Can we catch the crooks: Examining performance metrics of match-fixing association football players

Submitted by

Susanne Ellens

This thesis is submitted in fulfilment of the requirements for the award of

Masters by Research

Supervisor:	Dr Kevin Ball
Co Supervisor:	Professor Robert Aughey

November 2019

Institute for Health and Sport (IHES)

Victoria University, Melbourne, Australia

Abstract

Match-fixing is the process where the result of a sporting contest or game situation is deliberately manipulated for the personal material gain of one or more parties involved in that activity. Match-fixing is a serious problem in football affecting the integrity of the game. While indicators such as betting patterns have been used to identify match-fixing cases, there are still many that go undetected and even those that are exposed are difficult to prosecute due to a lack of hard evidence. Electronic performance and tracking systems can potentially assist in both identification and evidence- development actions by detecting unusual changes in a players' movement behaviour on the pitch. The purpose of this research was to examine whether performance metrics derived from players' positional x and y coordinates can detect match-fixing behaviour in football. Six different performance metrics have been examined and were used to create player performance profiles. The player performance profiles have been compared with standardized mean differences and were analysed with Approximate entropy (ApEn) analysis and different recursive partitioning techniques. Results show that match-fixing behaviour influenced defensive fixing players' performance metrics during a football game. Positional performance metrics were most associated with fixing behaviour and showed substantial differences compared to normal behaviour. Fixing players moved forward on the pitch and kept more distance towards the position-specific centroid. The altered movement pattern resulted in more spread of play in the lateral direction suggesting fixing players are stretching the defence to create space. Further studies should investigate the use of a wider range of fixing scenarios of numerous games to further develop the match-fixing detection framework. The findings of this thesis can be beneficial, not only for integrity purposes of the football related society, but also for a wider spectrum of team sports using electronic performance and tracking systems to measure player performance. These findings provide insights to player performance metrics underpinning match-fixing behaviour for defence players which can possibly assist in providing supporting evidence to prosecute match-fixing players. Further, it provides scientific knowledge to create a match-fixing detection approach which covers both betting and non-betting related match-fixing.

Student declaration

"I, Susanne Ellens, declare that the Master by Research thesis entitled [title of thesis] is no more than 60,000 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: Susanne Ellens [redacted] Date: 30/11/2019

Acknowledgments

First of all I want to thank my parents. Mam & dad without your love, support and faith in me I would have never even been in Australia! Thanks for always being there for me, even though you are 16.566 km away. Tim & Max, thanks for listening to my presentations through skype (sometimes even at midnight) and providing me with suggestions for my thesis. Also thanks to my whole family who have supported me from far far away.

Next I want to acknowledge my supervising team. Kev and Rob, this journey would have been impossible without you guys! Thanks for taking me on board of your ship and offer me the chance to stay in AUS to do my Masters. I hope we will be sailing for years to come☺. Also thanks to Steph & Darran, you guys helped me get settled and started in Australia, you should be included in everyone’s starting-in-Australia package!

Last but not least, a hand full of people who have assisted me in writing and brain breaks. Mick, thanks for being the krullenbol you are, your love, laziness and irresistible risotto made my thesis much more pleasant (also thanks for correcting a Fariety of my ongoing spelling mistakes). Vinneh, thanks for keeping me of my work with coffee, music and gym breaks, they were much needed. Alex & Rhi, thanks for keeping my suunglish up to date.

Abbreviations

ApEn	Approximate Entropy
EPTS	Electronic Performance and Tracking Systems
ESSA	European Sports Security Association
FIFA	Fédération Internationale de Football Association
GPS	Global Positioning System
HIR	High Intensity Running
ICSS	International Centre for Sports Security
IMFTF	INTERPOL Match-Fixing Task Force
LIR	Low Intensity Running
LPS	Local Positioning System
MIR	Medium Intensity Running
NBA	National Basketball Association
UEFA	Union of European Football Associations
UN	United Nations
VAR	Video assistant referee
VU	Victoria University

Publications & Presentations

Sections of this thesis will be submitted for publication and/or were presented at relevant scientific conference.

SCIENTIFIC CONFERENCE PRESENTATIONS

Ellens, S., Duthie, G., Aughey, R., Robertson, S., Serpiello, F., Varley, M., Evans, N., Ball, K. (2018). Concurrent validity of LPS, GPS and positioning systems to measure covered player area against a validated computer vision model (FIFA EPTS). European College of Sport Science, Dublin, Ireland.

Ellens, S., Aughey, R., Ball, K. (2019). Data-mining in football: can we catch the crooks? Institute for Health and Sport conference, Victoria University, Melbourne, Australia. (People choice award, best presentation).

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Match fixing in sport – considered as a process where an athlete intentionally illegally alters performance to permit others to win (Maennig, 2005) is a serious problem in association football (referred to as football for this thesis). Match-fixing is affecting the integrity of the game and can be traced back to the Olympic games of 388 B.C. (Weeber, 1991). Each year millions of football games are played where it is possible to fix a match for non-betting and betting related purposes (FIFA Communications Division, 2007).

The global interest in match-fixing is rising with current advanced technologies allowing for more depth investigations into the match fixing world (Heilemann, 2014). Match-fixers need to come up with new innovative methods to fix games, which makes the process of recognising and detecting match-fixing even harder.

The different measures to combat match-fixing have been studied from forensic economic and organizational perspectives which focussed mainly on monitoring betting activities to identify betting related match fixing cases, education and awareness training (Reade, 2014; Wolfers, 2006; UEFA, 2018; FIFA, 2018b). To date, no form of integrity system is available that covers both betting and non-betting related match-fixing. As such, there are still many match-fixing cases that go undetected and even those that are exposed are difficult to prosecute due to a lack of hard evidence.

A limitation of the current match fixing detection approaches is the lack of usage of player performance metrics. Player performance metrics are potentially important when looking into betting and non-betting related match-fixing, because players are the ones playing the game and can deliberately influence (fix) it. Player performance metrics allow the analysis of the performance of match-fixing players on field, which is currently unexplored in the literature. Exploring player performance is a crucial element in detecting match-fixing, as player performance deemed fixed needs to deviate from normal performance. There are currently no studies exploring the influence of match-fixing play on player performance metrics, creating a gap in the match-fixing literature.

Electronic performance and tracking systems (EPTS) are employed by clubs and leagues globally providing player performance metrics. The use of EPTS provides the ability to analyse and describe both location and movements performed by players on the pitch and EPTS are therefore commonly used in sports as football, hockey and Australian football (Sampaio & Maçãs, Measuring tactical behaviour in football, 2012; Macutkiewicz & Sunderland, 2011; Hiscock, Dawson, Heasman, & Peeling, 2012). Electronic performance and tracking systems can be used to derive player performance metrics which are representative of the movement behaviour and performance of football players and teams. A match fixing detection framework based on these performance metrics can potentially assist in both identification and evidence- development actions by detecting unusual changes in a player's performance on the pitch.

The aim of this thesis is to develop a match fixing detection framework based on performance metrics derived of an electronic performance tracking system to assist in providing evidence for match-fixing. The database used for this research is part of a study evaluating the validity of EPTS in football conducted by VU and FIFA. In the study, full pitch matches with known fixing scenarios were played by professional and semi-professional players. As such, this database provides a unique opportunity to investigate player performance in fixed games.

This research may be the first to investigate the effect of match-fixing on football players' performance, analysing their performance metrics during a game. The results of these studies will be beneficial, not only for integrity purposes of the football related society, but also for a wider spectrum of team sports who are using electronic performance tracking systems to measure player performance.

Chapter 2: Literature review

Relatively few studies have investigated match-fixing from a player performance perspective (Hill, 2010). In general, these studies are based on the role players have in match fixing (Boniface, et al., 2012; Haberfeld & Sheehan, 2013; Nowy & Breuer, 2017), or on the use of performance enhancing substances by players to fix a match (Gorse & Chadwick, 2011; Paoli & Donati, 2013; Maennig, 2002). Furthermore, different types of match-fixing cases have been researched (Masters, 2015; Maennig, 2005; Bricknell, 2015) and studied the involvement of gambling and betting in match fixing, but none investigated player performance related to match-fixing. Player performance related to match-fixing is hard to identify because it needs to deviate from a player his normal performance. Understanding if a player his performance is indeed related to deliberate different performance or an inherent different performance is important to deviate match-fixers from non-match-fixers. Deviating deliberate and inherent different performances make the process of identifying match-fixing based on player performance even harder and research into player performance related to match-fixing important.

The different strategies to combat match-fixing have been studied from forensic economic and organizational perspectives. Forensic economics uses economic models to forecast match outcome and compare their outcome with those forecasted by betting operators to detect match fixing (Reade & Akie, 2013; Forrest & McHale, 2015; Wolfers, 2006). Big organisations like FIFA and UEFA have different strategies to cope with match fixing focussing mainly on monitoring betting activities, and education and awareness training (UEFA, 2018; FIFA, 2018b). The management of match-fixing in sport clubs is based on the same principles (Sport and Recreation Ministers, 2011). According to the available literature there is currently no analytical tool available to detect match-fixing based on performance.

