Spatiotemporal Analysis of Australian Rules Football

Bartholomew Spencer

Institute for Health and Sport (IHeS)

Victoria University, Melbourne, Australia

Principal Supervisor: Assoc. Prof. Sam Robertson

Associate Supervisors: Assoc. Prof. Stuart Morgan, Prof. John Zeleznikow

Thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

April 2019

Abstract

Player tracking data has previously been used to quantify movement profiles in the Australian Football League (AFL), however little research exists into its use to measure the spatial interactions of players. This thesis presents new methodologies for measuring the spatial interactions and occupancy of players in team sports. Global positioning systems (GPS) and local positioning systems (LPS) spatiotemporal datasets were sourced from training sessions, Under-18s matches and elite-level AFL matches. Datasets were consolidated with play-by-play transactions to infer ball position. An initial pilot study investigated the relative importance of traditional performance indicators to inform the focus of later studies. Subsequent chapters investigated the relative phase of inter- and intra-team player couples and multiple approaches to the measurement of the spatial control of individuals. Gaussian mixture models (GMM) were used to estimate the density of player groups in order to analyse changes in congestion throughout a match. Player motion models fit on player displacements were combined with a measure of field equity to value the passing decisions of players. A new approach to player motion models was developed by fitting the weighted distributions of player commitment to contest events. The resultant models more realistically explain player behaviour in proximity to the ball. The models were used to measure the spatial control of teams, from which the spatial characteristics of passes in the AFL were extracted. Passes were clustered into three distinct styles. In the final chapter of this thesis, the models developed in the preceding sections are used to develop a new decision-making model. The expected outcomes of a player's passing options are modelled through consideration of field equity, spatial control, kicking variance and possession outcomes. Using this model, passing decisions from the 2017 and 2018 AFL seasons were analysed. In contrast to previous studies, the value of a player's decision is measured relative to their options, rather than to an increase in possession expectation. This thesis aims to derive insights into player movement behaviour in Australian football. Furthermore, the novel spatial metrics developed in this thesis have applications in player recruitment, coaching, and performance analysis.

Student Declaration

I, Bartholomew Spencer, declare that the PhD thesis entitled "Spatiotemporal Analysis of Australian Rules Football" is no more than 100,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work.

Signature: [redacted] Date: 30/04/2019

List of Abbreviations

A	Acceleration
AF	Australian Football
AFL	Australian Football League
BIC	Bayesian Information Criterion
DV	Decision Value
EO	Expected Outcome
EPV	Expected Possession Value
FE	Field Equity
GMM	Gaussian Mixture Models
GPS	Global Positioning System
HCP	High Coordination Pairs
KDE	Kernel Density Estimation
LCP	Low Coordination Pairs
LPS	Local Positioning System
MSE	Mean Squared Error
PA	Port Adelaide
PDF	Probability Density Function
RF	Random Forest
RR	Reachable Region
RT	Richmond Tigers
SD	Standard Deviation
V	Velocity
WB	Western Bulldogs

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Chapter Overview

This chapter outlines the research scope (Section 1.1) and outcomes (Section 1.2) of this thesis. Section 1.3 outlines the thesis structure and provides an overview of the research contained in each chapter. Finally, Section 1.4 provides the reader with an introduction to Australian Rules football, providing context for the analysis contained in the thesis.

1.1 RESEARCH SCOPE AND SIGNIFICANCE

Spatiotemporal data refers to data that has a spatial and temporal component. In the context of sport, a primary source of spatiotemporal data is player-tracking data collected by wearable tracking devices or computer vision technologies. While this data has been collected in the Australian Football League (AFL) since 2005 (Le Grand, 2007; Wisbey, et al., 2008), limited research has been conducted into its applications. In other team sports, research into its applications has been growing each year (Gudmundsson & Horton, 2017). In this thesis, methods for analysing spatiotemporal data in Australian Rules football are presented.

A review of spatiotemporal analysis in team sports that involved non-trivial computation by Gudmundsson and Horton (2017) posed a number of open questions that warrant additional research in this space. These questions are topics that have yet to be approached by previous studies (Gudmundsson & Horton, 2017). The initial approach of this thesis was to address these topics in the context of Australian Rules football. Where possible, analysis and applications were demonstrated that can be transferred to other invasion team sports, rather than specific to Australian football. In doing so, the aim of this thesis was to focus on the development of novel spatiotemporal analytical methodologies, rather than findings that are specific to Australian football. The topics of focus are as follows:

> **Topic 1.** *The function modelling player motion used in dominant region computations has often been simple for reasons of tractability or convenience. Factors such as physiological constraints of the players and a priori momentum have been ignored. A motion function that faithfully models player movement and is tractable for computation is an open problem.* (Gudmundsson & Horton, 2017)
>
> **Topic 2.** *The existing tools for determining whether a player is open to receive a pass only consider passes made along the shortest path between passer and receiver and where the ball is moving at constant velocity. The development of more realistic model that allows for aerial passes, effects of ball-spin, and variable velocities is an interesting research question.* (Gudmundsson & Horton, 2017)
>
> **Topic 3.** *The definition of spatial pressure in Taki et al.* (2000) *is simple and does not model effects such as the direction the player is facing or the direction of pressuring opponents, both of which would intuitively be factors that ought to be considered. Can a model that incorporates these factors be devised and experimentally tested?* (Gudmundsson & Horton, 2017)

These topics were addressed in the topics of player spatial occupancy and interactions. Spatiotemporal datasets provide researchers with a rich source of information on player locations. One of the key advantages that these datasets have over performance indicators is information on all players. Within the AFL, performance indicators are recorded on-ball events such as kicks, handballs and spoils. Not only do traditional performance indicators lack context (Lucey, et al., 2013), the majority measure on-ball statistics (that is, events that involve the ball in some way). In soccer, for example, it is estimated that the average player spends less than 5% of the match in possession of the ball (Fernandez & Bornn, 2018; Carling, et al., 2007). Hence, analysing player behaviour regardless of possession may yield insights that have applications in performance analysis. Measuring the occupancy and interactions of players is one way that off-ball behaviour can be quantified.

1.2 RESEARCH OUTCOMES

The primary outcomes of this thesis are new methods for measuring the spatial occupancy of players in invasion team sports. This includes metrics to summarise player density (see Chapter 5) and individual player occupancy with consideration of momentum and orientation (see Chapters 6, 7 and 8). These studies build upon previous work that considered space as continuous (e.g., Fernandez & Bornn, 2018; Brefeld et al., 2018), but introduces a new approach that fits empirical player behaviour to models, rather than arbitrary distributions (Fernandez & Bornn, 2018) or displacements regardless of movement context (Gudmundsson & Wolle, 2010; Horton, et al., 2015; Brefeld, et al., 2018).

An applied outcome of the latter studies is a greater understanding of the kicking in the AFL. Kicking has been previously examined using manually collected variables that are often discrete or categorical, such as distance bands (i.e., 0 – 20 metres, 20 – 40 metres, > 40 metres). For examples of this research, see Bedford and Shembri (2006) and Robertson et al., (2019). The metrics used in this thesis will analyse kicking using continuous metrics. This allows for a greater understanding of player decision-making and behaviour prior to passing. Furthermore, in Chapter 3 it was identified that measuring a team's playing style is difficult using performance indicators. A secondary outcome of the analysis presented in the final chapters of this thesis is examples of defining the types of passing styles that exist in the AFL.

1.3 THESIS OUTLINE

The research presented in this thesis is structured into eight chapters (excluding introduction). Each chapter is a paper that has been published or presented at an academic conference. At the end of each chapter is a brief discussion section that links the presented research to the overall themes of this thesis. All conference proceedings and papers are undergoing or have undergone peer review. Levels of continuity vary between chapters; however, the themes remain consistent. Furthermore, different datasets were used in different studies; hence, the scope of each chapter varies. Note that because each chapter was written as an independent paper, there may be repetition between chapters. Each chapter concludes with discussion section linking the chapter's findings to the overall thesis themes.

Chapter 2 is a review of the literature relevant to this thesis. The primary topics of focus are the validity of wearable technologies, statistical analysis in Australian football and approaches to the measurement of spatial occupancy and interactions of players in team sports. Within each topic, the current gaps in knowledge that will be addressed in this thesis are noted. The focus of this review is to identify the methodologies of these topics, rather than their current applications. Overall, there is a strong methodological focus in this thesis. Where applications are demonstrated, the intention is to exemplify the approach.

Chapter 3 is an initial pilot study of performance indicators in the AFL. AFL performance indicators have been the primary data used in most analytical studies, given the availability of data. To refine the focus of applications in later chapters, the relative importance of performance indicators is examined in models explaining match outcomes. While not spatiotemporal analysis, the objective here is to identify aspects of AFL performance that warrant further research using spatiotemporal datasets. Additionally, differences in team profiles and the relationship to success are explored. An objective is to demonstrate the inadequacy of performance indicators that lack context.

Chapter 4 is, to my knowledge, the first study in Australian football that utilises spatiotemporal datasets of both teams. The use of player-tracking data is exemplified in the measurement of co-coordinative behaviour. Relative phase of player couples is measured via the phase angles for angular velocity and acceleration. This builds upon the work of Morgan and Williams (2012) by measuring the relative phase of inter-team player couples. The findings of this study are used to cluster the coordination of player couples using *k*-means analysis.

Chapter 5 represents the beginnings of work into the measurement of spatial occupancy in the AFL. Gaussian mixture models (GMM) are used to estimate the density of player groups throughout a match. Additionally, a metric for measuring overall congestion is presented. This approach is limited to measuring density without consideration of player orientation and motion. Hence, this chapter serves as a brief introduction to the topic before more advanced methodology is presented in later chapters.

In Chapter 6 an initial decision-making model is presented. Spatial occupancy in this chapter is measured using discretely bound player motion models, where bounds are equivalent to the maximum displacements observed by players in velocity and time intervals. A team's spatial dominance is measured on a continuous scale, where dominance at a location is modelled using player motion models whose size depends on proximity to the ball.

Identifying the limitations of the motion model in Chapter 6, a new approach to the topic of player motion is presented in Chapter 7. Spatial occupancy is estimated via player commitment models which quantify the probability players would reposition to forthcoming contests. This forecast of future behaviour can be used to estimate the pressure they would apply to future passing contests. These measures are used to analyse passing types following player marks.

Chapter 7 presents an alternative decision-making model using the player motion described in Chapter 6. In addition to occupancy, equity and kicking variance, multiple passing outcomes are considered for each of a player's passing options. Kicking decisions for individuals and teams are analysed across the 2017 and 2018 AFL seasons. Future applications and evolutions of decision-making work in the AFL are discussed.

Finally, Chapter 8 summarises the work of this thesis and discusses future directions of spatiotemporal analysis. The research presented in this thesis is linked to the research topics identified in section 1.1 and discuss the additional contributions beyond said topics.

1.4 INTRODUCTION TO AUSTRALIAN RULES FOOTBALL

This section summarises the rules and terminology of Australian Rules football to provide context for the research presented in this thesis.

1.4.1 Overview

Australian Rules football is a popular team sport played in Australia. There are multiple leagues, the most important being the national league known as the Australian Football

League (AFL). Gameplay consists of two teams of 22 players, 18 of whom are on the field at any time and the remaining four are interchange players. Notably, teams can make up to 90 interchanges throughout a match. Gameplay occurs over four approximately 20-minute quarters.

The primary objective of the sport is to outscore the opponent by scoring *goals* and *behinds*, worth six and one point respectively. Goalposts are situated at either end of the field. Kicking the ball through the inner most pair of goalposts awards a goal. Moving the ball between the outer goalposts awards a behind.

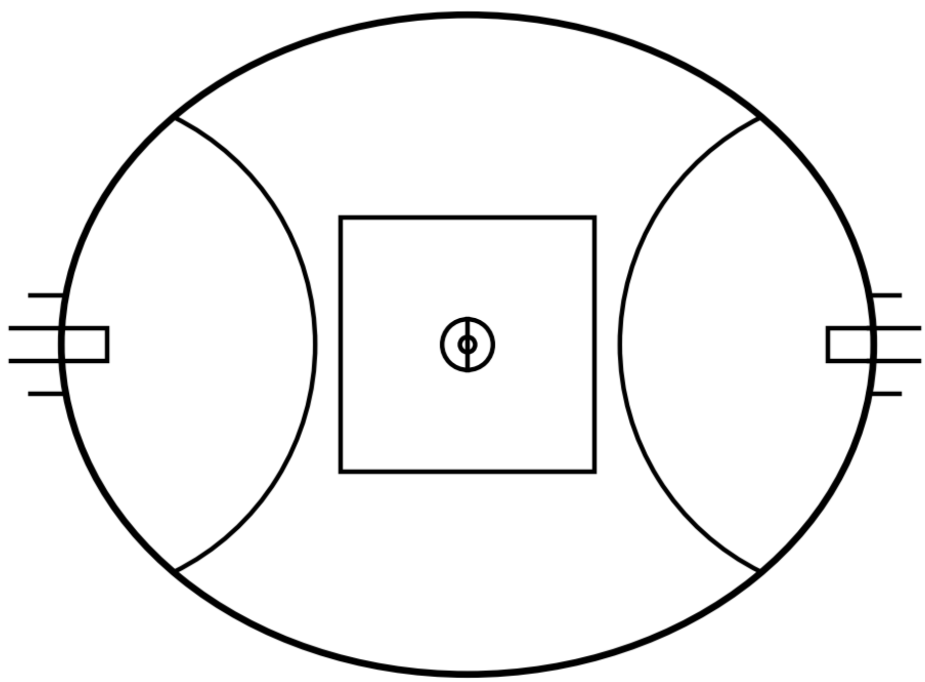


Figure 1-1. Australian football playing field.

The sport is played on an oval field of varying dimensions. Data used in this thesis were primarily collected at Docklands Stadium, Melbourne. The Docklands Stadium playing field is approximately 160 metres long and 130 metres wide (Fig. 1-1). The field is partitioned into three regions: two 50-metre arcs at each end of the field and a central region. Within the central region is the centre square, the middle of which is where play resumes via a centre bounce following a goal. If the last score was a behind, possession is handed to the defending team and play is resumed from the defenders' goalposts.

The large field dimensions combined with a lack of offside rule result in congested gameplay. Gameplay will typically consist of players following a pack, rather than holding positions as is the case in some other sports. This results in dense gameplay in which players have frequent interactions. As a result, the spatial interactions and occupancy of players is particularly interesting.

1.4.2 Terminology

The following section provides definitions for the Australian football terminology used in this thesis.

Behind A score worth one point. Awarded when the ball is moved between the outer goalposts, or if the ball is moved through the inner goalposts by method other than a kick from the attacking team.

Contest An event in which more than one player is attempting to win possession of the ball.

Contested Mark A mark awarded to a player who was under pressure from one or more opponents.

Disposal A pass (either a handball or kick) to give away possession.

Field Equity A measure of possession expectation given the current pressure phase and location of the ball. Specifics are described in O'Shaughnessy (2006).

Goal A score worth six points, awarded for kicking the ball through the inner goal posts.

Handball A pass executed by a player's hands, rather than a kick. A legal handball requires a player to punch the ball out of their hand. The common passing technique in other ball sports (e.g., Rugby Union) is considered an illegal disposal in Australian football.

Inside 50 Moving the ball into the 50-metre arc around the attacking goalposts.

Mark Awarded when a player receives the ball on the full after a kick that has travelled at least 15-metres. After a mark, there is a zone around the marking player that cannot be entered by opposition players. The marking player can then take their time before kicking or playing on.

Mark Play-On Playing on following a mark.

Possession Receiving and maintaining possession of the ball before a disposal.

Spoil Knocking the ball away from an opponent who is attempting a mark.

While there are many more definitions in Australian football, the above are important to the research contained in this thesis. More detailed definitions are located at the AFL website[1].

[1] https://www.afl.com.au/news/2017-12-28/stats-glossary-every-stat-explained

Chapter 2: Literature Review

Chapter Overview

This chapter introduces literature relevant to the collection and analysis of data in the AFL. While player-tracking technology has existed for a number of years in Australian Rules football, there has been minimal research into its applications beyond movement analysis (Foreman, et al., 2012). Through analysis of existing work, gaps in existing knowledge that could be addressed via spatiotemporal player tracking data are identified.

Current approaches to sports analytics in Australian Rules football are covered to emphasise the limited spatiotemporal analysis in this space. For other team sports, spatiotemporal analysis techniques that are relevant to the themes of this literature are discussed.

The review contains literature on the validity of wearable tracking devices (Section 2.1), Australian Rules football (Section 2.2) and spatiotemporal analysis in team sports (Section 2.3).

2.1 VALIDITY OF WEARABLES

Multiple studies have researched the validity of wearable global positioning system (GPS) and local positioning system (LPS) devices. Edgecomb and Norton (2006) compared the distance recordings of GPS systems in Australian Rules football, concluding that they overestimate distances by 4.8%. Subsequent literature documented the variability between individual devices (Jennings, et al., 2010) and inconsistencies in real-time (Aughey & Falloon, 2010). Inconsistencies of up to 10% were theorised to be the cause of limitations of in-play algorithms offered by GPS software (Aughey & Falloon, 2010). GPS devices recording at 10 Hz have been found to be up to six times more accurate at detecting velocity changes than 5 Hz devices (Varley, et al., 2012). However, 10 Hz devices have inaccuracies in accelerations above $4m/s^2$ (Akenhead, et al., 2014) and an increase to 15 Hz doesn't improve the validity of measurements (Johnston, et al., 2014). Despite these findings, 5 Hz recordings have proven adequate for most movements (Portas, et al., 2010; Hurst & Sinclair, 2013).

LPS wearable devices have been found to have adequate validity in indoor and outdoor settings (Sathyan, et al., 2012; Ogris, et al., 2012; Serpiello, et al., 2018). A recent study by Hoppe et al. (2018) compared GPS and LPS technologies for measuring distances in team sports, finding that 20 Hz LPS devices have superior validity and reliability compared to 10 Hz GPS (Hoppe, et al., 2018).

The data used in this thesis are sourced from 10 Hz LPS and GPS devices. While research has shown overestimations for distances, 10 Hz devices are generally considered to be adequate for most purposes (Akenhead, et al., 2014; Johnston, et al., 2014; Hoppe, et al., 2018). Furthermore, wearable devices produce more accurate and easier to analyse data than alternative optical technologies in Australian Rules football (Edgecomb & Norton, 2006).

An alternative source of player tracking data is optical tracking computer vision (CV) systems (Gudmundsson & Horton, 2017). While data from these sources may be more accurate than wearable technologies, these tracking systems can be limited by infrastructure and capture spaces (Barris & Button, 2008). An advantage of these systems is ball tracking (Thomas, et al., 2017), however player collisions and obstruction can limit their accuracy (Gudmundsson & Horton, 2017). In a study comparing CV, GPS and LPS tracking systems compared to VICON measurements, it was found that LPS systems were superior for measuring position (23±7 cm) than CV (56±16 cm) and GPS (96±49 cm) (Linke, et al., 2018). Furthermore, LPS (0.25±0.06 m/s) and GPS (0.28±0.07 m/s) were superior for measuring speed compared to CV (0.41±0.08 m/s) (Linke, et al., 2018). Finally, total distance covered during small sided games had lower errors for GPS (2.2%) and CV (2.7%) than LPS (4.0%) (Linke, et al., 2018).

It should be noted that the methods presented in this thesis can be employed on player tracking data from all sources. The only requirement for the reproduction of these methods is data in the (x, y, t) format. Models were developed with this transferability in mind.

2.2 AUSTRALIAN RULES FOOTBALL

2.2.1 Movement Analysis

Movement analysis has been a common theme in Australian football literature since before the development of wearable technologies. Early research introduced position profiles, however results were limited by the need for manual recordings (Jacques & Pavia, 1974; McKenna, et al., 1988; Dawson, et al., 2004).

The introduction of wearable GPS devices produced more significant literature. Position profiles have been developed based on match demands (Wisbey, et al., 2008; Wisbey, et al., 2010). These are defined by speed zones, continuous efforts, accelerations and surge demands for each position. Nomadic players were found to work harder than forwards and defenders (Wisbey, et al., 2010). Heasman et al. (2011) noted inconsistencies in the data used in previous match demand studies, thus reproduced movement profiles using consistent tracking devices. The findings of this study reflected those of previous research (Heasman, et al., 2011). GPS devices have been used to measure the movement demands of different leagues (Brewer, et al., 2010), types of movements (Coutts, et al., 2010) and between matches (Gray & Jenkins, 2010; Kempton, et al., 2013). Neville et al. (2012) noted the lack of GPS analysis in traditional performance analysis, proposing the use of GPS data to monitor player recording, training conditions and longitudinal performance using speed and distance data. Sullivan et al. (2014) analysed activity profiles in relation to play time and outcome, finding that the physical demands of the losing team are higher than those of the winning. Gronow et al. (2014) analysed the movement profiles of teams while in offensive and defensive phases, finding that speeds while on offense were lower for successful teams, and during defensive phases, successful teams spent greater time at higher speeds than unsuccessful teams. Overall, successful teams had greater possession (Gronow, et al., 2014).

Player movement data in AFL has been used to develop training drill classification systems in terms of their physical requirements (Corbett, et al., 2017a). The physical requirements of training drills were represented by high intensity running per minute, drill duration and high-intensity running as a percentage of total distance (Corbett, et al., 2017a). Three systems were developed to compare drills – *k*-means clustering measured the similarity between drills, z-scores compared drills to match conditions and a 'specificity index' was calculated from the z-scores (Corbett, et al., 2017a).

More recently, player movement data was analysed using linear mixed models to identify factors that affect performance (Ryan, et al., 2017). It was found that a high number of rotations, playing against strong opponents or winning results in small to moderate increases in total running distance (Ryan, et al., 2017). Furthermore, player involvements were found to have a relationship with meterage per minute (Corbett, et al., 2017b). In this study, Generalised Linear Mixed Models were used to identify relationships between involvements and movement features (Corbett, et al., 2017b). These included meterage per minute, high intensity running (> 14.4 km/h) per minute and very high intensity running (> 25 km/h) per minute. Weak relationships were noted between skilled involvements and high intensity running per minute and very high intensity running per minute (Corbett, et al., 2017b).

Despite the prevalence of movement analysis literature in Australian Rules football, Foreman et al. (2012) questioned the use of GPS devices in the sport, stating that their current use neglects positional data and doesn't improve performance. Neville et al. (2012) shared a similar sentiment, noting the literatures' emphasis on speed rather than location. Hence, there exists opportunities to explore the use of analytical techniques that consider player locations and interactions, rather than summarised movement data.

2.2.2 Statistical Analysis in the AFL

A variety of match statistics are collected for each AFL match (CIA, Champion Data Pty Ltd). These include play-by-play match event data (known as *transactions*) that contain approximate field positions (Jackson, 2016). Most analysis of match statistics has been applied in the prediction of match outcomes (e.g., Stefani & Clarke, 1992; Clarke, 1993; Bailey, 2000; Bailey & Clarke, 2004; Jackson, 2017), however there are studies that consider possession outcomes (e.g., Meyers et al., 2006; O'Shaughnessy, 2006; Jackson, 2016; Ryall, 2008) and player performance statistics (e.g., Stewart et al., 2007; Robertson et al., 2015; Robertson et al., 2016).

Match Prediction

Computer forecasting has been found to outperform human tipsters in the AFL (Clarke, 1992). A predictive model that considered team ability, team interactions and a common home ground advantage was found to outperform tipsters in the 1991-1992 AFL seasons (Clarke, 1992). The continued development of predictive models has been a research focus.

A dynamic home ground advantage has been shown to improve match predictions (Stefani & Clarke, 1992). Matches played between 1980 and 1989 were analysed, finding that the home ground advantage was greatest for West Coast Eagles (Stefani & Clarke, 1992). Notably, this was the only team from Western Australia that played in these seasons. Home ground advantage in the AFL was hypothesised to be due to travel fatigue (physiological effect), fans (psychological effect) and familiarity of playing conditions (tactical effect) (Stefani & Clarke, 1992).

Team level match statistics have been used to identify arbitrage opportunities in AFL betting markets (Bailey, 2000). Past margins, turnovers and inside 50s were found to be important (Bailey, 2000). While predictions were based on team-level statistics, Bailey (2000) suggested that player-level statistics should be examined to identify if a team is fielding a weaker side. Predictive models have incorporated player information. The frequency of interactions between players (e.g., a kick from player A to player B) have been measured and input into a linear predictive model that predicts match margins based on a team's players (Sargent & Bedford, 2013). More recently, player rating and injury data were used to simulate results in AFL (Jackson, 2017). Player ratings measure the value of on-ball contributions as the change in field equity between the beginning and end of a player's involvement (Jackson, 2016). The 2017 AFL season was simulated assuming each player had a 1/18 probability of an injury, the length and severity of which were modelled from empirical injury data (Jackson, 2017). The model successfully predicted the result of future matches in 72% of cases (Jackson, 2017).

In-game Predictions

In-game predictions have also been researched. Ryall (2011) incorporated interchange data and score margins to track the winning probability of teams through an AFL match. This was measured with linear and binary logistic regression models (Ryall, 2011). Continuous time Markov chains have been used to analyse transitions between match events (Meyer, et al., 2006). Resultant transition matrices can be used to predict the next event, including distance and time between events (Meyer, et al., 2006). This study analysed four events from the 2004 AFL season, hence limited their transitions to seven common events (Behind, Ball-Up Bunce, Centre Ball Up, Handball, Kick In, Kick, Throw In) (Meyer, et al., 2006). Expanding this methodology to more detailed transactions (e.g., kicking type) would require further data. In O'Shaughnessy (2006), *match equity* is the probability of a team winning at the current moment. A component of this is *field equity* which computes the expected result of the current possession, given the location and qualitative pressure of the current transaction (O'Shaughnessy, 2006). Jackson (2016) smoothed the field equity metric from O'Shaughnessy (2006) with regression splines and used this to value player contributions through a match. The value of a possession depends upon the number of players involved in the possession and the change in field equity between the beginning and end of the possession (Jackson, 2016). A player's rating is the sum of their possession contributions (Jackson, 2016). These contributions were found to be highly correlated with score differential (r = 0.96) (McIntosh, et al., 2018). These studies represent the beginnings of research in the AFL that incorporated spatial and temporal data. However, it should be noted that the spatial data in these studies is manually collected and no tracking data of individual players was used. Consideration of teammate and opponent locations may improve the accuracy of these metrics and has been researched in other team sports (e.g., Cervone, et al., 2016).

Goal Kicking

As the primary method of scoring points in AFL, multiple studies have analysed goal kicking. A study by Clarke and Norman (1998) examined the mathematics behind the decision-making of defenders, identifying situations in which it is preferable for defenders to rush a behind to prevent a potential goal. Supposing a given number of decision epochs or 'stages' remaining in a match and the score margin, this model identified scenarios where rushing a behind (a situation where a defender carries the ball over their own goal line) is preferable. It was found that rushing a behind is preferable in tight margins with enough stages remaining to score a goal (when on the losing team) or when there are not enough stages remaining for another score (when on the winning team) (Clarke & Norman, 1998). For large margins, rushing a behind is only preferable if greater than 20 stages remain (Clarke & Norman, 1998).

Since the 2007 season, 61% of shots on goal have been successful in the AFL (Andreson, et al., 2018). The addition of spatial data provides a greater understanding of goal kicking accuracy (Bedford & Shembri, 2006; Andreson, et al., 2018). Kicking accuracy is lower at greater distances and angles from the goal posts for all shots at goal (Bedford & Shembri, 2006), as well as for set shots at goal (Andreson, et al., 2018). Additionally, it has been shown that there is not a statistically significant difference (p >

0.05) between the accuracy of winning and losing teams (Andreson, et al., 2018). While this analysis incorporated positional information, data were partitioned into zones based on manual assessment of position (Bedford & Shembri, 2006; Andreson, et al., 2018). The addition of precise spatiotemporal data introduces new metrics, such as computing spatial pressure via density. Extracting this information from tracking data rather than recording it manually removes the possibility of subjective or inaccurate data. It has previously been shown that multiple factors affect the result of kicks in the AFL (Robertson, et al., 2019).

AFL Performance Indicators

More recently, the importance of match statistics has been assessed via post-hoc match predictions, or explanations. Stewart et al. (2007) produced regression models to identify the importance of 51 performance indicators. Variable importance was used to summarise player contributions across a season of the AFL (Stewart, et al., 2007). The least important variables were removed to produce a final model consisting of 12 performance indicators – kicks, long kicks (> 40 m), kicks to contests, kicks to open players, kicks to opponents, ball-up clearances, centre bounce clearances, bounces (while running with the ball), knocks, handballs and ball gets (Stewart, et al., 2007). It is noted that six of the 12 important performance indicators related to passes (five kicking and one handball). Differences in performance indicators between quarter outcomes have been examined, finding that skill involvements (e.g., disposals, kicks, marks) were greater in winning quarters (Sullivan, et al., 2013). This study also found that high speed running (>19.8 km/h), sprints and peak speeds were higher in losing quarters, and that there was increased physical activity in quarters with small score margins (Sullivan, et al., 2013). Noting the former studies' use of linear approaches, Robertson et al. (2016a) used 17 performance indicators to explain match outcomes using logistic regression and decision trees. This study found that there are multiple winning performance indicator profiles (Robertson, et al., 2016a). The study's use of relative performance indicators provides match context to the reported values. The most important variables in the models were kick and goal conversation values relative to opposition (Robertson, et al., 2016a).

Performance indicators have applications in the measurement of individual player performance. Sullivan et al. (2014) used physical activity profiles and player statistics to analyse player performance. The objective was to identify what factors contribute to coaches' subjective perception of player performance (Sullivan, et al., 2014). Stepwise multiple regression analysis revealed that skill-based performance characteristics (e.g., player rank, kicks, handballs, bounces) accounted for 42% of variance in coaches' perception of player performance (Sullivan, et al., 2014). Using the performance indicators exemplified in previous studies, Robertson et al. (2016b) assessed the within-team distribution of player performances to identify the optimal makeup of player skill. A model explaining match outcome found that only eight features contributed meaningfully to the model (Robertson, et al., 2016b). These were predominantly related to goals and disposals (Robertson, et al., 2016b). In general, it can be seen that offensive performance indicators relating to goal scoring and passing have been most important to predictive models (Robertson, et al., 2016a; Robertson, et al., 2016b; Stewart, et al., 2007; Sullivan, et al., 2014).

AFL performance indicators have also been used to explain ladder position of teams (Woods, 2016) and to analyse changes in game-play (Woods, et al., 2016). Hit-outs, clearances and inside 50s were significantly associated with final ladder position ($p < 0.05$) across an AFL season (Woods, 2016).

2.3 SPATIOTEMPORAL ANALYSIS IN TEAM SPORTS

There has been limited spatiotemporal analysis in team sports outside of basketball and soccer (Gudmundsson & Horton, 2017). The transfer of methodologies exemplified in these sports to other team sports was suggested as an open research topic by Gudmundsson and Horton (2017). In this section, literature relevant to the topics of spatial interactions and occupancy are discussed. The research presented in this thesis will continue the development of techniques in these topics.

2.3.1 Team Spatial Metrics

Recent applications of spatiotemporal tracking data have described the collective behaviour of teams (Memmert, et al., 2017). Metrics have been developed to summarise the spatial occupancy of teams in terms of their dispersion, surface area and width, reducing the complexity of spatiotemporal datasets (Memmert, et al., 2017). These metrics were developed for use in soccer (Memmert, et al., 2017). Notably, most literature in this space has used datasets from basketball and soccer (Gudmundsson & Horton, 2017). Hence, there exists many opportunities to transfer these techniques to similar team sports (Gudmundsson & Horton, 2017).

A team's *surface area* (or *playing space*) describes the amount of space that a team occupies at a specific point in time (Frenken, et al., 2011). Computing a team's surface area involves fitting a convex hull around its players (Frenken, et al., 2011). The area within the convex hull is the sum of the cross products of each vertex (known as the shoelace formula).

The team *centroid* is the centre position of a team's players (Memmert, et al., 2017). This has been calculated by taking the mean *x*-, *y*- position of all players (Frenken, et al., 2011), of all players excluding the goalkeeper (Frenken, et al., 2011), a weighted average considering player proximity to the ball (Clemente, et al., 2014) and of the centre between the farthest players (Lames, et al., 2010). From the centroid, team *dispersion* can be measured as the average dispersion between a team's centroid and its players (Clemente, et al., 2012a).

These metrics have been used to analyse many aspects of soccer. Some examples include differences in offensive and defensive phases (Vilar, et al., 2013; Castellano & Casamichana, 2015; Castellano, et al., 2013; Yue, et al., 2008; Clemente, et al., 2013), pitch sizes (Frenken, et al., 2011; Frencken, et al., 2013), small sided games (Aguiar, et al., 2015; Folgado, et al., 2014) and the behaviour of different cohorts (Clemente, et al., 2012a; Goncalves, et al., 2013).

While these metrics have revealed insights into the tactical behaviour of teams, they do so at a macro level. It has been suggested that summarising spatial information lacks important contextual information (Bialkowski, et al., 2014). For example, team surface area can be influence by outliers as it does not consider the density of playing groups. Furthermore, a team's centroid provides minimal information on formations. Density may be a more appropriate measure of a team's spatial formation.

2.3.2 Coordinative Behaviour

The use of spatiotemporal data allows for the analysis of player movements beyond those involved directly with the ball. Building on previous work that explored relations in squash (McGarry, et al., 1999; McGarry, et al., 2002) and tennis (Pault & Zanone, 2005; Lames, 2006), Bourbousson and colleagues used phase relations, derived from the lateral and longitudinal position of players and team centroids, to examine spatiotemporal coordination in basketball between players (Bourbousson, et al., 2010a) and teams (Bourbousson, et al., 2010b). The relative phase of couples was computed using a Hilbert transformation, which measures the phase difference between two time-series (Pault & Zanone, 2005). Furthermore, defensive pressure in basketball promotes different coordinative behaviour amongst teams (Leite, et al., 2014). The relative phase of footballers has been measured via distance to team centroids (Sampaio & Macas, 2012). Morgan and Williams (2012) suggested the existence of coordinative behaviour not limited to proximity, measuring the coordination between player couples using the relative phase of acceleration and angular velocity in field hockey. Relative phase has also been computed between player couplings with the ball in futsal (Travassos, et al., 2012). It was found that defending teams had stronger phase relations with the ball than attacking teams (Travassos, et al., 2012).

Analysing the synchronisation between large groups has been achieved using cluster phase analysis (Frank & Richardson, 2010). Duarte et al. (2013) assessed the synchronisation of two teams in an English Premier League match using cluster phase analysis. This was achieved by collecting player displacements and it was found that teams had greater synchrony in the longitudinal direction (Duarte, et al., 2013). More recently, cluster phase analysis has been used to analyse the synchronisation of players by playing position and match phase (Lopez-Felip, et al., 2018). It was found that team synchrony was higher when in a defensive phase (Lopez-Felip, et al., 2018).

While inter-team pairings have been explored in team sports (Bourbousson, et al., 2010b), this has previously been measured by proximity. Hence, there has been limited research into inter-team coordinative behaviour in team sports.

In Australian Rules football, Sargent and Bedford (2013) analysed passing frequency between player couples to predict the outcomes of matches based on team synergies, however no spatiotemporal data was used in this study.

2.3.3 Motion Models

Player motion models of varying complexity have been introduced to model the interaction of velocity, acceleration and orientation on a player's future displacements. Taki and Hasegawa (2000) modelled player motion using movement and acceleration vectors but considered acceleration as a fixed variable and did not consider deceleration. An alternative motion equation that added a resistive force to decrease acceleration was proposed by Fujimaru and Sugihara (2005). Using this model, a player can reach any point in a radius (r) around a centre point (x) as follows:

$$x = x_o + \frac{1 - e^{-\alpha t}}{\alpha} \cdot v_0 \qquad r = v_{max} \cdot \frac{1 - e^{-\alpha t}}{\alpha}$$

where x_o is the player's starting position, v_0 and v_{max} are their velocity and maximum velocity respectively and α is the magnitude of resistance. Fujimaru and Sugihara (2005) estimated α and v_{max} to be 1.3 and 7.8 m/s respectively. In their review of spatiotemporal analysis in team sports literature, Gudmundsson and Horton (2017) consider this model a more realistic approximation of motion than the model in Taki and Hasegawa (1998).

Noting the limitations of physics-cased motion equations, Gudmundsson and Wolle (2010) produced motion models fit on the observed displacements of players within time and velocity bands. The (x, y) co-ordinates of player displacements (relative to orientation) were extracted for player movements over whole-second periods. While three smoothing process were exemplified, smoothing via Kernel density estimation was found to produce performance closest to being optimal (Horton, 2013). A recent study detailed the process of modelling player displacements and compared its performance to physics-based equations for player motion (Brefeld, et al., 2018). Player movement was grouped into speed categories – stand (< 1 km/h), walk (1 – 7 km/h), jog (7 – 14 km/h), run (14 – 20 km/h) and sprint (> 20 km/h). Grouping speed and velocity into bands reduced model complexity and computation time (Brefeld, et al., 2018). Compared to physics-based models, the density of player displacements produces more realistic measures of future player movement (Brefeld, et al., 2018; Horton, 2013).

More recently, deep learning has been used to model player movement relative to player locations. This process increases the dimensionality of player behaviour models (Le, et al., 2017), hence is less interpretable for decision makers. Deep imitation learning was used to model average player behaviour in response to a team's movements in soccer (Le, et al., 2017). Consideration of dynamic playing position produced more accurate models (Le, et al., 2017). Increasing the dimensionality and complexity of player motion models requires large amounts of data. The deep imitation learning model from this study was fit on 100 soccer matches (Le, et al., 2017). While these models measure player behaviour with consideration of teammate positions and match phase, this data requirement limits its implementation.

Player motion in this thesis will be fit on significantly less data, hence deep learning was not used. Displacement-based motion models were used as a starting point to the measurement of player motion in AFL. These models lack movement context; hence a new approach was developed to consider the context of player displacements.

2.3.4 Spatial Occupancy

The region in which an individual can reach earlier than any other individual is known as their dominant region (Taki & Hasegawa, 2000; Gudmundsson & Horton, 2017). Dominant regions are similar to Voronoi tessellations (Okabe, et al., 1992), but typically consider player orientation and motion. While the addition of motion results in more realistic regions of control (Taki & Hasegawa, 2000), some studies have produced dominant regions from player proximity alone (e.g., Rein et al., 2017; Cervone et al., 2016b).

