 3D biomechanical data has revealed previously unknown inter-relationships among variables for different aspects of a lofted instep kick. Interpretation of the factors has shown in detail, the complex inter-relationships that exist. An understanding of these relationships could prove useful to coaches when instructing, and maybe useful in the development of coaching programmes related to the lofted instep kick.

REFERENCES

Anderson, D. and Sidaway, B. (1994) Coordination changes associated with practice of a soccer kick. *Research Quarterly for Exercise and Sport* **65,** 93-99.

Barfield, W., Kirkendall, D. and Yu, B. (2002) Kinematic instep kicking differences between elite female and male soccer players. *Journal of Sports Science and Medicine* **1,** 72-79.

Browder, K., Tant, C. and Wilkerson, J. (1991) A three-dimensional kinematic analysis of three kicking techniques in female soccer players. In C. Tant, P. Patterson and S. York (Eds). *Biomechanics in Sports IX: Proceedings of the 9th International Symposium on Biomechanics in Sports.* ISBS, Ames, Iowa, 95-100.

Elliott, B. (2001) Biomechanics of sport. In F. Pyke (Ed). *Better Coaching: Advanced Coaches Manual.* Human Kinetics, Champaign, Illinois, USA, 171-180.

Hair, J., Anderson, R., Tatham, R. and Black, W. (1998) *Multivariate Data Analysis.* Prentice-Hall, New Jersey, USA.

Hodges, N., Hayes, S., Horn, R. and Williams, A. (2003) Changes in coordination, control and outcome as a result of extended practice on a novel soccer skill. *Ergonomics.* **48**, 1672-1685.

Isokawa, M. and Lees, A. (1988) A biomechanical analysis of the instep kick motion in soccer. In T. Reilly, A. Lees, K. Davids and W. Murphy (Eds). *Science and Football.* E & FN Spon, London, UK, 449-455.

Kim, J.-O. and Mueller, C. (1978) *Factor Analysis: Statistical Methods and Practical Issues.* Sage Publications, London, UK.

Lees, A. (2002) Technique analysis in sports: A critical review. *Journal of Sports Sciences.* **20,** 813- 828.

Lees, A. (2003) Biomechanics applied to soccer skills. In T. Reilly and A. Williams (Eds). *Science and Soccer.* Routledge, London, UK, 109-119.

Lees, A., Kershaw, L. and Moura, F. (2005) The three-dimensional nature of the maximal instep kick in soccer. In T. Reilly, J. Cabri and D. Araujo (Eds), *Science and Football V*. Routledge, London, UK, 64-69.

Lees, A. and Nolan, L. (1998) The biomechanics of soccer: A review. *Journal of Sports Sciences* **16,** 211-234.

Lees, A. and Nolan, L. (2002) Three-dimensional kinematic analysis of the instep kick under speed and accuracy conditions. In T. Reilly, W. Spinks and A. Murphy (Eds). *Science and Football IV*. E & FN Spon, London, UK, 16-21.

Levanon, J. and Dapena, J. (1998) Comparison of the kinematics of the full-instep and pass kicks in soccer. *Medicine and Science in Sports and Exercise* **30,** 917-927.

Prassas, S., Terauds, J. and Nathan, T. (1990) Three-dimensional kinematic analysis of high and low trajectory kicks in soccer. In M. Nosek, D. Sojka, W. Morrison and P. Susanka (Eds). *Proceedings of the VIIIth International Symposium of the Society of Biomechanics in Sports.* Connex, Prague, Czechoslovakia, 145-149.

Shan, G. and Westerhoff, P. (2005) Full-body kinematics characteristics of the maximal instep soccer kick by male soccer players and parameters related to kick quality. *Sports Biomechanics* **4,** 59-72.

Stevens, J. (1996) *Applied Multivariate Statistics for the Social Sciences.* Lawrence Erlbaum Associates, New Jersey, USA.