Section 2.1 of this literature review will focus on defining the problem and scope of match fixing globally. It will first explore the origin of match-fixing followed by its different types and people involved. Sections 2.2 and 2.3 focus on the behaviour that actually constitutes match fixing and reasoning to participate. It is important to know why players participate in match-fixing to be able to combat match fixing, without reasons to participate, match-fixing would not exist. The last part of this section will focus on explaining the process fixers and players go through to conduct a fix. Section 2.4 reviews match fixing detection systems and the main processes used to prevent match-fixing from happening. Section 2.5 will focus on player performance and the different techniques used to assess it, followed by section 2.6 which summarizes the available knowledge on match-fixing and gaps in knowledge to be addressed in this thesis.

2.1 Match-fixing

2.1.1 Introduction to match-fixing

2.1.1.1 Origin of match-fixing

To get an understanding of what match-fixing in sport actually is, the origin of match fixing will first be reviewed. Match-fixing originates from corruption in sport, dating to the Olympic games of 388 B.C. when boxer Eupolos of Thessalia successfully bribed three opponent players (Weeber, 1991). Not only athletes were involved in corruption in sport, records of corruption in sports management are dating to 12 B.C. where the father of an Olympic wrestling athlete bribed the father of the opponent athlete to secure his sons victory (Maennig, 2005). According to Gymnasticus, work of Philostratus of Athens, corruption was not only present in the Olympic Games, trainers and athletes from all over the world became obsessed with profit rather than athletic excellence, they bought and sold victory (Decker, 1995, p. 152).

The way corruption is described depends on the definer's perspective. From an economists' perspective corruption is "the abuse of entrusted power for private gain" (Transparency International, 2018). Whereas from a political perspective it is seen as "behaviour which deviates from the formal duties of a public role because of private-regarding (personal, close family, private clique) pecuniary or status gains" (Nye, 1967). Both of the definitions cover the main scope of corruption as described in the Oxford English Dictionary "Dishonest or fraudulent conduct by those in power" but none of them leave space for sporting contest related corruption. In this thesis a narrower definition will be used to describe corruption in sport. Corruption in sport is considered as ''any illegal, immoral or unethical activity that attempts to deliberately distort the result of a sporting contest for the personal material gain of one or more parties involved in that activity" (Gorse & Chadwick, 2010).

2.1.1.2 Defining match-fixing

Many types of conduct, including match fixing, fall under the above outlined definition of corruption in sport. Match fixing in sports occurs when an athlete intentionally alters performance to permit others to win (Maennig, 2005). A fix can be performed by just one player or by multiple people in the team, usually without the knowledge of teammates. Match fixing includes conduct of fixing match results, spread of points or tanking (withdrawal) (Sport and Recreation Ministers, 2011). Match fixing is not only performed by athletes, it is also a behaviour by sporting officials who perform their tasks with the objectives and moral values of others (e.g. the relevant club) in mind (Maennig, 2005).

Other types of conduct exist that comprehends and encompasses corruption in sport such as doping (use of performance-enhancing substances), inside information (misuse of information for betting purposes) and sabotage (actions by one team disadvantaging the other team) (Gorse & Chadwick, 2011; Graycar, 2015; Svensson, 2005; Preston & Szymanski, 2000), however these are beyond the scope of this thesis.

2.1.1.3 Scope of match-fixing

Match-fixing in sport is part of a big criminal circuit and has been supported by wagered on sport of roughly $2 trillion a year as estimated by the International Centre for Sports Security (ICSS). Seven percent ($140 billion) of this immense amount of money is estimated by ICSS to be laundered by criminal circuits in football. Match-fixing generates big sums of money, yet at the same time the chances of being caught are excessively low (Feltes, 2013). As such, match-fixing is a common way for criminal organisations to fund criminal activities and launder dirty money (Feltes, 2013; FATF, 2009).

Match-fixing is happening all over the world. Between February 2019 and June 2019, 49 countries have been reported for match-fixing in football according to the INTERPOL weekly bulletin (INTERPOL, 2019a). The blue coloured countries in Figure 2-1 represents countries where investigations or sentences for match-fixing have been reported in the media between February 2019 and June 2019. It is difficult to estimate the

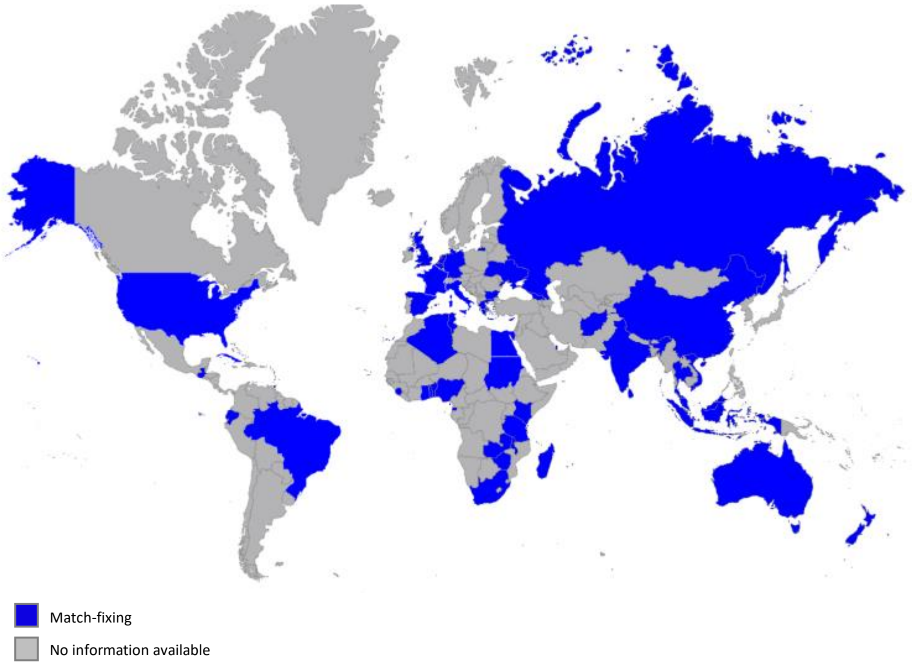

Figure 2-1 World map showing countries that have been reported in the media for match-fixing or illegal betting in football between February 2019 and June 2019.

exact number of countries involved in match fixing, as in most of the match-fixing cases hundreds of people from different countries are involved. Operation VETO of Europol (European Union's Law Enforcement Agency) investigated match-fixing in football and discovered a criminal network covering 425 people from more than 15 different countries and 380 matches suspicious of match-fixing (EUROPOL, 2013).

There are also many cases where match-fixing is not proven, but these matches seem to be fixed. The Champions League group stages in 2011 for example, where Lyon defeated Dinamo Zagreb by 7-1, where a 7-goal victory was needed for Lyon to proceed to the next round at the expense of Ajax (Williamson, 2011). This could seem suspicious but could also be explained by other factors as great mental and physical pressure. This is what makes detecting match-fixing and collecting hard evidence so difficult.

2.1.2 Betting related match fixing

Match-fixing can be divided into two categories, betting related and non-betting related match-fixing (Gorse & Chadwick, 2011). Betting related match fixing is based on the manipulation of results to secure a betting outcome and gain financial profit for those involved. Figure 2-2 is an overview of betting related match-fixing. The figure shows how illegal money is spread by criminal organisations to their collaborators of different countries and how the collaborators make the final step in arranging fixed matches involving teams and referees. The criminal organisations are the big overarching group and the money source, they do not directly get in touch with teams or referees on the field to arrange a fixed match. Collaborators in different countries are needed to act as intermediary in arranging fixed matches between the criminal organisations and teams and referees. The collaborators get the orders and money of the criminal organisations to arrange fixed matches and direct the orders and money towards the teams and referees to settle the fix.

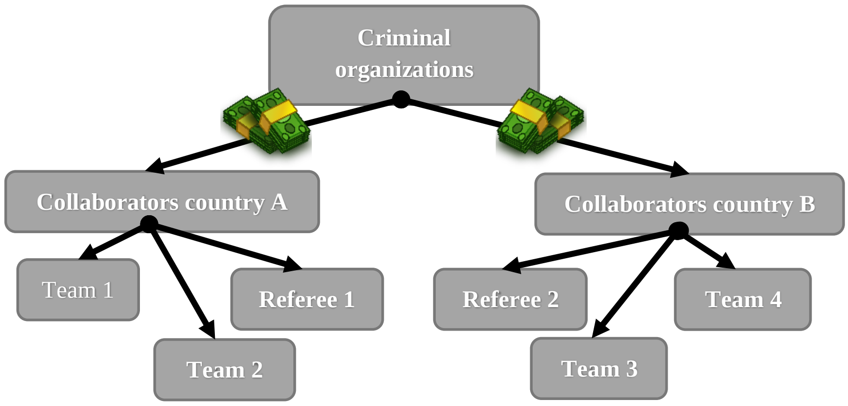

Figure 2-2 Overview of betting related match-fixing.