The concept of dominant regions was introduced by Taki et al. (1996). The model proposed in this study was used to evaluate teamwork in soccer (Taki & Hasegawa, 1998; Taki & Hasegawa, 2000). A more realistic motion model that considers deceleration by Fujimaru and Sugihara (2005) has been used to produce dominant regions in more recent studies. In soccer, studies of dominant regions found a difference in the regions of offensive and defensive teams (Fonseca, et al., 2012) and that successful offensive phases had thinner dominant regions than unsuccessful phases (Ueda, et al., 2014). More recently, Voronoi tessellations with no consideration of player motion were used to assess the value of soccer passes relative to a team's spatial control (Rein, et al., 2017). Voronoi tessellations can be combined with field equity metrics to derive the relative value of a player's occupancy (Cervone, et al., 2016b). In Cervone et al. (2016b), the value of a court location was calculated as the frequency in which players occupied said location.

The computational complexity of dominant regions was addressed by Nakanishi et al. (2009) who demonstrated a method for approximating regions they called *reachable polygonal regions* (RPR). Dominant regions are approximated via the intersections of player bands using RPR, rather than computing which player is dominant for each field location (Nakanishi, et al., 2009). Using RPR, Gudmundsson and Wolle (2010) produced dominant regions using player displacements, rather than physics-based motion models. The regions from this study were used to automate the analysis of passing in soccer (Horton, et al., 2015; Chawla, et al., 2017). When classifying passes on a qualitative scale, it was unclear if dominant regions were important to the classification algorithms (Chawla, et al., 2017).

Common amongst these approaches is consideration of spatial occupancy at a fixed point in time. It has been suggested that dominant regions do not correlate to player contributions and could be improved by using weighted dominant regions that consider field locations or proximity to the ball (Fujimura & Sugihara, 2005). Should the application of spatial occupancy be in respect to possession outcomes, occupancy need be considered relative to the ball and passing outcomes.

A recent study by Fernandez and Bornn (2018) suggested that space is continuous, hence occupancy should be measured as such. Occupancy was measured using bivariate normal distributions to produce a smooth surface of control (Fernandez & Bornn, 2018). Furthermore, occupancy was valued relative to a team's desire to occupy regions given the current ball location (Fernandez & Bornn, 2018). While this dynamic approach to occupancy is more logical, player motion in this study is fit with arbitrary bounds, rather than empirical data. Hence, there exists opportunities to apply realistic motion models in the measurement of spatial occupancy.

2.3.5 Player Roles and Team Formations

There have been a number of studies that identify the formations of teams. Within this topic, a focus of many studies is the assignment of dynamic playing roles.

In Lucey et al. (2014b), role assignment in basketball was used to identify role-swapping. While players will have a defined role, players will dynamically swap roles during plays in basketball (Lucey, et al., 2013). Player role was automatically classified by a supervised model fit on manual role assignment by human experts (Lucey, et al., 2013). A shooter is more likely to be open (i.e., low pressure) if role-swaps precede a three-point shot (Lucey, et al., 2014b). The approach used in Lucey et al. (2014b) has been used to cluster team formations into six groups (Bialkowski, et al., 2014). Automated labelling of the formations in Bialkowski et al. (2014) were found to align with manual labelling by experts. In Wei et al. (2015), the player role assignment process from Lucey et al. (2013) was used as a pre-processing step to the prediction of player possession.

Clemente et al. (2012b) introduced a method for quantifying team formations by summing the triangulated regions created by player positions. This was later refined in Clemente et al. (2015) with the inclusion of tactical positions, determined by longitudinal position thresholds. A team’s total attack and defense regions were computed from the triangulated team formations (Clemente, et al., 2015). Whilst simplistic, these approaches simplify Voronoi-like approaches and were able to discriminate playing roles using thresholds that aligned with expert assessment (Clemente, et al., 2012b; Clemente, et al., 2015).

Another approach to role assignment is consideration of player alignment to opponents (Gudmundsson & Horton, 2017). In basketball, a Hidden Markov Model was used to pair defenders with attackers based on a combination of distances between defenders, the ball and the hoop (Franks, et al., 2015).

2.3.6 Field Equity and Possession Values

Expected possession value (EPV) or field equity metrics assign a quantitative value to the current phase of play that describes its expected outcome. EPV metrics can be considered a short-term forecast of the current possession chain. These metrics can be used to value the contextual contributions of individual players (Cervone, et al., 2014; Jackson, 2016). From this, a greater understanding of player performance is derived and player decision-making can be valued relative to changes in EPV (Cervone, et al., 2014).

O'Shaughnessy (2006) introduced *field equity* in AFL. The field equity of each location on an AFL oval is equal to the average next score within a radius around said location, grouped by the current possession source (e.g., loose, uncontested) (O'Shaughnessy, 2006). This metric was later smoothed using regression splines (Jackson, 2016). The inputs to the AFL field equity metric are limited to possession source and field location (Jackson, 2016; O'Shaughnessy, 2006). Player contributions are measured by changes in equity (Jackson, 2016). Similar metrics have been developed in other team sports, including ice hockey (Kaplan, et al., 2014; Routley, 2015) and Rugby League (Kempton, et al., 2015). Given the richness of player tracking datasets, these measures could be improved with the addition of spatiotemporal player data, as has been the case in other team sports such as basketball (Cervone, et al., 2016a).

Lucey et al. (2014a) calculated the *Expected Goal Value* (EGV) in soccer using a linear regression model that included variables derived from spatial formations. The addition of spatiotemporal data increased the predictive power of stochastic processes, resulting in the *Expected Contribution* (EC) in Ultimate Frisbee (Weiss & Childers, 2014). The objective of these studies is to value specific player actions, providing context to possession involvements (Weiss & Childers, 2014; Lucey, et al., 2014a; Cervone, et al., 2016a).

Chang et al. (2014) used a least-squares regression function to compute the *Effective Shot Quality* (ESQ) of shot attempts in basketball. This model used information on the shot location and proximities of defenders to compute the expected result of said attempt, relative to the league average (Chang, et al., 2014). From this, a player can be benchmarked relative to shot expectations (Chang, et al., 2014).

A generalised EPV has been computed in the NBA using a stochastic process model and spatiotemporal data (Cervone, et al., 2016a). This model considers a player's discrete (e.g., passing) and continuous actions (e.g., moving to the left), referred to as macro- and micro-transitions respectively (Cervone, et al., 2016a). From the EPV, player decisions are evaluated relative to the average player using EPV-added (Cervone, et al., 2014). Evaluating decisions relative to an increase in EPV is logical, however this approach ignores whether there were alternative options that would have had a greater increase in EPV. This is in part because NBA EPV values the current possession, rather than computing future EPVs of teammates (Cervone, et al., 2016a). A new approach to player decisions would be to value decisions relative to available options. To do so would require forecasting the EPV of a player's options, increasing the computational complexity of the model.

2.3.7 Event Prediction

The behavioural features and spatial information of players have been used to predict short-term actions such as passing. In basketball, a multiclass conditional random field was used by Yue et al. (2014) to predict whether a play will retain possession, pass to a teammate or shoot for goal. The model utilises spatial features of player locations but does not quantify the likelihood of each action being successful (Yue, et al., 2014). Maheswaren et al. (2014) analysed rebounds in basketball to predict whether a rebound will be successful based on the shot attempt, rebound location and the height of the ball, accurately predicting 75% of events.

In soccer, predicting the future location of the ball (Kim, et al., 2010) and predicting who will have possession over short periods (Wei, et al., 2015) have been researched. The former used player trajectories to output a density array in which high density locations are locations that the ball is likely to be moved to (Kim, et al., 2010). In Wei et al. (2015), an augmented-Hidden Conditional Random Field takes observational features to predict who will possess the ball over various intervals. Over two-second intervals, the model was accurate in 99.25% of samples (Wei, et al., 2015).

Chapter 3: Clustering Team Profiles in the Australian Football League Using Performance Indicators[2]

Chapter Overview

This chapter serves as a pilot study to identify the significance of offensive and defensive performance indicators in the AFL. Findings from this study are used to inform the focus of later spatiotemporal studies. It was found that offensive performance indicators have greater importance in models that explain match outcome. Furthermore, of the primary two methods of ball movement in the AFL, kicks were more important than handballs.

This chapter includes an introduction to the topic of performance indicators (Section 3.1), discussion of the methodology for clustering team profiles (Section 3.2), summary of results (Section 3.3) and discussion and conclusions of the findings (Sections 3.4; 3.5). The findings of this chapter are discussed in the context of the themes of this thesis in a final discussion section (Section 3.6). Material from this chapter was presented at Mathsport 2016 (Spencer, et al., 2016).

[2] Spencer, B., Morgan, S., Zeleznikow, J., & Robertson, S. (2016). Clustering team profiles in the Australian Football League using performance indicators. In *Proceedings of the 13th Australasian Conference on Mathematics and Computers in Sport, Melbourne*.

[redacted]

[redacted] [redacted]

[redacted]

3.6 THESIS DISCUSSION

This chapter analysed non-spatiotemporal datasets. The objective was to identify important events in Australian football, analysis of which will be applications of methodology presented in later studies. Offensive variables were found to be more important than defensive ones, with respect to match outcome. Furthermore, kicking relate variables are more important than handball relative variables, hence kicking was determined to be a main focus of analysis undertaken later in the thesis. Finally, contested possessions were more important than uncontested possessions. Further analysis of contests and congestions is required to understand the importance of space in Australian football.

The methodology presented in this chapter serves as an introduction to the topic of team playing styles. The measurement of a team's playing style has applications in coaching and tactical analysis. Results of the style matchups may be an indication that playing styles quantified by AFL performance indicators are responsive. A limitation of this work is the inability to differentiate tactical behaviour (or game plans) and gameplay dictated (or in response to) opposition behaviour. Regardless, the ability for coaching staff to identify components of opposition playing styles in response to their own team's style may be beneficial when preparing for matches. If possible, future work into playing styles should aim to filter out components of gameplay that are in response to opposition behaviour.

A notable finding of this study was that team profiles defined by the clustering of performance indicators do not predict success. That is, teams with similar profiles do not share similar levels of success in most cases. A possible explanation for this is the lack of context of performance indicators, which has been identified as a limitation of these data sources (Lucey, et al., 2013; Bialkowski, et al., 2014). Spatiotemporal data can provide context to match events (Lucey, et al., 2013). The measurement of performance not captured by traditional performance indicators, in addition to analysing performance indicators in context, will result in a greater understanding of team performance in Australian football.

Chapter 4: Modelling Within-team Relative Phase Couplings Using Position Derivatives in Australian Rules Football[4]

Chapter Overview

Chapter 4 is the first study contained in this thesis that analyses spatiotemporal datasets in Australian Rules football. Building upon the work of Morgan and Williams (2012), the relative phase of intra- and inter-team player couples is explored.

Previous studies of coordinative behaviour are presented in Section 4.1. For the remainder of the chapter, methodology is outlined (Section 4.2), results are presented (Section 4.3) and discussed (Section 4.4; 4.5). The content of this chapter was published in a sports issue of Mathematical and Computer Modelling of Dynamical Systems (Spencer et al., 2017a).

[4] Spencer, B., Robertson, S., & Morgan, S. (2017). Modelling within-team relative phase couplings using position derivatives in Australian Rules football. *Mathematical and Computer Modelling of Dynamical Systems*, 23(4), 372-383. https://www.tandfonline.com/doi/abs/10.1080/13873954.2017.1336732

MATHEMATICAL AND COMPUTER MODELLING OF DYNAMICAL SYSTEMS, 2017
https://doi.org/10.1080/13873954.2017.1336732

ARTICLE

Modelling within-team relative phase couplings using position derivatives in Australian rules football

Bartholomew Spencer [a,b], Sam Robertson [a,b] and Stuart Morgan [c]

[a]Institute of Sport, Exercise and Active Living, Victoria University, Melbourne, Australia; [b]Western Bulldogs Football Club, Australia; [c]Department of Rehabilitation, Nutrition and Sport, LaTrobe University, Bundoora, Australia

ABSTRACT

Several approaches to the modelling of interpersonal movement coordination in sports, inspired by dynamical systems, have leveraged relative proximity to fixed ground points, such as the court midline to represent the phasic characteristics of movement in competition. While these approaches are useful in highly constrained sports such as tennis and squash, Australian football (AF) is played on a much larger playing area (approximately 150 m × 100 m) and is characterized by a 'rolling scrum' of interpersonal contests. Consequently, a different approach to modelling pairwise movement coordination is required. We propose a method that encodes interpersonal movement coordination using relative phase properties derived from angular velocity and acceleration. We demonstrate that these properties encode the level of temporal alignment of changes in running speed and direction between player pairs. This approach is illustrated using exemplar data from AF and explores net pairwise movement coordination within and between teams, and as a function of match duration.

ARTICLE HISTORY
Received 15 March 2016
Accepted 27 May 2017

KEYWORDS
Relative phase; player interactions; systems; modelling; Australian rules football

1. Introduction

Previous work has described the coordinated behaviour between individuals in net/wall sports such as squash [1,2], and later in tennis [3,4], and team sports such as basketball [5,6] and soccer [7–9]. In each case, these studies explored the periodicity of spatial relations between players with respect to some global feature in the playing area. McGarry and colleagues were the first to consider the spatio-temporal relationship between players in squash as a dynamical system expressed by transitions between phases of stability and instability introduced by perturbations [1]. In that study, radial distances from the T location at the centre of the squash court were measured, and the successive movements of each player to and from that location formed the periodic function. In this example, proximity to the T over time provided an excellent and compact representation of the features of a rally. Similarly in tennis, Lames considered distance from the court midline and baseline as a representation of the periodic movement of players in a baseline rally, alternately moving laterally to retrieve a shot, then to the centre in preparation to return the next shot [4]. These ideas were further extended to team sports by Bourbousson and colleagues, who examined spatio-temporal coordination in basketball between player dyads [5] and teams [6]. In that work, the phase relations were derived from the lateral and longitudinal positions of players and team centroid, respectively.

Relative phase itself provides an insight to the degree of movement coordination between players, or groups of players by using the centroid. In tennis and squash, the phase coupling

CONTACT Stuart Morgan s.morgan@latrobe.edu.au

compresses information about the transition from a stable rally-mode where neither player has any distinct advantage, to a transient point of instability where the climax of the point is evident. However, for team sports on much larger playing areas, the position of players in relation to some global pitch feature such as the midline is meaningless. For example, Australian rules football is played on areas that are frequently greater than 150 m × 100 m, with 36 players on field at any moment. This results in a complex set (rolling scrum) of local interactions between players that are largely unrelated to the relative positions on the pitch.

Recently, Morgan and Williams [10] quantified within-team phase relations between player pairings in football (soccer) where the coordinative features were derived from acceleration and angular velocity. Acceleration can be considered a coordinative feature since players who are in-phase can be thought of changing speed (either accelerating or decelerating) in unison. Similarly, angular velocity can be considered a coordinative feature since playing pairs that are in-phase are understood to be changing direction in the same range at the same time. In both cases, anti-phase states indicate that the players within a coupling are changing speed or direction at equal rates of opposite direction (i.e. one player slows at the same rate that another player speeds, or one player turns left at the same rate another player turns right). This method permits coordinated player couplings to be described independently of the pitch location, which is a well-suited approach to large-field team sports.

Our contribution in this paper is to introduce a new method for modelling inter-player coordination in large-field team sports where the absolute positions of players to the field are not important. This paper will also explore the coordinated features of inter- and intra-team player couplings in the context of Australian rules, an invasion team sport consisting of 2 sides of 18 players competing with the objective of scoring goals (worth 6 points) and behinds (worth 1 point). Furthermore, we will confirm previous results presented in [10] and examine the temporal characteristics of these relationships throughout the duration of an Australian football (AF) match.

2. Methodology

2.1. *Data set*

Player recordings from the TAC Cup Grand Final, the premier Under-18 AF league in Victoria, Australia, were collected. All data for the match (played between the Oakleigh Chargers and Eastern Ranges) were provided by Catapult (Catapult Innovations, Melbourne, Australia). The data set consisted of individual local positioning system (LPS) recordings for each player for the match, collected via Catapult's wearable LPS tracking system, Clearsky. A total of 41 individual player files were used, with 5 missing due to recording complications on match day. Of the five missing players, four were members of the Eastern Ranges; hence, Oakleigh was designated as the primary team for analysis. Upon inspection, it was found that Oakleigh's only missing player was an interchange player, hence minimizing the influence of this absence on the final findings. For this particular match, Oakleigh defeated Eastern with a score of 74 (10 goals and 13 behinds) to 61 (9 goals and 7 behinds).

Each participant's recording included time and calibrated location. Players were classified in three primary playing positions, Midfield, Forward and Defender, based on the labels provided by team squad sheets submitted on match day. Data were de-identified for the purposes of the study with ethical permission granted to complete the study by the relevant human ethics committee.

2.2. *Data preprocessing*

The raw data were recorded at 10 Hz; however, the beginning of the individual player recordings was not synchronized. The recording system provided a global system time, and all data measurements included a timestamp corresponding to the global time. To resolve the temporal misalignment, the data were down-sampled to 1 Hz and temporally aligned to the nearest decisecond. The down-sampling process involved sampling the first recording at each second for all players to ensure maximal temporal alignment. In total, each player's data set consisted of approximately 6600 rows (approximately 110 min) including time, and x- and y-location.

Game interruptions between playing periods, and the event where a player interchanged out of the competition, were inferred from the raw movement of players. Recordings for players located outside the pitch boundary were removed from the data such that only those players who were on the field of play at any given time were included in the analysis. Furthermore, only players who began the match on field were included in the analysis.

2.3. *Phase angles*

Quantifying phase relations involves the derivation of a performance attribute (i.e. velocity) and plotting this in relation to its rate of change. Phase angles are then derived from the slope of a point to its origin. Graphical representations of these relations are commonly referred to as phase portraits (see Figure 1(a,b)). In this case, phase portraits were estimated by velocity and acceleration, and angular displacement and angular velocity. The equations for estimating phase angles are described below.

2.4. *Phase angles for acceleration*

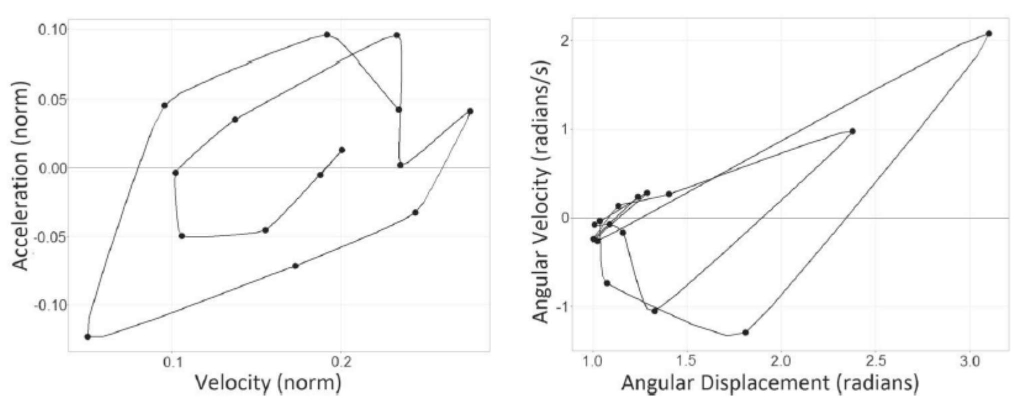
Velocity (V_i) was derived from the raw player position data and normalized as follows, where instantaneous velocity and velocity minima and maxima are represented as V, $V_{\min}$ and $V_{\max}$, respectively:

$$V_i = \frac{V - V_{\min}}{V_{\max} - V_{\min}} \tag{1}$$

Acceleration (A_i) was derived from the normalized velocity:

$$A_i = \frac{V_i - V_{i-1}}{\Delta t} \tag{2}$$

The phase angle (ϕ_i) for acceleration were found as a function of acceleration and velocity:



(a) Acceleration (b) Angular velocity

Figure 1. Phase portraits for (a) acceleration and (b) angular velocity.

$$\phi_i = \tan^{-1}\left(\frac{A_i}{V_i}\right) \tag{3}$$

2.5. *Phase angles for angular velocity*

Angular displacement (θ_i) was calculated as the inner product of consecutive 1-s movement vectors, **a** and **b**:

$$\theta_j = 1 + \cos^{-1}\left(\frac{\mathbf{a}\cdot\mathbf{b}}{||\mathbf{a}||\cdot||\mathbf{b}||}\right) \tag{4}$$

Angular velocity (ω_j) follows from θ_j:

$$\omega_j = \frac{\theta_j - \theta_{j-1}}{\Delta t} \tag{5}$$

Finally, the phase angle for angular velocity (ϕ_j) was derived as a function of angular displacement (θ_j) and angular velocity (ω_j). Note that since angular displacement values spread very near to zero would result in temporally unstable phase angles, the raw angular displacements (θ_i) are offset by a value of 1. All phase angles were converted to degrees.

$$\phi_j = \tan^{-1}\left(\frac{\omega_j}{\theta_j}\right) \tag{6}$$

2.6. *Relative phase*

Phase angles for acceleration (ϕ_i) and angular velocity (ϕ_j) were derived for both teams across all moments of play. Inter-team (n = 153) and intra-team (n = 270) couplings were compiled for the included on-field players. Pairwise relative phase (ε) was calculated for each player couple as the difference between phase angles for acceleration and angular velocity.

$$\varepsilon_{\mathrm{AB}} = \phi_{\mathrm{playerA}} - \phi_{\mathrm{playerB}} \tag{7}$$

Previous work in dynamical systems in sports (e.g. Ref. [4]) refers to in-phase states where the relative phase is near zero, and anti-phase states where the relative phase is near ±180. The former would be true if two players were both accelerating or decelerating at equal relative rates, while the latter would be true if one player was accelerating at the same relative rate as the other was decelerating. In practical terms, both instances represent examples of coordinated behaviour, where the movements are directly coupled even though the sign of the change is inverted. Further, this logic applies equally to angular velocity, where it is the temporally coupled behaviour that we want to encode with relative phase rather than the direction itself. Therefore, we fold the tails of the relative phase distributions in such a way that moments of strong anti-phase coupling are transformed to the centre of a zero-based distribution, and in-phase and anti-phase states have equal coordinative value (see Equation 8).

$$\varepsilon' = k - (\varepsilon - k), \quad \text{where } k = \begin{cases} 90 & \varepsilon > 90 \\ -90 & \varepsilon < -90 \\ \varepsilon & -90 \le \varepsilon \le 90 \end{cases} \tag{8}$$

2.7. *Clustering*

The mean and standard deviations for both acceleration and angular velocity relative phase data were calculated for each possible player coupling. Since the transformed relative phase values are centred around the mean, central tendency is not a helpful index for the between-group comparison of coordinated behaviour (since the mean relative phase will always be approximately zero). Therefore, we reason that standard deviation provides a good estimate of the level of phase coupling between players. Where the standard deviation is higher, a greater portion of moments will be characterized by out-of-phase behaviour, and where the standard deviation is lower, we reason that a greater portion of transformed relative phase moments are nearer to zero indicating greater coupling.

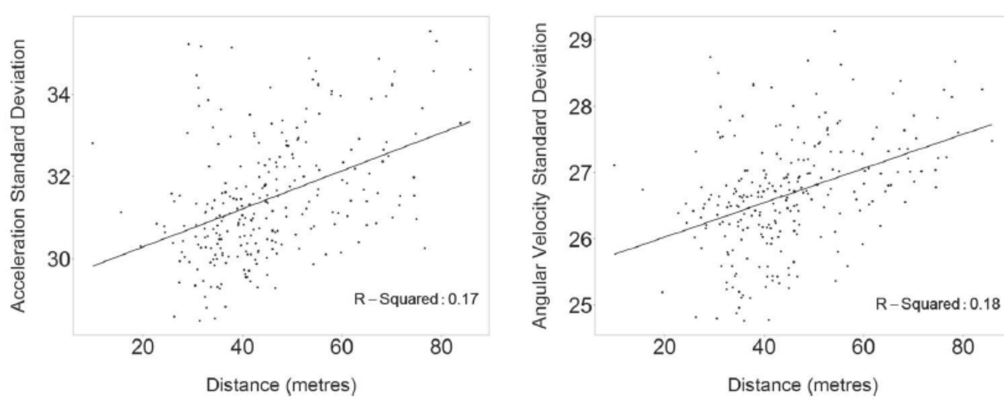
Given 18 players are allowed on each team, there are 630 possible permutations of player couplings at any moment. In order to reduce the complexity of this analysis, pairings were then grouped as either intra- or inter-team. Intra-team couplings consisted of all possible parings between players on the same team, and inter-team couplings consisted of all possible pairings within a team. *k*-Means clustering was then used to separately generate clusters of similar player couples based on their proximity in a two-dimensional space of acceleration and angular velocity relative phase standard deviations for the relevant permuted couplings. *k*-Means was chosen as an arbitrary method for identifying groups of high and low coordination player couples to aid in the analysis and visualization of results. These data are presented in Figures 3 and 4. *k*-Means works by positioning *k* centroids repeatedly until the means diverge and has been previously used to visualize sporting styles (e.g. see Refs. [11,12]). *k*-Clusters were chosen based on a sum of squares plot.

2.8. *Temporal analysis*

Box plots of relative phase by time period (AF games consist of four 25 min quarters) were compiled for inter- and intra-team couplings to compare mean coupling as a function of the time of the game.

3. Results

A plot of relative phase standard deviations and player proximities for all permutation of player pairs is presented in Figure 2. This plot exhibits the characteristics of a positive relationship between relative phase behaviour and physical proximity on the playing area, confirming that the proximity relationships described in Ref. [10] can be observed in other large-field team sports. This indicates that players who are closer to each other are more likely to move in concert, both in terms of changes in velocity (Figure 2(a)), and in change in direction of movement (Figure 2(b)).



(a) Acceleration (b) Angular velocity

Figure 2. Mean relative phase standard deviations for acceleration (a) and angular velocity (b) by mean Euclidian pairwise distance.

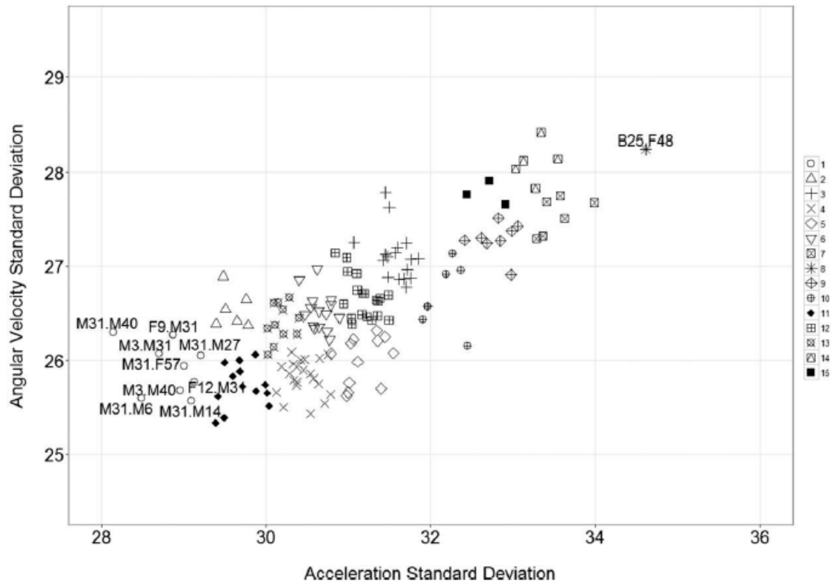


Figure 3. Angular velocity and acceleration relative phase standard deviations by cluster for *intra*-team couplings. Labels for player role are included for members of the LCP cluster (8) and the HCP cluster (1).

3.1. *Intra-team coordination*

k-Means clustering for the intra-team pairings with 15 clusters is presented in Figure 3. The pairs were labelled according to their playing role and shirt number, which results in a unique label for each player. The player role labels are *back/defender* (B), *midfielder* (M) and *forward/attacker* (F). Additionally, the mean standard deviations for acceleration and angular velocity relative phase were computed over all parings in each cluster, and the mean values are presented in Table 1.

From within the intra-team clusters, cluster 8 can be defined as the least coupled set of player pairs, while cluster 1 can otherwise be defined as the most coupled set of player pairs where the

Table 1. Mean acceleration and angular velocity cluster standard deviations for *intra*- and *inter*-team clusters.

	Intra-team clusters		*Inter*-team clusters	
Cluster ID	Acceleration SD	Angular velocity SD	Acceleration SD	Angular velocity SD
1[a]	28.84	25.92	33.04	27.57
2	29.59	26.54	30.17	25.43
3	31.58	27.11	32.63	26.94
4[c]	30.45	25.82	29.13	25.20
5	31.17	26.00	30.59	26.30
6[d]	30.65	26.52	34.92	28.33
7	33.54	27.54	31.81	26.64
8[b]	34.61	28.24	30.53	26.89
9	32.80	27.29	33.53	28.26
10	32.16	26.68	31.32	27.13
11	29.71	25.72	30.97	25.78
12	31.18	26.66	34.20	27.54
13	30.18	26.39	31.19	26.56
14	33.26	28.11	29.89	26.22
15	32.69	27.77	32.17	27.65

SD: Standard Deviation.
[a]High-coordination pairs for intra-team clusters.
[b]Low-coordination pairs for intra-team clusters.
[b]Low-coordination pairs for intra-team clusters.
[c]High-coordination pairs for inter-team clusters.
[d]Low-coordination pairs for inter-team clusters.

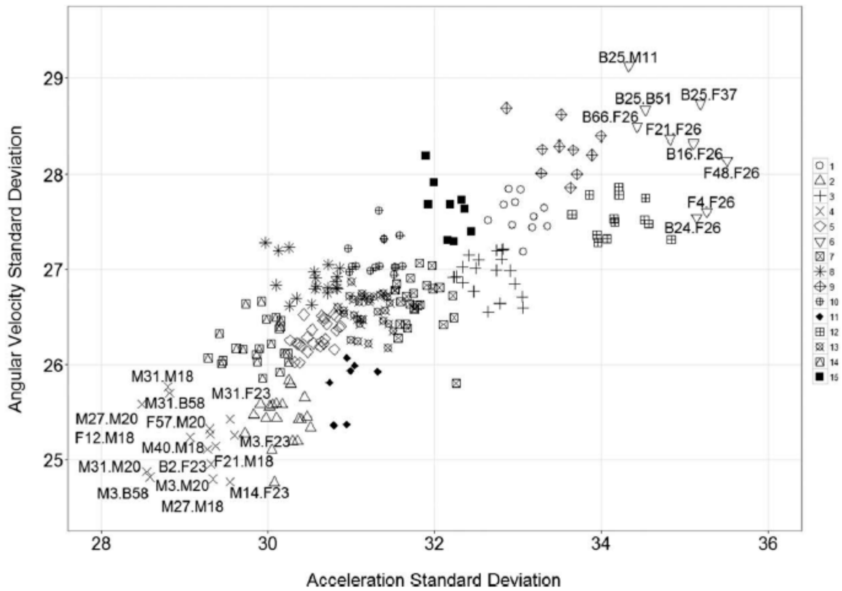


Figure 4. Angular velocity and acceleration relative phase standard deviations by cluster for *inter*-team couplings. Labels for player role are included for members of the LCP cluster (6) and the HCP cluster (4).

variability in phase coupling is lowest. For the purpose of comparison, we refer to cluster 1 as high-coordination pairs (HCP) and cluster 8 as low-coordination pairs (LCP).

It is worth noting that each of the player pairs observed in the HCP cluster contained at least one midfield player (M), and most of the pairs were composed of two midfielders. In contrast, the player pairs observed in the LCP cluster commonly featured pairings of players who exhibit the greatest median distance apart, indicating these are pairs of players who are positioned at opposite ends of the playing field.

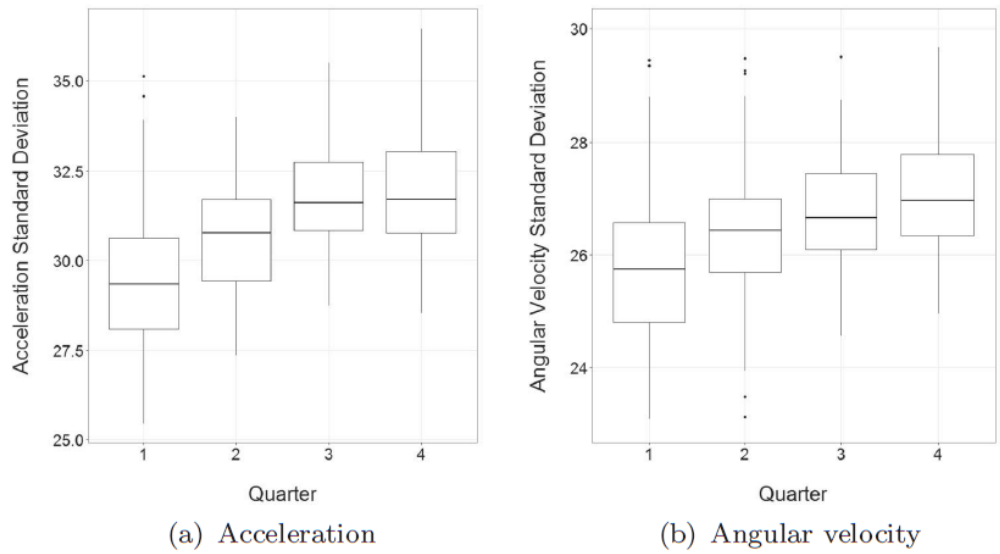
3.2. *Inter-team coordination*

k-Means clustering for the inter-team pairings with 15 clusters is presented in Figure 4. Analysis of the intra-team clusters revealed that cluster 6 can be defined as the least coupled set of player pairs, while cluster 4 can otherwise be defined as the most coupled set of player pairs. As previously described, we refer to cluster 4 and cluster 6 as HCP and LCP, respectively.

As was previously observed in the intra-team pairings, the HCP consist predominantly of midfield positions; however, they are more frequently paired with defender/back or forward players when compared to intra-team HCP. Of note, Oakleigh player M31 was part of four of the 15 HCP inter-team pairings. Inter-team LCP are largely composed of pairings of positions who are positioned at opposite ends of the field.

3.3. *Temporal analysis*

Box plots showing mean intra-team relative phase angle standard deviations by game quarter are presented in Figure 5. The mean standard deviations over each quarter increased systematically for the acceleration relative phase (Figure 5a) and angular velocity relative phase (Figure 5(b)). These plots indicate that the coordinative behaviour represented by coupled changes in running speed and direction of movement decreases systematically over the course of the AF match.



(a) Acceleration (b) Angular velocity

Figure 5. Relative phase standard deviation box plots for acceleration (a) and angular velocity (b) by game quarter.

3.4. *Player coupling exemplars*

This modelling approach encodes multidimensional movement characteristics between player pairs and can be used make pairwise classifications of players as strongly- or weakly coupled. In Figures 6 and 7, we present 30-s sequences of game play where the relative phase is derived from continuous angular displacement and angular velocity data for a weakly coupled and strong-coupled pair, respectively. While the exemplar shown in Figure 6 exhibits no obvious coordination, Figure 7 illustrates an example where the paired players are highly coordinated. Similarly, the

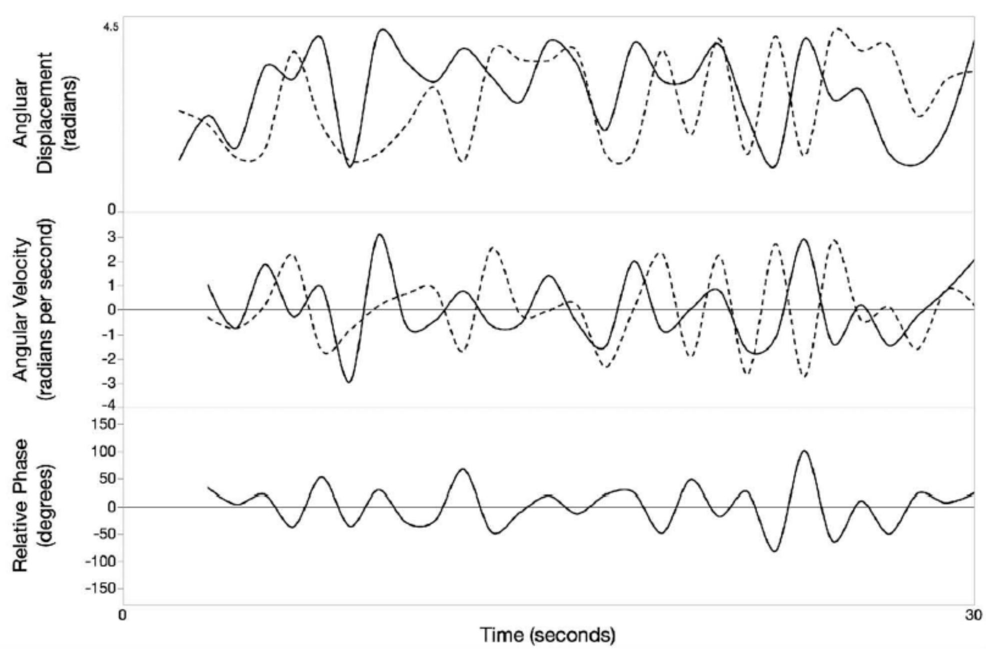


Figure 6. Exemplar angular displacement, angular velocity and relative phase for a LCP player coupling (solid = B25, dashed = F48).