2.1.2.1 Betting related match fixing involving teams

The match fixing scandal in 2013 involving the Southern Stars (Victorian Premier League football team) is an example of betting related match fixing. Six individuals were involved in this match-fixing incident, including four players, the coach and the collaborating match fixer Mr Segaran Gsubramaniam (Nettleton & Chong, 2013). Mr Gsubramaniam is said to be the collaborator between match-fixers in Hungary and Malaysia and the Southern Stars football team. Mr Gsubramaniam fixed matches of the Southern Stars by instructing the participating players to fix the match in favour of his match-fixers betting outcome. A player of the Southern Stars not involved in the match-fixing incident stated that his teammates' performance spoke abundance and displayed lack of intensity (Kerr, 2016). This is an example indicating that match-fixing performance can deviate from normal performance.

Another betting related match-fixing example involving teams is the 2011 Scommessopoli scandal in Italy (Forrest, 2012; Reade, Detecting corruption in football, 2014) which involved the professional second, third and fourth divisions (Serie B, Lega Pro Prima Divisione and Lega Pro Seconda Divisione). The scandal involved illegal criminal organisations fixing betting results of relatively minor games and Coppa Italia

(National Cup) matches for betting purposes. The fixed matches were usually at the end of the season (once the winner of the competition was decided) to reduce the risk of being discovered.

2.1.2.2 Betting related match fixing involving referees

The previous examples do not involve referees to fix a match, but betting related match fixing can also involve referees as seen in the friendly match of Malaysia versus Syria in 2009 (Perumal, Righi, & Piano, 2014). The match was fixed by Wilson Raj Perumal by arranging and instructing a Kenyan referee to fix the match by three goals. The referee indeed officiated the match by assigning three penalties to Malaysia. Malaysia won the match by 4-1 and the fix was set.

2.1.3 Non-betting related match-fixing

Non-betting related match fixing is based on the manipulation of results for sporting related matters, such as a league or tournament victory over a rival (Boeri & Severgnini, 2013). Figure 2-3 is an overview of non-betting related match-fixing. The figure shows the interaction between teams and how the pressure and probability of winning a match can influence the fixing of match results. If the pressure of a match for two teams is different, e.g. team A is high, team A needs to win to reach a certain position in their competition, and the match pressure for team B is low, not at stake of losing or winning anything, the chances of a match-fix happening are high. The teams can weigh their probability of winning a match against each other and can end up deciding to fix the match to secure the needed win for team A. The main reason to fix these matches is to allow a team to reach a certain position in their competition or to win a championship (e.g. promotion to a higher league, avoiding relegation, qualification for the Champions League).

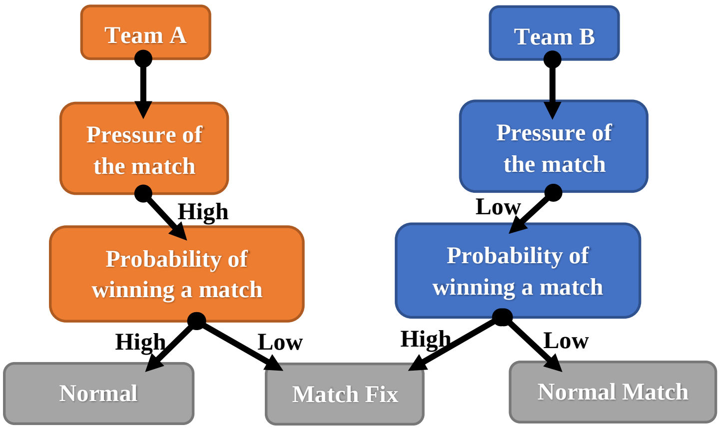

Figure 2-3 overview of non-betting related match-fixing

2.1.3.1 Non-betting related match-fixing performed by teams

Non-betting related match-fixing can be performed in several ways. The fix can be performed by teams and referees. A fix performed by a team is seen in the alleged non-betting related match-fixing Australian football match between Melbourne and Richmond in round 18, 2009 (Ralph, 2011). Melbourne was in the position to receive a priority draft pick if they lost the match against Richmond. A priority draft pick is an additional draft pick assigned to the poorest performing teams to improve future on-field performance. The coach of Melbourne deemed underperforming of his team to receive a priority draft pick more beneficial than winning. Melbourne used different tactics than usual during the game, such as moving the key defenders into the forward-line, to make sure Richmond won the match. Richmond eventually won the match by a four-point difference.

2.1.3.2 Non-betting related match-fixing performed by referees

The non-betting related fix can also be performed by referees as seen in the Calciopoli scandal in 2006 in Italy (Boeri & Severgnini, 2013). Calciopoli involved the professional first and second divisions (Serie A and Serie B). The large match fixing system was brought to light by tapping phone conversations for a doping investigation in the first division. The scandal involved selecting favourable referees by club managers of teams such as A.C. Milan, Juventus and Lazio. The club managers succeeded in selecting favourable referees by exerting pressure on officials of the football federation and on referees. The scandal resulted in 78 fixed matches out of 380 matches in the 2004-2005 Italian Championship.

As match-fixing is a prevalent problem all over the world it makes research into these illegal, immoral and unethical activities in football important to keep the integrity of the game and the spirit of fair play.

2.2 Why participate in match-fixing?

There is a lack of research in football investigating the behaviours that contribute to players being involved in match fixing (Gorse, Chadwick, & Byers, 2014). The following section reviews this limited body of literature.

2.2.1 Financial gain

There are a variety of reasons players are motivated to fix matches. A major reason to participate in betting related match-fixing is financial gain, as the salaries of the majority of the football players are fairly low (Reade, 2014), and the risk of being caught is very low due to legal loopholes or lack of evidence. Nonetheless, there are also many examples of well-paid footballers being involved in match-fixing (Hill, 2010). The low risk of being caught versus the high chances of making a big financial profit thanks to the possibility to bet on all outcomes possible, makes match-fixing attractive for players and organised crime (Haberfeld & Sheehan, 2013).

Another reason for players to participate in match-fixing is threat of the criminal organizations (Forrest, 2013). Criminal organisations will force players to fix matches by threatening them to harm family/close friends. Furthermore, once the players are involved in match-fixing the criminal organisations will force players to continue fixing games by threatening them to reveal their earlier match fixing transgressions. So once involved in match fixing, it's hard to escape. Veli Sezgin, chairman of Turkish football club Akçaabat Sebatspor, tried to thwart one of these criminal organisations by fitting his goal keeper a tape recorder when he heard about attempts to bribe his team (Hill, 2010). Veli successfully recorded a match fixing attempt and the fix failed. However, Veli got shot four months later. Threats from criminal organisations can occur when football players do not make the fix they promised towards their criminal organisation and get in debt, they owe the criminal organisation money for not fulfilling their fix. A scenario similar to this happened in 2011 in Italy where a player, Marco Paolini, of Serie B side Cremonese tried to pay off a big amount of gambling debts by organising defeats for his bookies (Richardson, 2012). But his team unexpectedly won a lot of matches which forced him to take devastating measures, Paolini drugged his teammates with sleeping pills to decrease the performance of his team and fulfil the arrangement with his bookies.

2.2.2 Ladder positioning

Another reason for players to participate in betting and non-betting related match-fixing is to reach a certain position in a competition or tournament ladder. Teams participate in fixing the outcome of the ladder to obtain competitive advantage of a different position to for example avoid relegation, avoid an opponent player or for qualification purposes.

Fixing a match to avoid relegation is likely to happen by a team in the bottom of their competition ladder in the end of the competition, where a match can financially be meaningless for their opponent. This to reach a certain position in the ladder to prevent relegation to a lower league in combination with financial reasons. For example, the difference in revenue for a team playing in the 1st or 2nd division in England is estimated to be £60 million (Hill, 2010, pp. 116-117), making the financial pressure of relegation high. The end of the season in the third lowest division in the Czech Republic 2007 involved fixing a match to avoid relegation (Numerato, 2016). Two teams fixed a match whereof one was positioned in the middle of the ladder and the other team in the bottom of the ladder and at stake of relegation. The coach of the team at stake of relegation made sure his team would win the game by making an arrangement (fix) with the team in the middle of the ladder. On game day, the team in the middle of the ladder showed up without three key players and indeed lost the game.

Fixing a match to avoid an opponent player can happened in a tournament. A tournament can be structured in a way that the rank of a team in the group stages determines their opponent in the next round. If a team has already qualified it may take other factors into account when playing the last match, such as the opponent they will face next (Chater, Arrondel, Gayant, & Laslier, 2018). The team can play according to their desired rank in the ladder which results in the desired opponent (Figure 2-4).

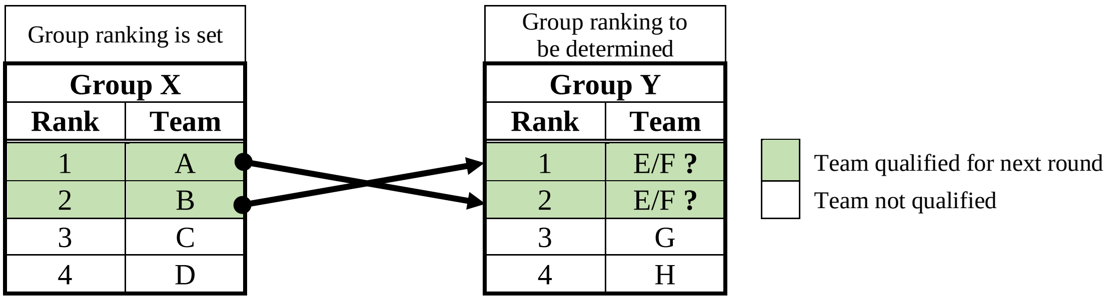

Figure 2-4: The group ranking of group X is set whereas the group ranking of group Y still needs to be determined by a last match. Team A (nb.1 of group X) opponent player for the next round will be the number 2 of Group Y, team B (nb.2 of group X) opponent will be the number 1 of Group Y. As the ranking in group Y still needs to be determined, the teams can play their last match according to a match-result (lose, win or draw) resulting in their desired rank in the ladder and opponent team (where possible).

The women’s football tournament at the 2012 Olympics involved match-fixing of the Japanese team to reach a certain position in the ladder (Rogers, 2012), this to avoid 480 kilometres of travel. World Champion Japan intentionally drew its final group game in Cardiff against the weakest team in the competition. Japan intentionally drew to position themselves second in their group to avoid 480 kilometres of travel to Glasgow to play the quarter final and instead remain in Cardiff.

The matches to reach a certain position in the competition ladder are a high potential for betting and non-betting related match-fixing for the teams not in the run to reach a certain position. Teams that have already qualified for a tournament or of which the match is financially meaningless, meaning they are not in the run for the 'big prize money' or at stake of relegation (Reade, 2014), can deem underperforming more beneficial than winning and can make money by fixing the game for betting purposes. Teams can also make a non-betting related match-fix, situations like this happen where for example a fix for a draw will qualify both teams for the next round of the competition or tournament (Chater, Arrondel, Gayant, & Laslier, 2018). Another situation for a non-betting related match-fix can be the elimination of a disliked team at stake of relegation by making an opponent also at stake of relegation win (Numerato, 2016).

2.3 How is a "fix" performed?

There are different ways a fix can be performed. This section reviews the process fixers and players go through to make a successful fix happen based on reports of match fixing cases.

The start of every fix for the match fixers is to get in touch with the players (including referees). It is challenging for match fixers to find potential targets as players live and work in a close environment where unusual people and behaviour do not go unseen. The process fixers need to get through without attracting attention to the player or creating suspicion is something like seduction (Boles & Garbin, 1984). Fixers will try to get close to players by trying to become friends with them, use prostitutes, use the same hotel as the player or be involved in their club at an official position. Another method match-fixers use is the use of agents, better known as "runners" (Hill, 2010; Reade, Detecting corruption in football, 2014). These runners try to become friends with a player, figure out if the player is a target for bribes and get the player in touch with the match fixers.

The next stage is the way they approach the player, just getting in touch with the player does not immediately result in a successful fix. The fixer needs to be mindful in what he says and how he acts to each player as every person is different. After approaching the player(s) it is time for the fixer to let the player(s) know what the fix needs to be. If the match fixer wants to bet for match result, the fixer needs to wait to see what the odds are doing as most of the bets being placed starting from 2 hours before the game. If the odds are for his chosen team to win, the fixer will get the best profit when the team loses. In the short period of time before the game starts, the fixer now needs to communicate his desired fix/match result with his players without attracting attention to him nor his players and make sure the players understood the required result. If that is all set, it is now up to the players to deliver the result.

Sometimes a fixer can make a bigger profit by betting on performance indicator sub-categories as team covering most distance, number of corners or player to shoot at the crossbar (Hill, 2010, pp. 32-33), this is better known as spot-fixing. To accomplish a successful fix of sub-categories a fixer does not only ask the players to lose/win a game but change the way they play. Spot-fixing is a phenomenon also present in different sports as cricket where they concede runs in a specific over (Cairns, 2013) or tennis where they throw a certain set (Schnitzer, 2017).

There are many different tactics/ways players use to deliver the required fix result. Playing with less effort or playing differently is playing behaviour that can easily be performed by players and is a big part of the match-fixing strategies. There are certain common ways players fix a game. Every single player (from goalkeeper to forwards) on the pitch can be involved in match-fixing, see Figure 2-5A for an overview of the player positions. According to Rafiq Saad's confession from the Royal Malaysian Police files (Hill, 2010) these are common fix strategies for players: ***referees:*** giving away penalties. ***Goalkeepers*****:** Leave the area as often as possible so the goal is clear. ''Goalkeeper drops the ball: he could catch it, but he just pats it away.'' ***Defence:*** Left and right back will not assist the sweeper when he is being attacked, see Figure 2-5B, and vice versa → defence will not play all out and allow the attackers to get by. According to a player ‘‘playing a stupid offside was the best tactic: You go charging up the field to play an offside and the forward goes charging through’’. The Suicide Pass (Borristow, Bernard, Pendlebury, & Greaves, 1960), used by goalies and defenders: ‘‘The ball is placed by a defender/goalie too far away for the goalkeeper/defender to clear it or gather it, but near enough to the opposing forward for him to nip in and score a gift goal.’’ ***Forward:*** Keep the ball as long as possible or move the ball straight to the opponent to allow them to take away the ball. Miss goal opportunities by kicking straight to the keeper or completely miss the goal. These are all factors that need to be taken into consideration when looking for match-fixing.

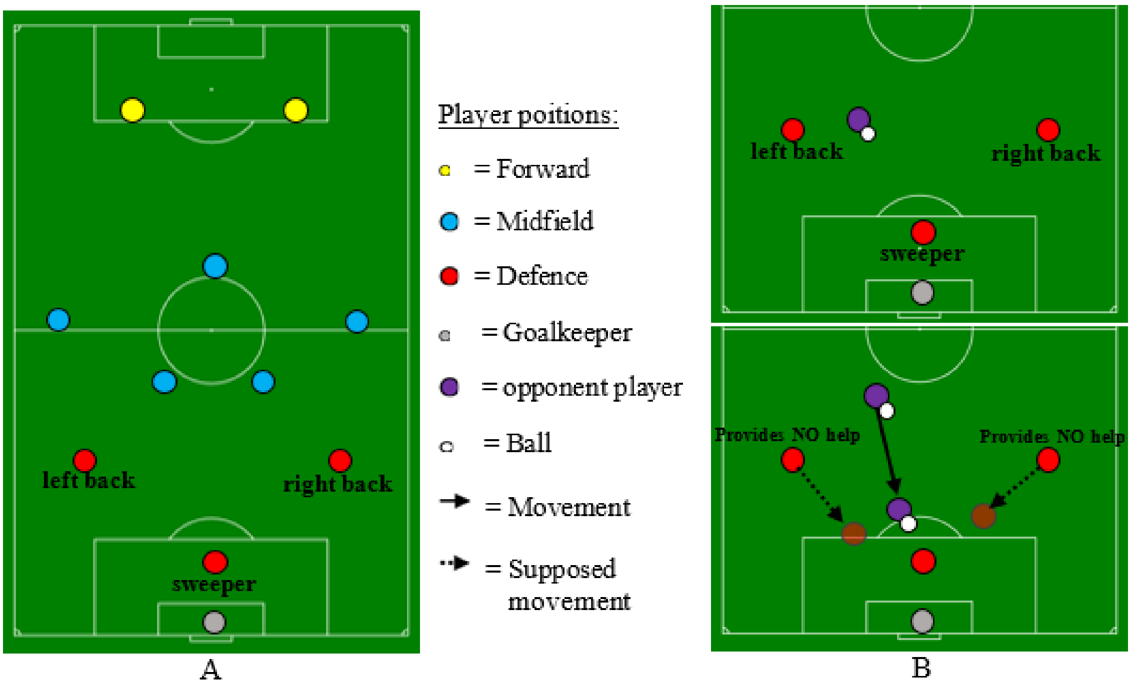

Figure 2-5 A) Overview of player positions of one team on the pitch. Note this is a 3-5-2 line-up, team formations can be different as there is a lot of variety possible. B) Fixing scenario, left and right back won’t assist sweeper when he’s under attack by an opponent player.

After the fix took place and the desired result was delivered by the players, it’s time for fixers to pay their players. The payment is a tricky part, players are mostly well-known figures whose life is closely monitored by media and the football world. The payment needs to be done without attracting attention to the player or creating suspicion. A Malaysian reporter, Johnson Fernandez, described in a report that match fixers bought winning lottery tickets for their match fixing players so that the players got money that seemed legit (Fernandez J. , 1993). After this has been done, the fix is complete and it is time to plan the next fix.

2.4 Current processes to combat match-fixing

The world of football has different processes to combat match-fixing. Deterring and dealing with match-fixing in football is complex and involves highly efficient networks for information sharing between lots of different operators such as law enforcers, governments, betting operators and sporting organizations. In this section the combatting processes are reviewed.