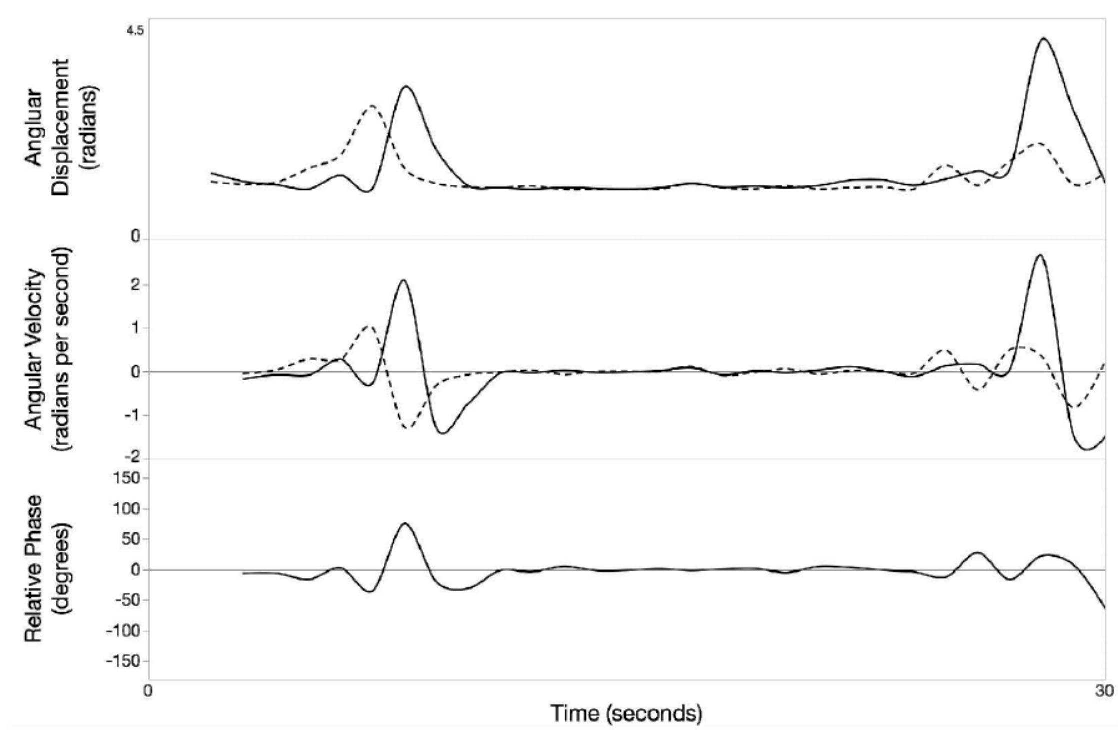


Figure 7. Exemplar angular displacement, angular velocity and relative phase for a HCP player coupling (solid = M31, dashed = F57).

same exemplar data are used to generate Figures 8 and 9, showing acceleration coupling between the same players. Here also, it is evident in Figure 9 that the players are changing speed in a highly coordinated manner, and the subsequent relative phase representation highlights both the general state of coordination and the brief moment at approximately 6 s where the velocity changes are momentarily misaligned.

4. Discussion

This paper demonstrates a method of estimating pairwise player coupling that is suitable for team sport games, particularly games played on large playing areas. Previous approaches to estimating inter-player coordination leveraged movement features relative to fixed locations in the playing area, such as the middle 'T' in squash (e.g. Ref. [1]). While this approach is helpful in small-court games, where the coordination of movement between players is functionally related to their physical location, the same rule may not apply in large field games. AF is played on very large surfaces and is characterized by a high number of interpersonal exchanges that are functionally independent of their location on the playing area. For instance, mid-field players can engage in a direct contest at any location on the ground, and therefore, representing their interactive movement features in a dynamical systems framework using relative phase as the principal measure of coordination requires a new approach. We studied a spatio-temporal data set in AF and illustrated exemplar results using a position-independent approach.

We demonstrated a method where phase angles are derived from separate phase portraits expressing velocity and acceleration, and angular displacement and angular velocity, respectively. The angle from the origin of a phase portrait to the momentary acceleration or angular velocity value (in polar coordinates) can be regarded as a momentary state of movement. The pairwise

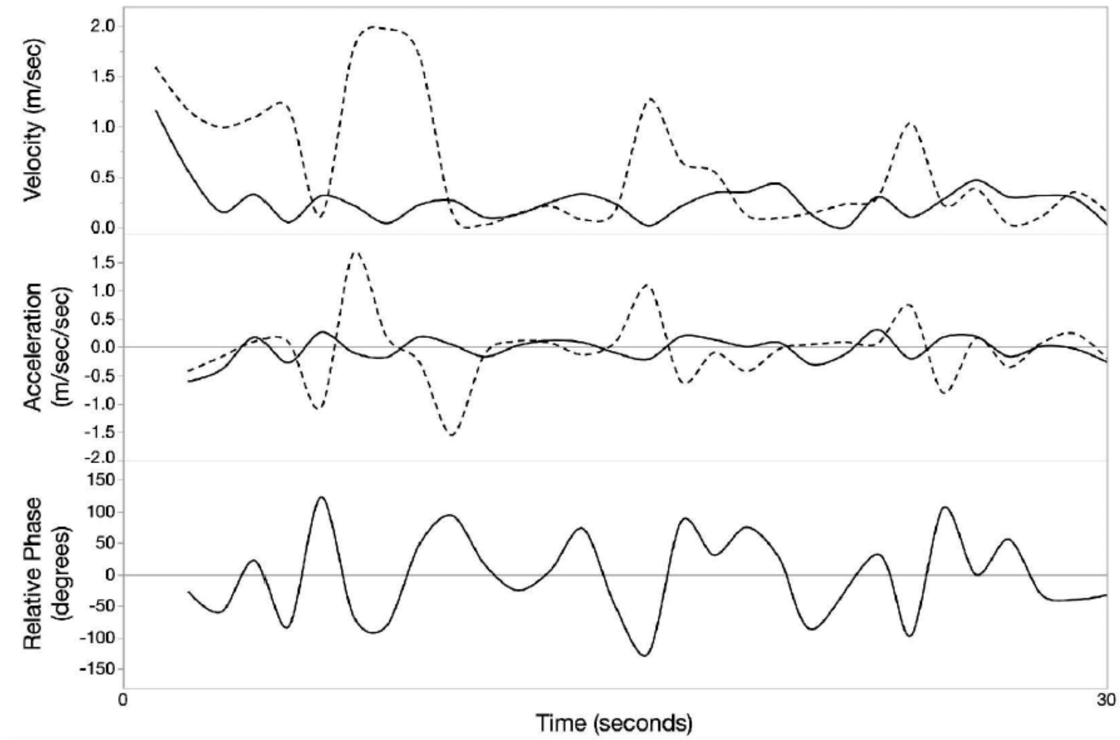


Figure 8. Exemplar velocity, acceleration and relative phase for a LCP player coupling (solid = B25, dashed = F48).

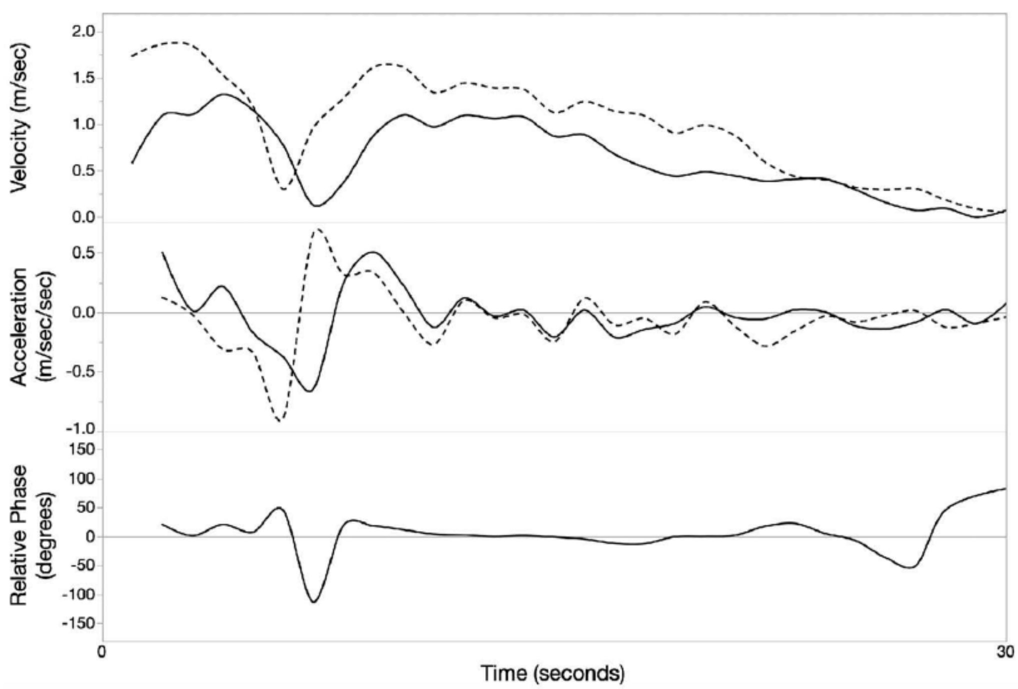


Figure 9. Exemplar velocity, acceleration and relative phase for a HCP player coupling (solid = M31, dashed = F57).

time-series comparison of these continuous states between players can then be used to estimate the degree of movement coordination.

The first observation we made is that there is an inverse correlation between interpersonal coordination and the distance between players. Players that are on average closer to each other are also more tightly coupled, both in terms of their changes in acceleration, and changes in angular velocity. In many respects, this is an obvious characteristic in team sports on large playing areas. Players who are further from the locus of gameplay would have no reason to move in concert with those distant players who are directly involved in the movement of the ball. This should be regarded as a sanity check and evidence of the contextual validity of our approach.

Analysis of the temporal characteristics of relative phase angles revealed that the standard deviation for the relative phase angles for both acceleration and angular velocity increases in later quarters. This indicates a potential diminishing degree structure within teams as time progresses. The decrease in coupling observed over the course of the match is also evidence that fatigue may impact the interpersonal coordination between players. This result could be explained by the proposition that as players fatigue, the variability in their capacity or willingness to run with an intensity that matches other teammates may be increased. This would predictably result in lower coupling results on that dimension.

In representing coordinative behaviour as a dynamical system, one needs to consider possible control parameters that cause a stable system to shift into some alternate state. Lames has previously asserted that the players in a tennis match could be considered as two sub-systems, strongly coupled by features of the game (such as the size and dimensions of the court, and the net-based format of the game), and characteristics of the game play (e.g. a baseline rally) [4]. In many respects, AF is divorced from the structural constraints of tennis, with a large, open field of play, and without the inherent periodic construct of a net-based sport where the ball travels back and forth between players. Nevertheless, if we can consider an AF match as a system with many degrees of freedom, represented by many paired sub-systems, the stability and phase characteristics of those subsystems may be driven by match-related control parameters, such as tactical constraints, fatigue, skill-level and role-based constraints. These constraints could be considered control parameters in the sense that they cause the characteristics of pairwise behaviour to change. We may infer evidence of fatigue in the systematic degrading of coupled movements through the course of an AF match in our data. Further research should explore whether this trend is a robust feature of other matches. Similarly, tactical constraints (which are not considered here) are a potential control parameter. Consider the instance that a coach directs his/her team to play a zonal defence, where players defend spatial areas rather than specific opponent players. Such a scenario would predictably result in a substantive shift in the state of the pairwise subsystems.

Most previous system-based representations of movement coordination in sport have focussed on small court games such as tennis [4], squash [1] and basketball [5,6] (notwithstanding other work in soccer: e.g. Refs. [7–9]). In each instance, the order parameters considered are derived from the absolute location of individuals, or group centroids, in the playing area. This is a suitable approach to small court games, and one might argue further that even in soccer (which is played on a much larger field), role-based constraints on players assert some predictable regularity in both position and relative movement. For instance, a right-side winger will rarely wander beyond the unilateral spatial domain specific to that role. In those sports, regularity in role-based movement applies, and order parameters derived from absolute position may be helpful. In AF, however, the field-based location of players is mostly divorced from player roles (with only a few exceptions). In considering immediate derivatives of movement such as acceleration and angular velocity as order parameters, it is possible to represent coordinated behaviour in a way that is decoupled from the absolute location of players on the field of play. Importantly, this approach is not inconsistent or mutually exclusive from position-based parameters. Indeed, it may provide additional insight in a dynamical systems framework to model the combination of position and movement parameters in court and small-field games.

5. Conclusion

Our aim in this work has been twofold. Principally, we aimed to demonstrate a method for quantifying the characteristics of interpersonal movement coordination between pairs of players in large-format team sports. We extend previous innovations that use a dynamical systems framework, and relative phase, to model interpersonal movement in small-court games (where the location of players is the order parameter), and present an alternative approach that leverages derivatives of interpersonal movement that are independent of pitch location. Second, we aimed to demonstrate the observable features of pairwise coordination within and between teams in AF. This analysis revealed that movement coordination is higher amongst midfield players within the same team and direct opponents between teams, and that some aspects of interpersonal movement coordination may diminish throughout matches, perhaps due to fatigue. This approach should enable future work to learn more about potential control parameters that may direct a dynamical systems representation of large-format team sports.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] T. McGarry, M.A. Kahn, and I.M. Franks, *On the presence and absence of behavioural traits in sport: an example from championship squash match-play*, J Sports Sci 17 (1999), pp. 297–311. doi:10.1080/026404199366019

[2] T. McGarry, D.I. Anderson, S.A. Wallace, M.D. Hughes, and I.M. Franks, *Sport competition as a dynamical self-organizing system*, J Sports Sci. 20 (2002), pp. 771–781. Available at doi:10.1080/026404102320675620, pMID: 12363294.

[3] Y. Pault and P.G. Zanone, *A dynamical analysis of tennis: concepts and data*, J Sports Sci 10 (2005), pp. 1021–1032. doi:10.1080/02640410400021682

[4] M. Lames, *Modelling the interaction in game sports - relative phase and moving correlations*, J. Sports Sci. Med. 5 (2006), pp. 556–560.

[5] J. Bourbousson, C. Seve, and T. McGarry, *Space-time coordination dynamics in basketball: part 1. Intra- and inter-couplings among player dyads.*, J Sports Sci 28 (2010), pp. 339–347. doi:10.1080/02640410903503632

[6] J. Bourbousson, C. Seve, and T. McGarry, *Space-time coordination dynamics in basketball: part 2. The interaction between the two teams*, J Sports Sci 28 (2010), pp. 349–358. doi:10.1080/02640410903503640

[7] J.F. Gréhaigne, D. Bouthier, and B. David, *Dynamic-system analysis of opponent relationships in collective actions in soccer*, J Sports Sci 15 (1997), pp. 137–149. doi:10.1080/026404197367416

[8] K. Davids, D. Araujo, and R. Shuttleworth, *Applications of dynamical systems theory to football*, Sci. football V (2005), pp. 537–550.

[9] W. Frenken, K. Lemmink, N. Delleman, and C. Visscher, *Oscillations of centroid position and surface area of soccer teams in small-sided games*, Eur. J. Sport Sci. 11 (2011), pp. 215–223. doi:10.1080/17461391.2010.499967

[10] S. Morgan and M. Williams, *Using relative phase of position derivatives to quantify within-team player couplings*, in 11th Australiasian Conference on Mathematics and Computers in Sport, MathSport, Melbourne, 2012, pp. 129–138.

[11] L. Gyarmati, H. Kwak, and P. Rodriguez, *Searching for a unique style in soccer*, in 2014 KDD Workshop on Large-Scale Sports Analytics, New York City, 2014.

[12] J. Sampaio, T. McGarry, J. Calleja-González, S.J. Sáiz, X.S. i del Alcázar, and M. Balciunas, *Exploring game performance in the national basketball association using player tracking data*, PLoS One 10(2015). doi:10.1371/journal.pone.0132894

Modelling Within-team Relative Phase Couplings Using Position Derivatives in Australian Rules Football

Abstract

Several approaches to the modelling of inter-personal movement coordination in sports, inspired by dynamical systems, have leveraged relative proximity to fixed ground points, such as the court midline to represent the phasic characteristics of movement in competition. While these approaches are useful in highly constrained sports such as tennis and squash, Australian football is played on a much larger playing area (approximately 150m x 100m) and is characterised by a “rolling scrum” of interpersonal contests. Consequently, a different approach to modelling pairwise movement coordination is required. We propose a method that encodes inter-personal movement coordination using relative phase properties derived from angular velocity and acceleration. We demonstrate that these properties encode the level of temporal alignment of changes in running speed and direction between player pairs. This approach is illustrated using exemplar data from Australian football and explores net pairwise movement coordination within and between teams, and as a function of match duration.

4.1 INTRODUCTION

Previous work has described the coordinated behaviour between individuals in net/wall sports such as squash (McGarry, et al., 1999; McGarry, et al., 2002), and later in tennis (Pault & Zanone, 2005; Lames, 2006), and team sports such as basketball (Bourbousson, et al., 2010a; Bourbousson, et al., 2010b) and soccer (Grehaigne, et al., 1997; Davids, et al., 2005; Frenken, et al., 2011). In each case these studies explored the periodicity of spatial relations between players with respect to some global feature in the playing area. McGarry and colleagues were the first to consider the spatiotemporal relationship between players in squash as a dynamical system expressed by transitions between phases of stability and instability introduced by perturbations (McGarry, et al., 1999). In that study, radial distances from the T location at the centre of the squash court were measured, and the successive movements of each player to and from that location formed the periodic function. In this example, proximity to the T over time provided an excellent and compact representation of the features of a rally. Similarly, in tennis, Lames considered distance from the court midline and baseline as a representation of the periodic movement of players in a baseline rally, alternately moving laterally to retrieve a shot, then to the centre in preparation to return the next shot (Lames, 2006). These ideas were further extended to team sports by Bourbousson and colleagues, who examined spatiotemporal coordination in basketball between player dyads (Bourbousson, et al., 2010a), and teams (Bourbousson, et al., 2010b). In that work, the phase relations were derived from the lateral and longitudinal positions of players and team centroid respectively.

Relative phase itself provides an insight to the degree of movement coordination between players, or groups of players by using the centroid. In tennis and squash, the phase coupling compresses information about the transition from a stable rally-mode where neither player has any distinct advantage, to a transient point of instability where the climax of the point is evident. However, for team sports on much larger playing areas, the position of players in relation to some global pitch feature such as the midline is meaningless. For example, Australian Rules football is played on areas that are frequently greater than 150 m x 100 m, with 36 players on field at any moment. This results in a complex set (rolling scrum) of local interactions between players that are largely unrelated to the relative positions on the pitch.

Recently, Morgan and Williams (Morgan & Williams, 2012) quantified within-team phase relations between player pairings in football (soccer) where the coordinative features were derived from acceleration and angular velocity. Acceleration can be considered a coordinative feature since players that are in-phase can be thought of changing speed (either accelerating or decelerating) in unison. Similarly, angular velocity can be considered a coordinative feature since playing pairs that are in-phase are understood to be changing direction in the same range at the same time. In both cases, anti-phase states indicate the players within a coupling are changing speed or direction at equal rates of opposite direction (i.e. one player slows at the same rate that another player speeds, or one player turns left at the same rate another player turns right). This method permits coordinated player couplings to be described independently of the pitch location, which is a well-suited approach to large-field team sports.

Our contribution in this paper is to introduce a new method for modelling inter-player coordination in large-field team sports where the absolute positions of players to the field are not important. This paper will also explore the coordinated features of inter- and intra-team player couplings in the context of Australian Rules, an invasion team sport consisting of two sides of 18 players competing with the objective of scoring goals (worth six points) and behinds (worth one point). Furthermore, we will confirm previous results presented in Morgan and Williams (2012) and examine the temporal characteristics of these relationships throughout the duration of an Australian football match.

4.2 METHODOLOGY

4.2.1 Dataset

Player recordings from the TAC Cup Grand Final, the premier Under-18 AF league in Victoria, Australia were collected. All data for the match (played between the Oakleigh Chargers and Eastern Ranges) were provided by Catapult (Catapult Innovations, Melbourne, Australia). The dataset consisted of individual local positioning system (LPS) recordings for each player for the match, collected via Catapults wearable LPS tracking system, Clearsky. A total of 41 individual player files were used, with five missing due to recording complications on match day. Of the five missing players, four were members of the Eastern Ranges, hence Oakleigh was designated as the primary team for analysis. Upon inspection it was found that Oakleigh's only missing player was an interchange player, hence minimising the influence of this absence on the final findings. For this particular match Oakleigh defeated Eastern with a score of 74 (10 goals and 13 behinds) to 61 (9 goals and 7 behinds).

Each participant's recording included time and calibrated location. Players were classified in three primary playing positions, Midfield, Forward, and Defender, based on the labels provided by team squad sheets submitted on match day. Data was deidentified for the purposes of the study with ethical permission granted to complete the study by the relevant human ethics committee.

4.2.2 Data Pre-processing

The raw data were recorded at 10 Hz, however, the beginning of the individual player recordings were not synchronised. The recording system provided a global system time, and all data measurements included a timestamp corresponding to the global time. To resolve the temporal misalignment, the data were down-sampled to 1 Hz and temporally aligned to the nearest deci-second. The down-sampling process involved sampling the first recording at each second for all players to ensure maximal temporal alignment. In total, each player's dataset consisted of approximately 6600 rows (approximately 110 minutes) including time, and x- and y-location.

Game interruptions between playing periods, and the event where a player interchanged out of the competition were inferred from the raw movement of players. Recordings for players located outside the pitch boundary were removed from the data, such that that only those players who were on the field of play at any given time were included in the analysis. Furthermore, only players who began the match on field were included in the analysis.

4.2.3 Phase Angles

Quantifying phase relations involves the derivation of a performance attribute (i.e., velocity) and plotting this in relation to its rate of change. Phase angles are then derived from the slope of a point to its origin. Graphical representations of these relations are commonly referred to as phase portraits (see Figure 4-1a and Figure 4-1b). In this case, phase portraits were estimated by velocity and acceleration, and angular displacement and angular velocity. The equations for estimating phase angles are described below.

4.2.4 Relative Phase Angles for Acceleration

Velocity (Vi) was derived from the raw player position data and normalised as follows, where instantaneous velocity and velocity minima and maxima are represented as V, V_{min} and V_{max} respectively:

$$V_i = \frac{V - V_{min}}{V_{max} - V_{min}} \tag{1}$$

Acceleration (A_i) was derived from the normalised velocity:

$$A_i = \frac{V_i - V_{i-1}}{\Delta t} \tag{2}$$

The phase angle (φ_i) for acceleration were found as a function of acceleration and velocity:

$$\varphi_i = \tan^{-1}\left(\frac{A_i}{V_i}\right) \tag{3}$$

4.2.5 Relative Phase Angles for Angular Velocity

Angular displacement (θ_j) was calculated as the inner product of consecutive one-second movement vectors, a and b:

$$\theta_j = 1 + \cos^{-1}\left(\frac{a \cdot b}{\|a\| \cdot \|b\|}\right) \tag{4}$$

Angular velocity (ω_j) follows from θ_j:

$$\omega_j = \frac{\theta_j - \theta_{j-1}}{\Delta t} \tag{5}$$

Finally, the phase angle for angular velocity (φ_j) was derived as a function of angular displacement (θ_j) and angular velocity (ω_j). Note that since angular displacement values spread very near to zero would result in temporally unstable phase angles, the raw angular displacements (θ_i) are offset by a value of 1. All phase angles were converted to degrees.

$$\varphi_j = \tan^{-1}\left(\frac{\omega_j}{\theta_j}\right) \tag{6}$$

4.2.6 Relative Phase

Phase angles for acceleration (φ_i) and angular velocity (φ_j) were derived for both teams across all moments of play. Inter-team (n = 153) and intra-team (n = 270) couplings were compiled for the included on-field players. Pair-wise relative phase (∈) was calculated for each player couple as the difference between phase angles for acceleration and angular velocity.

$$\in_{AB} = \varphi_{playerA} - \varphi_{playerB} \qquad (7)$$

Previous work in dynamical systems in sports (e.g., Lames, 2006) refers to in-phase states where the relative phase is near zero, and anti-phase states where the relative phase is near ±180°. The former would be true if two players were both accelerating or decelerating at equal relative rates, while the latter would be true if one player was accelerating at the same relative rate as the other was decelerating. In practical terms, both instances represent examples of coordinated behaviour, where the movements are directly coupled even though the sign of the change is inverted. Further, this logic applies equally to angular velocity, where it is the temporally-coupled behaviour that we want to encode with relative phase rather than the direction itself. Therefore, we fold the tails of the relative phase distributions in such a way that moments of strong anti-phase coupling are transformed to the centre of a zero-based distribution, and in-phase and anti-phase states have equal coordinative value (see Equation 8).

$$\in' = k - (\in - k)\,, where\ k = \begin{cases} 90 & \in > 90 \\ -90 & \in < -90 \\ \in & -90 \leq \in \leq 90 \end{cases} \qquad (7)$$

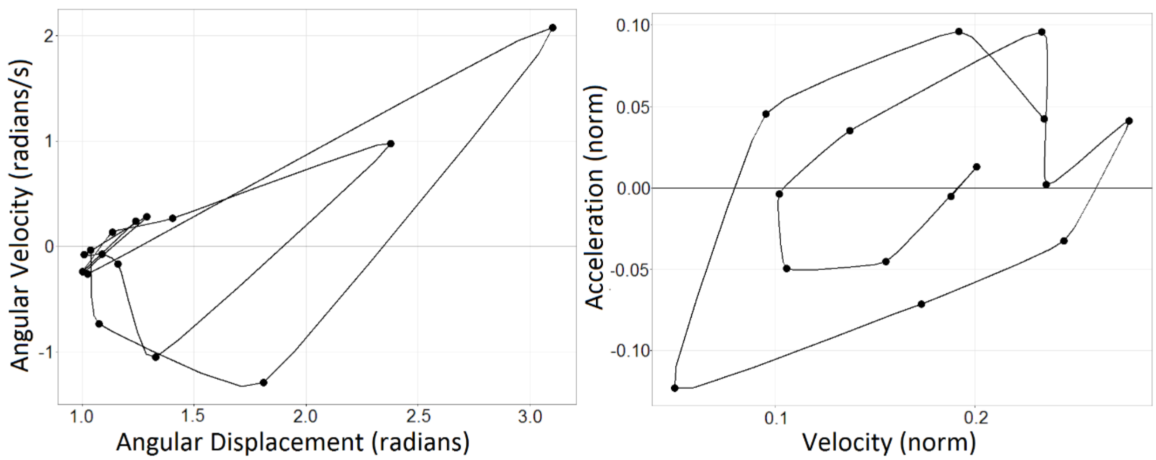


Figure 4-1. Phase portraits for (a) acceleration, and (b) angular velocity.

4.2.7 Clustering

The mean and standard deviations for both acceleration and angular velocity relative phase data were calculated for each possible player coupling. Since the transformed relative phase values are centred around the mean, central tendency is not a helpful index for the between-group comparison of coordinated behaviour (since the mean relative phase will always be approximately zero). Therefore, we reason that standard deviation provides a good estimate of the level of phase coupling between players. Where the standard deviation is higher, a greater portion of moments will be characterised by out-of-phase behaviour, and where the standard deviation is lower, we reason that a greater portion of transformed relative phase moments are nearer to zero indicating greater coupling.

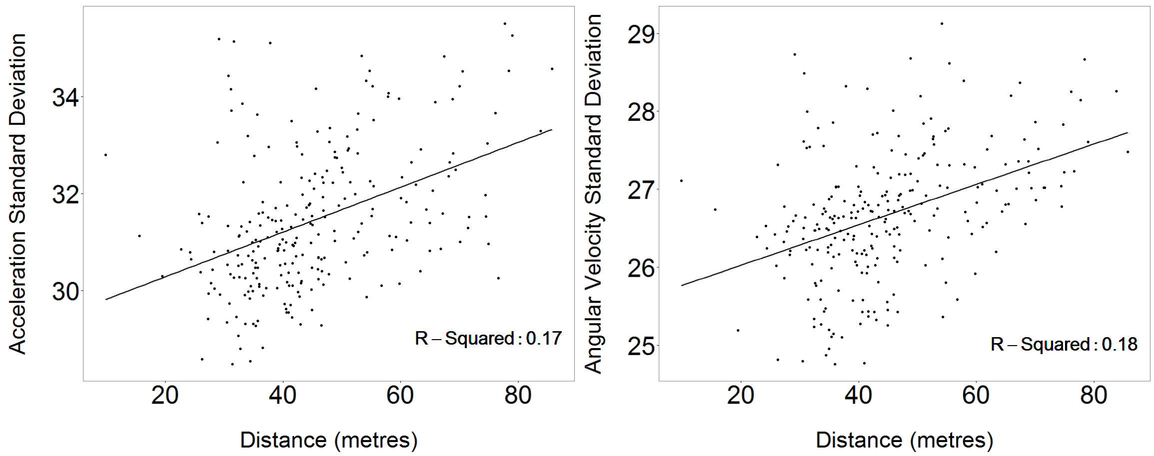


Figure 4-2. Mean relative phase standard deviations for acceleration (a), and angular velocity (b) by mean Euclidean pair-wise distance.

Given 18 players are allowed on each team, there are 630 possible permutations of player couplings at any moment. In order to reduce the complexity of this analysis, pairings were then grouped as either intra-, or inter-team. Intra-team couplings consisted of all possible parings between players on the same team, and inter-team couplings consisted of all possible pairings within a team. k-means clustering was then used to separately generate clusters of similar player couples based on their proximity in a 2-dimensional space of acceleration and angular velocity relative phase standard deviations for the relevant permuted couplings. *k*-means was chosen as an arbitrary method for identifying groups of high and low coordination player couples to aid in the analysis and visualisation of results. These data are presented in Figures 4-3 and 4-4. *k*-means works by positioning *k* centroids repeatedly until the means diverge and has been previously used to visualise sporting styles (for examples see Gyarmati, et al., 2014; Sampaio, et al., 2015). *k* clusters were chosen based on a sum of squares plot.

4.2.8 Temporal Analysis

Box-plots of relative phase by time period (AF games consist of four 25-minute quarters) were compiled for inter- and intra-team couplings to compare mean coupling as a function of the time of the game.

4.3 RESULTS

A plot of relative phase standard deviations and player proximities for all permutation of player pairs is presented in Figure 4-2. This plot exhibits the characteristics of a positive relationship between relative phase behaviour and physical proximity on the playing area, confirming the proximity relationships described in Morgan and Williams (2012) can be observed in other large-field team sports. This indicates that players who are closer to each other are more likely to move in concert, both in terms of changes in velocity (Figure 4-2a), and in change in direction of movement (Figure 4-2b).

4.3.1 Intra-team Coordination

k-means clustering for the intra-team pairings with 15 clusters, are presented in Figure 4-3. The pairs were labelled according to their playing role and shirt number, which results in a unique label for each player. The player role labels are back/defender (B), midfielder (M), and forward/attacker (F). Additionally, the mean standard deviations for acceleration and angular velocity relative phase were computed over all parings in each cluster, and the mean values are presented in Table 4-1.

Table 4-1. Mean acceleration and angular velocity cluster standard deviations for intra- and inter-team clusters.

Cluster ID	*Intra*-Team Clusters		*Inter*-Team Clusters	
	Acceleration s.d.	Angular Velocity s.d.	Acceleration s.d.	Angular Velocity s.d.
1[a]	28.84	25.92	33.04	27.57
2	29.59	26.54	30.17	25.43
3	31.58	27.11	32.63	26.94

4[c]	30.45	25.82	29.13	25.20
5	31.17	26.00	30.59	26.30
6[d]	30.65	26.52	34.92	28.33
7	33.54	27.54	31.81	26.64
8[b]	34.61	28.24	30.53	26.89
9	32.80	27.29	33.53	28.26
10	32.16	26.68	31.32	27.13
11	29.71	25.72	30.97	25.78
12	31.18	26.66	34.20	27.54
13	30.18	26.39	31.19	26.56
14	33.26	28.11	29.89	26.22
15	32.69	27.77	32.17	27.65

[a] *High Coordination Pairs for Intra-team clusters.*
[b] *Low Coordination Pairs for Intra-team clusters.*
[c] *High Coordination Pairs for Inter-team clusters.*
[d] *Low Coordination Pairs for Inter-team clusters.*

From within the intra-team clusters, cluster 8 can be defined as the least coupled set of player pairs, while cluster 1 can otherwise be defined as the most coupled set of player pairs where the variability in phase coupling is lowest. For the purpose of comparison, we refer to cluster 1 as High Coordination Pairs (HCP), and cluster 8 as Low Coordination Pairs (LCP).

It is worth noting that each of the player pairs observed in the HCP cluster contained at least one midfield player (M), and most of the pairs were comprised of two midfielders. In contrast, the player pairs observed in the LCP cluster commonly featured pairings of players who exhibit the greatest median distance apart, indicating these are pairs of players who are positioned at opposite ends of the playing field.

4.3.2 Inter-team Coordination

k-means clustering for the inter-team pairings with 15 clusters, are presented in Figure 4-4. Analysis of the intra-team clusters revealed that cluster 6 can be defined as the least coupled set of player pairs, while cluster 4 can otherwise be defined as the most coupled set of player pairs.

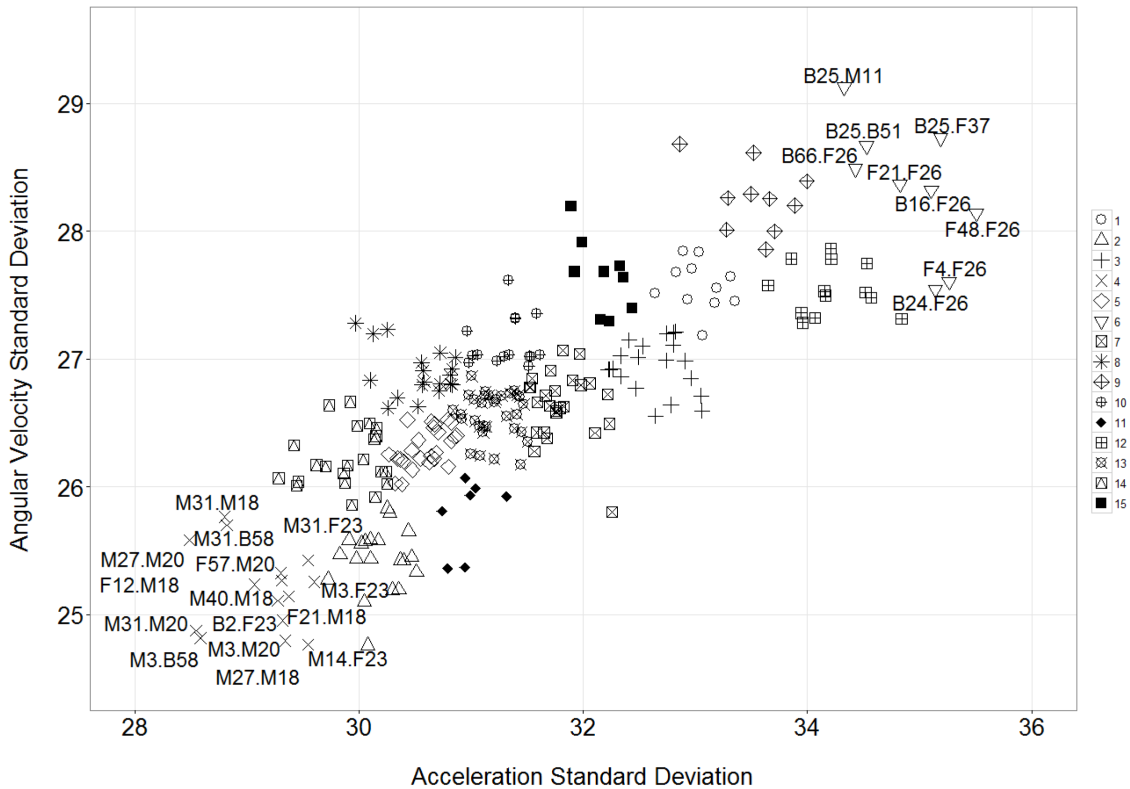


Figure 4-3. Angular velocity and acceleration relative phase standard deviations by cluster for intra-team couplings. Labels for player role are included for members of the LCP cluster (8) and the HCP cluster (1).

As previously described, we refer to cluster 4 and cluster 6 as High Coordination Pairs (HCP) and Low Coordination Pairs (LCP) respectively. As was previously observed in the intra-team pairings, the HCP consist predominantly of midfield positions, however they are more frequently paired with defender/back or forward players when compared to intra-team HCP. Of note, Oakleigh player M31 was part of four of the 15 HCP inter-team pairings. Inter-team LCP are largely comprised of pairings of positions who are positioned at opposite ends of the field.

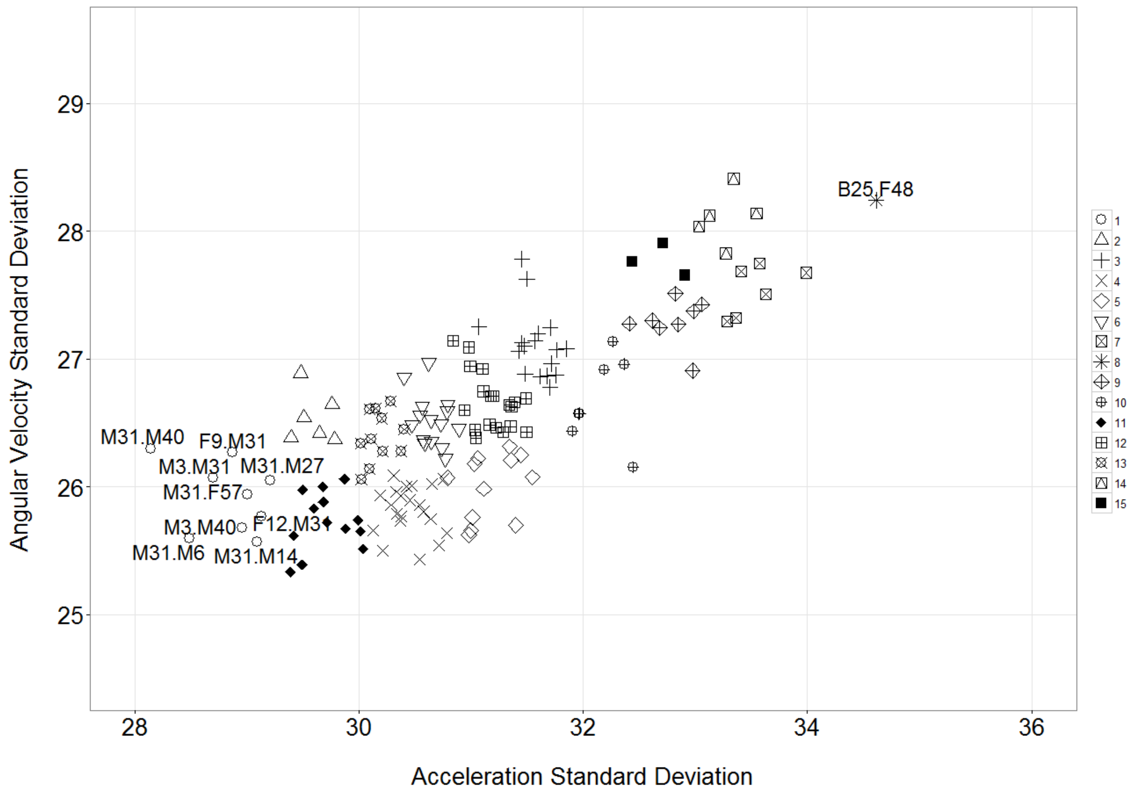


Figure 4-4. Angular velocity and acceleration relative phase standard deviations by cluster for inter-team couplings. Labels for player role are included for members of the LCP cluster (6) and the HCP cluster (4).