2.4.1 Fraud detection systems

Forensic economics is evolving in the sports world, using economic models to detect corruption in sport (Wolfers, 2006). As most of the match-fixing cases are connected to gambling or betting (Wen-Bin Lin & Mei-Yen Chen, 2015; Gorse & Chadwick, 2011), economic models in combination with bookmakers data can be used to predict match outcomes and detect possible fixed matches. Figure 2-6 is an overview of a fraud detection system. The economic model, M_e, predict match outcome based on available relevant sports information such as strength of the teams and latest match outcomes (Reade, 2014). The forecasted match outcome of the economic model is ought to be the same as that of bookmakers, M_b, so that $M_e = M_b$ based on the assertion that both models are time synced and that most relevant information is used by both bookmakers and economic models in forecasting match outcome (Reade & Akie, Using Forecasting to Detect Corruption in International Football, 2013). If corrupt activity is present in a match to fix the match result, the match outcome will be altered, M_{fix}. Bookmaker odds will follow the exchange prices which will be much higher for the arranged fix (Hanson & Oprea, 2009). This results in deviating odds between the economic model (which does not include exchange information) and bookmaker odds $M_e \neq M_b + M_{fix}$. As such, deviating odds between the economic model and bookmakers is the criterion for an alert to report irregular (and potentially suspicious) activity in betting markets.

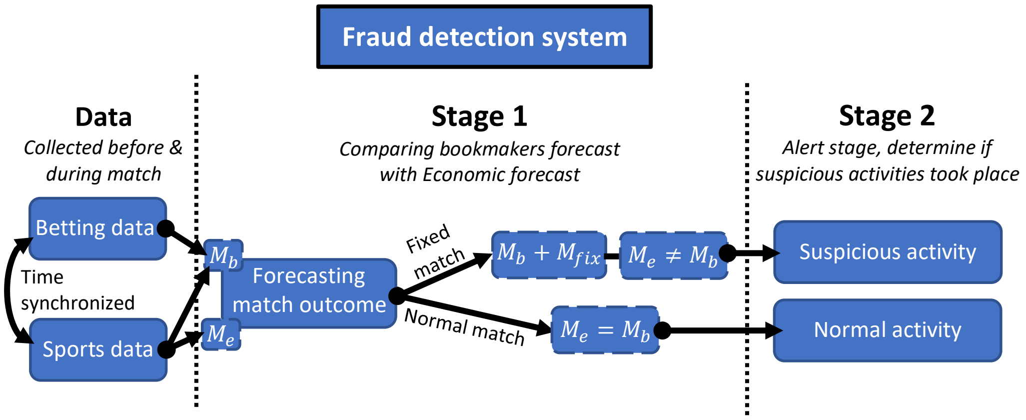

Figure 2-6: Overview of a fraud detection system. M_e is the economic forecasting model based on sports data. M_b is the bookmakers forecasting model based on sports data and betting data. M_{fix} is the influence of a fixed match on forecasting models.

Lots of research into match-fixing in football involves investigating economic models to forecast match outcome (Reade & Akie, 2013; Hvattum & Arntzen, 2010; Reade, 2014), which are also used in other sports as basketball (Wolfers, 2006) and tennis (Rodenberg & Feustel, 2014). The economic models used were variations of Poisson distributions and regression models. For further details into the economic models see appendix A, an overview of used economic models to forecast match results and goals scored.

Sporting organizations are working closely with international betting integrity monitoring agents who use fraud detection systems, as described above, to detect suspicious betting patterns. An example of a successful match fixing alert of a fraud detection system is the 2013 match fixing case in the Victorian Premier League, Australia (My Football, 2013). Football Federation Australia (FFA) worked together with betting integrity monitoring agents who released an alert for the presence of suspicious betting patterns in the

Victorian Premier League. An investigation followed which resulted in the arrest of ten people who were alleged to have been involved in the manipulation of match results.

Fraud detection systems are of great value for combatting betting related match-fixing. But fraud detection systems have its limitations, it cannot detect non-betting related match fixing and the monitoring of betting is limited by the covered betting operators.

2.4.2 Overarching organizations

The cooperation of organizations is key in the fight against match-fixing. As in most of the match-fixing cases hundreds of people from different countries are involved (as explained in section 2.1), information sharing between the organizations from all over the world is important to build a tactical information platform, permitting countries to anticipate to threats and taking measures for an extensive and unified approach to combat match-fixing. The global INTERPOL Match-Fixing Task Force (IMFTF), fight corruption and match-fixing in sport by working together with law enforcement agencies around the world for sharing information and investigation (INTERPOL, 2019a). To date 84 countries are members of the IMFTF, sharing a global network of investigators who share information (INTERPOL, 2019b). INTERPOL also offers global training and awareness with a focus on match manipulation and irregular and illegal betting. An ongoing operation executed by INTERPOL codenamed SOGA (Soccer Gambing) aimed to identify criminal organisations behind illegal football betting, resulted in more than 12,500 arrests between 2007-2016 and closure of 3,400 illegal gambling places which handled bets worth around $6.4 billion (INTERPOL, 2016). SOGA involved regional support units, officers from INTERPOL National Central Bureaus and specialist units from across the region who all worked together with global information sharing and resulted in the success of the operation.

The governing body of world football, FIFA (Fédération Internationale de Football Association), and the administrative body of European football, UEFA (Union of European Football Associations), work together with INTERPOL to fight against match-fixing, but also have their own approaches. Both organisations have a zero-tolerance policy towards match-fixing and their own match-fixing reporting platform (FIFA, 2018b; UEFA, 2018). UEFA and FIFA also monitor betting activities (as explained in paragraph 2.4.1), both work together with Sportradar Integrity Services to identify match-fixing attempts with monitoring suspicious betting activity around the world (FIFA, 2017b; UEFA, 2019). UEFA also joined forces with European Union's law enforcement agency (Europol) in 2014 (EUROPOL, 2014) and recently signed an information sharing agreement with European Sports Security Association (ESSA) (UEFA, 2018), both to strengthen their campaign to combat match-fixing in European football. Both FIFA and UEFA their main focus is to raise awareness in the whole football world, education is key in this element. FIFA and UEFA have an education program (UEFA, 2018; FIFA, 2018b) to bring match-fixing to the attention of players, coaches and referees, so people involved in football can recognise the problem earlier, resist against it and report it.

2.4.3 On field technology

On field technology in football can also make a difference in match-fixing and help prevent corrupt behaviour. Goal-line technology and VAR (video assistant referee) (FIFA, 2014; FIFA, 2017a) make the execution of corrupt behaviour more difficult. The goal line technology determines whether the complete ball has crossed the goal line and was 100% a goal. The goal line technology prevents corrupt referees from giving away a faulty goal. Video assistant referee is used during a game to constantly check for obvious and clear errors related to four match changing situations. These match changing situations are goals and offences leading to a goal, penalty decisions and offences leading to a penalty, direct red card incidents and mistaken identity of a player. The VAR team communicates with the on-field referee when a clear mistake has happened in one of the four match changing situations (FIFA, 2017a), which makes match-fixing based on decisions more difficult for corrupt referees.

Several different measures have been taken to prevent, detect and analyse match fixing in football, which makes the probability of detecting match fixing higher and corrupt activity less likely (Reade, 2014). But to date, there is no form of integrity system available to cover both betting and non-betting related match-fixing (Gorse & Chadwick, 2011).

2.5 Player performance

Match analyses can be easily undertaken with currently available technologies, such as electronic performance and tracking systems (EPTS), which allow the measurement of player performance. Such metrics can also potentially be implied in an integrity system to cover both betting and non-betting related match-fixing.

Player performance metrics are potentially important when looking into match-fixing, as players are the ones that actually play the game and can influence it. Electronic performance and tracking systems is a technology employed by clubs and leagues globally providing player performance metrics. The use of EPTS provides the ability to analyse and describe both location and movements performed by players on the pitch. Player performance metrics can identify the physiological and tactical demands of a sport and create an understanding of the demands of position specific roles. Such metrics can assist coaches in identifying good and bad performance of players or teams based on the identified physiological and tactical demands. In this section EPTS and derived player performance metrics describing the movement behaviour of football players and teams will be reviewed.

2.5.1 Electronic performance and tracking systems

Electronic performance tracking systems primarily track the position of a player but can also be used in combination with other microelectronics (e.g. gyroscopes, accelerometers) (football-technology FIFA, 2018). The three commonly used types of EPTS in football for assessing player performance are global positioning systems (GPS), local positioning systems (LPS) and optical positioning systems (Buchheit, et al., 2014; Varley, Fairweather, & Aughey, 2012; Frencken, Lemmink, & Delleman, 2010).

Global positioning system is a space-based radionavigation system (U.S. Government, 2008). Radio signals are broadcasted by GPS satellites from space with their location, status and precise time. The radio signals are received on earth by a GPS device which calculates its distance from the satellite based on time of arrival of the broadcasted radio signal. When a GPS device knows its distance from at least four satellites it can geometrically determine its own location on earth in three dimensions. The use of a GPS device in football does not require any further material or people. A downside of the use of GPS in football is the need for a direct connection between GPS device and satellite. If the GPS device is not able to receive radio signals from the satellite due to any obstruction, it loses connection and information. As football matches are mostly played inside a stadium, the high concrete stands and, if present, the roof of the stadium interfere with the direct connection between satellites and GPS device, which make the use of a GPS device difficult.

Local positioning system is a radio-frequency system using a locally deployed infrastructure (Sathyan, Shuttleworth, Hedley, & Davids, 2012). The working principles of LPS is similar to that of GPS, the difference is that radio signals of LPS are being broadcasted from locally deployed devices at know locations to measure the range to the LPS receiver. As LPS uses locally deployed devices rather than satellites used by GPS, it does not have the obstruction problem interfering with the direct connection between LPS device and receiver and can be used indoor or in stadia without trouble.

Optical positioning systems are a camera-based (Gozse, 2015). Image processing algorithms are used on captured videos to identify the position of objects and players based on their particular features. Optical positioning system is based on a locally (like LPS) placed camera and can be used indoor or in stadia without trouble. A downside of optical positioning systems are tracking occlusions which happen if the optical position system fails to identify an object or player, resulting in data loss.

The agreement between these different systems have been evaluated by different studies (Buchheit, et al., 2014; Linke, Link, & Lames, 2018; Harley, Lovell, Barnes, Portas, & Weston, 2011; Randers, et al., 2010; Siegle, Stevens, & Lames, 2013). These studies showed that LPS demonstrated the highest position accuracy followed by optical positioning systems and GPS. It showed that position accuracy had the biggest error with high running speeds in combination with fast changes of direction. Speed and acceleration errors were the lowest in LPS followed by GPS and optical positioning systems, where the error increased when speed increased. Overall, the studies reported large between system differences where the magnitude was related to the variable of interest, implying that comparing of results using different systems should be done with caution.

2.5.2 Performance metrics

Performance metrics are derived from EPTS tracked position data and map the player and team performance to single values to describe the movement behaviour of football players and teams (Memmert, Lemmink, & Sampaio, 2017; Low, et al., 2019). The use of different physiological and tactical performance metrics have been reviewed (Table 2-1) these studies allowed the development of different performance profiles according to playing positions.

Table 2-1: Performance metrics for performance analysis based on position data

Performance metric	Method	Description	Reference
Space control	Voronoi	Use of Voronoi diagrams to model space control	(Taki & Hasegawa, 2000; Fujimura & Sugihara, 2005; Fonseca, Milho, Travassos, & Araújo, 2012)
	Heatmap	Use of heat maps to model movement behaviour in space	(Brooks, Kerr, & Guttag, Using Machine Learning to Draw Inferences from Pass Location Data in Soccer, 2016; Machado, et al., 2017)
Distance from team centroid	Euclidian metrics	Average distance of a player towards his team centroid	(Frencken, Lemmink, Delleman, & Visscher, 2011; Sampaio & Maçãs, Measuring tactical behaviour in football, 2012; Gonçalves, Figueira, Maçãs, & Sampaio, 2013; Bourbousson, Sève, & McGarry, 2010; Romero Clavijo, Corrêa, & Trindade Pinheiro Menuchi, 2017)
Team formation	Distance	Average dispersion of a team in the width and length of a pitch	(Folgado, Lemmink, Frencken, & Sampaio, 2014; Castellano, Álvarez, Figueira, Coutinho, & Sampaio, 2013; Bartlett, Button, Robins, Dutt-Mazumder, & Kennedy, 2012)
Speed profile	Distance over time	Time and distance of a player measured in different speed zones	(Mohr, Krustrup, & Bangsbo, 2003; Suarez-Arrones, et al., 2014; Di Salvo, et al., Performance Characteristics According to Playing Position in Elite Soccer, 2007; Clemente, Couceiro, Martins, Ivanova, & Mendes, 2013)

Voronoi diagrams are used to analyse the dominant region of players. The dominant region is seen as the area on the pitch that a player can reach faster than any other player on the pitch under the conditions of same speed and reaction (Voronoi, 1907). Creating a dominant region and delimiting this space of other players is important in football for shooting purposes as players who cluster together are less likely to score (Kapidžić, Mejremić, Bilalić, & Bečirović, 2010). Next to the fact that space control is important for shooting purposes, it is also one of the main skills in football (Kannekens, Elferink-

Gemser, & Visscher, 2011). Voronoi diagrams are used in football to model this space control of players to evaluate teamwork (Taki & Hasegawa, 2000; Fujimura & Sugihara, 2005).

Heatmaps are used to review how a match develops from a player position perspective (Machado, et al., 2017). A heatmap shows where the players are concentrated during the match and can show preferred sides of the pitch at specific times of the match. The ability to map player position offers an insight into the preferred movement patterns of players and creates a fingerprint of a team's behaviour that can be used to identify a specific team with 87% accuracy (Brooks, Kerr, & Guttag, Using Machine Learning to Draw Inferences from Pass Location Data in Soccer, 2016). As such, heatmaps are a valuable tool to represent a team's/player's movement behaviour.

The use of distance between players and the team centroid (team centroid is the average position of all players excluding goalkeeper) provide valuable tactical behaviour of players in football (Frencken, Poel, Visscher, & Lemmink, 2012; Bartlett, Button, Robins, Dutt-Mazumder, & Kennedy, 2012) and the intra-team coordination (Sampaio & Maçãs, Measuring tactical behaviour in football, 2012). It describes the interaction between players that creates team behaviour (Frencken, Lemmink, Delleman, & Visscher, 2011). Players present more regularity in movement variations with their own position (defender/midfielder/forward) specific centroid (Gonçalves, Figueira, Maçãs, & Sampaio, 2013). The position specific centroid is the average position of players of a specific position. Player distance to position specific centroid of central defenders and midfielders are highly predictable (Memmert, Lemmink, & Sampaio, 2017; Gonçalves, Figueira, Maçãs, & Sampaio, 2013), indicating that these players positioning on the pitch are highly predictable. The predictability of player distance to position specific centroid of forwards tends to be weaker, which is explained by the need of forwards to be less predictable when playing.

The team dispersion is a metric which can be expressed by the stretch index and length and width of a team (Bourbousson, Sève, & McGarry, 2010; Frencken, Lemmink, Delleman, & Visscher, 2011). The stretch index is the average distance of all players from the team centroid. The length and width of the team is the distance between the two players furthest apart along the width and length of the pitch. Regularity in the values of length and width of a team have shown that tactical behaviour during a match tends to be repeated (Castellano, Álvarez, Figueira, Coutinho, & Sampaio, 2013). The dispersion of a team is used for tactical performance of team formation and indicates how a team is using the space by stretching and compressing the team (Folgado, Lemmink, Frencken, & Sampaio, 2014).

Speed profiles are used in football to quantify physical demands of the game (Stagno, Thatcher, & Van Someren, 2007). The playing position (defence, midfield, forward) is a key factor in understanding the physical profile of a player (Di Salvo, et al., Performance Characteristics According to Playing Position in Elite Soccer, 2007) and allowed the development of different speed profiles according to the playing positions. A speed profile describes the distance covered/time spend in different speed zones. For example, central defenders cover the least distance during a match and demonstrate the most time spend in low-intensity running, while midfielders cover the largest distance and least time spend in low-intensity running compared to other positions (Clemente, Couceiro, Martins, Ivanova, & Mendes, 2013; Mohr, Krustrup, & Bangsbo, 2003; Di Salvo, et al., Performance Characteristics According to Playing Position in Elite Soccer, 2007). The greatest distance covered by midfielders can be explained because they act as a link between defence and forwards (Reilly & Thomas, 1976). Greater distance is covered and more time is spend in high intensity running and sprinting in the first than in the second half (Mohr, Krustrup, & Bangsbo, 2003; Reilly & Thomas, 1976), high intensity running is also less in the last 15min of a game due to fatigue (Mohr, Krustrup, & Bangsbo, 2003). Position specific speed profiles help to understand a player's specific performance and quantify physical demands of the game.

Nonlinear analysis is needed when analysing performance metrics from an approach of system complexity (Low, et al., 2019). Player interactions in football are complex and player position data are often noisy and non-linear. Complex situations are characterised by non-linear analysis by taking into account the time series, thus dynamic nature, of player's interactions which is not possible by linear analysis (Stergiou, Buzzi, Kurz, & Heidel, 2004). Non-linear analysis in football studied predictability and synchronisation of performance metrics (Gonçalves, Figueira, Maçãs, & Sampaio, 2013; Memmert, Lemmink, & Sampaio, 2017). Approximate Entropy (ApEn) is commonly used to quantify the regularity (predictability) of a performance metric (Pincus, 1991). The regularity expresses the systems complexity, lower ApEn values have a higher regularity and thus a lower complexity. A down side of the use of ApEn, between-study comparisons are difficult to perform, ApEn can be influenced by the use of different data sampling frequencies, vector lengths and tolerance levels. Predictability of performance metrics compare data with itself, while synchronization of performance metrics measure similarity between data. Relative phase is one of the methods used to quantify similarity (synchronisation) between data where values close to zero refer to data in synchronisation (Folgado, Duarte, Fernandes, & Sampaio, 2014). Little is still known about the most appropriate method for analysing synchronisation, more research is needed to analyse advantages and disadvantages of the different methods.

Research into performance metrics of fixing players in a fixed match has not been conducted, it can potentially generate insights if a match-fixing player's movement behaviour differs from its normal movement behaviour and assist in both identification and evidence- development actions by detecting unusual changes in a player's performance on the pitch.

2.6 Summary of the literature review

To summarise, match-fixing is a large problem in football affecting the sport on every level. Different measures have been taken to prevent match fixing from occurring, but to date, no measure is fairly effective and there is no form of integrity system available that covers both betting and non-betting related match-fixing. The fact that no measure is fairly effective in detecting and observing match fixing makes the chances of being caught very low due to lack of evidence. The low chances of being caught due to lack of evidence plus the high chances of making a big amount of money is reasoning for players to participate in match-fixing.