4.3.3 Temporal Analysis

Box-plots showing mean intra-team relative phase angle standard deviations by game quarter are presented in Figure 4-5. The mean standard deviations over each quarter increased systematically for the acceleration relative phase (Figure 4-5a), and angular velocity relative phase (Figure 4-5b). These plots indicate that the coordinative behaviour represented by coupled changes in running speed and direction of movement decreases systematically over the course of the AF match.

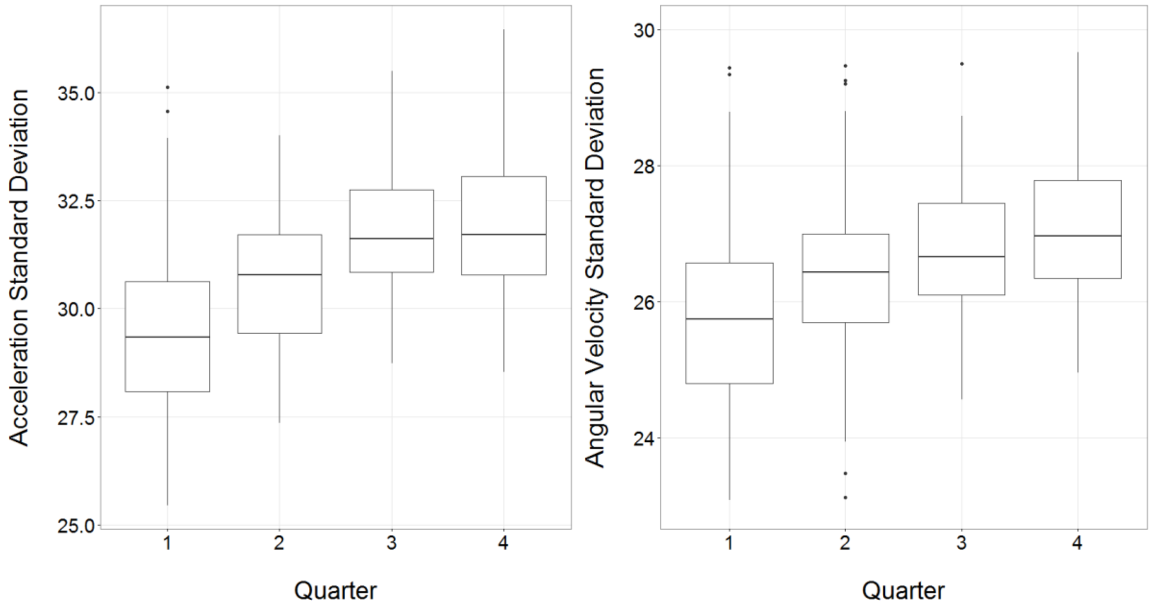


Figure 4-5. Relative phase standard deviation box plots for acceleration (a) and angular velocity (b) by game quarter.

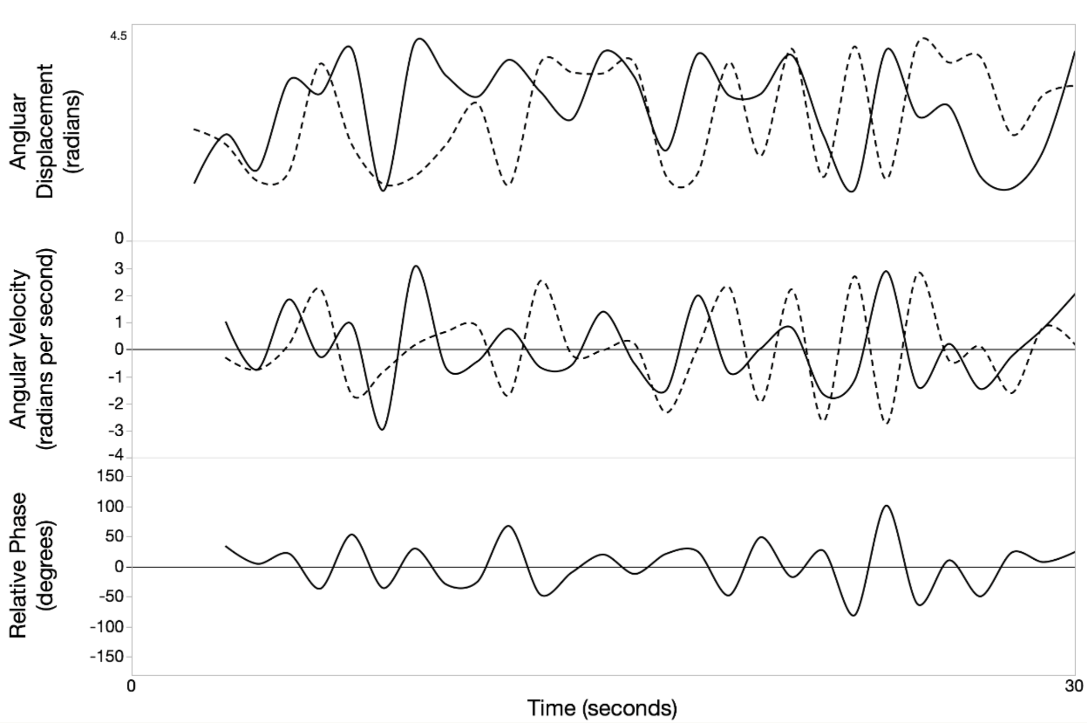


Figure 4-6. Exemplar angular displacement, angular velocity and relative phase for a LCP player coupling (solid = B25, dashed = F48).

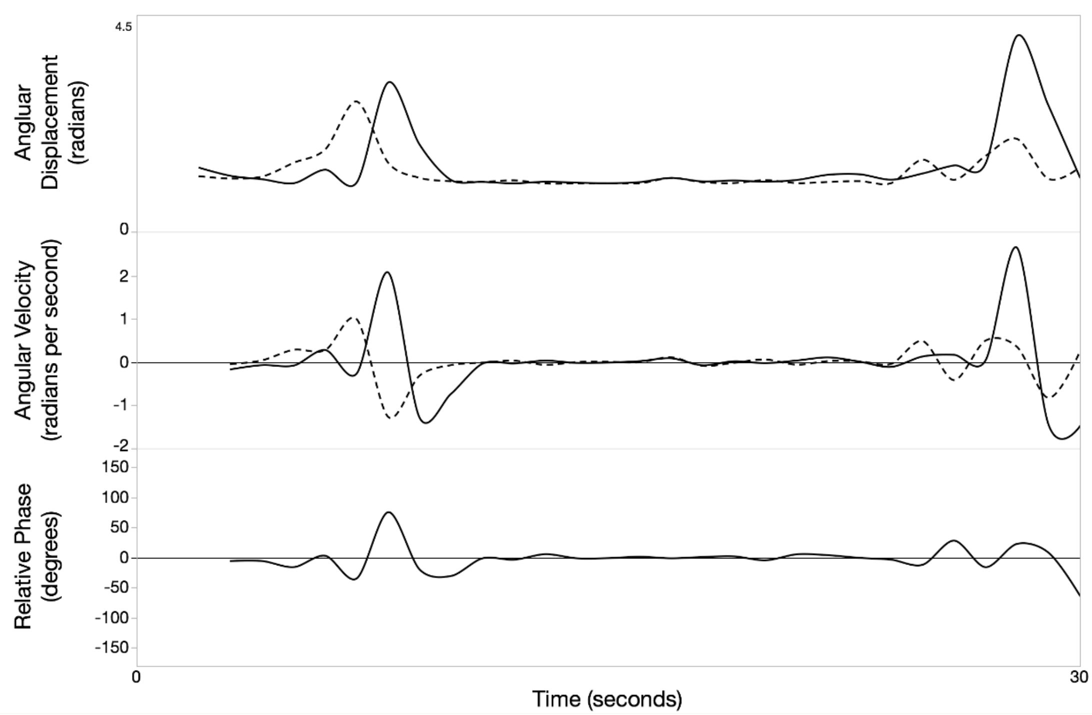


Figure 4-7. Exemplar angular displacement, angular velocity and relative phase for a HCP player coupling (solid = M31, dashed = F57).

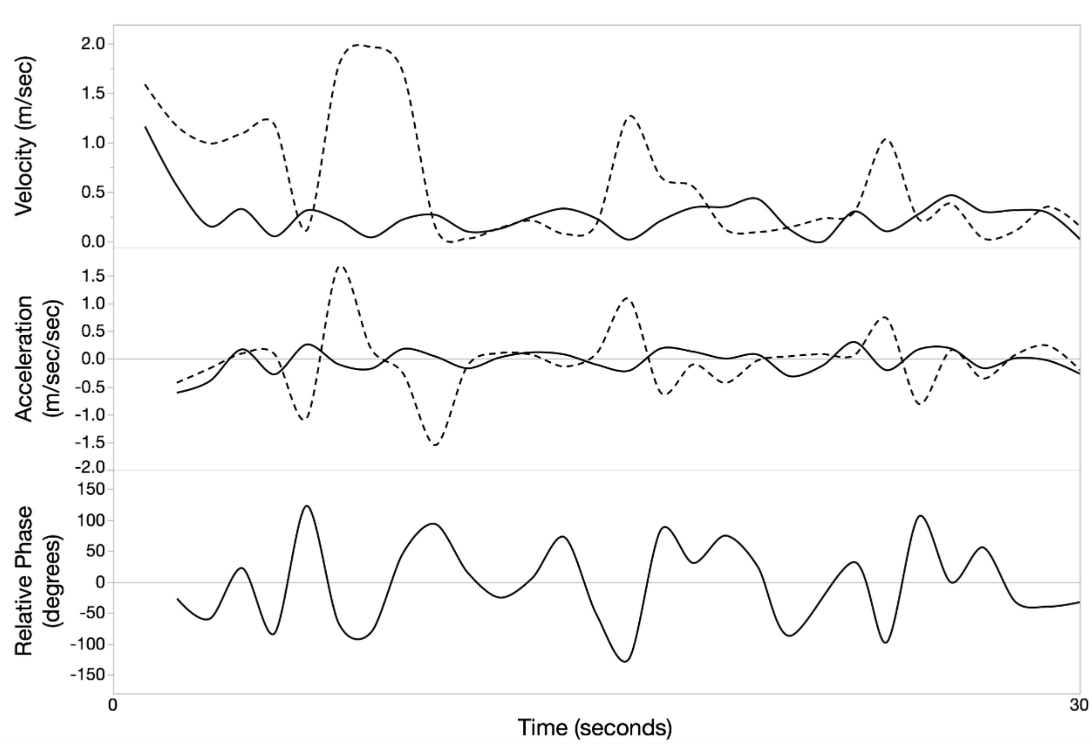


Figure 4-8. Exemplar velocity, acceleration and relative phase for a LCP player coupling (solid = B25, dashed = F48).

4.3.4 Player Coupling Exemplars

This modelling approach encodes multi-dimensional movement characteristics between player pairs and can be used make pair-wise classifications of players as strongly-, or weakly-coupled. In Figures 4-6 and 4-7 we present 30-second sequences of game play where the relative phase is derived from continuous angular displacement and angular velocity data for a weakly-coupled and strong-coupled pair respectively. While the exemplar shown in Figure 4-6 exhibits no obvious coordination, Figure 4-7 illustrates an example where the paired players are highly coordinated. Similarly, the same exemplar data are used to generate Figures 4-8 and 4-9 showing acceleration coupling between the same players. Here also, it is evident in Figure 4-9 that the players are changing speed in a highly coordinated manner, and the subsequent relative phase representation highlights both the general state of coordination, and the brief moment at approximately 6-seconds where the velocity changes are momentarily misaligned.

4.4 DISCUSSION

This paper demonstrates a method of estimating pair-wise player coupling that is suitable for team sport games, particularly games played on large playing areas. Previous approaches to estimating inter-player coordination leveraged movement features relative to fixed locations in the playing area, such as the middle “T” in squash (McGarry, et al., 1999). While this approach is helpful in small-court games, where the coordination of movement between players is functionally related to their physical location, the same rule may not apply in large field games. Australian Football is played on very large surfaces and is characterised by a high number of inter-personal exchanges that are functionally independent of their location on the playing area. For instance, mid-field players can engage in a direct contest at any location on the ground, and therefore, representing their interactive movement features in a dynamical systems framework using relative phase as the principal measure of coordination requires a new approach. We studied a spatiotemporal dataset in Australian Football and illustrated exemplar results using a position-independent approach.

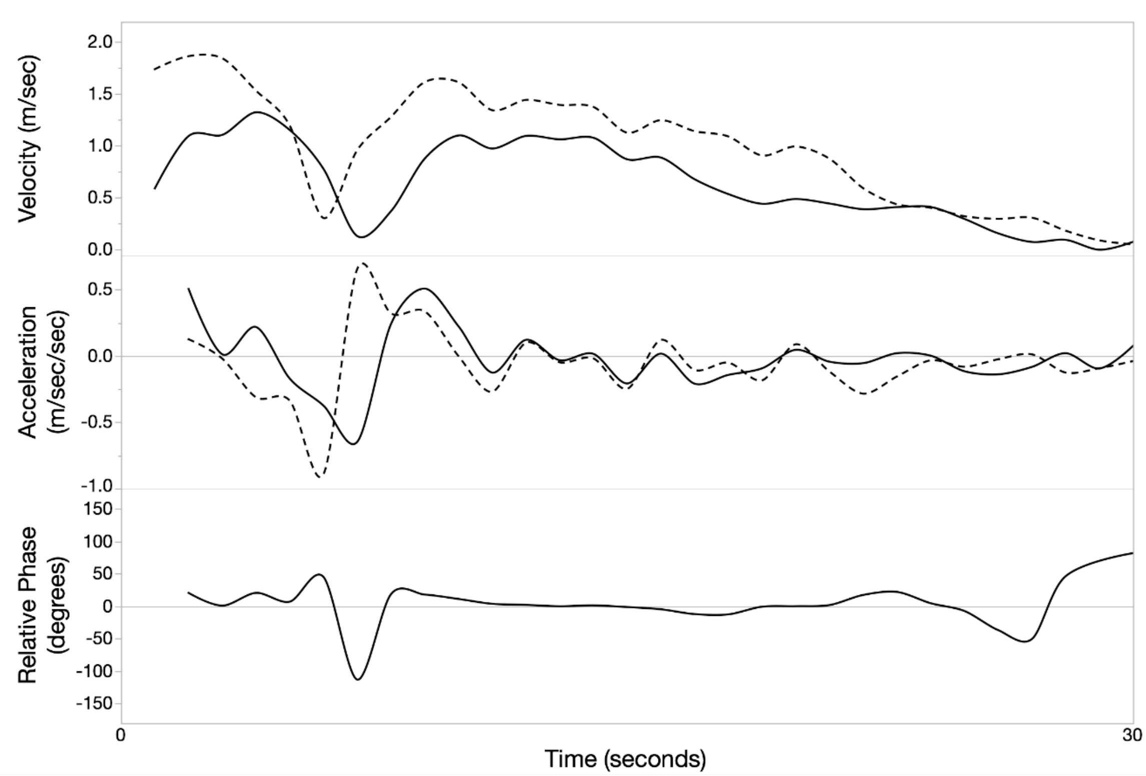


Figure 4-9. Exemplar velocity, acceleration and relative phase for a HCP player coupling (solid = M31, dashed = F57).

We demonstrated a method where phase angles are derived from separate phase portraits expressing velocity and acceleration, and angular displacement and angular velocity respectively. The angle from the origin of a phase portrait to the momentary acceleration or angular velocity value (in polar coordinates) can be regarded as a momentary state of movement. The pair-wise time-series comparison of these continuous states between players can then be used to estimate the degree of movement coordination.

The first observation we made is that there is an inverse correlation between inter-personal coordination and the distance between players. Players that are on average closer to each other, are also more tightly coupled, both in terms of their changes in acceleration, and changes in angular velocity. In many respects this is an obvious characteristic in team sports on large playing areas. Players who are further from the locus of gameplay would have no reason to move in concert with those distant players who are directly involved in the movement of the ball. This should be regarded as a sanity-check, and evidence of the contextual validity of our approach.

Analysis of the temporal characteristics of relative phase angles revealed that the standard deviation for the relative phase angles for both acceleration and angular velocity increases in later quarters. This indicates a potential diminishing degree structure within teams as time progresses. The decrease in coupling observed over the course of the match is also evidence that fatigue may impact the interpersonal coordination between players. This result could be explained by the proposition that as players fatigue the variability in their capacity or willingness to run with an intensity that matches other teammates may be increased. This would predictably result in lower coupling results on that dimension.

In representing coordinative behaviour as a dynamical system, one needs to consider possible control parameters that cause a stable system to shift into some alternate state. Lames has previously asserted that the players in a tennis match could be considered as two sub-systems, strongly coupled by features of the game (such as the size and dimensions of the court, and the net-based format of the game), and characteristics of the game play (e.g. a baseline rally) (Lames, 2006). In many respects, Australian football is divorced from the structural constraints of tennis, with a large, open field of play, and without the inherent periodic construct of a net-based sport where the ball travels back and forth between players. Nevertheless, if we can consider an AF match as a system with many degrees of freedom, represented by many paired sub-systems, the stability and phase characteristics of those subsystems may be driven by match-related control parameters, such as tactical constraints, fatigue, skill-level and role-based constraints. These constraints could be considered control parameters in the sense that they cause the characteristics of pair-wise behaviour to change. We may infer evidence of fatigue in the systematic degrading of coupled movements through the course of an AF match in our data. Further research should explore whether this trend is a robust feature of other matches. Similarly, tactical constraints (which are not considered here) are a potential control parameter. Consider the instance that a coach directs his/her team to play a zonal defence, where players defend spatial areas rather than specific opponent players. Such a scenario would predictably result in a substantive shift in the state of the pair-wise subsystems.

Most previous systems-based representations of movement coordination in sport have focussed on small court games such as tennis (Lames, 2006), squash (McGarry, et al., 1999), and basketball (Bourbousson, et al., 2010a; Bourbousson, et al., 2010b) (notwithstanding other work in soccer: e.g. Grehaigne, et al., 1997; Davids, et al., 2005; Frenken, et al., 2011). In each instance the order parameters considered are derived from the absolute location of individuals, or group centroids, in the playing area. This is a suitable approach to small court games, and one might argue further that even in soccer (which is played on a much larger field), role-based constraints on players assert some predictable regularity in both position and relative movement. For instance, a right-side winger will rarely wander beyond the unilateral spatial domain specific to that role. In those sports, regularity in role-based movement applies, and order parameters derived from absolute position may be helpful. In Australian football, however, the field-based location of players are mostly divorced from player roles (with only a few exceptions). In considering immediate derivatives of movement such as acceleration and angular velocity as order parameters, it is possible to represent coordinated behaviour in a way that is decoupled from the absolute location of players on the field of play. Importantly, this approach is not inconsistent, or mutually exclusive from position-based parameters. Indeed, it may provide additional insight in a dynamical systems framework to model the combination of position and movement parameters in court and small-field games.

4.5 CONCLUSION

Our aim in this work has been twofold. Principally, we aimed to demonstrate a method for quantifying the characteristics of inter-personal movement coordination between pairs of players in large-format team sports. We extend previous innovations that use a dynamical systems framework, and relative phase, to model interpersonal movement in small-court games (where the location of players is the order parameter) and present an alternative approach that leverages derivatives of interpersonal movement that are independent of pitch location. Secondly, we aimed to demonstrate the observable features of pair-wise coordination within and between teams in Australian football. This analysis revealed that movement coordination is higher amongst midfield players within the same team, and direct opponents between teams, and that some aspects of inter-personal movement coordination may diminish throughout matches, perhaps due to fatigue. This approach should enable future work to learn more about potential control parameters that may direct a dynamical systems representation of large-format team sports.

4.6 THESIS DISCUSSION

One limitation of this study is its reliance on AFL playing position classifications. Positions are dynamic and, in general, less restrictive than those in other invasion team sports (Jackson, 2016). Future work on the coordinative behaviour of Australian footballers could utilise information on dynamic playing positions (for example Jackson, 2016) to see if this yields more applicable results.

Chapter 5: Measuring Player Density in Australian Rules Football Using Gaussian Mixture Models[5]

Chapter Overview

Chapter 5 represents the first approach to measuring the spatial occupancy of players in the AFL. This chapter is a brief introduction and serves as a bridge towards more advanced methodology of later chapters.

This chapter contains introduction (Section 5.1), methods (Section 5.2), results (Section 5.3) and discussion and conclusion (Section 5.4) sections. Finally, the results of this chapter are discussed in the context of this thesis (Section 5.5). This chapter was presented at the 5th International Congress on Complex Systems in Sports (Spencer, et al., 2017b).

[5] Spencer, B., Morgan, S., Zeleznikow, J., & Robertson, S. (2017). Measuring player density in Australian Rules football using Gaussian mixture models. *Complex Systems in Sport, International Congress Linking Theory and Practice*. https://www.frontiersin.org/books/Complex_Systems_in_Sport_International_Congress_Linking_Theory_and_Practice/1381

5.5 THESIS DISCUSSION

The work presented in this chapter represents the first approach to the topic of spatial occupancy of players in the AFL. This study serves as an introduction to the topic and leads into the methodology of later studies.

This simplistic approach measures the density of player groups in the AFL, however can be applied to any team sport. To do so would require evaluation of GMM parameters (in particular, the number of components) for various field sizes and player numbers.

This approach has notable limitations. The biggest limitation is a lack of consideration of player motion and orientation. As a result, specific applications are limited. The trade-off is computability. GMM density fitting is considerably faster than more advanced spatial occupancy methods presented in the remaining sections of this thesis.

Chapter 6: A Method for Evaluating Player Decision-Making in the Australian Football League[6]

Chapter Overview

In this chapter an initial decision-making model is presented. Underpinning the model is a player motion model fit on the displacements of Australian footballers and an equity model described in O'Shaughnessy (2006) and Jackson (2016). See Section 2.2.2 of the Literature Review for an overview of these methods.

The approach is placed in the context of existing literature (Section 6.1) and described in Section 6.2. The results of this model are presented in Section 6.3. The chapter concludes with a discussion of the findings (Section 6.4) and final conclusions (Section 6.5). Finally, the results of this chapter are discussed in the context of the overall thesis (Section 6.6). This study was presented at Mathsport 2018 (Spencer, et al., 2018).

[6] Spencer, B., Bedin, T., Farrow, D., & Jackson, K. (2018). A method for evaluating player decision-making in the Australian Football League. *Mathsport 2018*

[redacted]

6.6 THESIS DISCUSSION

The work presented in this chapter builds upon the initial measurement of player density, as shown in Chapter 5. The addition of orientation and motion in the spatial occupancy model addresses limitations identified in a review of spatiotemporal research in team sports by Gudmundsson and Horton (2017). The simplistic player motion model in this chapter considers the effects of player motion and orientation; however, one shortcoming of the model is that it considers each displacement in a player's reachable region to be equally likely. It would be more realistic to model player motion with a distribution that measures the likelihood of each displacement. Recently, player motion has been modelled this way (Brefeld, et al., 2018; Fernandez & Bornn, 2018). Whether these models are applicable to Australian football will be explored in future chapters.

Combining the spatial influence of individual players allows for the calculation of a team's dominance. Dominance becomes a measure of spatial pressure – if dominance is low at a player's location, this indicates high opposition presence. Previous measures of spatial pressure have not considered the orientation and motion of pressuring opponents (Gudmundsson & Horton, 2017). The model presented in this chapter considered the orientation and motion of opponents. One limitation of this model is the use of discrete bounds for player motion. Logically, a player is not able to reach the limits of their observed motion in all situations. Hence, fitting the distribution of player displacements would improve this model. This concept is explored in Chapters 7 and 8.

This study represents the first research into the quantitative measurement of player decision-making in Australian football. Measuring player decision-making provides additional information about player performance. This has applications in performance analysis, player recruitment and coaching (Cervone, et al., 2014). Partitioning the components of a pass (i.e., risk, reward and Decision Value) provides a greater understanding into player performance. For example, it is possible to differentiate a player with poor execution but great decision-making, from a player with great execution but poor decision-making (who may be frequently choosing low risk passes compared to the former player).

Chapter 7: Fitting Motion Models to Contextual Player Behaviour[7]

Chapter Overview

This chapter presents a new player motion model, fit on player commitment to contest events. A player is said to commit to a contest if they reposition to participate in said contest. Participation is determined by a two-metre radius around the contest location. This represents a new approach to the topic of player motion and builds upon ideas of motion discussed in Chapter 6. The outputs of this chapter will be applied in decision-making analysis in Chapter 8.

This chapter includes an introduction (Section 7.1), overview of methods (Section 7.2**Error! Reference source not found.**) and the presentation and discussion of results (Sections 7.3; 7.4; 7.5). The results of this chapter are discussed in the context of the overall thesis (Section 7.6). This study was accepted for presentation at the 12th International Symposium on Computer Science in Sport which will be held in Moscow, Russia from 8-10 July 2019. Presented papers will be published by Springer in the Advances in Intelligent Systems and Computing series.

[7] Spencer, B., Jackson, K., Robertson, S. (2019). Fitting motion models to contextual player behaviour. *International Symposium on Computer Science in Sport*
https://iacss2019.ru/

Fitting Motion Models to Contextual Player Behaviour

Abstract

The objective of this study was to incorporate contextual information into the modelling of player motion. This was achieved by combining the distributions of forthcoming passing contests that players committed to and those they did not. The resultant array measures the probability a player would commit to forthcoming contests in their vicinity. Commitment-based motion models were fit on 46220 samples of player behaviour in the Australian Football League. It was found that the shape of commitment-based models differed greatly to displacement-based models for Australian footballers. Player commitment arrays were used to measure the spatial occupancy and dominance of the attacking team. The spatial characteristics of pass receivers were extracted for 2934 passes. Positional trends in passing were identified. Furthermore, passes were clustered into three components using Gaussian mixture models. Passes in the AFL are most commonly to one-on-one contests or unmarked players. Furthermore, passes were rarely greater than 25 m.

7.1 INTRODUCTION

The measurement of a player's spatial occupancy can reveal insights into space, congestion and passing opportunities. While early research into spatial occupancy considered players as fixed objects, recent iterations of Voronoi-like dominant regions have incorporated the effects of player motion (Gudmundsson & Horton, 2017; Brefeld, et al., 2018). Underlying these approaches is limited consideration of the continuous nature of space. Should the application of spatial occupancy involve possession outcomes, space should be considered relative to the ball.

Recent studies have addressed this concept. Fernandez and Bornn (2018) measured the spatial dominance of teams by representing a player's influence as a bivariate normal distribution. The result considers the continuous nature of space but is not fit on empirical data. Brefeld et al. (2018) fit player motion models on the distribution of a player's observed displacements but did not consider the context of those displacements (i.e., the current possession location). Logically, the amount of spatial dominance a team exhibits over a location need be measured relative to how players would control said space if the ball were moved to that location.

In this study we present a method of fitting player motion models with consideration of displacement context. Models are fit on player commitment to passing contests, rather than raw displacements. Resultant models measure the probability a player would contest a pass to locations in their vicinity. We demonstrate the applications of these models in the analysis of kicking in the Australian Football League (AFL).

7.2 METHODS

Ball tracking is not commercially available in AFL; however, ball location can be inferred from play-by-play data. Player motion models are proposed as an adequate forecast of future behaviours in the absence of precision ball tracking. Hence, the objective of this study was to model player motion with consideration of the context of player displacements, without increasing their dimensionality beyond consideration of location, velocity and time.

7.2.1 Data and Pre-processing

LPS player-tracking data (x, y, t) were collected from the 2017 and 2018 AFL seasons. Tracking data (10 Hz) were consolidated with play-by-play event data (known as transactions). Transactions are recorded to the nearest second, hence are assumed to occur at the beginning of the second when combined with LPS datasets. Player orientation and velocity were calculated from the tracking data under the assumption that players were oriented in the direction of their movement. For analysis, passes that begin with and ended with a mark were extracted (*mark-to-mark* passes). This constraint ensured that location could be inferred. A *mark* is awarded when a) a player catches a kick on the full, and b) the kick travelled at minimum of 15 m.

7.2.2 Possession Contests

Commitment models are fit on player participation to forthcoming passing contests. Passing contests are pass events in which more than one player attempts to win the ball. In the AFL datasets, events that fit this criterion are *contested marks* and *spoils* transactions. The former refers to a pass caught by a player while under pressure and the latter relates to a marking attempt in which the ball is knocked away by an opponent. Passing contests are henceforth referred to as contests.

7.2.3 Modelling Process

Each contest involves two events of interest: the pass that preceded the contest and the contest transaction. The timestamps of these events are referred to as t_p and t_c respectively. When referring to a player's commitment we are referring to the likelihood a player will commit to a forthcoming contest, given their position and momentum at t_p. The commitment modelling process is as follows:

1. Player momentum and position at t_p and the ball's travel time, or *time-to-point*, are recorded. The latter is simply $t_c - t_p$.

2. For each player, compute the relative location of the contest. This relative location is considered a potential player displacement. The relative location is as follows:

$$\theta = cos^{-1}\left(\frac{\overrightarrow{AB}\cdot\overrightarrow{BC}}{\|\overrightarrow{AB}\|\cdot\|\overrightarrow{BC}\|}\right) \quad (1)$$

$$(x, y) = (d \cdot \cos\theta\,, d \cdot \sin\theta) \quad (2)$$

where *AB* is the player's movement vector, *BC* is the displacement vector to the contest and *d* is the Euclidean distance between the player and the contest.

3. If the Euclidean distance between the player and the contest is less than two meters at t_c, player commitment (*C*) is recorded as 1 (hence, the player realized the potential displacement), else if greater than two meters, commitment is recorded as 0.
4. The dataset is partitioned into *commitment* and *no commitment* sets along *C*.
5. Distribution of both datasets is estimated via Kernel density estimation (KDE) with Gaussian kernels. Datasets are four-dimensional, containing the relative contest location (*x*, *y*), player velocity (*v*) and ball time-to-point (*t*).
6. The distributions are combined, weighted according to event frequency, using the following function:

$$p_i(x, y, v, t) = \frac{wf_{C=1}(x,y,v,t)}{wf_{C=1}(x,y,v,t)+(1-w)f_{C=0}(x,y,v,t)} \quad (3)$$

where $f_{C=1}$ and $f_{C=0}$ are the distributions, and *w* is the weight.

The two-meter threshold for player commitment (step 3) was chosen as an adequate distance after discussion with AFL analysts. Individual distributions represent the density of contests that were committed to ($f_{C=1}$) and those that were ignored ($f_{C=0}$). By combining the distributions (Eq. 3) the resulting variable (p_i) measures the probability that a new sample (given x, y, v, t) belongs to the commitment distribution. The resultant array measures a player's spatial influence. A player's influence is a forecast of their behaviors in respect to a forthcoming passing contest.

7.2.4 Spatial Metrics

We measure the spatial influence of a team as the sum of the influence of its players:

$$Inf(x,y) = \sum_{i=1}^{18} Pr_i \tag{4}$$

and dominance is the proportion of space a team owns at a location:

$$Dom_a(x,y) = \frac{Inf_a(x,y)}{Inf_a(x,y) + Inf_o(x,y)} \tag{5}$$

7.2.5 Passing Analysis

Commitment models have previously been used to analyse decision-making in the AFL (Spencer, et al., In Review). In this study, commitment models are used to analyse characteristics of passes. Mark-to-mark passes were extracted from the transactional dataset. The kicking distance (metres), spatial dominance, influence and *equity* of passes were recorded. AFL field equity (FE) is a measure of the value of space described in Jackson (2016). The equity of a pass is the change in FE between the passer and receiver ($equity = FE_{receiver} - FE_{passer}$). Metrics were analysed at different field locations. Spearman correlation coefficient was used to assess the relationship between metrics and the distance between the receiver and the attacking goals. To define passing types, characteristics of passes were clustered via Gaussian mixture models, with the number of components chosen via the elbow method (Madhulatha, 2012).

7.3 RESULTS

An example output visualizing the spatial dominance and influence of an attacking team is presented in Fig. 7-1, where areas of darker green represent higher dominance.

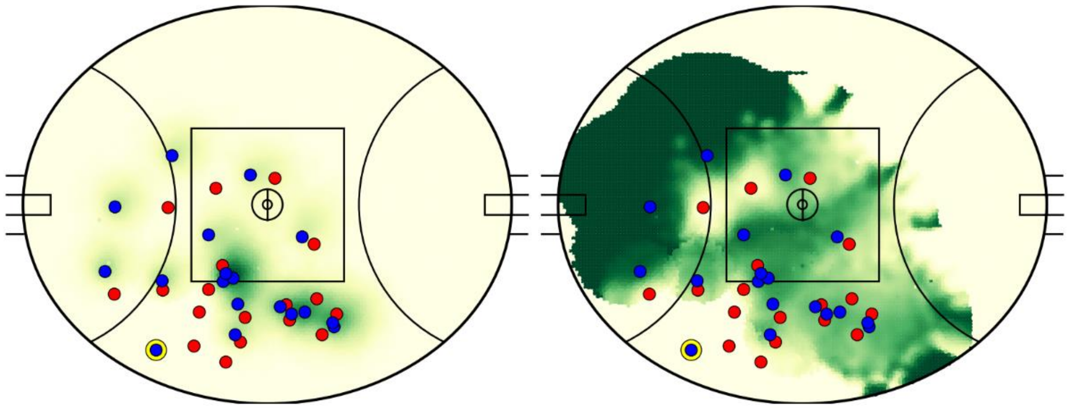


Figure 7-1. Example output of spatial dominance for the attacking team (in blue). The player with possession is circled in yellow (towards the lower boundary).

7.3.1 Commitment Models

Player commitment behaviour was recorded for 46220 samples. The C = 1 and C = 0 datasets consisted of 6392 and 39828 samples (w = 0.14). Fig. 7-2 visualizes commitment models for two velocities for t = 2 s. These are compared to motion models fit on player displacements (as in Brefeld, et al., 2018). Fitting displacements (Fig. 7-2b, Fig. 7-2d) suggests players are unlikely to reorient, hence are insufficient for modelling behaviour to forthcoming contests.

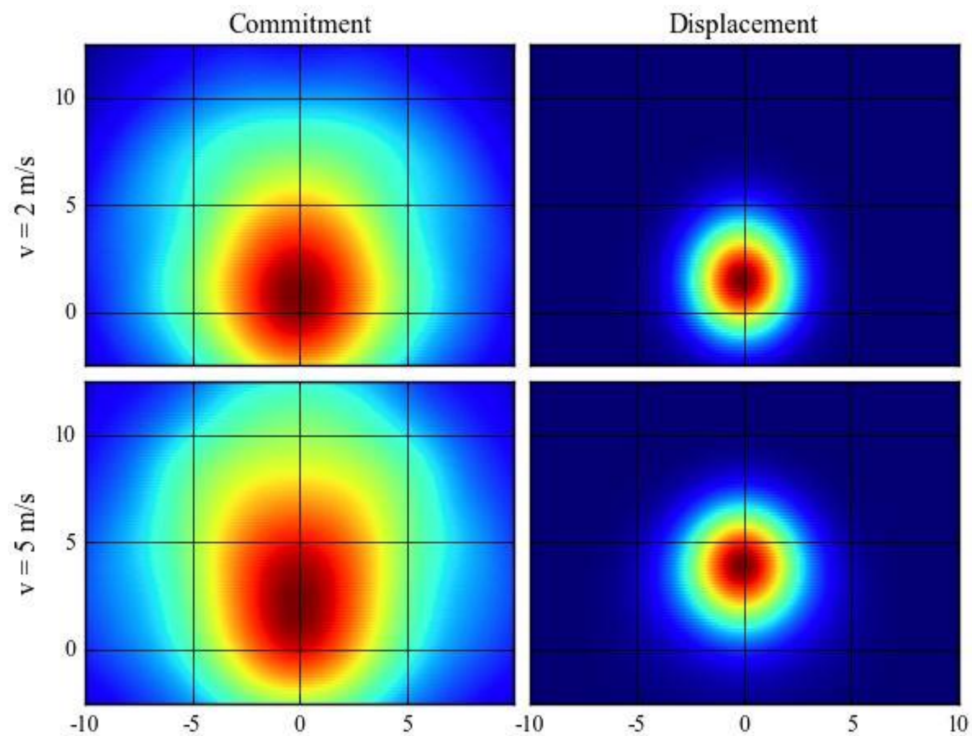


Figure 7-2. Player commitment (left) and displacement (right) motion models for v = 2 m/s (top) and v = 5 m/s (bottom). Density represents the probability of making a displacement.

7.3.2 Passing Analysis

A total of 2934 passes were analysed. Two-dimensional distributions of passing features are presented in Fig. 7-3. Dominance of passes is bimodal. The dominance and influence of receivers was recorded and smoothed by field location (Fig. 7-4). There is a trend towards passes to lower dominance receivers towards the attacking goal. Furthermore, influence of receivers is high in the in the forward 50 region. This is indicative of kicks to congested groups, rather than individual players. Minimal correlation was found between the distance to objective and both dominance ($\rho = 0.05$, $p < 0.01$) and influence ($\rho = -0.08$, $p < 0.01$).

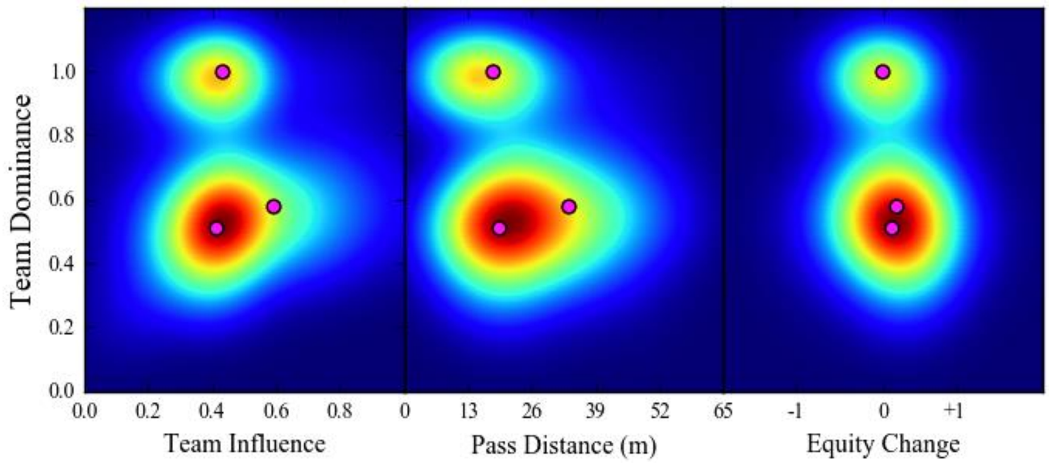


Figure 7-3. Distributions (estimated via KDE) of (a) Influence, (b) Distance and (c) Equity relative to Dominance. GMM Component means are presented as magenta points in the 2D plots.

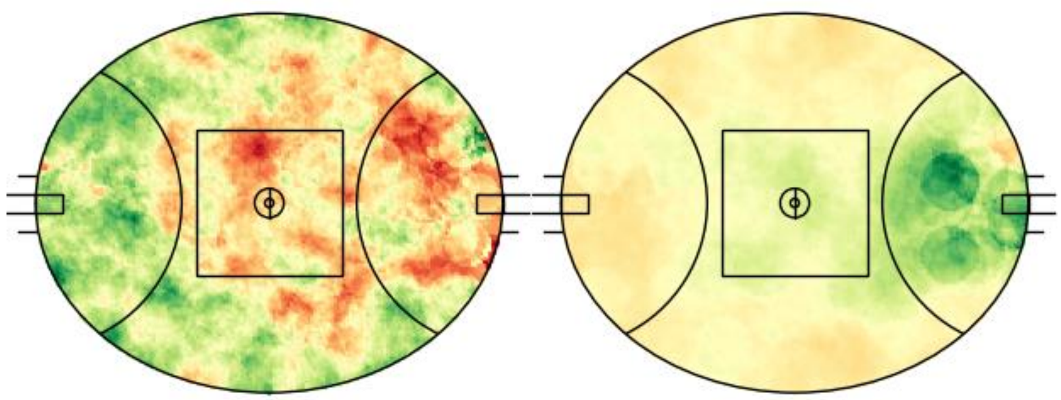


Figure 7-4. Smoothed spatial dominance (left) and influence (right) of pass receivers. Attacking team is moving left to right. High dominance and influence are indicated by darker green regions.

7.3.3 Passing Clusters

Passes were clustered via GMM into three components. Component means are visualized in two-dimensions in Fig. 7-3. Characteristics of the components are presented in Table 7-1. Component 1 represents a medium-range pass to a group of players in congestion (*influence* > 0.5, *dominance* < 1.0), component 2 is a short-range pass to an open player (*dominance* = 1.0) and component 3 is a short-range pass to a one-on-one contest (*influence* < 0.5, *dominance* < 1.0).

Table 7-1. The weight and means of Gaussian mixture model components.

Variable	Component 1	Component 2	Component 3
Weight	0.43	0.24	0.33
Dominance (%)	0.58	1.00	0.51
Influence	0.59	0.43	0.41
Distance (m)	33.3	17.9	19.4
Equity	0.09	0.00	0.06

7.4 DISCUSSION

This study presented a method for fitting player motion models with consideration of the context of player displacements. This was achieved via the fitting of participation to forthcoming events, rather than to observed player displacements, representing a new approach to player motion models. Additionally, the models in this study fit the distribution of samples in four-dimensions, choosing to consider velocity and time as continuous rather than categorical as in Brefeld et al. (2018).

It was observed that commitment models suggest a higher likelihood of reorientation than motion models fit on player displacements (see Fig. 7-2). In particular, displacement-based models forecast very few repositions in the negative *y*- axis. Observation of player commitment behaviours suggest reorientation is possible in all directions. The low probability of reorientation in displacement-based motion models is likely due to the nature of gameplay in AFL. The large field size and typical gameplay result in players frequently following the ball, rather than holding formations. Hence, for the analysis demonstrated in this study, motion models fit on player displacements are inadequate for describing future behaviour.

Commitment models are fit on behaviour to the next possession, hence are limited to applications that consider short-term behaviour. At higher velocities, the spread of a player's influence increases and the shape changes (see Fig. 7-2). These considerations do not affect the applications presented in this study. It should be noted that commitment models were fit on 46220 samples which is roughly equivalent to the number of one-second displacements a player would make in a single match. As a result, these models may be less smooth than motion models fit on displacements (Fig. 7-2). Bandwidth selection during the fitting process can be modified to account for this.

A noteworthy limitation of commitment models is a reliance on transactions of differing frequency to player-tracking datasets. As a result, transactions and player-tracking may be misaligned by up to one second. The generous commitment radius of two metres deals with this to an extent, however higher frequency transactions would reduce the noise of resultant models.

Studies analysing passing in the AFL have previously utilized discrete passing features and manually collected data (e.g., Robertson, et al., 2019). The computation of spatial features presents continuous metrics for passing analysis. Spatial dominance of receivers was found to be bimodal at dominance of an equal contest (dominance = 0.5) and an open player (dominance = 1.0). It was noted that passes to open players were rarely greater than 25 m. There is an indication that the spatial characteristics of receivers differs by region, despite minimal correlation between these metrics and a player's distance to the goalposts. In particular, the influence of receivers was higher in the forward 50 region than elsewhere. This is indicative of a pass to a congested group of players. Furthermore, early results show that receiver dominance is higher in the defensive 50 region, indicative of risk aversion in defensive positions. These results may be explained by team formations. Players have more space to work with when a team has possession in their defensive 50.

This space decreases as the ball is moved towards the attacking goalposts, hence players become more congested.

Analysis of the spatial characteristics of passing produced three passing clusters. While the equity of all components was minimal, the short-range pass to an open player (component 2) had a mean equity of 0.00, hence does not typically improved a team's scoring chance. This may be a pass to stall play in the absence of better options. The low mean passing distance of components 2 and 3 (< 20 m) suggests a tendency to execute short-range passes.

While the analysis in this study has focused on on-ball possessions, measures of spatial occupancy have applications in off-ball analysis. Fernandez and Bornn (2018) utilized similar methodology to analyse space creation of off-ball actions in soccer. Future applications of spatial occupancy should continue the development of these topics.

7.5 CONCLUSION

A new method for measuring player spatial occupancy was exemplified in this study. The occupancy of Australian footballers was estimated via the probability they would reposition to forthcoming passes contests. When compared to displacement-based motion models in Australian football, commitment models were found to be a better representation of contextual player behavior. Resultant commitment models were used to describe the kicking landscape of AFL footballers, finding that passes were frequently to one-on-one contests or open players. Furthermore, long kicks are infrequent and there is a significance number of passes around the minimum marking distance.

7.6 THESIS DISCUSSION

The commitment-based motion model presented in this study is an evolution of the player motion model presented in Chapter 6. By considering the probability of commitment at each location in a player's vicinity, this model is a more accurate representation of player behaviour than displacement-based models. The difference in shape of these models compared to the models from Brefeld et al. (2018) (Fig. 7-2) supports this belief for the applications presented in this study. It is likely that the decision-making model from Chapter 6 would be improved with commitment-based motion models. This will be the objective of Chapter 8.

Passing and passing contests were identified as important events in Chapter 2 (Spencer, et al., 2016). Analysis of passing in this chapter revealed interesting insights into mark-to-mark passes in the AFL. It was identified that team profiles defined by performance indicators struggled to discriminate winning teams (Chapter 2). The development of passing styles, defined by unsupervised clustering, is a step towards the development of team playing styles measured with spatiotemporal datasets.

The LPS player tracking systems used in this chapter are Catapult Clearsky devices[8]. The validity of these units has been the subject of a previous study in which they were found to have adequate validity compared to optical systems (Serpiello, et al., 2018). Further information on the validity of LPS devices can be found in Section 2.1 of this thesis.

[8] https://www.catapultsports.com/products/clearsky-t6

Chapter 8: Modelling the Quality of Player Passing Decisions in Australian Rules Football Relative to Risk, Reward and Commitment[9]

Chapter Overview

In this chapter, the decision-making model from Chapter 6 is revisited. The addition of the motion model developed in Chapter 7 produces more realistic measures of spatial control. This chapter concludes the work that has been developed in preceding chapters.

This chapter consists of an introduction (Section 8.1), related work (Section 8.2), methods (Section 8.3), results (Section 8.4), discussion (Section 8.5) and conclusions (Section 8.6). The content of this chapter is under review in an indexed Q1 journal.

[9] Spencer, B., Jackson, K., Bedin, T., & Robertson, S. (In Review). Modelling the quality of passing decisions in Australian Rules football relative to risk, reward and commitment.

frontiers

Modelling the quality of player passing decisions in Australian Rules football relative to risk, reward and commitment

Bartholomew Spencer[1*], Karl Jackson[2], Tim Bedin[2], Sam Robertson[1]

[1]Institute for Health & Sport (IHES), Victoria University, Melbourne, Victoria, Australia

[2]Champion Data Pty Ltd, Melbourne, Victoria, Australia

*** Correspondence:**
Bartholomew Spencer
bartholomew.spencer@vu.edu.au

Keywords: Motion models, Spatiotemporal, Decision-making, Team sports, Australian Rules football, Player tracking

Abstract

The value of player decisions has typically been measured by changes in possession expectations, rather than relative to the value of a player's alternative options. This study presents a mathematical approach to the measurement of passing decisions of Australian Rules footballers that considers the risk and reward of passing options. A new method for quantifying a player's spatial influence is demonstrated through a process called commitment modelling, in which the bounds and density of a player's motion model are fit on empirical commitment to contests, producing a continuous representation of a team's spatial ownership. This process involves combining the probability density functions of contests that a player committed to, and those they did not. Spatiotemporal player tracking data was collected for AFL matches played at Docklands Stadium in the 2017 and 2018 seasons. It was discovered that the probability of a player committing to a contest decreases as a function of their velocity and of the ball's time-to-point. Furthermore, the peak density of player commitment probabilities is at a greater distance in front of a player the faster they are moving, while their ability to participate in contests requiring re-orientation diminishes at higher velocities. Analysis of passing decisions revealed that, for passes resulting in a mark, opposition pressure is bimodal, with peaks at spatial dominance equivalent to no pressure and to a one-on-one contest. Density of passing distance peaks at 17.3 m, marginally longer than the minimum distance of a legal mark (15 m). Conversely, the model presented in this study identifies long-range options as have higher associated decision-making values, however a lack of passes in these ranges may be indicative of differing tactical behaviour or a difficulty in identifying long-range options.

1 Introduction

Team sport athletes are consistently presented with situations in which their decisions effect the immediate state of a game. These consist of overt on-ball decisions relating to passing or shooting, however also include off-ball actions such as occupation of a given space. Whilst previous works have quantified the impact of a decision on some measure of possession expectation (Cervone et al., 2014; Cervone et al., 2016; Jackson, 2016) or on measures of spatial control (Fernandez & Bornn, 2018), their value has typically been measured by the change in some metric or relative to a contextual mean. We believe the value of a player's decision should be quantified relative to the alternative options that were available. Although a pass may yield a positive increase in a team's scoring chance by x, the decision is by definition sub-optimal if alternatives exist that increase it by greater than x. By measuring a player's decision relative to their options, we can quantitatively attribute value to a player's decision-making abilities, further decoupling components of a player's performance.

The expected possession value (EPV) metric considers spatiotemporal data, match phase and player behaviours to quantify possession outcomes in basketball (Cervone, et al., 2014; Cervone, et al., 2016). Computing the change in EPV between possessions assigns a value to player possession contributions. A player's decision is valued relative to the tendencies of other players in the same situation, producing a player's EPVA (EPV-added over replacement) as the sum of a player's EPV-added ($EPV_{end} - EPV_{start}$) across all possessions. In Jackson (2016), Australian Rules footballers ranking points are the sum of their possession contributions, valued relative to the event and location, an extension of the measure of field equity developed in O'Shaughnessy (2006). Similar to Cervone et al., (2014), player contributions are measured relative to mean outcomes and a player is deemed to be a good decision maker if their involvement improved their team's field equity, a measurement of scoring chance relative to match phase and possession location. In Horton et al., (2015), football passes were labelled qualitatively using machine learning algorithms with quantitative inputs, learnt from manual labelling of passing quality by sporting professionals. The inclusion of player dominant regions, a method of bounding a player's spatial ownership via consideration of player momentum, suggests the quality of a pass has some dependence on a team's spatial control.

Common amongst these studies is the valuation of player decisions with respect to some change in possession expectation. Another approach would be to value decisions relative to alternative options, however, modelling this problem presents unique challenges. While quantifying a decision after the fact can be done by measuring the change in a given objective, each option available to a player has an accompanying probability of success. Multiple studies have measured the risk of passes in football. In Szczepanski and McHale (2016), the success of a pass depended upon the skill of a player and their teammates, field position of the pass location and destination, and pressure. The latter was approximated dependent on a player's typical playing positions and time between passes, rather than consideration of opponent locations due to an absence of player tracking data. Power et al., (2017) measured the risk and reward of passing options using spatiotemporal tracking data, where the risk of a pass considers player velocity, defender proximity and momentum, and possession statistics and the reward of a pass is the probability that the pass will result in a shot on goal. From their measure of risk, the risk tendencies and completion rates of players were analysed. Our recent work in AFL produced measures of risk and reward via discrete player motion models and measures of future possession expectations respectively (Spencer, et al., 2018).

In this study we value a player's passing decisions through consideration of the risk and reward of their options. We measure the risk of a pass through modelling of individual and team spatial control, and reward via a measure of field equity detailed in Jackson (2016). We present a new method for modelling spatial control via probabilistic modelling of player commitment to contests with consideration of their momentum. This process, referred to as commitment modelling, produces player motion models that more realistically represent player behaviour based on their proximity to important events. We use the resultant decision-making model to analyse characteristics of player decision-making, its predictability, and distributions of risk taking within teams.

2 Related Work

2.1 Motion Models

There exist many methods for representing a player's spatial occupancy. One common approach, particularly in football, is that of Voronoi tessellations which bound a player's owned space as the space in which they could occupy before any other player. Simple applications of this approach do not consider player orientation, velocity, or individual physical capabilities (e.g. Fonseca, et al., 2012). Taki and Hasegawa (2000) produced variations incorporating a player's orientation, velocity, but assumed consistent acceleration. Fujimaru and Sugihara (2005) proposed an alternative motion equation, adding a resistive force that decreases velocity. This approach involved a generalised formula that more realistically represented a player's inability to cover negative space if moving at speed. Gudmundsson and Wolle (2010) individualised these models, fitting a player's dominant region from observed tracking data.

Underlying these models is an assumption that spatial ownership is binary. That is, each location on the field is owned completely by a single player, determined by the time it would take them to reach said location, henceforth referred to as their time-to-point. Through observations of contests, we propose that ownership of space is continuous. For a given location, if the time-to-point of the ball is greater than the time-to-point of at least two players, then no single player owns the space completely. This distinction is important if we wish to quantify spatial occupancy (and its creation) relative to the ball, given its time-to-point, as we need to account for changes in field formations that could occur between possessions.

Recent papers have addressed this. The density of playing groups was explored with Gaussian mixture models in Spencer et al. (2017). Spencer et al. (2018) produced a smoothed representation of a team's control using non-probabilistic player motion models fit on observed tracking data. While a team's ownership was expressed on a continuous scale, the use of motion models with discrete bounds may result in unrealistic estimations of a player's influence (Brefeld et al, 2018). Fernandez and Bornn (2018) measured a player's influence area using bivariate normal distributions that considered a player's location, velocity, and distance to the ball. The result is a smoothed surface of control in which a team's influence over a region is continuous, however the size of a player's influence is within a selected range, rather than learnt from observed movements. Recently, Brefeld et al. (2018) fit player motion models on the distribution of observed player movements, utilising these probabilistic models to produce more realistic Voronoi-like regions of control. In the interest of computing time, two-dimensional models were produced for different speed and time bands, hence the resultant models are not continuous in all dimensions.

Given its contested and dynamic nature, a continuous representation of space control is preferable (e.g. Fernandez and Bornn, 2018; Spencer, et al., 2018). Furthermore, a player logically exhibits greater control over space in which they are closer, hence we develop probabilistic motion models in this paper. When probabilistic models are fit on the entirety of a player's movements (as in Brefeld et al., 2018), we find that the probability of player reorientation is underestimated. In decision-making modelling, our interest is in measuring the contest of space that would occur if the ball were kicked to said space. Hence to represent this realistically, it is important to fit the distribution of player movements observed under similar circumstances. We model a player's behaviour when within proximity of contests. We achieve this via a procedure we call *commitment modelling*, where we fit the distribution of player commitment to contests in four dimensions (velocity, time, and x- and y- field position). The result is a realistic representation of player behaviours when presented with the opportunity to participate in a contest.

3 Materials and Methods

3.1 Data and Pre-Processing

Spatiotemporal player tracking data was collected from the 2017 and 2018 AFL seasons. Data were collected by local positioning system (LPS) wearable Catapult Clearsky devices (Catapult Sports, Melbourne, Australia), situated in a pouch positioned between the players' shoulder blades. Positional data in the form of Cartesian coordinates was recorded at a frequency of 10 Hz for all 44 players. To ensure consistent tracking and field dimensions, analysed matches were limited to those played at Docklands Stadium, Melbourne. Play-by-play transactional data (i.e., match events such as kicks, marks, and spoils, and their associated meta-data) were manually collected by Champion Data (Champion Data Pty Ltd, Melbourne, Australia). These events are henceforth referred to as transactions. Consolidation of transaction and tracking data was used to infer ball position from possession, as ball tracking data is not available in Australian Rules football. Transactions were recorded to the nearest second, hence it was assumed they occurred at the beginning of a second when matched to 10 Hz tracking data. If the location of one or more players was lost during a passage of play, said passage was omitted from the analysis. In total, data from 60 matches was used in this study.

A player's velocity and orientation were calculated from raw positional data. It was assumed players were oriented in the direction of their movement, hence orientation was extracted from consecutive tracking samples (i.e., a player's orientation was recorded as the angle formed by consecutive tracking samples, relative to the positive y-axis). A player's change in orientation was considered as the angle between two vectors, $\overrightarrow{AB}$ and $\overrightarrow{BC}$, where A, B, and C are the player's three most recent positions, and the angle describes the change in orientation between positions B and C (Equation 1). The same process was used to calculate the location of an event relative to a player (where A and B are a player's previous and current position, and C is the location of interest). Velocity, recorded in metres/second, was calculated as the Euclidean distance between a player's current position and their position, one second prior.

$$\theta = \cos^{-1}\left(\frac{\overrightarrow{AB}\cdot\overrightarrow{BC}}{\|\overrightarrow{AB}\|\cdot\|\overrightarrow{BC}\|}\right) \tag{1}$$

In this study, only player decisions following a mark were included, given that a mark provides the player with time to make an informed decision. In Australian Rules football, a mark is a kick greater than 15 m that is received by a player on the full (i.e., without bouncing). To locate the destination of a player's kick following their mark, the next transaction must also be a mark. If the next possession following a kick is not a mark, we are unable to reliably locate the intended target, given a reliance on transactions to infer ball position.

3.2 Commitment Modelling

For analysis purposes, a contest was defined as a transaction following a pass in which at least one player from each team was involved and the ball location (for both the preceding kick and the receive) could be inferred from the consolidated datasets. In this study, the contest transaction types were spoils and contested marks. The former is an attempted pass that was physically prevented by the opposition and the latter is a mark in which multiple players attempted to receive the ball. For each contest, interest related to two moments – the pass that preceded the contest and the contest itself. For each moment, the time (t_p and t_c respectively) and field formation (position, orientation, and velocity of all on-field players) were recorded. A player was considered as having committed to a contest if their Euclidean distance from the location of the contest was less than two metres at t_c. Using a player's position at t_p

and their commitment (recorded as a binary value), a model was developed that quantified the probability a player would commit to a contest across a continuous space within their vicinity.

For each contest, we record player's velocity, orientation, and position, and define the time between t_p and t_c as the ball's time-to-point. For each player, compute the relative location of the contest to player orientation and position. If the Euclidean distance between said player's position at time t_c and the contest location is ≤2 m, set their commitment to 1, else commitment is set to 0 if the distance is >2 m. A player's velocity, commitment, the ball's time-to-point, and the relative x- and y- co-ordinates of the contest are recorded. Given that options are only considered in a 60 m radius of the kicker, the maximum repositioning time available to a player never exceeds four seconds, hence it is unlikely that a player can relocate more than 30 m in this period. In the interest of computation time, player commitment behaviour is only recorded for players within 35 m of the contest locations.

The data was separated by the binary commitment variable, and kernel density estimation (KDE) used to estimate their probability density functions (PDFs). KDE is a form of data smoothing in which the PDF of a dataset is estimated, the form of which depends on the chosen kernel function and bandwidth inputs (Silverman, 1986). KDE has previously been used in motion model studies by Brefeld and colleagues (2018) who produced motion models on the distribution of a player's observed movements, regardless of context. In this study Gaussian kernel functions were used and bandwidth was set to 1.5, chosen after experimentation of different values. Datasets were four-dimensional, containing player velocity (m/s), ball time-to-point (s), and the relative x- and y- co-ordinate of the contest (m).

Individually, these distributions represent the density of the data-sets in four dimensions. If a player's positional information and the ball location is known, the probability they will commit to a contest at location x is as follows:

$$\Pr(x) = \frac{wf_{c=1}}{wf_{c=1} + (1 - w)f_{c=0}} \tag{2}$$

where w is a weighting factor equal to the size of the commitment dataset divided by the total number of samples, and $f_{c=1}$ and $f_{c=0}$ are the PDFs for the datasets where commitment = 1 and commitment = 0 respectively. A player's commitment probability ($\Pr(x)$) considers their position relative to x, their velocity, and the ball's time-to-point. Ball time-to-point to a location is equal to the distance between the ball and the location, divided by ball velocity. Ball velocity was estimated as 18.5 m/s after manually timing kicks from two quarters of a single AFL match and taking the average, however we note that this is a rough estimation as distances were estimated from manually recorded transactions. This represents a novel method for combining the distributions of two datasets of unequal sample size, where the resultant metric quantifies the probability that a new point belongs to each distribution. The combination of these distributions in a 2D space is illustrated in Figure 1. The resultant distributions can be calculated for a player's position, providing a distribution of the likelihood of their repositioning to each location, such that we derive a representation of their spatial influence comparable to that of traditional motion models.

**** INSERT FIGURE 1 ABOUT HERE ****

3.3 Decision-Making Model

Following a pass, the ball can be received on the full, resulting in a mark, or can be received after a bounce, in which case a mark is not awarded. Hence, each of a player's passing options has four possible outcomes – successful passes in which a teammate receives the ball before (A) or after (B) it bounces, and unsuccessful passes in which an opponent does the same (C and D respectively). For each option, we calculate the probability (p) and value (e) of each event (Equation 3). As we consider players to be moving objects who exhibit spatial influence over locations not at their present position, the player with the ball could theoretically kick to any location within a radius equal to their maximum kicking distance. The typical maximum range of elite footballers has been found to be between 55 and 63 m (Ball, 2008c), hence the kicking radius in this study is set to 60 m. While some locations are likely sub-optimal choices, we calculate the expected outcome (EO) of each location within said radius. The EO for a location, x, is as follows:

$$EO(x) = p_A(x)e_a(x) + p_B(x)e_a(x) - p_C(x)e_o(x) - p_D(x)e_o(x) \tag{3}$$

where e_a and e_o are the field equity values for the attacking team and their opponent respectively. Derivation of field equity in AFL has been the focus of previous studies (O'Shaughnessy, 2006; Jackson, 2016).

From the EO of a pass, we calculate the value of a decision (referred to as the decision value or DV) as the EO of the pass that was executed, divided by the maximum EO contained in a player's kicking range (EO_{opt}):

$$DV(x) = \frac{EO(x)}{EO_{opt}} \tag{4}$$

The EO of a pass will be negative if the equity at its target location is negative. For a decision with negative EO, the associated DV will likewise be negative. For a DV <-1, we set DV to -1.

3.3.1 Outcome Probabilities

For a given location, a team's spatial influence (INF) is the sum of the influence of its players:

$$INF(x) = \sum_{i=1}^{18} Pr_i(x) \tag{5}$$

where Pr_i is the commitment probability array for player i, from Equation 2. An attacking team's influence is a measure of the commitment of its players. From the influence of each team, we calculate the attacking team's spatial dominance (DOM) as:

$$DOM_a(x) = \frac{INF_a(x)}{INF_a(x) + INF_o(x)} \tag{6}$$

where $INF_a(x)$ and $INF_o(x)$ are the influence of the attacking team and their opponent at x.

The attacking team's dominance at x is the proportion of space they own. Logically, greater spatial dominance translates to a higher chance of a successful pass. Given that dominance is a relative measure, it is possible for a team to have high dominance over a location where influence is low. In such a case, while the probability of a successful pass is high due to their dominance, the probability that their players will reach the location is low, hence such a location is likely a poor passing location. To account for this, we calculate the probability of a successful mark (p_A and p_C from Equation 3) as a team's dominance multiplied by their influence.

$$p(x) = DOM(x) \times INF(x) \tag{7}$$

Given that a team's desired outcome is a successful pass resulting in a mark, this probability (Equation 7) is of particular importance when analysing a pass. We refer to p_A as the *risk* of a pass, where higher values indicate a safer passing option.

If a pass does not result in a mark, the probability that either team would win the ball is simply equal to their dominance (p_B and p_D from Equation 3).

3.3.2 Kicking Variance

Given imperfect accuracy of kicks, there is a chance that a kick will not reach its intended target. To incorporate this variance, we represent the likely target of a kick using a 2D Gaussian distribution with covariance equal to 5% of the kicking distance. The modified EO of a kick is equal to the summed product of the kicking Gaussian's PDF and the raw EO values contained in its radius:

$$EO_{mod}(x) = \sum_{i \in S} EO(i)f(i) \tag{8}$$

where *S* is the set of integer co-ordinates in a radius around *x* equal to 5% of the Euclidean distance between the ball and *x*.

3.4 Statistical Analysis

For each analysed event, the optimal pass is identified as the pass to a teammate within a 60 m radius of the kicker whose EO is highest. The characteristics of the pass that was made and the pass identified as being optimal were extracted for all kicks that were preceded and resulted in a mark across the analysed matches (see Table 1 for a list of variables and definitions). We refer to the pass that was made as the *decision* and the pass identified as the optimal option as the *alternative* (note that if the decision was optimal it will be equal to the alternative). Descriptive statistics (mean ± SD) were produced for all metrics. Spearman's correlation coefficient (ρ) was used to measure the correlation of decision-making metrics with location. KDE was used to fit the distribution of analysed variables, finding that the decision-making metrics are not normally distributed. The Mann-Whitney *U* test was used to assess differences between the characteristics of decisions and alternatives (Mann & Whitney, 1947).

We explore team level trends in decision-making by comparing two teams. Teams were selected by taking the teams with the highest samples who fit the following criteria – one team who finished in the top 8 (*Team A*) in both the 2017 and 2018 regular AFL playing seasons, and one team who finished in the bottom 10 in the same seasons (*Team B*). Participation in the play-off finals in AFL is between the top 8 teams, hence the choice of cut-off criteria. Furthermore, the distribution of team samples is heavily skewed, hence importance was placed on selecting teams with adequate sample sizes. Differences between team-level statistics were measured using the Mann-Whitney *U* test. Within-team decision-making is analysed for both teams. We fit the distribution of mean decision-making characteristics for each player on the team. All analyses were carried out in the Python programming language, using SciPy (Jones, et al., 2014) and the Scikit-learn (Pedregosa, et al., 2011) packages.

**** INSERT TABLE 1 ABOUT HERE ****

4 Results

4.1 Motion Models

Motion models were produced from 46220 instances of player commitment. Within the dataset there were 6392 instances of player commitment (Commitment = 1), and 39828 instances of no commitment (Commitment = 0), producing a weighting coefficient (w) of 0.14. Resultant motion models for four different player velocities for ball time-to-point of two seconds are visualised in Figure 2. Peak commitment probabilities occurred at 0.8 m for a velocity of 2 m/s (Figure 2a), 1.6m for 4 m/s (Figure 2b), 3.7 m for 6 m/s (Figure 2c), and 5.3 m for 8 m/s (Figure 2d). While density peaks at further distances as velocity increases, a negative correlation is revealed between player velocity (integers from 1 to 8 m/s) and peak commitment probabilities ($\rho = -0.80$ for $t = 2$ seconds), and between ball time-to-point (whole second integers from 1 to 4 seconds) and peak commitment probabilities ($\rho = -1$ for velocity = 4 m/s). At higher velocities, the probability that a player will commit to a contest decreases as the relative angle increases. For a velocity of 8 m/s or greater, player's exhibit minimal influence on space in the negative y- axis (i.e., behind their direction of orientation). As velocity increases, we also note that the shape of a player's commitment inverts.

**** INSERT FIGURE 2 ABOUT HERE ****

4.2 Decisions and Alternatives

A total of 2935 passes matched the selection criteria across 60 matches (48.9 ± 14.7 kicks per match). An example decision-making output is visualised in Figure 3. In this example, the kicker passes to a teammate positioned towards the boundary line in the defensive 50 m region, while the model identified three higher value passes to teammates positioned towards the centre of the field. Figure 4 presents the components that constitute EO calculations. Summarised characteristics of decisions and alternatives are presented in Table 2. The mean of all analysed variables was lower for decisions compared to alternatives and all differences were statistically significant (refer to Table 2).

**** INSERT FIGURE 3 ABOUT HERE ****

A very weak correlation was noted between vertical displacement from centre and DV of decisions ($\rho = 0.06$). Horizontal displacement from the attacking team's goal is positively correlated with DV ($\rho = 0.56$).

**** INSERT FIGURE 4 ABOUT HERE ****

The distributions of decision-making characteristics are presented in Figure 5. The distribution of dominance (Figure 5a) is bimodal, with peak density for decisions at DOM = 0.54 and a local maximum at DOM = 1.0. This global peak at 0.54 represents a contest between two teams that slightly favours the attacker, while the local maximum at 1.0 represents a kick to an area of absolute dominance. The distribution of alternatives is similarly bimodal, with a greater negative skew and density around absolute dominance. Influence of decisions (Figure 5b) reveals peak density at INF = 0.43, which is comparable to the average peak density of player commitment models (Figure 2). Density for risk peaks at 0.25 (Figure 5c). The shape of the distributions of EO for decisions and alternatives are different, with decisions exhibiting peak density at EO = 0.14 (Figure 5d), and minimal density is noted at EO > 2, while alternatives are noted as having a greater range of EO values, with no notable density peak. DV follows a relatively normal distribution for decisions (Figure 5e) and distributions of kicking distance (Figure 5f) exhibit opposite skews (decisions are positively skewed, while alternatives negatively). Density of kicking distance for decisions is highest at 17.3 m, marginally longer than the 15 m minimum distance required for a legal mark. Small density peaks at 0.0 are observed for the dominance, influence, and risk of alternatives.

**** INSERT TABLE 2 ABOUT HERE ****

**** INSERT FIGURE 5 ABOUT HERE ****

4.3 Team-level Characteristics

The distributions of passing characteristics for two teams are presented in Figure 6 and the summary statistics in Table 3. There was minimal difference in the dominance, influence, risk, and distance of decisions between the two teams. The mean EO and DV for Team B are higher than those of Team A, however no differences were found to be statistically significant. While the shape of variable distributions is similar for both teams, it is noted that Team B exhibits a greater negative skew for EO, DV, and distance variables. Distributions of mean decision-making characteristics for players amongst both teams were found to be similar (Figure 7). While the differences between player-level standard deviations were not found to be statistically significant, the distributions for dominance and distance variance display visual differences.

**** INSERT TABLE 3 ABOUT HERE ****

**** INSERT FIGURE 6 ABOUT HERE ****

**** INSERT FIGURE 7 ABOUT HERE ****

5 Discussion

This study demonstrates a method for measuring characteristics of player pass decision-making in invasion team sports. Previous studies of player decisions have measured decisions relative to some current measure of possession expectation (e.g., Cervone et al., 2014), rather than relative to the value of alternative passes that were presented. While the former approach assigns value to a specific kick, relative measures of decision-making assign value to individual decisions. Similar to the distinction between player accuracy and shot difficulty (e.g., Chang, et al., 2014), assigning value to player decision-making presents greater insights into individual player performance. The adoption of decision-making evaluation in combination with measurements of accuracy and risk would allow for targeted coaching and recruitment, as well as defining categories of player tactical behaviour.

A major component of the decision-making modelling were player motion models, fit on the weighted distributions of player commitment to contests. While previous studies have developed probabilistic motion models with arbitrary bounds (Fernandez and Bornn, 2018) or from a player's observed displacements (Brefeld et al., 2018), the commitment modelling approach demonstrated in this study fits player behaviour with consideration of movement context, representing a new approach to the measurement of a player's spatial influence. Furthermore, the models are parameterized through the fitting of density in four dimensions (with consideration of a player's velocity, time and x- and y- co-ordinates), presenting a continuous representation of player commitment. A notable finding of the motion models is that commitment peaks are of lower density for higher velocity and time values. That is, players are overall less likely to commit to an upcoming contest if the ball is further away (hence, a high time-to-point) or if they are moving at high velocities. This finding is logical and may be explained by a desire to simply corral an opponent or reposition for future involvements, rather than participate in the immediate transaction. As with alternative motion models, we found that a player's influence in the negative y-axis (i.e., behind them) degrades as their speed increases. While models fit on player commitments more realistically measure their likelihood to occupy future space, the models only consider a player's position and momentum, not teammate locations. A player's participation in a contest logically has some dependence on the position of their teammates, hence attempts to incorporate may produce more realistic models.

A key finding in this study are the novel insights into the decision-making and passing tendencies of Australian Rules footballers. Previous studies have identified the importance of kicking in the AFL (Stewart, et al., 2007; Robertson, et al., 2016) but there has been minimal work into describing the kicking landscape at elite levels at a transactional level (e.g., distance, level of pressure), despite studies on the biomechanics of kicking in Australian Rules football (e.g., Ball, 2008a; Ball, 2008b). This study found that kicks resulting in a mark are most commonly short, with a density peak at 17.3 metres (mean = 25 m), marginally longer than the minimum distance required for a legal mark. This could be the result of tactical behaviour, or indicative of the ease in which close options can be identified due to lower visual obstruction. Furthermore, successful marks are most often to players in one on one contests or to players who are completely open (as suggested by the bimodal distribution of passing dominance and the density peaks of risk), which may be indicative of risk aversion, however more research is required to understand individual player behaviour.

In contrast to player decisions, the optimal alternative passes that were identified by the model presented in this study were long distance kicks, less frequently to unmarked individuals. While the distribution of dominance was similarly bimodal for alternatives, the peak at absolute dominance (DOM = 1.0) was less intense than for decisions. The higher density for passes to areas of dominance between 0.5 (a 50/50 contest) and 1.0 suggests kicks to areas in which multiple teammates have an opportunity to receive the ball. This is reinforced by the distribution of influence for alternatives (Figure 5b) where more density is noted for influence above 0.5 compared to decisions. Long-range passes having higher associated values (EO and DV) is logical due to the inclusion of AFL field equity, in which the value of space increases as the distance and angle to the goalposts decreases (Jackson, 2016; Figure 4c). The contrast in distances between decisions and alternatives (Figure 5f) could be due to several factors, such as a difficulty for players to identify long-range options (due to visual obstruction and lower decision-making time, for example) or an underestimation of kicking accuracy by the model. Due to the unavailability of precision ball tracking in AFL, this study used an arbitrary measurement of kicking accuracy. Should more detailed transactional data or LPS ball tracking become available, it is believed that kicking accuracy could be modelled from empirical data. The density peaks at values towards 0 for dominance, influence, and risk can be explained by situations in which all passing options are positioned in areas of negative field equity (e.g., field formations in the defensive 50 m area), resulting in an optimal decision being a kick to an area of no spatial dominance (hence, no negative associated equity). This is a common problem with models that use equity-based rewards, where moving the ball backwards is often associated with a reduction in equity.

Team level analysis revealed that the less successful team in the 2017/2018 season had higher average DV than the more successful team. Furthermore, while within-team distribution of player averages were similar, the player variance of DV was more positively skewed for the less successful team. Of particular interest is the finding that the less successful team executed passes of higher value, potentially suggestive of a difference in playing styles. These findings could be suggestive of tactical behaviour that warrants further research into player and team-level decision-making, particularly consideration of contextual information such as match conditions, score deficits, and tactical styles.

6 Conclusion

This work represents the beginning of ongoing research into player decision-making in the AFL. The decoupling of player decision-making from overall player performance allows for a more precise understanding of player ability that has applications in coaching and scouting. Underlying the decision-making model is a player motion model fit on the combined distributions of relative contest locations that were committed to, and those that were not. The resulting motion model quantifies the probability that a player would commit to a contest, given their velocity, orientation, and past behaviours. It was found that player commitment decreases as a function of velocity and available time, offering insights into the commitment behaviour of players. Analysis of passes revealed that players typically execute short kicks that are most commonly to teammates in one-on-one or unmarked situations, resulting in a bimodal distribution of passing dominance. Conversely, the mathematical model presented in this paper identifies long-range options as having higher expected value, given the inclusion of field equity which rewards possession closer to the goalposts. This mismatch in decisions could be due to the ease in which short-range options can be identified and executed compared to long-range options. Differences in decision-making variables between two analysed teams suggests a need for expanded datasets and research into player decision-making with consideration of match context.

7 Conflict of Interest

Authors Karl Jackson and Tim Bedin were employed by Champion Data. All other authors declare no competing interests.

8 Ethics Statement

Players provided their data and written and informed consent to the commercial provider as part of their collective bargaining agreement. Ethical approval for this study was granted by Victoria University's Human Research Ethics Committee (VU HREC 24514).

9 Author Contributions

BS, KJ, SR: Conceptualisation; BS, KJ: Methodology; TB: Data extraction; BS, TB: Data cleaning; BS: Data analysis & visualisation; BS, SR: Writing; BS, SR, KJ, TB: Drafting

10 References

Ball, K. (2008a). Biomechanical considerations of distance kicking in Australian Rules football. *Sports Biomechanics*, *7*(1), 10-23.

Ball, K. (2008b). Foot interaction during kicking in Australian rules football. In *Science and football VI* (pp. 62-66). Routledge.

Ball, K. (2008c). Use of weighted balls for improving kicking for distance. *Science and Football VI*, 285.

Brefeld, U., Lasek, J., & Mair, S. (2018). Probabilistic movement models and zones of control. *Machine Learning*, 1-21.