Player performance metrics are potentially important when looking into betting and non-betting related match-fixing, as players are the ones that actually play the game and can influence it. Player performance metrics are derived from electronic performance tracking systems (EPTS) which provide the ability to analyse and describe both location and movement behaviour performed by players on the pitch. Electronic performance tracking systems can be used to derive player performance metrics which are representative of the movement behaviour and performance of football players and teams. Player performance metrics give the opportunity to examine player performance representative of "normal" and suspicious "match-fixing" performance, which is a crucial element in detecting match-fixing, as player performance deemed fixed needs to deviate from normal performance. Player performance metrics can potentially generate insights if the movement behaviour of a match-fixing player differs from its normal movement behaviour.

An integrity system for the detection of unusual changes in a player's performance on the pitch based on performance metrics, can assist in providing evidence for match-fixing and potentially fill the current gap of lack of evidence against match fixing and coverage of both betting and non-betting related match-fixing.

Chapter 3: Aims

3.1 General AIM

The objective of this thesis was to determine if performance metrics derived from players' positional x and y coordinates can detect match-fixing behaviour in football.

3.2 Specific AIMS

1) To determine if physical performance metrics differ between normal and fixing play.
   - Do speed metrics differ?
   - Do distance metrics differ?
2) To determine if positional performance metrics differ between normal and fixing play.
   - Do centroid based metrics differ?
   - Does the Voronoi area metric differ?
   - Do heatmap metrics differ?
3) Determine if an integrity system can be developed based on performance metrics for the detection of potential match-fixing activity.

Chapter 4: Methods

4.1 Procedures

The database used for this research is part of a study evaluating the validity of EPTS in football conducted by Victoria University and FIFA. The database consisted of position data (x, y) from male sub-elite football players across nineteen competitive matches on an official-sized pitch (105 x 68 m) with official football rules. For matches to be included in this research it had to fulfil five criteria: 1) The match was played with match-fixing scenarios; 2) The teams had 11 players each side; 3) Teams did not change formation during a game, this will ensure results do not differ based on formation changes; 4) Match-fixing players had the same positional role so between player analyses were possible; 5) Match-fixing players played at least 40 minutes each half. This resulted in 2 matches being included for this research. A complete overview of the flow diagram of criteria for a match to be included in this research is presented in Figure 4-1. The two included matches consisted of a total of four external-defending match-fixing players, two in each match, where each match-fixing player has legally been instructed to fix certain minutes of the match without the knowledge of other teammates. Players with the same positional role as the match-fixing players have been used as control group.

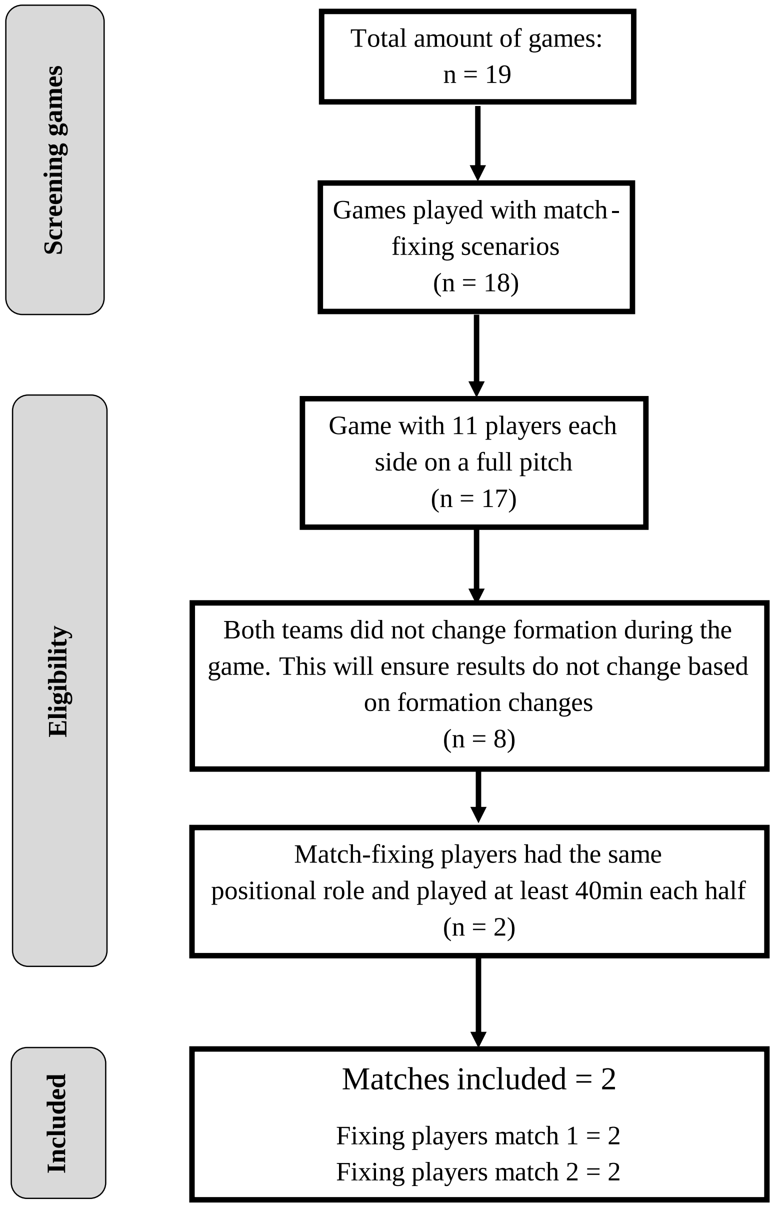

Figure 4-1 Flow diagram of match selection process

Different LPS systems for the two matches were used to collect player position data, X/Y coordinates in its raw format, both at a sample frequency of 20Hz for each player. The player position data was used in subsequent processing. This was due to the larger project whereby different systems were tested against a reference system. Due to a non-disclosure agreement, specific information cannot be given about the EPTS used. However, both systems have been proved to be valid. The goalkeepers participated in the games but they were excluded from the analysis as their movement behaviour is different from all other players and their positioning is restricted to a specific area.

4.2 Data processing

Each player position file was obtained directly from X/Y coordinates of the EPTS. Using these coordinates, each player position file was classified by their positional roles (defender, midfielder, forward) (Figure 4-2) as performed in previous tactical and physical performance-based research (Di Salvo, et al., Performance Characteristics According to Playing Position in Elite Soccer, 2007).

Player position		D1	D2	D3	M1	M2	M3	F
Game 1	Team 1	**1**	2	1	1	3	1	1
	Team 2	1	2	**1**	1	2	1	2
Game 2	Team 3	**1**	2	1	1	2	1	2
	Team 4	**1**	1	1	1	3	1	2

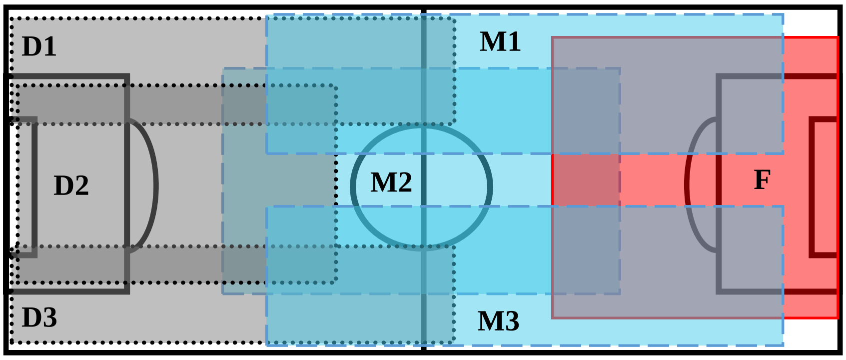

Figure 4-2 Overview of player positional roles. The table indicates the number of players each team had in each of the positions, the numbers coloured in *red* indicate the fixing players. *D1&D3*: External defenders; *D2*: Central defenders; *M1&M3*: External midfielders; *M2*: Central midfielders; *F*: Forwards.

Player performance profiles for the fixing players and players with the same positional role have been created of player positional data (x, y) by calculation of six performance metrics as outlined below. All performance metrics have been calculated per minute unless stated differently. All programs written for calculations and dedicated routines have been performed in MATLAB version R2018b (MathWorks, Inc., Massachusetts, USA). Results from the MATLAB programs were checked against the same calculations by hand and using Excel, with 100% agreement obtained.

4.2.1 Physical performance metrics

Different speed zones were used to get a detailed speed profile of a player (Mohr, Krustrup, & Bangsbo, 2003). The speed was classified into speed zones of: Low Intensity Running (LIR) 0.0-1.9 $m \cdot s^{-1}$; Medium Intensity Running (MIR) 1.9-3.8 $m \cdot s^{-1}$; High Intensity Running (HIR) 3.8-5.6 $m \cdot s^{-1}$; Sprint >5.6 $m \cdot s^{-1}$ (Casamichana, Castellano, & Castagna, 2012; Varley, Fairweather, & Aughey, 2012). The speed of each player was calculated as the Euclidean distance between consecutive pairs of coordinates (x, y), divided by the timestamp (1).

$$Speedplayer_k = \sum_{n=1}^{L} \frac{\sqrt{(Y_{n+1} - Y_n)^2 + (X_{n+1} - X_n)^2}}{T} \qquad k = 1, \dots, z \qquad (1)$$

Where z is the total amount of players and L is the total amount of timestamps from the position file of a player (k). Y and X are the y and x coordinates of a player k and T is the timestamp.