Cervone, D., D'Amour, A., Bornn, L., & Goldsberry, K. (2014, February). POINTWISE: Predicting points and valuing decisions in real time with NBA optical tracking data. In *Proceedings of the 8th MIT Sloan Sports Analytics Conference, Boston, MA, USA* (Vol. 28, p. 3)

Cervone, D., D'Amour, A., Bornn, L., & Goldsberry, K. (2016). A multiresolution stochastic process model for predicting basketball possession outcomes. *Journal of the American Statistical Association*, *111*(514), 585-599.

Chang, Y. H., Maheswaran, R., Su, J., Kwok, S., Levy, T., Wexler, A., & Squire, K. (2014, February). Quantifying shot quality in the NBA. In *Proceedings of the 8th Annual MIT Sloan Sports Analytics Conference. MIT, Boston, MA*.

Fernandez, J., & Bornn, L. (2018). Wide open spaces: A statistical technique for measuring space creation in professional soccer. In *Sloan Sports Analytics Conference*.

Fonseca, S., Milho, J., Travassos, B., & Araújo, D. (2012). Spatial dynamics of team sports exposed by Voronoi diagrams. *Human movement science*, *31*(6), 1652-1659.

Fujimura, A., & Sugihara, K. (2005). Geometric analysis and quantitative evaluation of sport teamwork. *Systems and Computers in Japan*, *36*(6), 49-58.

Jackson, K. (2016). *Assessing Player Performance in Australian Football Using Spatial Data* (Doctoral dissertation, PhD Thesis, Swinburne University of Technology).

Jones, E., Oliphant, T., & Peterson, P. (2014). {SciPy}: open source scientific tools for {Python}.

Gudmundsson, J., & Wolle, T. (2014). Football analysis using spatio-temporal tools. *Computers, Environment and Urban Systems*, *47*, 16-27.

Horton, M., Gudmundsson, J., Chawla, S., & Estephan, J. (2015, May). Automated classification of passing in football. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 319-330). Springer, Cham.

Pedregosa, Fabian, et al. "Scikit-learn: Machine learning in Python." *Journal of machine learning research* 12.Oct (2011): 2825-2830.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, 50-60.

O'Shaughnessy, D. M. (2006). Possession versus position: strategic evaluation in AFL. *Journal of sports science & medicine*, *5*(4), 533.

Power, P., Ruiz, H., Wei, X., & Lucey, P. (2017). Not all passes are created equal: Objectively measuring the risk and reward of passes in soccer from tracking data. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1605-1613). ACM.

Robertson, S., Back, N., & Bartlett, J. D. (2016). Explaining match outcome in elite Australian Rules football using team performance indicators. *Journal of sports sciences*, *34*(7), 637-644.

Silverman, B. W. (1986). Density estimation for statistics and data analysis. Chapman & Hall, London.

Spencer, B., Morgan, S., Zeleznikow, J., & Robertson, S. (2017) Measuring Player Density in Australian Rules Football Using Gaussian Mixture Models. In *Complex Systems in Sport, International Congress Linking Theory and Practice* (pp. 172-174).

Spencer, B., Bedin, T., Farrow, D., & Jackson, K. (2018). A method for evaluating player decision-making in the Australian Football League. In *Proceedings of 14th Australasian Conference on Mathematics and Computers in Sport* (pp. 7-12). Sunshine Coast, Australia: University of the Sunshine Coast

Stewart, M., Mitchell, H., & Stavros, C. (2007). Moneyball applied: Econometrics and the identification and recruitment of elite Australian footballers. *International Journal of Sport Finance*, *2*(4), 231-248.

Szczepański, Ł., & McHale, I. (2016). Beyond completion rate: evaluating the passing ability of footballers. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *179*(2), 513-533.

Taki, T., & Hasegawa, J. I. (2000, June). Visualization of dominant region in team games and its application to teamwork analysis. In *cgi* (p. 227). IEEE.

11 Figures and Tables

Table 1. Definitions of decision-making variables.

Variable	Definition
Dominance	The proportion of space owned by a team (see Equation 6)
Influence	A measure of spatial occupancy irrespective of opposition locations, equal to the summed commitment probabilities of a team's players (see Equation 5)
Risk	The likelihood of a successful pass resulting in a mark (see Equation 7)
Decision Value (DV)	The value of a player's passing decision, measured relative to the optimal decision available at the time of the pass (see Equation 4)
Expected Outcome (EO)	A numerical value describing the expected value of passing to a field position that considers the risk and reward of said pass (see Equation 3)
Distance	The Euclidean distance between two points. For a kick, distance is the Euclidean distance between the location of the kicker and of the receiver

Passing Decisions in Australian Football

Table 2. Mean values for decision-making variables between decisions and alternatives. Values are presented as Mean ± SD and all differences are statistically significant.

Variable	Decisions	Alternatives
Dominance	0.66 ± 0.23	0.75 ± 0.23
Influence	0.51 ± 0.27	0.63 ± 0.31
Risk	0.33 ± 0.19	0.47 ± 0.21
Expected Outcome	0.34 ± 0.46	2.11 ± 1.41
Decision Value	0.13 ± 0.42	0.78 ± 0.24
Distance	25.0 ± 11.8	42.7 ± 17.8

Table 3. Mean values for decision-making variables between Team A and Team B. *p*-values for differences are presented in Figure 5.

Variable	Team A	Team B
Dominance	0.66 ± 0.23	0.66 ± 0.23
Influence	0.52 ± 0.24	0.49 ± 0.24
Risk	0.34 ± 0.19	0.32 ± 0.17
Expected Outcome	0.29 ± 0.39	0.34 ± 0.42
Decision Value	0.08 ± 0.42	0.13 ± 0.43
Distance	24.3 ± 12.0	24.9 ± 11.6

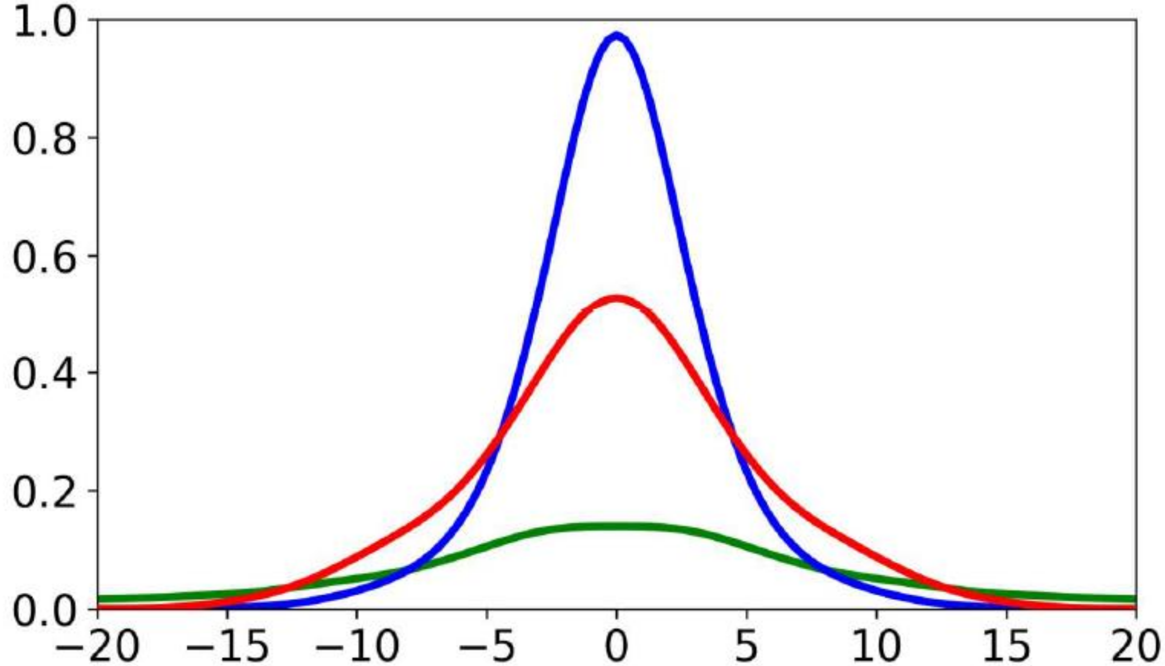


Figure 1. Two-dimensional representation of the commitment modelling process. The blue line represents the distribution for commitment values of 1 ($f_{c=1}$), and the green line represents the distribution for commitment values of 0 ($f_{c=0}$). The red line represents player influence ($\Pr(x)$), derived from the combined commitment distributions (see Equation 2). This exemplar represents a player's commitment probability across relative x- co-ordinates for a y- displacement of 1 m, velocity = 4 m/s, and time = 2 s. All co-ordinates are relative to player orientation.

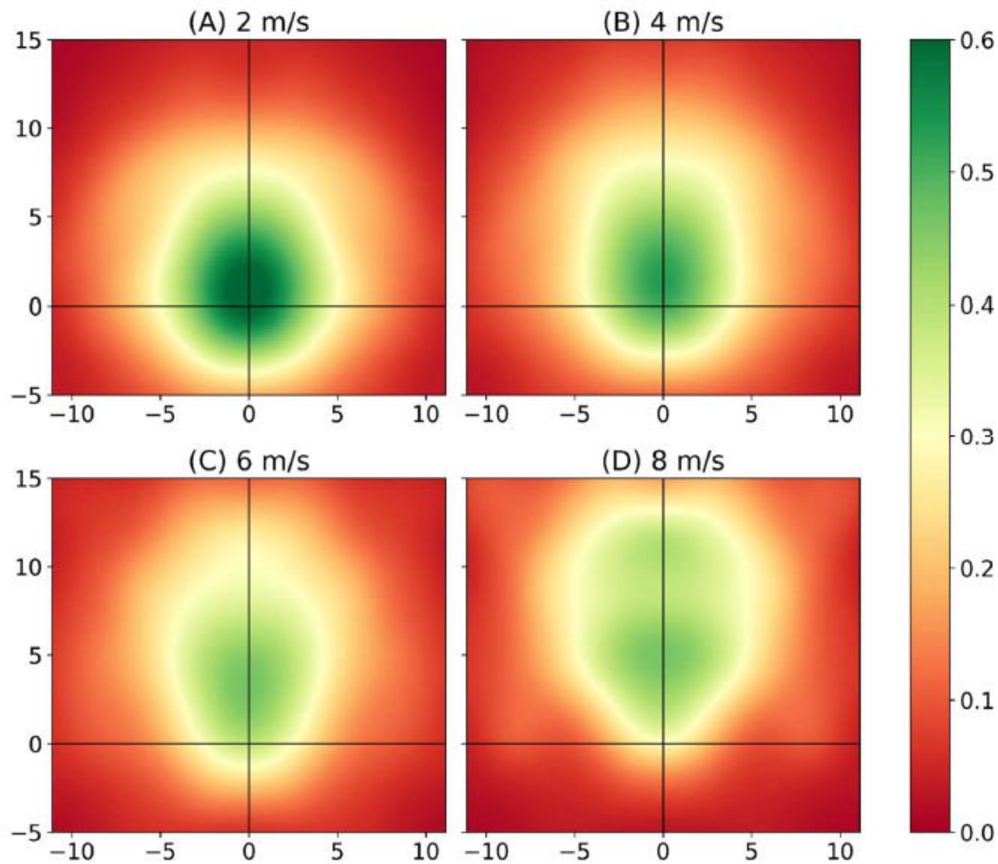


Figure 2. Motion models representing a player's area of influence whilst moving at various velocities for ball time-to-point of 2 s. Heatmap intensity is equivalent to the probability that a player (at the point of origin) would participate in a contest at relative x-, y- co-ordinates, as quantified by observed commitment behaviours.

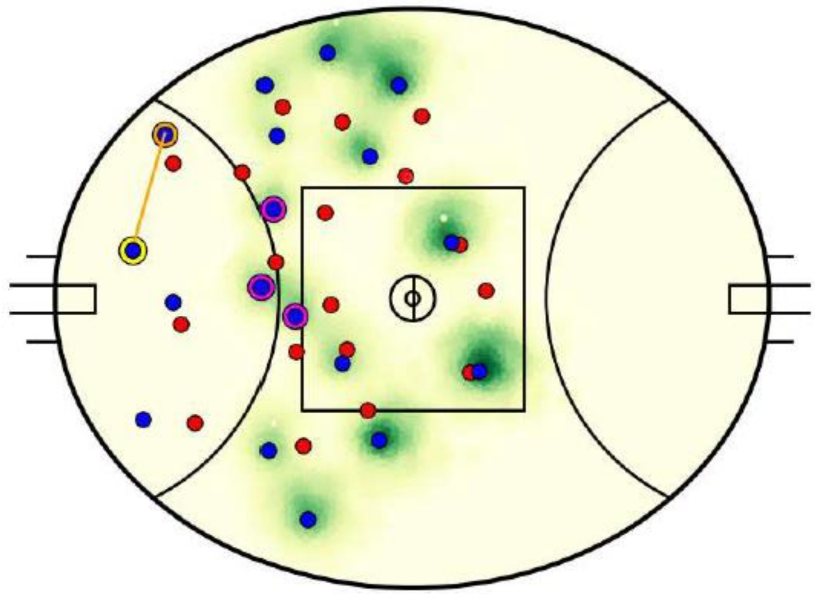


Figure 3. An example output of the decision-making model. The attacking team players are plotted in blue and their opponents in red. The kicker (circled in yellow) executed a pass along the orange line to the receiver (circled in orange). The model identified three higher valued passes (to players circled in magenta) towards the middle of the field that are within a 60 m radius of the kicker. The intensity of green correlates to the expected outcome of passes to each field position.

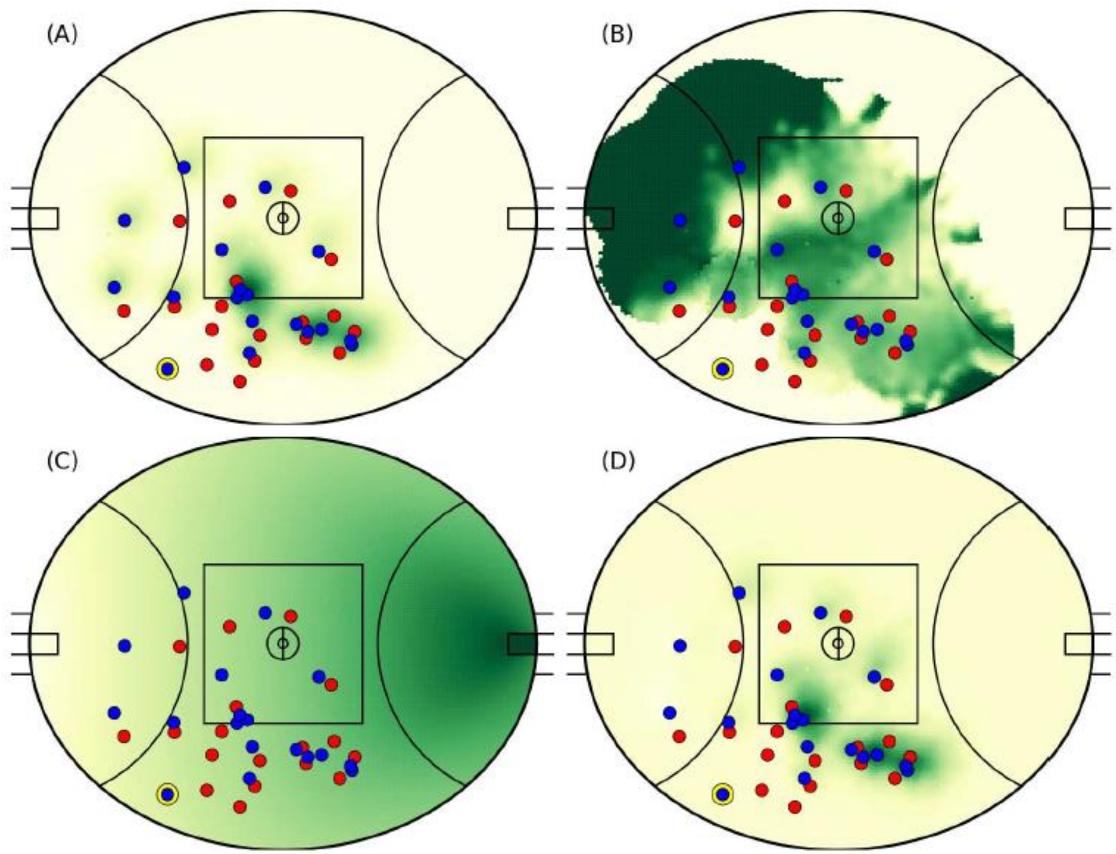


Figure 4. Team influence (A), dominance (B), field equity (C), and resultant expected outcomes (D) relative to the player in possession (circled in yellow, towards the lower boundary). High value space is represented as darker green regions. Team influence measures the spatial influence of the attacking team (whose players are in blue), while dominance measures their spatial ownership relative to the opposition (whose players are in red). All values are calculated relative to the player in possession. When complete, the model presented in this paper identifies two high value areas towards the centre square, both viable passing options (see D).

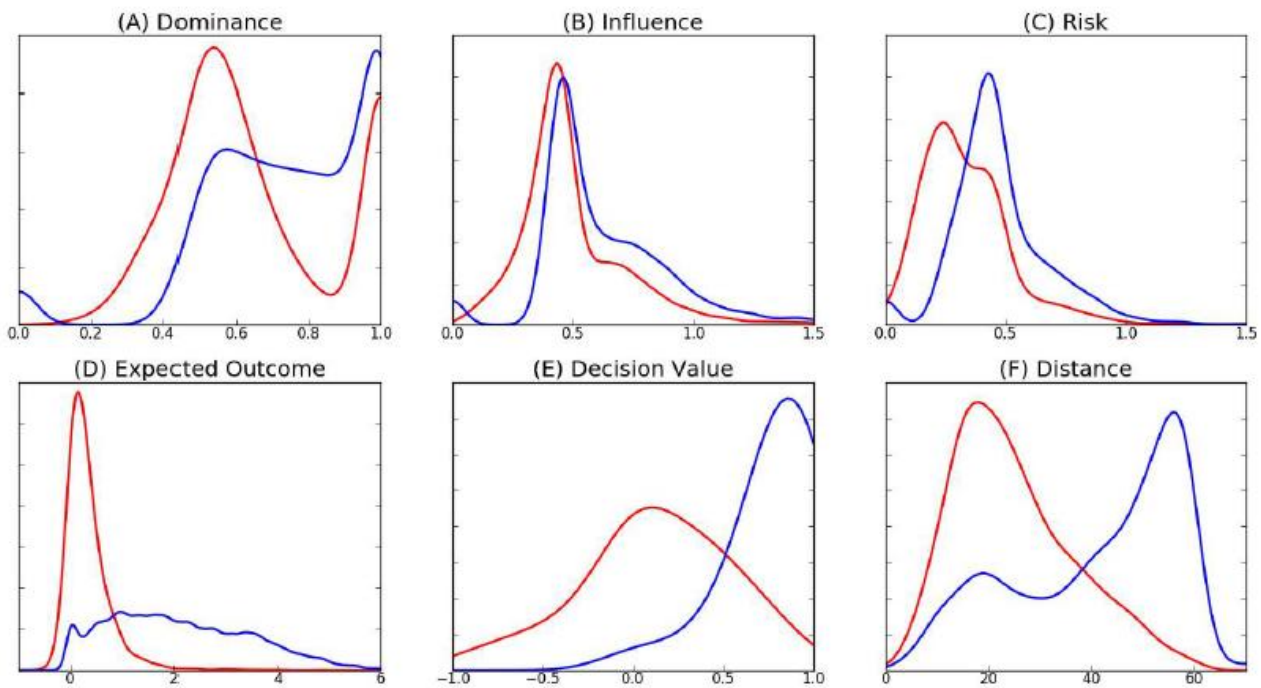


Figure 5. Distribution of decision-making characteristics for decisions (red) and alternatives (blue).

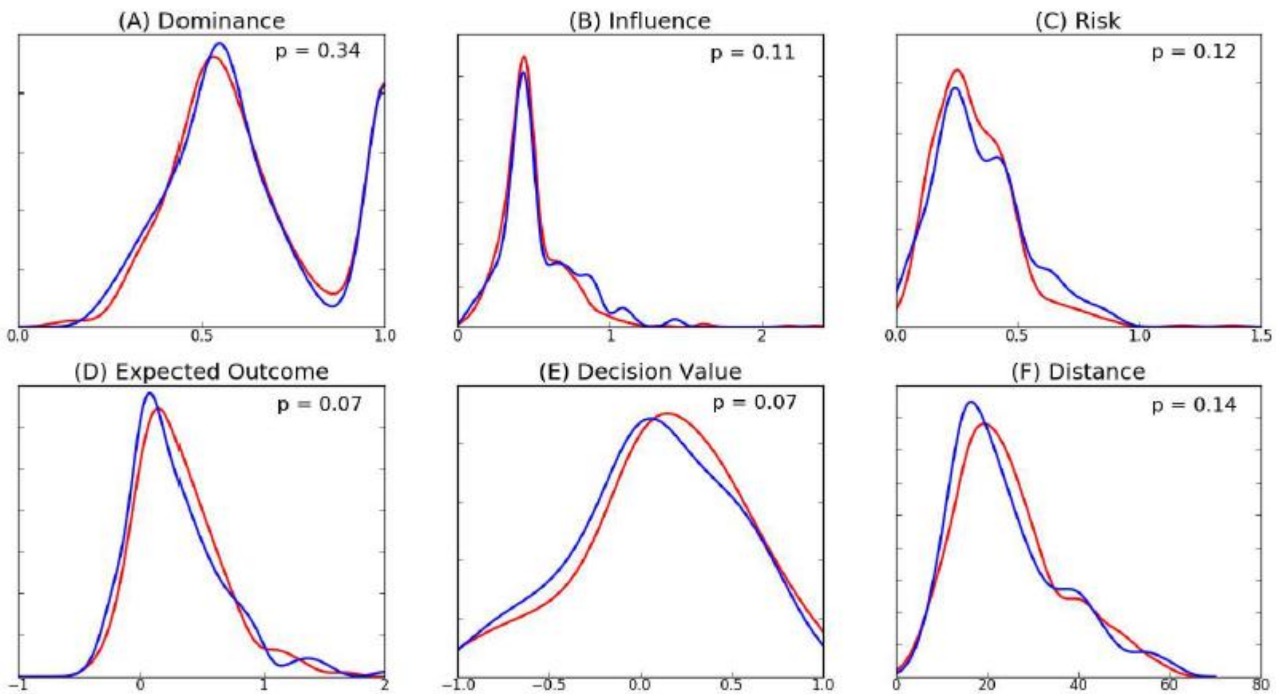


Figure 6. Distribution of team-level decision-making variables for Team A (blue) and Team B (red). Associated p-values (computed using the Mann-Whitney U test) are presented for each variable.

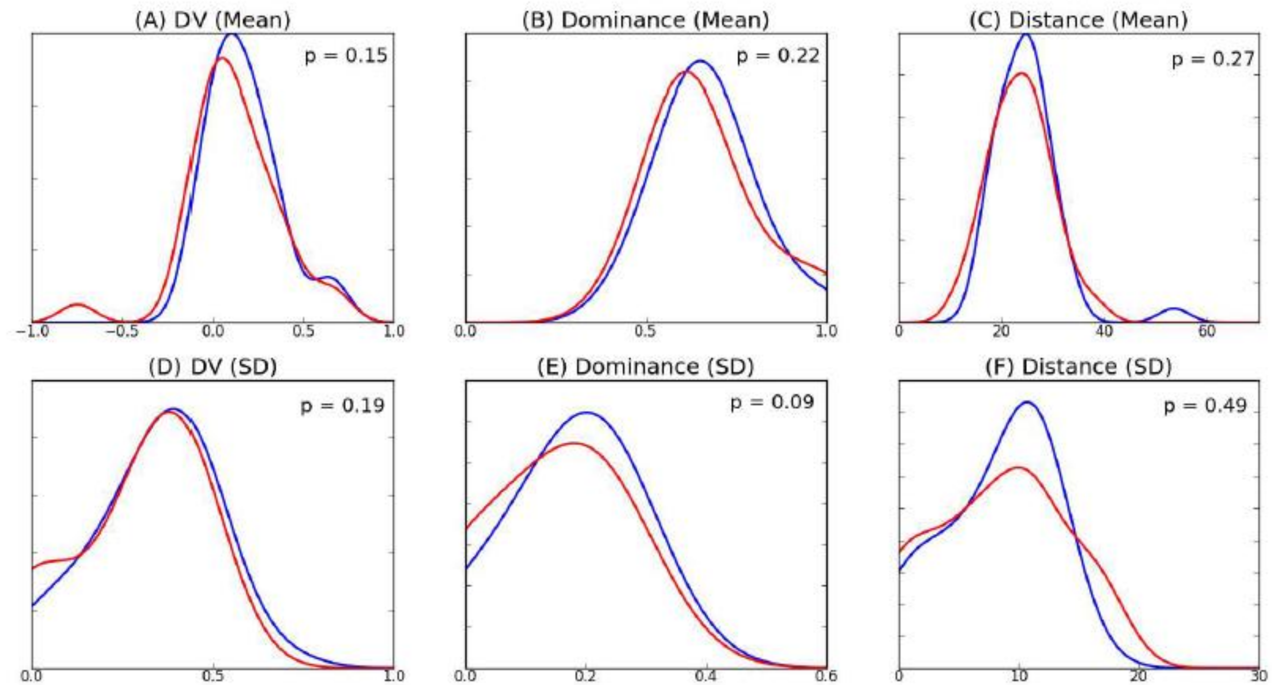


Figure 7. Within-team distributions for decision-making for Team A (blue) and Team B (red). Top row are the distributions for the mean decision-making abilities of players and the bottom row are variance.

Modelling the Quality of Player Passing Decisions in Australian Rules Football Relative to Risk, Reward and Commitment

Abstract

The value of player decisions has typically been measured by changes in possession expectations, rather than relative to the value of a player's alternative options. This study presents a mathematical approach to the measurement of passing decisions of Australian Rules footballers that considers the risk and reward of passing options. A new method for quantifying a player's spatial influence is demonstrated through a process called commitment modelling, in which the bounds and density of a player's motion model are fit on empirical commitment to contests, producing a continuous representation of a team's spatial ownership. This process involves combining the probability density functions of contests that a player committed to, and those they did not. Spatiotemporal player tracking data was collected for AFL matches played at Docklands Stadium in the 2017 and 2018 seasons. It was discovered that the probability of a player committing to a contest decreases as a function of their velocity and of the ball's time-to-point. Furthermore, the peak density of player commitment probabilities is at a greater distance in front of a player the faster they are moving, while their ability to participate in contests requiring re-orientation diminishes at higher velocities. Analysis of passing decisions revealed that, for passes resulting in a mark, opposition pressure is bimodal, with peaks at spatial dominance equivalent to no pressure and to a one-on-one contest. Density of passing distance peaks at 17.3 m, marginally longer than the minimum distance of a legal mark (15 m). Conversely, the model presented in this study identifies long-range options as have higher associated decision-making values, however a lack of passes in these ranges may be indicative of differing tactical behaviour or a difficulty in identifying long-range options.

8.1 INTRODUCTION

Team sport athletes are consistently presented with situations in which their decisions effect the immediate state of a game. These consist of overt on-ball decisions relating to passing or shooting, however also include off-ball actions such as occupation of a given space. Whilst previous works have quantified the impact of a decision on some measure of possession expectation (Cervone, et al., 2014; Cervone, et al., 2016a; Jackson, 2016) or on measures of spatial control (Fernandez & Bornn, 2018), their value has typically been measured by the change in some metric or relative to a contextual mean. We believe the value of a player's decision should be quantified relative to the alternative options that were available. Although a pass may yield a positive increase in a team's scoring chance by x, the decision is by definition sub-optimal if alternatives exist that increase it by greater than x. By measuring a player's decision relative to their options, we can quantitatively attribute value to a player's decision-making abilities, further decoupling components of a player's performance.

The expected possession value (EPV) metric considers spatiotemporal data, match phase and player behaviours to quantify possession outcomes in basketball (Cervone, et al., 2016a; Cervone, et al., 2014). Computing the change in EPV between possessions assigns a value to player possession contributions. A player's decision is valued relative to the tendencies of other players in the same situation, producing a player's EPVA (EPV-added over replacement) as the sum of a player's EPV-added (EPVend – EPVstart) across all possessions. In Jackson (2016), Australian Rules footballers ranking points are the sum of their possession contributions, valued relative to the event and location, an extension of the measure of field equity developed in O'Shaughnessy (2006). Similar to Cervone et al.

(2014), player contributions are measured relative to mean outcomes and a player is deemed to be a good decision maker if their involvement improved their team's field equity, a measurement of scoring chance relative to match phase and possession location. In Horton et al. (2015), football passes were labelled qualitatively using machine learning algorithms with quantitative inputs, learnt from manual labelling of passing quality by sporting professionals. The inclusion of player dominant regions, a method of bounding a player's spatial ownership via consideration of player momentum, suggests the quality of a pass has some dependence on a team's spatial control.

Common amongst these studies is the valuation of player decisions with respect to some change in possession expectation. Another approach would be to value decisions relative to alternative options, however, modelling this problem presents unique challenges. While quantifying a decision after the fact can be done by measuring the change in a given objective, each option available to a player has an accompanying probability of success. Multiple studies have measured the risk of passes in football. In Szczepanski and McHale (2016), the success of a pass depended upon the skill of a player and their teammates, field position of the pass location and destination, and pressure. The latter was approximated dependent on a player's typical playing positions and time between passes, rather than consideration of opponent locations due to an absence of player tracking data. Power et al. (2017) measured the risk and reward of passing options using spatiotemporal tracking data, where the risk of a pass considers player velocity, defender proximity and momentum, and possession statistics and the reward of a pass is the probability that the pass will result in a shot on goal. From their measure of risk, the risk tendencies and completion rates of players were analysed. Our recent work in AFL produced measures of risk and reward via discrete player motion models and measures of future possession expectations respectively (Spencer, et al., 2018).

In this study we value a player's passing decisions through consideration of the risk and reward of their options. We measure the risk of a pass through modelling of individual and team spatial control, and reward via a measure of field equity detailed in Jackson (2016). We present a new method for modelling spatial control via probabilistic modelling of player commitment to contests with consideration of their momentum. This process, referred to as commitment modelling, produces player motion models that more realistically represent player behaviour based on their proximity to important events. We use the resultant decision-making model to analyse characteristics of player decision-making, its predictability, and distributions of risk taking within teams.

8.2 RELATED WORK

8.2.1 Motion Models

There exist many methods for representing a player's spatial occupancy. One common approach, particularly in football, is that of Voronoi tessellations which bound a player's owned space as the space in which they could occupy before any other player. Simple applications of this approach do not consider player orientation, velocity, or individual physical capabilities (e.g. Fonseca, et al., 2012). Taki and Hasegawa (2000) produced variations incorporating a player's orientation, velocity, but assumed consistent acceleration. Fujimaru and Sugihara (2005) proposed an alternative motion equation, adding a resistive force that decreases velocity. This approach involved a generalised formula that more realistically represented a player's inability to cover negative space if moving at speed. Gudmundsson and Wolle (2010) individualised these models, fitting a player's dominant region from observed tracking data.

Underlying these models is an assumption that spatial ownership is binary. That is, each location on the field is owned completely by a single player, determined by the time it would take them to reach said location, henceforth referred to as their time-to-point. Through observations of contests, we propose that ownership of space is continuous. For a given location, if the time-to-point of the ball is greater than the time-to-point of at least two players, then no single player owns the space completely. This distinction is important if we wish to quantify spatial occupancy (and its creation) relative to the ball, given its time-to-point, as we need to account for changes in field formations that could occur between possessions.

Recent papers have addressed this. The density of playing groups was explored with Gaussian mixture models in Spencer et al. (2017). Spencer et al. (2018) produced a smoothed representation of a team's control using non-probabilistic player motion models fit on observed tracking data. While a team's ownership was expressed on a continuous scale, the use of motion models with discrete bounds may result in unrealistic estimations of a player's influence (Brefeld, et al., 2018). Fernandez and Bornn (2018) measured a player's influence area using bivariate normal distributions that considered a player's location, velocity, and distance to the ball. The result is a smoothed surface of control in which a team's influence over a region is continuous, however the size of a player's influence is within a selected range, rather than learnt from observed movements. Recently, Brefeld et al. (2018) fit player motion models on the distribution of observed player movements, utilising these probabilistic models to produce more realistic Voronoi-like regions of control. In the interest of computing time, two-dimensional models were produced for different speed and time bands, hence the resultant models are not continuous in all dimensions.

Given its contested and dynamic nature, a continuous representation of space control is preferable (e.g. Fernandez and Bornn, 2018; Spencer, et al., 2018). Furthermore, a player logically exhibits greater control over space in which they are closer, hence we develop probabilistic motion models in this paper. When probabilistic models are fit on the entirety of a player's movements (as in Brefeld et al., 2018), we find that the probability of player reorientation is underestimated. In decision-making modelling, our interest is in measuring the contest of space that would occur if the ball were kicked to said space. Hence to represent this realistically, it is important to fit the distribution of player movements observed under similar circumstances. We model a player's behaviour when within proximity of contests. We achieve this via a procedure we call commitment modelling, where we fit the distribution of player commitment to contests in four dimensions (velocity, time, and x- and y- field position). The result is a realistic representation of player behaviours when presented with the opportunity to participate in a contest.

8.3 MATERIALS AND METHODS

8.3.1 Data and Pre-processing

Spatiotemporal player tracking data was collected from the 2017 and 2018 AFL seasons. Data were collected by local positioning system (LPS) wearable Catapult Clearsky devices (Catapult Sports, Melbourne, Australia), situated in a pouch positioned between the players' shoulder blades. Positional data in the form of Cartesian coordinates was recorded at a frequency of 10 Hz for all 44 players. To ensure consistent tracking and field dimensions, analysed matches were limited to those played at Docklands Stadium, Melbourne. Play-by-play transactional data (i.e., match events such as kicks, marks, and spoils, and their associated meta-data) were manually collected by Champion Data (Champion Data Pty Ltd, Melbourne, Australia). These events are henceforth referred to as transactions. Consolidation of transaction and tracking data was used to infer ball position from possession, as ball tracking data is not available in Australian Rules football. Datasets were joined via universal timestamps present in both datasets. Transactions were recorded to the nearest second, hence it was assumed they occurred at the beginning of a second when matched to 10 Hz tracking data. If the location of one or more players was lost during a passage of play, said passage was omitted from the analysis. A total of 2236 passes across 60 matches were analysed in this study.

A player's velocity and displacement direction were calculated from raw positional data. Displacement direction was extracted from consecutive tracking samples (i.e., a player's displacement direction was recorded as the angle formed by consecutive tracking samples, relative to the positive y-axis). A player's change in displacement direction was considered as the angle between two vectors, $\overrightarrow{AB}$ and $\overrightarrow{BC}$, where A, B, and C are the player's three most recent positions, and the angle describes the change in displacement direction between positions B and C (Equation1). The same process was used to calculate the location of an event relative to a player (where A and B are a player's previous and current position, and C is the location of interest). Velocity, recorded in metres/second, was calculated as the Euclidean distance between a player's current position and their position, one second prior.

$$\theta = \cos^{-1}\left(\frac{\overrightarrow{AB} \cdot \overrightarrow{BC}}{\|\overrightarrow{AB}\| \cdot \|\overrightarrow{BC}\|}\right) \tag{1}$$

In this study, only player decisions following a mark were included, given that a mark provides the player with time to make an informed decision. In Australian Rules football, a mark is a kick greater than 15 m that is received by a player on the full (i.e., without bouncing). To locate the destination of a player's kick following their mark, the next transaction must also be a mark. If the next possession following a kick is not a mark, we are unable to reliably locate the intended target, given a reliance on transactions to infer ball position.

8.3.2 Commitment Modelling

For analysis purposes, a contest was defined as a transaction following a pass in which at least one player from each team was involved and the ball location (for both the preceding kick and the receive) could be inferred from the consolidated datasets. In this study, the contest transaction types were spoils and contested marks. The former is an attempted pass that was physically prevented by the opposition and the latter is a mark in which multiple players attempted to receive the ball. For each contest, interest related to two moments – the pass that preceded the contest and the contest itself. For each moment, the time (tp and tc respectively) and field formation (position, displacement direction, and velocity of all on-field players) were recorded. A player was considered as having committed to a contest if their Euclidean distance from the location of the contest was less than two metres at tc. Using a player's position at tp and their commitment (recorded as a binary value), a model was developed that quantified the probability a player would commit to a contest across a continuous space within their vicinity.

For each contest, we record player's velocity, displacement direction, and position, and define the time between tp and tc as the ball's time-to-point. For each player, compute the relative location of the contest to player displacement direction and position. If the Euclidean distance between said player's position at time tc and the contest location is ≤2 m, set their commitment to 1, else commitment is set to 0 if the distance is >2 m. A player's velocity, commitment, the ball's time-to-point, and the relative x- and y- co-ordinates of the contest are recorded. Given that options are only considered in a 60 m radius of the kicker, the maximum repositioning time available to a player never exceeds four seconds, hence it is unlikely that a player can relocate more than 30 m in this period. In the interest of computation time, player commitment behaviour is only recorded for players within 35 m of the contest locations.

The data was separated by the binary commitment variable, and kernel density estimation (KDE) used to estimate their probability density functions (PDFs). KDE is a form of data smoothing in which the PDF of a dataset is estimated, the form of which depends on the chosen kernel function and bandwidth inputs (Silverman, 1986). KDE has previously been used in motion model studies by Brefeld and colleagues (2018) who produced motion models on the distribution of a player's observed movements, regardless of context. In this study Gaussian kernel functions were used, and bandwidth was set to 1.5, chosen after experimentation of different values. Datasets were four-dimensional, containing player velocity (m/s), ball time-to-point (s), and the relative x- and y- co-ordinate of the contest (m).

Individually, these distributions represent the density of the data-sets in four dimensions. If a player's positional information and the ball location is known, the probability they will commit to a contest at location x is as follows:

$$\Pr(x) = \frac{wf_{c=1}}{wf_{c=1} + (1-w)f_{c=0}} \tag{2}$$

where w is a weighting factor equal to the size of the commitment dataset divided by the total number of samples, and $f_{c=1}$ and $f_{c=0}$ are the PDFs for the datasets where commitment = 1 and commitment = 0 respectively. A player's commitment probability ($\Pr(x)$) considers their position relative to x, their velocity, and the ball's time-to-point. Ball time-to-point to a location is equal to the distance between the ball and the location, divided by ball velocity. Ball velocity was estimated as 18.5 m/s after manually timing kicks from two quarters of a single AFL match and taking the average, however we note that this is a rough estimation as distances were estimated from manually recorded transactions. This represents a novel method for combining the distributions of two datasets of unequal sample size, where the resulting metric quantifies the probability that a new point belongs to each distribution. The combination of these distributions in a 2D space is illustrated in Figure 1. The resultant distributions can be calculated for a player's position, providing a distribution of the likelihood of their repositioning to each location, such that we derive a representation of their spatial influence comparable to that of traditional motion models.

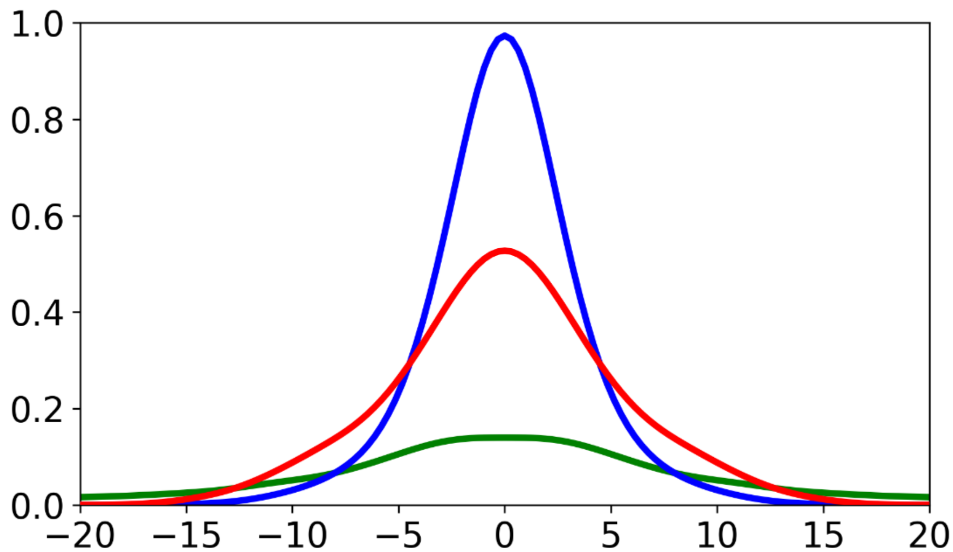


Figure 8-1. Two-dimensional representation of the commitment modelling process. The blue line represents the distribution for commitment values of 1 ($f_{c=1}$), and the green line represents the distribution for commitment values of 0 ($f_{c=0}$). The red line represents player influence ($Pr(x)$), derived from the combined commitment distributions (see Equation 2). This exemplar represents a player's commitment probability across relative x- co-ordinates for a y- displacement of 1 m, velocity = 4 m/s, and time = 2 s. All co-ordinates are relative to player displacement direction.

8.3.3 Decision-making Model

Following a pass, the ball can be received on the full, resulting in a mark, or can be received after a bounce, in which case a mark is not awarded. Hence, each of a player's passing options has four possible outcomes – successful passes in which a teammate receives the ball before (A) or after (B) it bounces, and unsuccessful passes in which an opponent does the same (C and D respectively). For each option, we calculate the probability (p) and value (e) of each event (Equation 3). As we consider players to be moving objects who exhibit spatial influence over locations not at their present position, the player with the ball could theoretically kick to any location within a radius equal to their maximum kicking distance. The typical maximum range of elite footballers has been found to be between 55 and 63 m (Ball, 2008c), hence the kicking radius in this study is set to 60 m. While some locations are likely sub-optimal choices, we calculate the expected outcome (EO) of each location within said radius. The EO for a location, x, is as follows:

$$EO(x) = p_A(x)e_a(x) + p_B(x)e_a(x) - p_C(x)e_o(x) - p_D(x)e_o(x) \tag{3}$$

where e_a and e_o are the field equity values for the attacking team and their opponent respectively. Derivation of field equity in AFL has been the focus of previous studies (O'Shaughnessy, 2006; Jackson, 2016).

From the EO of a pass, we calculate the value of a decision (referred to as the decision value or DV) as the EO of the pass that was executed, divided by the maximum EO contained in a player's kicking range (EOopt):

$$DV(x) = \frac{EO(x)}{EO_{opt}} \tag{4}$$

The EO of a pass will be negative if the equity at its target location is negative. For a decision with negative EO, the associated DV will likewise be negative. For a DV <-1, we set DV to -1.

8.3.4 Outcome Probabilities

For a given location, a team's spatial influence (INF) is the sum of the influence of its players:

$$INF(x) = \sum_{i=1}^{18} Pr_i(x) \tag{5}$$

where Pr_i is the commitment probability array for player i, from Equation 2. An attacking team's influence is a measure of the commitment of its players. From the influence of each team, we calculate the attacking team's spatial dominance (DOM) as:

$$DOM_a(x) = \frac{INF_a(x)}{INF_a(x) + INF_o(x)} \tag{6}$$

where $INF_a(x)$ and $INF_o(x)$ are the influence of the attacking team and their opponent at x.

The attacking team's dominance at x is the proportion of space they own. Logically, greater spatial dominance translates to a higher chance of a successful pass. Given that dominance is a relative measure, it is possible for a team to have high dominance over a location where influence is low. In such a case, while the probability of a successful pass is high due to their dominance, the probability that their players will reach the location is low, hence such a location is likely a poor passing location. To account for this, we calculate the probability of a successful mark (p_A and p_C from Equation 3) as a team's dominance multiplied by their influence.

$$p(x) = DOM(x) \times INF(x) \tag{7}$$

Given that a team's desired outcome is a successful pass resulting in a mark, this probability (Equation 7) is of particular importance when analysing a pass. We refer to p_A as the risk of a pass, where higher values indicate a safer passing option.

If a pass does not result in a mark, the probability that either team would win the ball is simply equal to their dominance (p_B and p_D from Equation 3).

8.3.5 Kicking Variance

Given imperfect accuracy of kicks, there is a chance that a kick will not reach its intended target. To incorporate this variance, we represent the likely target of a kick using a 2D Gaussian distribution with covariance equal to 5% of the kicking distance. The modified

EO of a kick is equal to the summed product of the kicking Gaussian's PDF and the raw EO values contained in its radius:

$$EO_{mod}(x) = \sum_{i \in S} EO(i) f(i) \tag{8}$$

where S is the set of integer co-ordinates in a radius around x equal to 5% of the Euclidean distance between the ball and x.

8.3.6 Statistical Analysis

For each analysed event, the optimal pass is identified as the pass to a teammate within a 60 m radius of the kicker whose EO is highest. The characteristics of the pass that was made and the pass identified as being optimal were extracted for all kicks that were preceded and resulted in a mark across the analysed matches (see Table 1 for a list of variables and definitions). We refer to the pass that was made as the decision and the pass identified as the optimal option as the alternative (note that if the decision was optimal it will be equal to the alternative). Descriptive statistics (mean ± SD) were produced for all metrics. Spearman's correlation coefficient (ρ) was used to measure the correlation of decision-making metrics with location. KDE was used to fit the distribution of analysed variables, finding that the decision-making metrics are not normally distributed. The Mann-Whitney U test was used to assess differences between the characteristics of decisions and alternatives (Mann & Whitney, 1947).

We explore team level trends in decision-making by comparing two teams. Teams were selected by taking the teams with the highest samples who fit the following criteria – one team who finished in the top 8 (Team A) in both the 2017 and 2018 regular AFL playing seasons, and one team who finished in the bottom 10 in the same seasons (Team

B). Participation in the play-off finals in AFL is between the top 8 teams, hence the choice of cut-off criteria ensured one team who participated in the finals, and one team from the cohort who did not make the finals group. Furthermore, the distribution of team samples is heavily skewed, hence importance was placed on selecting teams with adequate sample sizes. This skew in team samples is due to this study's focus on matches played at a single stadium, hence teams who more frequently played matches at this stadium appear more frequently in the dataset. Differences between team-level statistics were measured using the Mann-Whitney U test. Within-team decision-making is analysed for both teams. We fit the distribution of mean decision-making characteristics for each player on the team. All analyses were carried out in the Python programming language, using SciPy (Jones, et al., 2014) and the Scikit-learn (Pedregosa, 2011) packages.

8.4 RESULTS

8.4.1 Motion Models

Motion models were produced from 46220 instances of player commitment. Within the dataset there were 6392 instances of player commitment (Commitment = 1), and 39828 instances of no commitment (Commitment = 0), producing a weighting coefficient (w) of 0.14. Resultant motion models for four different player velocities for ball time-to-point of two seconds are visualised in Figure 8-2. Peak commitment probabilities occurred at 0.8 m for a velocity of 2 m/s (Figure 8-2a), 1.6m for 4 m/s (Figure 8-2b), 3.7 m for 6 m/s (Figure 8-2c), and 5.3 m for 8 m/s (Figure 8-2d). While density peaks at further distances as velocity increases, a negative correlation is revealed between player velocity (integers from 1 to 8 m/s) and peak commitment probabilities ($\rho = -0.80$ for $t = 2$ seconds), and between ball time-to-point (whole second integers from 1 to 4 seconds) and peak commitment probabilities (ρ = -1 for velocity = 4 m/s). At higher velocities, the probability that a player will commit to a contest decreases as the relative angle increases. For a velocity of 8 m/s or greater, player's exhibit minimal influence on space in the negative y- axis (i.e., behind their displacement direction). As velocity increases, we also note that the shape of a player's commitment inverts.

Table 8-1. Definitions of decision-making variables.

Variable	Definition
Dominance	The proportion of space owned by a team (see Equation 6)
Influence	A measure of spatial occupancy irrespective of opposition locations, equal to the summed commitment probabilities of a team's players (see Equation 5)
Risk	The likelihood of a successful pass resulting in a mark (see Equation 7)
Decision Value	The value of a player's passing decision, measured relative to the optimal decision available at the time of the pass (see Equation 4)
Expected Outcome	A numerical value describing the expected value of passing to a field position that considers the risk and reward of said pass (see Equation 3)
Distance	The Euclidean distance between two points. For a kick, distance is the Euclidean distance between the location of the kicker and of the receiver

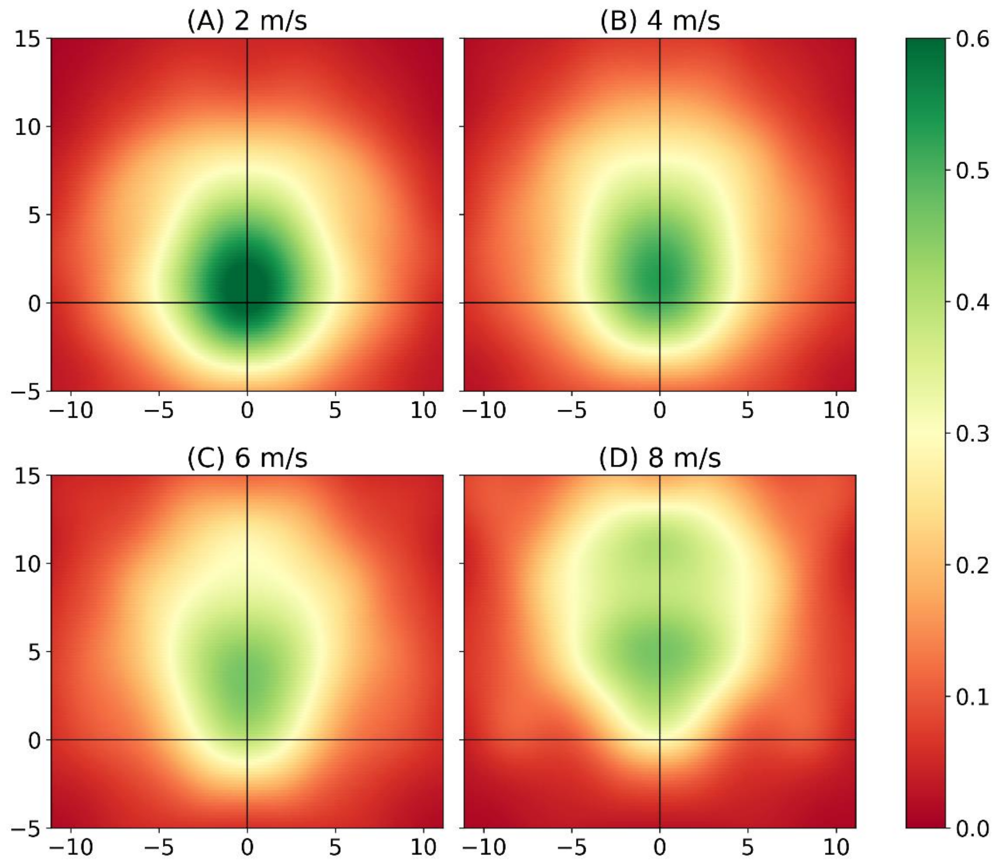


Figure 8-2. Motion models representing a player's area of influence whilst moving at various velocities for ball time-to-point of 2 s. Heatmap intensity is equivalent to the probability that a player (at the point of origin) would participate in a contest at relative x-, y- co-ordinates, as quantified by observed commitment behaviours.

8.4.2 Decisions and Alternatives

A total of 2935 passes matched the selection criteria across 60 matches (48.9 ± 14.7 kicks per match). An example decision-making output is visualised in Figure 8-3. In this example, the kicker passes to a teammate positioned towards the boundary line in the defensive 50 m region, while the model identified three higher value passes to teammates positioned towards the centre of the field. Figure 8-4 presents the components that constitute EO calculations. Summarised characteristics of decisions and alternatives are presented in Table 8-2. The mean of all analysed variables was lower for decisions compared to alternatives and all differences were statistically significant (refer to Table 8-2).

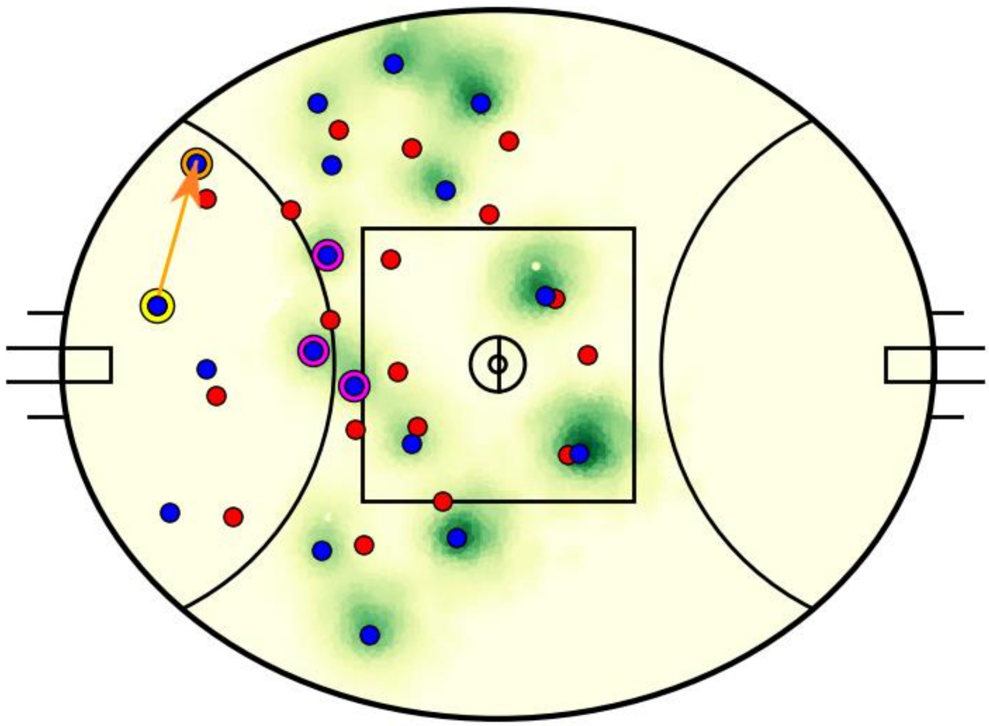


Figure 8-3. An example output of the decision-making model. The attacking team players are plotted in blue and their opponents in red. The kicker (circled in yellow) executed a pass along the orange line to the receiver (circled in orange). The model identified three higher valued passes (to players circled in magenta) towards the middle of the field that are within a 60 m radius of the kicker. The intensity of green correlates to the expected outcome of passes to each field position.

A very weak correlation was noted between vertical displacement from centre and DV of decisions ($\rho = 0.06$). Horizontal displacement from the attacking team's goal is positively correlated with DV ($\rho = 0.56$).

The distributions of decision-making characteristics are presented in Figure 8-5. The distribution of dominance (Figure 8-5a) is bimodal, with peak density for decisions at

DOM = 0.54 and a local maximum at DOM = 1.0. This global peak at 0.54 represents a contest between two teams that slightly favours the attacker, while the local maximum at

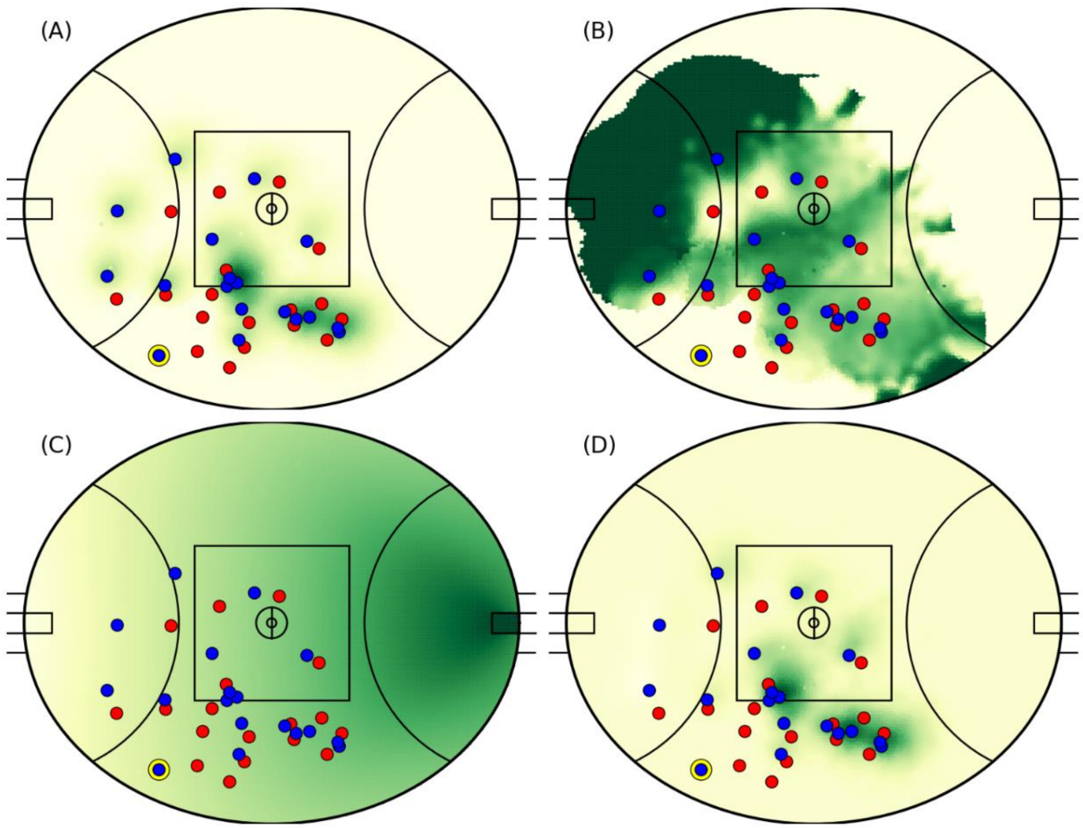


Figure 8-4. Team influence (A), dominance (B), field equity (C), and resultant expected outcomes (D) relative to the player in possession (circled in yellow, towards the lower boundary). High value space is represented as darker green regions. Team influence measures the spatial influence of the attacking team (whose players are in blue), while dominance measures their spatial ownership relative to the opposition (whose players are in red). All values are calculated relative to the player in possession. When complete, the model presented in this paper identifies two high value areas towards the centre square, both viable passing options (see D).

1.0 represents a kick to an area of absolute dominance. The distribution of alternatives is similarly bimodal, with a greater negative skew and density around absolute dominance. Influence of decisions (Figure 8-5b) reveals peak density at INF = 0.43, which is comparable to the average peak density of player commitment models (Figure 8-2). Density for risk peaks at 0.25 (Figure 8-5c). The shape of the distributions of EO for decisions and alternatives are different, with decisions exhibiting peak density at EO = 0.14

(Figure 8-5d), and minimal density is noted at EO > 2, while alternatives are noted as having a greater range of EO values, with no notable density peak. DV follows a relatively

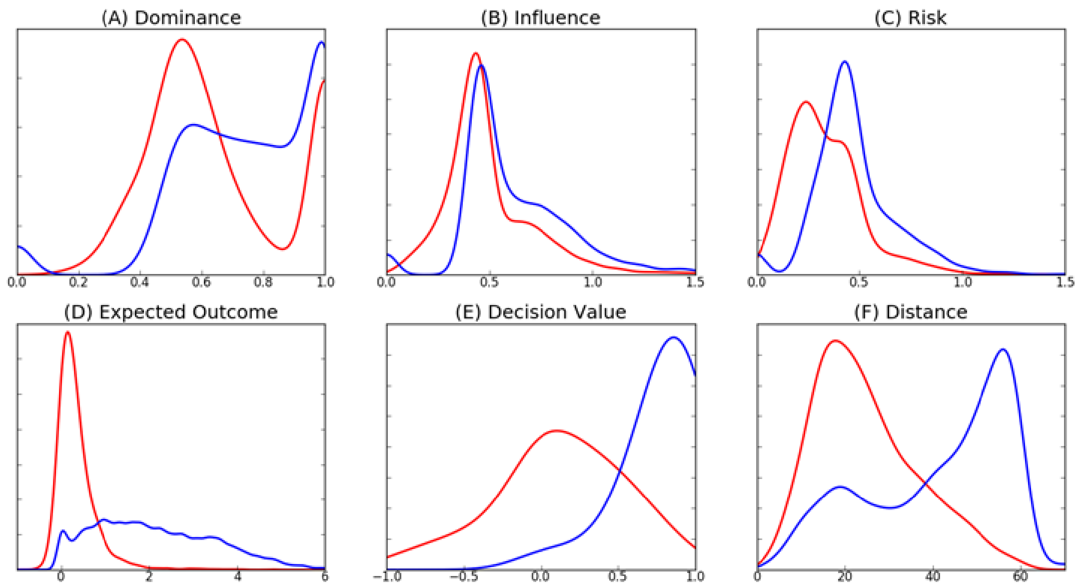


Figure 8-5. Distribution of decision-making characteristics for decisions (red) and alternatives (blue).

normal distribution for decisions (Figure 8-5e) and distributions of kicking distance (Figure 8-5f) exhibit opposite skews (decisions are positively skewed, while alternatives negatively). Density of kicking distance for decisions is highest at 17.3 m, marginally longer than the 15 m minimum distance required for a legal mark. Small density peaks at 0.0 are observed for the dominance, influence, and risk of alternatives.

8.4.3 Team-level Characteristics

The distributions of passing characteristics for two teams are presented in Figure 8-6 and the summary statistics in Table 8-3. There was minimal difference in the dominance, influence, risk, and distance of decisions between the two teams. The mean EO and DV for Team B are higher than those of Team A, however no differences were found to be statistically significant. While the shape of variable distributions is similar for both teams, it is noted that Team B exhibits a greater negative skew for EO, DV, and distance variables. Distributions of mean decision-making characteristics for players amongst both teams were

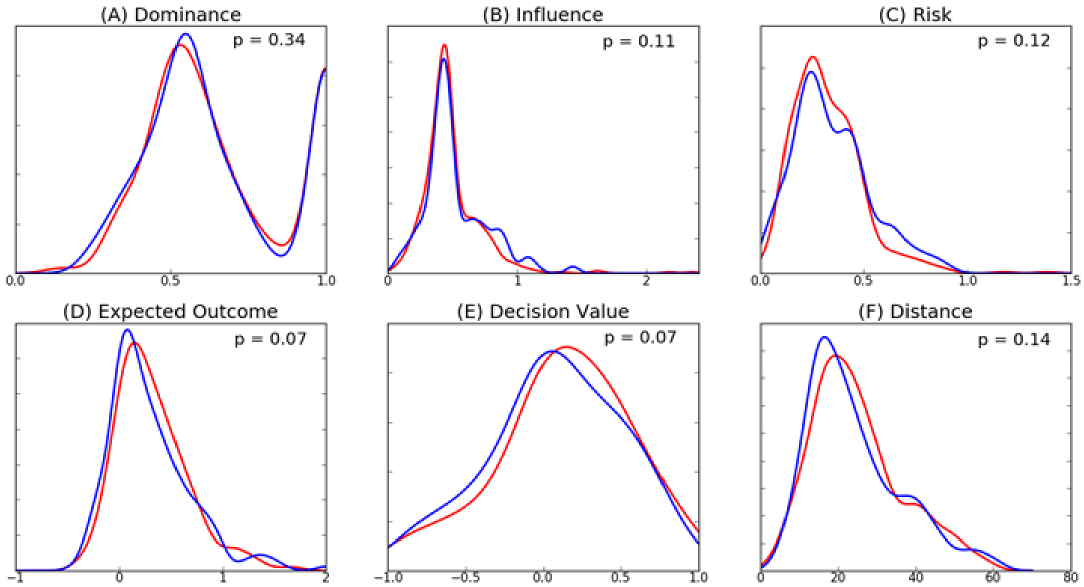


Figure 8-6. Distribution of team-level decision-making variables for Team A (blue) and Team B (red). Associated p-values (computed using the Mann-Whitney U test) are presented for each variable.

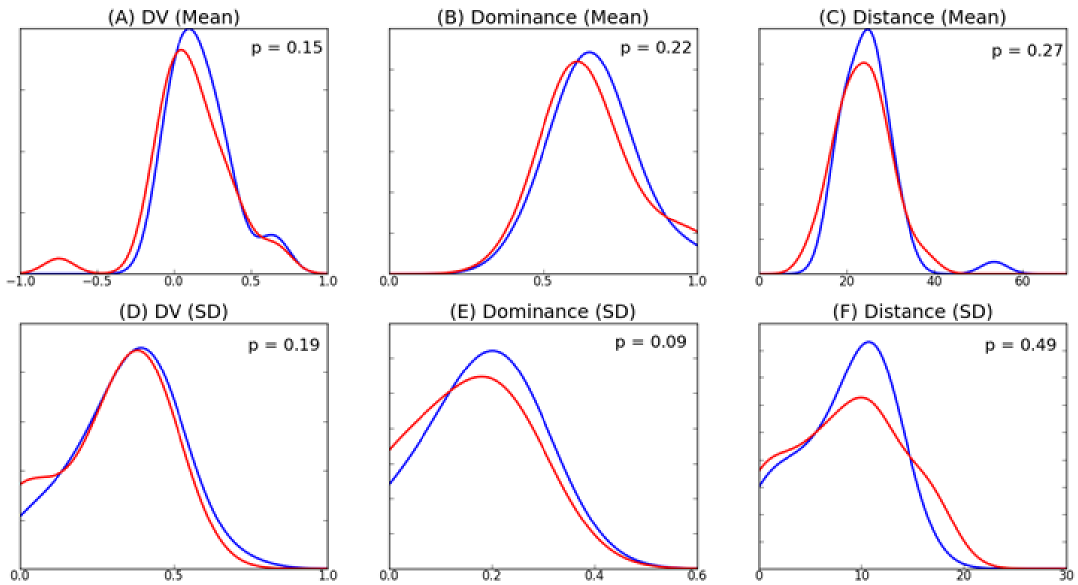


Figure 8-7. Within-team distributions for decision-making for Team A (blue) and Team B (red). Top row are the distributions for the mean decision-making abilities of players and the bottom row are variance.

found to be similar (Figure 8-7). While the differences between player-level standard deviations were not found to be statistically significant, the distributions for dominance and distance variance display visual differences.

Table 8-2. Mean values for decision-making variables between decisions and alternatives. Values are presented as Mean ± SD and all differences are statistically significant.

Variable	Decisions	Alternatives
Dominance	0.66 ± 0.23	0.75 ± 0.23
Influence	0.51 ± 0.27	0.63 ± 0.31
Risk	0.33 ± 0.19	0.47 ± 0.21
Expected Outcome	0.34 ± 0.46	2.11 ± 1.41
Decision Value	0.13 ± 0.42	0.78 ± 0.24
Distance	25.0 ± 11.8	42.7 ± 17.8

Table 8-3. Mean values for decision-making variables between Team A and Team B. p-values for differences are presented in Figure 5.

Variable	Team A	Team B

Dominance	0.66 ± 0.23	0.66 ± 0.23
Influence	0.52 ± 0.24	0.49 ± 0.24
Risk	0.34 ± 0.19	0.32 ± 0.17
Expected Outcome	0.29 ± 0.39	0.34 ± 0.42
Decision Value	0.08 ± 0.42	0.13 ± 0.43
Distance	24.3 ± 12.0	24.9 ± 11.6

8.5 DISCUSSION

This study demonstrates a method for measuring characteristics of player pass decision-making in invasion team sports. Previous studies of player decisions have measured decisions relative to some current measure of possession expectation (e.g., Cervone et al., 2014), rather than relative to the value of alternative passes that were presented. While the former approach assigns value to a specific kick, relative measures of decision-making assign value to individual decisions. Similar to the distinction between player accuracy and shot difficulty (e.g., Chang, et al., 2014), assigning value to player decision-making presents greater insights into individual player performance. The adoption of decision-making evaluation in combination with measurements of accuracy and risk would allow for targeted coaching and recruitment, as well as defining categories of player tactical behaviour.

A major component of the decision-making modelling were player motion models, fit on the weighted distributions of player commitment to contests. While previous studies have developed probabilistic motion models with arbitrary bounds (Fernandez & Bornn, 2018) or from a player's observed displacements (Brefeld, et al., 2018), the commitment modelling approach demonstrated in this study fits player behaviour with consideration of movement context, representing a new approach to the measurement of a player's spatial influence. Furthermore, the models are parameterized through the fitting of density in four dimensions (with consideration of a player's velocity, time and x- and y- co-ordinates), presenting a continuous representation of player commitment. A notable finding of the motion models is that commitment peaks are of lower density for higher velocity and time values. That is, players are overall less likely to commit to an upcoming contest if the ball is further away (hence, a high time-to-point) or if they are moving at high velocities. This finding is logical and may be explained by a desire to simply corral an opponent or reposition for future involvements, rather than participate in the immediate transaction. As with alternative motion models, we found that a player's influence in the negative y-axis (i.e., behind them) degrades as their speed increases. While models fit on player commitments more realistically measure their likelihood to occupy future space, the models only consider a player's position and momentum, not teammate locations. A player's participation in a contest logically has some dependence on the position of their teammates, hence attempts to incorporate may produce more realistic models.

A key finding in this study are the novel insights into the decision-making and passing tendencies of Australian Rules footballers. Previous studies have identified the importance of kicking in the AFL (Stewart, et al., 2007; Robertson, et al., 2016a) but there has been minimal work into describing the kicking landscape at elite levels at a transactional level (e.g., distance, level of pressure), despite studies on the biomechanics of kicking in Australian Rules football (e.g., Ball, 2008a; Ball, 2008b). This study found that kicks resulting in a mark are most commonly short, with a density peak at 17.3 metres (mean = 25 m), marginally longer than the minimum distance required for a legal mark.

This could be the result of tactical behaviour, or indicative of the ease in which close options can be identified due to lower visual obstruction. Furthermore, successful marks are most often to players in one on one contests or to players who are completely open (as suggested by the bimodal distribution of passing dominance and the density peaks of risk), which may be indicative of risk aversion, however more research is required to understand individual player behaviour.

In contrast to player decisions, the optimal alternative passes that were identified by the model presented in this study were long distance kicks, less frequently to unmarked individuals. While the distribution of dominance was similarly bimodal for alternatives, the peak at absolute dominance (DOM = 1.0) was less intense than for decisions. The higher density for passes to areas of dominance between 0.5 (a 50/50 contest) and 1.0 suggests kicks to areas in which multiple teammates have an opportunity to receive the ball. This is reinforced by the distribution of influence for alternatives (Figure 8-5b) where more density is noted for influence above 0.5 compared to decisions. Long-range passes having higher associated values (EO and DV) is logical due to the inclusion of AFL field equity, in which the value of space increases as the distance and angle to the goalposts decreases (Jackson, 2016; Figure 8-4c). The contrast in distances between decisions and alternatives (Figure 8-5f) could be due to several factors, such as a difficulty for players to identify long-range options (due to visual obstruction and lower decision-making time, for example) or an underestimation of kicking accuracy by the model. Due to the unavailability of precision ball tracking in AFL, this study used an arbitrary measurement of kicking accuracy. Should more detailed transactional data or LPS ball tracking become available, it is believed that kicking accuracy could be modelled from empirical data. The density peaks at values towards 0 for dominance, influence, and risk can be explained by situations in which all passing options are positioned in areas of negative field equity (e.g., field formations in the defensive 50 m area), resulting in an optimal decision being a kick to an area of no spatial dominance (hence, no negative associated equity). This is a common problem with models that use equity-based rewards, where moving the ball backwards is often associated with a reduction in equity.

Team level analysis revealed that the less successful team in the 2017/2018 season had higher average DV than the more successful team. Furthermore, while within-team distribution of player averages were similar, the player variance of DV was more positively skewed for the less successful team. Of particular interest is the finding that the less successful team executed passes of higher value, potentially suggestive of a difference in playing styles. Future research into player and team-level decision-making should consider contextual information such as match conditions, score deficits, and tactical styles. Despite these differences in the mean and standard deviation of team-level metrics, we note that the differences were not statistically significant in all cases ($p > 0.05$). Compared to the league-wide averages, the greatest differences experience by either team were of Team A’s DV and EO. Given that the decision-making model is developed from league-wide averages, this may suggest that Team A executes passes at a level above the league average. The reward component is fit on the average equity gain, given field location and pressure, hence it is possible that individual teams equity gains may have significant variation. Future research into the decision-making of Australian footballers should consider differences in outcomes to identify if there is a difference in the execution of passes between teams. That is, do certain teams outperform the mathematical averages of this decision-making model?

8.6 CONCLUSION

This work represents the beginning of ongoing research into player decision-making in the AFL. The decoupling of player decision-making from overall player performance allows for a more precise understanding of player ability that has applications in coaching and scouting. Underlying the decision-making model is a player motion model fit on the combined distributions of relative contest locations that were committed to, and those that were not. The resulting motion model quantifies the probability that a player would commit to a contest, given their velocity, displacement direction, and past behaviours. It was found that player commitment decreases as a function of velocity and available time, offering insights into the commitment behaviour of players. Analysis of passes revealed that players typically execute short kicks that are most commonly to teammates in one-on-one or unmarked situations, resulting in a bimodal distribution of passing dominance. Conversely, the mathematical model presented in this paper identifies long-range options as having higher expected value, given the inclusion of field equity which rewards possession closer to the goalposts. This mismatch in decisions could be due to the ease in which short-range options can be identified and executed compared to long-range options.

Chapter 9: Discussion and Conclusions

Chapter Overview

In this chapter, the outputs and results of this thesis are discussed.

This chapter contains a summary of thesis research (Section 9.1) and contributions to knowledge (Section 9.2). The limitations (Section 9.3) and future research directions (Section 9.4) are discussed.

9.1 RESEARCH SUMMARY

In this section, the research outputs of chapters contained in this thesis are briefly summarised, grouped by theme.

9.1.1 Coordinative Behaviour

Chapter 4 analysed coordinative behaviour of Australian footballers. The relative phase of intra- and inter-team player couples was measured. The phase angles for angular velocity and acceleration were examined. It was found that high coordinative couples involved midfield players. Furthermore, coordination degrades in later quarters.

9.1.2 Spatial Occupancy

Chapter 5 marked the beginning of the thesis' investigation into space in Australian football. Gaussian mixture models were used to estimate the density of player groups. This served as an approximation of congestion, however did not discriminate between teams. Hence, this work informed a change in the focus of future work on spatial occupancy.

In chapter 6, a variation of the motion models presented in Gudmundsson and

Wolle (2010) and Brefeld et al. (2018) were used to identify the spatial influence of players. Observed displacements were bound by smoothed *egg*-shaped ellipses. In theory, a player is able to reposition to any location contained in these bounds. When applied across teams, the spatial ownership (or dominance) of the attacking team can be measured on a continuous scale (via overlapping influence).

Chapter 7 and 8 develop a more realistic measure of spatial occupancy. In contrast to the model from chapter 6, the *commitment models* developed in chapter 7 represent the amount of space a player can reach as a distribution. The outputs describe the probability a player would reach locations in their vicinity, dependent on their velocity and available time.

9.1.3 Player Decision-Making

Two models for measuring player decisions were presented. The value of a player's decision is measured relative to their available options. It was found that Australian footballers tend to execute short-range passes which are frequently low value comparative to available alternatives. Furthermore, in Chapter 7 it was observed that the spatial features of receiving players differs by field region.

9.2 CONTRIBUTION TO KNOWLEDGE

This thesis has contributed new approaches to measurements of player behaviour and spatial occupancy. The applications of these approaches have produced insights into the behaviour of Australian footballers.

9.2.1 Spatial Occupancy

Gudmundsson and Horton (2017) suggested that the modelling of player motion has often been simple and has ignored the physiological constraints and momentum of players. In this thesis, multiple methods of incorporating player momentum into the modelling of player motion were exemplified. This was taken one step further with the addition of contextual information of a player's displacement (chapter 7).

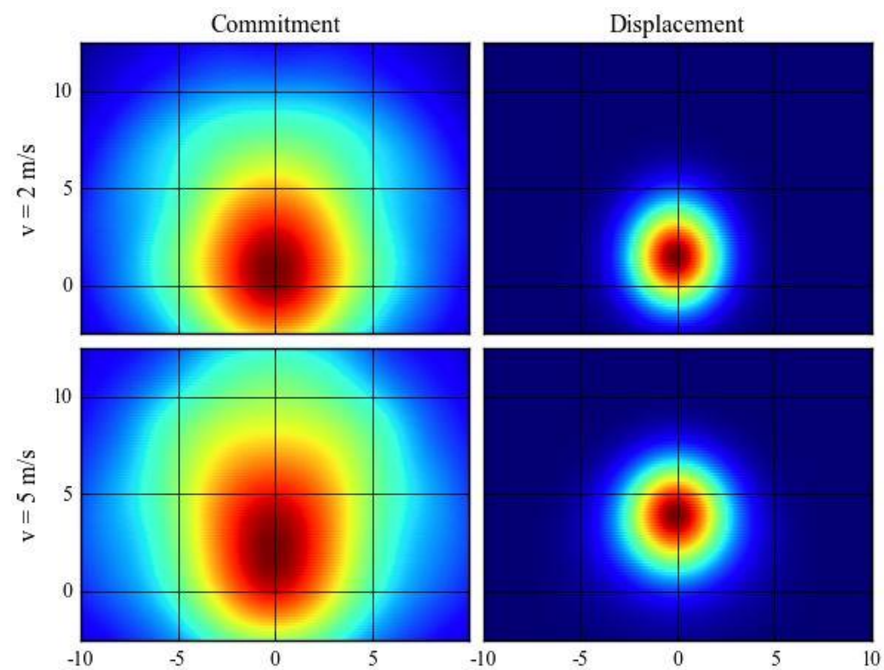


Figure 9-1. Comparison of commitment (left) and displacement (right) motion models for two velocities.

A new approach to the modelling of player motion was presented in Chapter 7. Commitment models measure the likelihood that a player will reposition to forthcoming contests. These models improve upon existing models of player motion should the application require the measurement of space in respect to the ball. Early research into player motion models suggested contextual information could improve upon existing approaches (Fujimura & Sugihara, 2005). Displacement-based continuous models have not considered the context of player displacements (e.g., Gudmundsson & Wolle, 2010; Brefeld, et al., 2018). Commitment models are compared to displacement-based models in

Figure 9-1. The difference in density reveals that player displacements are dependent on the location of the ball.

In contrast to a recent study that measured space relative to the ball (Fernandez & Bornn, 2018), commitment models are fit on empirical player data. The result is a more dynamic model that reflects a player's historical behaviour. Notably, the shape of the models produced by this process do not fit a normal two-dimensional Gaussian (Fig. 9-1), hence the Gaussian process described in Fernandez and Bornn (2018) would be inadequate for application in Australian football.

Early measures of spatial pressure were simple and did not considered orientation of players (Gudmundsson & Horton, 2017). This has recently been addressed by Fernandez and Bornn (2018), although for different applications. The models presented in this thesis output a player's spatial influence. From influence, dominance is calculated as the proportion of space a team owns. These metrics provide information about the spatial pressure exhibited by individuals which has been suggested as an area that requires more research (Gudmundsson & Horton, 2017).

9.2.2 Decision-making

The decision-making metrics presented in Chapters 7 and 9 represent a new approach to the measurement of player decisions. Previous studies have measured player decisions relative to change in possession expectations (e.g., Cervone et al., 2014). The approach in this thesis was to measure a decision relative to a player's available options. While a player's decision may have increased a possession expectation, if there existed alternatives that would have increased it by a greater amount then that decision was sub-optimal. Measuring decisions in this manner requires possession expectations to be forecast for available options, rather than computed for the immediate possession as in previous studies (e.g., Cervone et al., 2014). Doing so requires calculating the percentage chance of a pass being successful. This was estimated via a team's spatial dominance.

A component of the decision-making models was consideration of player momentum when computing possession values in their vicinity. Hence, the optimal receiving location of a player can be identified, given the location and momentum of opponents. Previous studies of player passing have only considered passes along the shortest path between the passer and receiver (Gudmundsson & Horton, 2017).

9.2.3 Contributions to AFL

The primary contributions of this thesis are the insights derived into Australian football. Prior to the research presented in this thesis, there had been no research in Australian football that utilised positional data of both teams. While player-tracking data had been used to summarise movement profiles of footballers (e.g., Wisbey, et al., 2010; Brewer, et al., 2010; Heasman, et al., 2011), no studies had used the positional data of players.

Coordinative Behaviour

The measurement of player coordinative behaviour via relative phase couplings revealed that high coordinative pairs of both intra- and inter-team couples involved at least one midfield player. Midfield players generally have higher work rates than other positions (Wisbey, et al., 2010), hence these findings are logical. Furthermore, this study extended the findings of Morgan and Williams (2012) with the addition of inter-team pairings. Inter-team pairings were proposed as a means of identifying players who are being marked by opponents. It was found that both intra- and inter- coordinative behaviour degraded throughout the analysed match.

Motion Models

The three approaches to spatial occupancy presented in this study provided insights into the behaviour of Australian footballers. The summarised density of players did not have a notable trend throughout a match of Australian football (chapter 5). However, there were notable differences between the density and its entropy during successful and unsuccessful possession chains.

Chapter 6 is one of the first studies to apply displacement-based models outside of soccer (for previous applications, see Gudmundsson & Wolle, 2010; Horton et al., 2015; Brefeld et al., 2018). A slight variation in bounding the player displacements was presented. Rather than using a convex hull or smoothed ellipses, partial ellipses were fit independently in the positive and negative *y*- axis. This process produces an approximation of player limits and is less computationally expensive than convex hull methods. Applying the models to Australian footballers produced logical insights – namely, that players are able to cover more forward space as velocity increases, and that reorientation to cover negative space is unlikely at high velocities.

Commitment modelling produced novel insights into player participation to contest transactions (i.e., contested marks and spoils). From the commitment dataset, it was found that approximately 14% of players within 30 metres of the forthcoming contest will reposition to participate in said contest. Further insights were similar to that of displacement-based models – players are less likely to commit to contests behind them as velocity increases. Interestingly, the peak commitment probabilities decrease as velocity and time increases. This suggests that players who have more time to are less likely to commit to the forthcoming contest. It is likely that they use this time to reposition for future possessions.

Player Decision-Making

This thesis presented two models for measuring player decision-making. The approaches had many similarities, with the key difference that the second model (Chapter 8) considered more passing outcomes. It was discovered that players infrequently execute decisions identified as high value. High value decisions were typically long range passes to open players, while common passes were short range. Possible reasons for this were discussed in Chapters 6, 7 and 8.

Spatial metrics were used to analyse player passes. Analysis of passing in the AFL has previously been limited to discrete performance indicators which are manually collected, hence have questions around their accuracy. The metrics developed in this thesis were able to analyse aspects of kicking on a continuous scale. Players most commonly passed to one-on-one contests or open players. Furthermore, passes were rarely longer than 25 metres. Spatial metrics were used to cluster passes into three clusters in Chapter 7. Finally, it was found that players are more risk averse in the defensive 50 region and more frequently pass to congested player groups in the forward 50 region. While these findings are logical, previous statistics have been unable to quantify this on a continuous scale.

9.2.4 Applications to Other Team Sports

While the methodology presented in this thesis were applied to Australian football, these methods could be applied to other invasion team sports for similar results. Research was approached with transferability in mind, ensuring that the methods used were not dependent on the behaviour or number of players involved.

Spatial occupancy, in particular, could be transferred to other team sports. It was shown that commitment models were more representative of player behaviour than displacement-based motion models in Australian football. It is likely that findings would be similar in other invasion team sports where players typically follow the ball, rather than hold positions. Basketball would be one example of this.

The simplistic density metric presented in Chapter 5 is applicable to all team sports. This process utilises Gaussian mixture models (GMM) to estimate density and proposed the use of the Bayesian Information Criterion to estimate the congestion of spatial formations. Research into appropriate GMM inputs would be required to adapt this methodology to different field and team sizes.

9.3 LIMITATIONS

The limitations of individual methodologies are discussed in their relevant chapters. The general limitations of this work are discussed below.

The accuracy of the findings presented in each chapter are dependent on the accuracy of player-tracking technology. It has been suggested that the 10 Hz GPS and LPS devices that were used to collect data are adequate for most purposes (Akenhead, et al., 2014; Johnston, et al., 2014; Hoppe, et al., 2018). Furthermore, it was found that LPS devices are more accurate than GPS devices (Hoppe, et al., 2018). Hence, it is assumed that accuracy of player-tracking technology will continue to improve over time. Methodology presented in these studies can be applied on future technologies that collect (x, y) data on players.

A further limitation is a reliance on transactional data for ball position and match events. Transactions are manually collected by Champion Data for each match. There has been minimal research into the accuracy of their data collection. Robertson et al. (2016b) analysed the reliability of performance indicators by comparing Champion Data's reported events with individually collected data, finding strong agreement between the two (intra-class correlation coefficient > 0.94 for all performance indicators). This paper did not investigate the accuracy of performance indicator constraints (e.g., pressure source) (Robertson, et al., 2016b). Should ball tracking become available, methodology could be developed that did not use transactional datasets. All methods presented in this thesis utilized transactions purely for event timestamps, rather than positional information which is also collected. This limits the research's exposure to the inaccuracies these datasets may present. However, the reliability of transaction time-stamps has not been researched.

The metrics developed in this thesis require opposition data to implement. In the AFL, access to opposition tracking data is limited. Hence, uptake of these metrics in the AFL is unlikely until teams are granted data on all teams.

Finally, a notable limitation of aspects of Chapters 6, 7 and 8 is their use of the AFL field equity metric. Field equity was produced in O'Shaughnessy (2006) and later smoothed in Jackson (2016). This metric was derived from datasets that predate the uptake of player-tracking technologies in Australian football. Hence, this metric does not consider the locations of teammates and opponents. An updated equity metric may improve the findings of these studies.

9.4 FUTURE WORK

There exist many opportunities for spatiotemporal analysis in Australian football.

Future work can revisit methodologies from early studies which were limited by data availability. Notably, the research presented in chapter 4 and 5 were conducted on a single match of Under-18s Australian football and a single within-club training match respectively.

A secondary output of the research in Chapter 7 was identify frequently occurring passing types. This was achieved via the unsupervised clustering of spatial characteristics. Research into the measurement of playing profiles in Chapter 3 found that team profiles based on performance indicators didn't discriminate winning profiles effectively (Spencer, et al., 2016). The measurement of distinct styles from spatiotemporal data is a step towards quantifying a team's playing style. This scope of this work will be extended to measure the frequently occurring styles of different aspects of Australian football.

There remain limitations in the decision-making models. These include assumptions of fixed ball velocity and limitations in AFL field equity. Continued research into player decision-making should continue to address the remaining limitations. Future studies will have a focus on incorporating spatial information of teammates and their opponents into the computation of AFL field equity. Furthermore, the current player motion models are fit on league-wide data. The next stage of commitment-based motion modelling should produce models for different playing positions. Through continued data collection, it will be possible to fit motion models on the data of individual players, reflecting differences in their movement behaviours.

An interesting topic for future study is in the use of decision-making metrics to identify tactical behaviour. As with the applied work on passing types in Chapter 7, there exist different types of player decisions. Understanding the relationship between player decisions and team tactics would provide a greater understanding of a team's playing *style*. Furthermore, the difference between player decisions and those identified as optimal by the decision-making model should be explored. Whether the trend towards short passes is due to tactical or psychological reasons is an interesting topic.

Without precise ball tracking, assumptions were required when modelling passing outcomes. While ball position can be inferred from consolidating transactional (play-by-play match events) and tracking datasets, this assumes straight line passes. This process was used to infer ball position in latter chapters. A similar process has been exemplified in small-sided football matches (Folgado, et al., 2014). The process from this study included recording the time and location of out-of-bounds events, hence ball position can be inferred whilst the ball is out of play (Folgado, et al., 2014). Regardless, this process is reliant on unrealistic expectations of ball paths. Hence, applying the decision-making models to other team sports that have spatiotemporal ball data will be a future research focus.

Finally, the applicability of processes developed in this thesis to coaching, performance analysis, and recruitment should be assessed. Validation work on the metrics from this thesis (e.g., Decision Value) is an important step in their implementation at a team level. These metrics provide insights into components of player performance that are not captured by current data collection processes.

Chapter 10: References

1. Aguiar, M. et al., 2015. Footballer's movement behaviour during 2-, 3-, 4- and 5-a-side small-sided games. *Journal of Sports Sciences*, 33(12), pp. 1259-1266.

2. Akenhead, R., French, D., Thompson, K. & Hayes, P., 2014. The acceleration dependent validity and reliability of 10Hz GPS. *Journal of Science and Medicine in Sport*, pp. 17(5), 562-566.

3. Andreson, D., Breed, R., Spittle, M. & Larkin, P., 2018. Factors affecting set shot goal-kicking performance in the Australian Football League. *Perceptual and Motor Skills*, 125(4), pp. 817-833.

4. Aughey, R. & Falloon, C., 2010. Real time versus post game GPS data in team sports. *Journal of Science and Medicine in Sport*, 13(3), pp. 348-349.

5. Bailey, M., 2000. Identifying arbitrage opportunities in AFL betting markets through mathematical modelling. *Proceedings of the Fifth Australian Conference on Mathematics and Computers in Sport*, pp. 37-42.

6. Bailey, M. & Clarke, S., 2002. Predicting the brownlow medal winner. *Proceedings of the Sixth Australian Conference on Mathematics and Computers in Sport*, pp. 56-62.

7. Ball, K., 2008a. Biomechanical considerations of distance kicking in Australian Rules football. *Sports Biomechanics*, 7(1), pp. 10-23.

8. Ball, K., 2008b. Foot interactions during kicking in Australian Rules football. *In Science and Football VI*, pp. 62-66.

9. Ball, K., 2008c. Use of weighted balls for improving kicking for distance. *In Science and Football VI*, p. 285.

10. Barris, S. & Button, C., 2008. A review of vision-based motion analysis in sport. *Sports Med*, 38(12), pp. 1025-1043.

11. Bedford, A. & Shembri, A., 2006. An analysis of goal-kicking accuracy in Australian Rules football. *Proceedings of the 8th Conference on Mathematics and Computers in Sport.*

12. Bialkowski, A. et al., 2014. Large scale analysis of soccer matches using spatiotemporal tracking data. *2014 IEEE International Conference on Data Mining.*

13. Bourbousson, J., Seve, C. & McGarry, T., 2010a. Space-time coordination dynamics in basketball: Part 1. intra- and inter-couplings among player dyads. *Journal of Sports Sciences,* Volume 28, pp. 339-347.

14. Bourbousson, J., Seve, C. & McGarry, T., 2010b. Space-time coordination dynamics in basketball: Part 2. the interaction between the two teams. *Journal of Sports Sciences,* Volume 28, pp. 349-358.

15. Brefeld, U., Lasek, J. & Mair, S., 2018. Probabilistic movement models and zones of control. *Machine Learning,* 108(1), pp. 127-147.

16. Breiman, L., 2001. Random forests. *Machine Learning,* Volume 45, pp. 5-32.

17. Brewer, C. et al., 2010. Movement pattern comparisons in elite AFL and sub-elite WAFL Australian Rules football games using GPS. *Journal of Science and Medicine in Sport*, 13(6).

18. Carling, C., Williams, A. & Reilly, T., 2007. Handbook of soccer match analysis: A systematic approach to improving performance. *Routledge.*

19. Castellano, J. et al., 2013. Identifying the effects from the quality of opposition in a football team positioning strategy. *International Journal of Performance Analysis in Sport*, 13(3), pp. 822-832.

20. Castellano, J. & Casamichana, D., 2015. What are the differences between first and second divisions of Spanish football teams?. *International Journal of Performance Analysis in Sport*, 15(1), pp. 135-146.

21. Cervone, D., Bornn, L. & Goldsberry, K., 2016b. NBA court realty. *MIT Sloan Sports Analytics Conference.*

22. Cervone, D., D'Amour, A., Bornn, L. & Goldsberry, K., 2014. POINTWISE: Predicting points and valuing decisions in real time with NBA optical tracking data. *In MIT Sloan Sports Analytics Conference.*

23. Cervone, D., D'Amour, A., Bornn, L. & Goldsberry, K., 2016a. A multiresolution stochastic process model for predicting basketball possession outcomes. *Journal of American Statistical Association,* 111(514), pp. 585-599.

24. Chang, Y. et al., 2014. Quantifying shot quality in the NBA. *In Proceedings of the 8th Annual MIT Sloan Sports Analytics Conference.*

25. Chawla, S., Estephan, J., Gudmundsson, J. & Horton, M., 2017. Classification of passes in football matches using spatiotemporal data. *ACM Transactions on Spatial Algorithms and Systems*, 3(2), p. 6.

26. Clarke, S., 1992. Computer and human tipping of AFL football - a comparison of 1991 results. *Proceedings of the First Conference on Mathematics and Computers in Sport.*

27. Clarke, S. & Norman, J., 1998. When to rush a 'behind' in Australian Rules football: a dynamic programming approach. *Journal of the Operational Research Society*, Volume 49, pp. 530-536.

28. Clemente, F., Couceira, M., Martins, F. & Korgaokar, A., 2012a. Automatic tools of match analysis: Inspecting soccer team's dispersion. *International Journal of Applied Sports Sciences*, 24(2), pp. 81-88.

29. Clemente, F. et al., 2013. Measuring tactical behaviour using technological metrics: case study of a football game. *International Journal of Sports Science & Coaching*, 8(4), pp. 723-739.

30. Clemente, F., Martins, F. & Couceiro, M., 2012b. Towards a new method to analyse soccer teams tactical behaviour: measuring the effective area of player. *Indian Journal of Science and Technology*, 5(12).

31. Clemente, F. et al., 2015. Developing a tactical metric to estimate the defensive area of soccer teams: The defensive player area. *Journal of Sports Engineering and Technology*, 229(3).

32. Clemente, F. et al., 2014. Intelligent systems for analyzing soccer games: The weighted centroid. *Ingenieria e Investigacion*, 34(3), pp. 70-75.

33. Corbett, D. et al., 2017a. Development of physical and skill training drill prescription systems for elite Australian Rules football. *Science and Medicine in Football.*

34. Corbett, D., Sweeting, A. & Robertson, S., 2017b. Weak relationships between stint duration, physical and skilled match performance in Australian football. *Frontiers in Psychology*.

35. Coutts, A. et al., 2010. Match running performance in elite Australian Rules football. *Journal of Science and Medicine in Sport,* Volume 13.

36. Davids, K., Araujo, D. & Shuttleworth, R., 2005. Applications of dynamical systems theory to football. *Science and Football V.*

37. Dawson, B. et al., 2004. Player movement patterns and game activities in the Australian Football League. *Journal of Science and Medicine in Sport,* 7(3), pp. 278-291.

38. Duarte, R. et al., 2013. Competing together: Assessing the dynamics of team-team and player-team synchrony in professional association football. *Human Movement Science*, 32(4), pp. 555-566.

39. Edgecomb, S. & Norton, K., 2006. Comparison of global positioning and computer-based tracking systems for measuring player movement distance during Australian football. *Journal of Science and Medicine in Sport,* 9(1), pp. 25-32.

40. Falgado, H., Bravo, J., Pereira, P. & Sampaio, J., 2019. Towards the use of multidimensional performance indicators in football small-sided games: the effects of pitch orientation. *Journal of Sports Sciences,* pp. 1064-1071.

41. Fernandez, J. & Bornn, L., 2018. Wide open spaces: A statistical technique for measuring space creation in professional soccer. *In: Sloan Sports Analytics Conference.*

42. Folgado, H., Lemmink, K., Frencken, W. & Sampaio, J., 2014. Length, width and centroid distance as measures of teams tactical performance in youth football. *European Journal of Sport Science,* 14(1), pp. 487-492.

43. Fonseca, S., Milho, J., Travassos, B. & Araujo, D., 2012. Spatial dynamics of team sports exposed by Voronoi diagrams. *Human Movement Science,* 31(6), pp. 1652-1659.

44. Foreman, K., Deegan, G. & Wigley, G., 2012. Global position systems in AFL, worthwhile or a waste of time?. *23rd Australasian Conference on Information Systems.*

45. Franks, A., Miller, A., Bornn, L. & Goldsberry, K., 2015. Characterizing the spatial structure of defensive skill in professional basketball. *The Annals of Applied Statistics,* 9(1), pp. 94-121.

46. Frank, T. & Richardson, M., 2010. On a test statistic for the Kuramoto order parameter of synchronization: An illustration for group synchronization during rocking chairs. *Physica D,* Volume 239, pp. 2084-2092.

47. Frencken, W., Van Der Plaats, J., Visscher, C. & Lemmink, K., 2013. Size Matters: Pitch dimensions constrain interactive team behaviour in soccer. *Journal of Systems Science and Complexity,* 26(1), pp. 85-93.

48. Frenken, W., Lemmink, K., Delleman, N. & Visscher, C., 2011. Oscillations of centroid position and surface area of soccer teams in small-sided games. *European Journal of Sports Science,* Volume 11, pp. 215-223.

49. Fujimura, A. & Sugihara, K., 2005. Geometric analysis and quantitative evaluation of sport teamwork. *Systems and Computers in Japan,* pp. 36(6), 49-58.

50. Goncalves, B., Figueira, B., Macas, V. & Sampaio, J., 2013. Effect of player position on movement behaviour, physical and physiological performances during an 11-a-side football game. *Journal of Sports Sciences,* 32(2), pp. 191-199.

51. Gray, M. & Jenkins, D., 2010. Match analysis and the physiological demands of Australian football. *Sports Medicine,* 40(4), pp. 347-360.

52. Green, S., 2012. Assessing the performance of Premier League goalscorers. *OptaPro Blog.*

53. Grehaigne, J., Bouthier, D. & David, B., 1997. Dynamic-system analysis of opponent relationships in collective actions in soccer. *Journal of Sports Sciences,* 15(2), pp. 137-149.

54. Gronow, D. et al., 2014. Team movement patterns with and without ball possession in Australian Football League players. *International Journal of Performance Analysis in Sport*, Volume 14, pp. 635-661.

55. Gudmundsson, J. & Horton, M., 2017. Spatio-temporal analysis of team sports. *ACM Computing Surveys*, 50(2), pp. 1-22.

56. Gudmundsson, J. & Wolle, T., 2010. Towards automated football analysis: Algorithms and data structures. *10th Australasian Conference on Mathematics and Computers in Sport.*

57. Gudmundsson, J. & Wolle, T., 2014. Football analysis using spatio-temporal tools. *Computers, Environment and Urban Systems,* Volume 47, pp. 16-27.

58. Gyarmati, L., Kwak, H. & Rodriguez, P., 2014. Searching for a unique style in soccer. *2014 KDD Workshop on Large-Scale Sports Analytics.*

59. Heasman, J., Dawson, B. & Lay, B., 2011. The relationship between GPS match activity profile and performance in the AFL. *Science and Football VII: The proceedings of the Seventh World Congress on Science and Football.*

60. Hoppe, M., Baumgart, C. P. T. & Freiwald, J., 2018. Validity and reliability of GPS and LPS for measuring distances covered and sprint mechanical properties in team sports. *PlosONE.*

61. Horton, M., 2013. Automated classification of passes in football. *Masters by coursework thesis, The University of Sydney.*

62. Horton, M., Gudmundsson, J., Chawla, S. & Estephan, J., 2015. Automated classification of passing in football. In: s.l.:Springer, Cham, pp. 319-330.

63. Hurst, H. & Sinclair, J., 2013. Validity and reliability of 5 Hz GPS for measurements of non-linear cycling distance and velocity. *International Journal of Sports Science and Engineering*, 7(1), pp. 11-16.

64. Jackson, K., 2016. Assessing player performance in Australian football using spatial data. *Doctoral Dissertation, PhD Thesis, Swinburne University of Technology.*

65. Jackson, K., 2016. Measuring the similarity between palyers in Australian football. *Mathematics and Computers in Sport,* Volume 13, pp. 9-14.

66. Jackson, K., 2017. Using player quality and injury profiles to simulate results in Australian football. *Mathsport*, p. 199.

67. Jacques, T. & Pavia, G., 1974. An analysis of the movement patterns of players in an Australian Rules League Football match. *Journal of Sports Med*, 5(10), pp. 10-21.

68. Jennings, D. et al., 2010. Variability of GPS units for measuring distance in team sports. *International Journal of Sports Physiology and Performance*, 5(4).

69. Johnston, R. et al., 2014. Validity and interunit reliability of 10 Hz and 15 Hz GPS units for assessing athlete movement demands. *The Journal of Strength & Conditioning Research,* 28(6), pp. 1649-1655.

70. Jones, E., Oliphant, T. & Peterson, P., 2014. *{SciPy}: open source scientific tools for Python.* s.l.:s.n.

71. Kaplan, E., Mongeon, K. & Ryan, J., 2014. A markov model for hockey: Manpower differential and win probability added. *INFOR: Information Systems and Operational Research,* 52(2), pp. 39-50.

72. Kempton, T., Kennedy, N. & Coutts, A., 2015. The expected value of possession in professional rugby league match play. *Journal of Sports Sciences*, 34(7), pp. 645-650.

73. Kempton, T. et al., 2013. Match-to-match variation in physical activity and technical skill measures in professional Australian football. *Journal of Science and Medicine in Sport*, 18(1), pp. 109-113.

74. Kim, K. et al., 2010. Motion fields to predict play evolution in dynamic sport scenes. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*

75. Kodinariya, T. & Makwana, P., 2013. Review on determining number of cluster in K-means clustering. *International Journal of Advance Research in Computer Science and Management Studies,* 1(6), pp. 90-95.

76. Lames, M., 2006. Modelling the interaction in game sports - relative phase and moving correlations. *Journal of Sports Science and Medicine,* Volume 5, pp. 556-560.

77. Lames, M., Ertmer, J. & Walter, F., 2010. Oscillations in football - order and disored in spatial interactions between the two teams. *International Journal of Sport Psychology,* 41(4), p. 85.

78. Le Grand, C., 2007. *AFL's hi-tech future within reach,* Adelaide: Australian.

79. Le, H., Carr, P. Y. Y. & Lucey, P., 2017. Data-driven ghosting using deep imitation learning. *MIT Sloan Sports Analytics Conference 2017.*

80. Leite, N. et al., 2014. Effect of defensive pressure on movement behaviour during an under-18 basketball game. *International Journa of Sports Medicine,* 35(9), pp. 743-748.

81. Linke, D., Link, D. & Lames, M., 2018. Validation of electronic performance and tracking systems EPTS under field conditions. *Plos ONE.*

82. Lopez-Felip, M., Davids, T., Frank, T. & Dixon, J., 2018. A cluster phase analysis for collective behavior in team sports. *Human movement science,* Volume 59, pp. 96-111.

83. Lucey, P. et al., 2013. Representing and discovering adversarial team behaviors using player roles. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.*

84. Lucey, P. et al., 2014a. "Quality vs Quantity": Improved shot prediction in soccer using strategic features from spatiotemporal data. *MIT Sloan Sports Analytics Conference.*

85. Lucey, P. et al., 2014b. Analyzing team movement in basketball using tracking data. *8th Annual MIT Sloan Sports Analytics Conference.*

86. Lucey, P. et al., 2013. Assessing team strategy using spatiotemporal data. *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* pp. 1366-1374.

87. Lupo, C., Condello, G. & Tessitore, A., 2012. Notational analysis of elite men's water polo related to specific margins of victory. *Journal of Sports Science & Medicine,* 11(1), p. 516.

88. Madhulatha, T., 2012. An overview of clustering methods. *ISOR Journal of Engineering,* 2(4), pp. 719-725.

89. Maheswaran, R. et al., 2014. The three dimensions of rebounding. *MIT Sloan Sports Analytics Conference 2014.*

90. Mann, H. & Whitney, D., 1947. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics,* pp. 50-60.

91. McGarry, T. et al., 2002. Sport competition as a dynamical self-organizing system. *Journal of Sports Sciences,* Volume 20, pp. 771-781.

92. McGarry, T., Kahn, M. & Franks, I., 1999. On the presence and absence of behavioural traits in sport: an example from championship squash match-play. *Journal of Sports Sciences,* Volume 17, pp. 297-311.

93. McIntosh, M., Kovalchik, S. & Robertson, S., 2018. Validation of the Australian Football League player ratings. *International Journal of Sports Science & Coaching,* 13(6), pp. 1064-1071.

94. McKenna, M., Patrick, J. & Sandstorm, E., 1988. Computer-video analysis of activity patterns in Australian Rules football. *Science and Football. London: Routledge.*

95. Memmert, D., Lemmink, K. & Sampaio, J., 2017. Current approaches to tactical performance analyses in soccer using position data. *Sports Medicine,* 47(1), pp. 1-10.

96. Meyer, D., Forbes, D. & Clarke, S., 2006. Statistical analysis of notational AFL data using continuous time Markov chains. *Journal of Sports Science & Medicine,* 5(4), pp. 525-532.

97. Morgan, S. & Williams, M., 2012. Using relative phase of position derivatives to quantify within-team player couplings. *11th Australasian Conference on Mathematics and Computers in Sport,* pp. 129-138.

98. Nakanishi, R., Maeno, J., Murakami, K. & Naruse, T., 2009. An approximate computation of the dominant region diagram for the real-time analysis of group behaviours. *Proceedings of the 13th Annual RoboCup International Symposium,* pp. 228-239.

99. Narizuka, T., Yamamoto, K. & Yamazaki, Y., 2014. Statistical properties of position-dependent ball-passing networks in football games. *Physica A: Statistical Mechanics and its Applications,* Volume 412, pp. 157-168.

100. Neville, J., Rowlands, D., Wixted, A. & James, D., 2012. Application of GPS devices to longitudinal analysis on game and training data. *Procedia Engineering,* Volume 34, pp. 443-448.

101. Ogris, G. et al., 2012. Accuracy of the LPM tracking system considering dynamic position changes. *Journal of Sports Sciences,* 30(14), pp. 1503-1511.

102. Okabe, A., Boots, B. & Sugihara, K., 1992. *Spatial tessellations concepts and applications of Voronoi diamgrams.* New York: Wiley & Sons.

103. O'Shaughnessy, D., 2006. Possession versus position: strategic evaluation in AFL. *Journal of Sport Science & Medicine,* 5(4), p. 533.

104. Pault, Y. & Zanone, P., 2005. A dynamical analysis of tennis: Concepts and data. *Journal of Sports Sciences,* 23(10), pp. 1021-1032.

105. Pedregosa, F., 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research,* pp. 2825-2830.

106. Portas, M., Harley, J., Barnes, C. & Rush, C., 2010. The validity and reliability of 1 Hz and 5 Hz global positioning systems for linear, multidirectional and soccer-specific activities. *International Journal of Sports Physiology and Performance,* 5(4), pp. 448-458.

107. Power, P., Ruiz, H., Wei, X. & Lucey, P., 2017. Not all passes are created equal: Objectively measuring the risk and reward of passes in soccer from tracking data. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1605-1613.

108. Rein, R., Raabe, D. & Memmert, D., 2017. Which pass is better? Novel approaches to assess passing effectiveness in elite soccer. *Human Movement Science,* Volume 55, pp. 172-181.

109. Richman, J. & Moorman, J., 2000. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology,* 278(6).

110. Robertson, S., Back, N. & Bartlett, J., 2016a. Explaining match outcome in elite Australian Rules football using team performance indicators. *Journal of Sports Sciences,* 34(7), pp. 637-644.

111. Robertson, S., Gupta, R. & McIntosh, S., 2016b. A method to assess the influence of individual player performance distribution on match outcome in team sports. *Journal of Sports Sciences,* 34(19), pp. 1893-1900.

112. Robertson, S., Spencer, B., Back, N. & Farrow, D., 2019. A rule induction framework for the determination of representative learning design in skilled performance. *Journal of Sports Sciences,* pp. 1-6.

113. Routley, K., 2015. A markov game model for valuing player actions in ice hockey. *Doctoral dissertation.*

114. Ryall, R., 2011. Predicting outcomes in Australian Rules football. *PhD Thesis, RMIT.*

115. Ryan, S., Coutts, A., Hocking, J. & Kempton, T., 2017. Factors affecting match running performance in professional Australian football. *International Journal of Sports Physiology and Performance,* 12(9), pp. 1199-1204.

116. Sampaio, J., Lago, C., Casais, L. & Leite, N., 2010. Effects of starting score-line, game location, and quality of opposition in basketball quarter score. *European Journal of Sport Science,* 10(6), pp. 391-396.

117. Sampaio, J. & Macas, V., 2012. Measuring tactical behaviour in football. *International Journal of Sports Medicine,* 33(5), pp. 395-401.

118. Sampaio, J. et al., 2015. Exploring game performance in the national basketball association using player tracking data. *PLoS ONE,* p. 10.

119. Sargent, J. & Bedford, A., 2013. Evaluating Australian Football League player contributions using interactive network simulation. *Journal of sports Science & Medicine,* 12(1), p. 116.

120. Sathyan, T., Shuttleworth, R., Hedley, M. & Davids, K., 2012. Validity and realibility of a radio positioning system for tracking athletes in indoor and outdoor team sports. *Behaviour Research Methods,* 44(4), pp. 1108-1114.

121. Schwarz, G., 1978. Estimating the dimension of a model. *The Annals of Statistics,* 6(2), pp. 461-464.

122. Serpiello, F. et al., 2018. Validity of an ultra-wideband local positioning system to measure locomotion in indoor sports. *Journal of Sports Sciences,* 36(15), pp. 1727-1733.

123. Silverman, B., 1986. Density estimatin for statistics and data analysis. *Chapman & Hall, London.*

124. Soni Madhulatha, T., 2012. An overview of clustering methods. *IOSR Journal of Engineering,* 2(4), pp. 719-725.

125. Spearman, W. et al., 2017. Physics-based modelling of pass probabilities in soccer. *Proceedings of the 11th MIT Sloan Sports Analytics Conference.*

126. Spencer, B., Bedin, T., Farrow, D. & Jackson, K., 2018. A method for evaluating player decision-making in the Australian Football League. *Proceedings of the 14th Australasian Conference on Mathematics and Computers in Sport,* pp. 7-12.

127. Spencer, B., Jackson, K., Bedin, T. & Robertson, S., In Review. Modelling the quality of player passing decisions in Australian Rules football relative to risk, reward and commitment.

128. Spencer, B., Morgan, S., Zeleznikow, J. & Robertson, S., 2016. Clustering team profiles in the Australian Football League using performance indicators. *In Proceedings of the 13th Australasian Conference on Mathematics and Computers in Sport, Melbourne.*

129. Spencer, B., Morgn, S., Zeleznikow, J. & Robertson, S., 2017. Measuring player density in Australian Rules football using Gaussian mixture models. *In: Complex Systems in Sport, International Congress Linking Theory and Practice,* pp. 172-174.

130. Stefani, R. & Clarke, S., 1992. Predictions and home advantage for Australian Rules football. *Journal of Applied Statistics,* Volume 19, pp. 251-261.

131. Stewart, M., Mitchell, H. & Stavros, C., 2007. Moneyball applied: Econometrics and the identification and recruitment of elite Australian footballers. *International Journal of Sport Finance,* 2(4), pp. 231-248.

132. Sullivan, C. et al., 2013. Match score affects activity profile and skill performance in professional Australian football players. *Journal of Science and Medicine in Sport.*

133. Sullivan, C. et al., 2014. Factors affecting match performance in professional Australian football. *International Journal of Sports Physiology and Performance,* 9(3), pp. 561-566.

134. Szczepanski, L. & McHale, I., 2016. Beyond completion rate: evaluating the passing ability of footballers. *Journal of the Royal Statistical Soceity: Series A (Statistics in Society),* pp. 179(2), 513-533.

135. Taki, T. & Hasegawa, J., 1998. Dominant region: a basic feature for group motion analysis and its application to teamwork evaluation in soccer games. *Videometrics VI,* Volume 3641, pp. 48-58.

136. Taki, T. & Hasegawa, J., 2000. Visualisation of dominant region in team games and its application to teamwork analysis. *Proceedings Computer Graphics International 2000 IEEE*, pp. 227-235.

137. Taki, T., Hasegawa, J. & Fukumura, T., 1996. Development of motion analysis sustem for quantitative evaluation of teamwork in soccer games. *Proceedings of the 3rd IEEE International Conference on Image Processing,* Volume 3, pp. 815-888.

138. Thomas, G. et al., 2017. Computer vision for sports: Current applications and research topics. *Computer Vision and Image UNderstanding.*

139. Travassos, B., Araujo, D., Duarte, R. & McGarry, T., 2012. Spatiotemporal coordination behaviors in futsal (indoor football) are guided by informational game constraints. *Human Movement Science,* Volume 31, pp. 932-945.

140. Ueda, F., Masaaki, H. & Hiroyuki, H., 2014. The causal relationship between dominant region and offense-defense performance–focusing on the time of ball acquisition. *Football Science,* Volume 11, pp. 1-17.

141. Varley, M., Fairweather, I. & Aughey, R., 2012. Validity and reliability of GPS for measuring instantaneous velocity during acceleration, deceleration and constant motion. *Journal of Sports Sciences,* 30(2), pp. 121-127.

142. Vilar, L. et al., 2013. Spatio-temporal constraints on decision-making during shooting performance in the team sport of futsal. *Journal of Sports Sciences,* 31(8), pp. 840-846.

143. Weiss, C. & Childers, S., 2014. Spatial statistics to evaluate player contribution in Ultimate. *MIT Sloan Sports Analytics Conference.*

144. Wei, X. et al., 2016. "The thin edge of the wedge": accurately predicting shot outcomes in tennis using style and context priors. *MIT Sloan Sports Analytics Conference.*

145. Wei, et al., 2015. Forecasting events using an augmented hidden conditional random field. *Proceedings of the 12th Asian Conference on Computer Vision.*

146. Wisbey, B., Montgomery, P. & Pyne, D., 2008. Quantifying changes in AFL player game demands using GPS tracking: 2008 AFL season. *Florey (ACT): FitSense Australia.*

147. Wisbey, B., Montgomery, P., Pyne, D. & Rattray, B., 2010. Quantifying movement demands of AFL football using GPS tracking. *Journal of Science and Medicine in Sport,* 13(5), pp. 531-536.

148. Woods, C., 2016. The use of team performance indicator characteristics to explain ladder position at the conclusion of the Australian Football League home and away season. *International Journal of Performance Analysis in Sport,* 16(3), pp. 837-847.

149. Woods, C., Robertson, S. & French Collier, N., 2016. Evolution of game-play in the Australian Football League from 2001 to 2015. *Journal of Sports Sciences,* 35(19), pp. 1879-1887.

150. Yue, Y. et al., 2014. Learning fine-grained spatial models for dynamic sports play prediction. *2014 IEEE International Conference on Data Mining,* pp. 670-679.

151. Yue, Z., Broich, H., Seifriz, F. & Mester, J., 2008. Mathematical analysis of a soccer game. Part 1: Individual and collective behaviours. *Sutdies in Applied Mathematics,* 121(3), pp. 223-243.