The total distance travelled in metres and percentage of total distance covered in each speed zone was calculated from the X/Y position data. The time spend in each speed zone was calculated from the timestamp data.

4.2.2 Positional performance metrics

Voronoi

A Voronoi diagram (Voronoi, 1907) was used to analyse the spatial distribution of players and define a dominant region for players. The Voronoi area is the area around a player that is closer to that player than to any other player on the pitch (Okabe, Boots, Sugihara, & Chiu, 2000). The Voronoi diagram allowed spatial partitioning of the pitch into cells (dominant regions), where each cell was associated with a player (Figure 4-3). For each player, at each second of the game using players' position coordinates (x, y), a dominant region was obtained from the Voronoi diagram (Fonseca, Milho, Travassos, & Araújo, 2012).

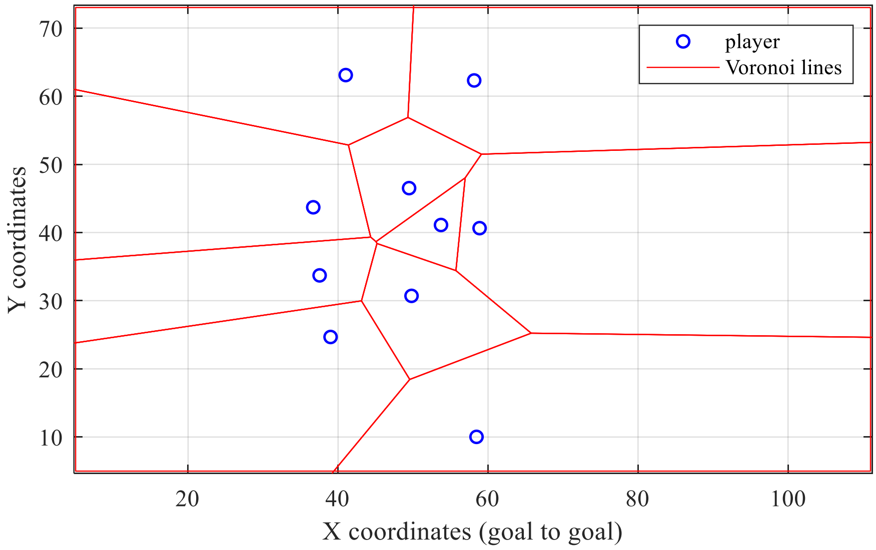

Figure 4-3: Example of a Voronoi diagram on a pitch generated of player position data of one team.

Centroid

The team centroid and three position-specific centroids (defender, midfielder, forward) were calculated for each timestamp using the players' coordinates (x, y). The team centroid was calculated as the mean position of all players of one team (Gonçalves, Figueira, Maçãs, & Sampaio, 2013). The three position specific centroids were calculated as the mean position of players of each of the three specific positions, e.g. defender centroid is the mean position of the defenders of one team.

Distance from centroid

The distance of each player from their team centroid and their position-specific centroid was calculated for each timestamp using the centroid data and player position data. The distance was calculated along the x-axis (longitudinal, goal to goal) and y-axis (Figure 4-4).

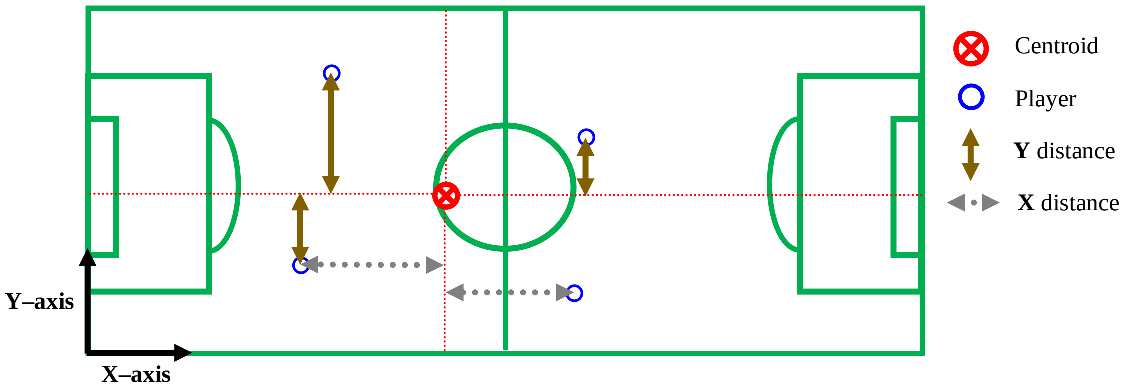

Figure 4-4: Example of the x-axis, y-axis and radial distance of a player from the centroid.

Stretch index

The team stretch index and position specific stretch index was calculated for each timestamp using the distance to centroid data. The team stretch-index was calculated separately along the x-axis and y-axis and was computed as the average distance in the x-axis or y-axis between each player and the team-centroid (Bourbousson, Sève, & McGarry, 2010). The position stretch-index in the x-axis and y-axis was computed as the average distance in the x-axis or y-axis between players of the same positional role and their position-specific centroid.

Heat map

A heat map of a player's position was made to analyse most used sections of the pitch by a player. To create a heat map, the pitch was divided into twenty-four different sections (6 longitudinal and 4 lateral) (Reilly & Williams, 1990). The zones were named based on the width and height of the zone (Figure 4-5). The percentage of time spend by a player in each section of the heatmap was calculated.

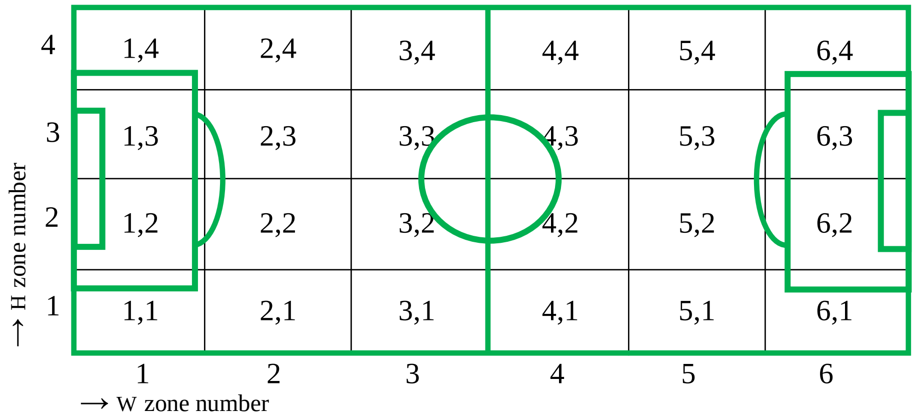

Figure 4-5: Pitch divided into a heatmap of twenty-four zones. The numbers inside a zone represent the zone number, the first number is the width (W) zone number and the second is the height (H) zone number.

4.2.3 Statistics

Approximate entropy (ApEn) was used to evaluate regularity in players' movement behaviour (Sampaio & Maçãs, Measuring tactical behaviour in football, 2012). The ApEn quantifies regularity in a time series by measuring the logarithmic likelihood that runs from patterns that are close (within tolerance r) to m contiguous observations and remain close (within the r) on following incremental comparisons (Pincus, 1991; Pincus & Goldberger, 1994). ApEn was calculated when the minimum length of data points exceeded 200, due to ApEn being sensitive to short data sets (n < 200) (Yentes, et al., 2013). The values used were 1.0 for vectors' length (m) and the tolerance (r) of 0.5 (Richman & Moorman, 2000; Gonçalves, Figueira, Maçãs, & Sampaio, 2013). The ApEn calculation produces a unit-less value ranging between 0-2, where values closer to zero represent higher regularity in the data (Pincus, 1991). The ApEn of the positional performance metrics (stretch-index, distance towards centroid and Voronoi area) were calculated for each fixing player and players with the same positional role during fixing play and normal play. ApEn values will indicate if fixing players influence the regularity of their performance metrics during fixing play.

Descriptive statistics (mean and standard deviation) were calculated for all performance metrics. Coefficient of variation was used to determine variation within a performance metric.

The values for distance towards team and positional centroid, stretch index in X and Y direction, Voronoi area, speed band metrics (time spend in second, distance travelled in meters and percentage of total distance covered) and ApEn values were compared between fixing and normal play via standardized mean differences, calculated with pooled variance and 90% confidence intervals (Hopkins, Marshall, Batterham, & Hanin, 2009). The used threshold values for the effect statistics were trivial 0.2, small 0.6, moderate 1.2, large 2, and very large >2 (Hopkins, Marshall, Batterham, & Hanin, 2009). The smallest worthwhile differences were estimated by multiplying the standardized units by 0.2. Uncertainty in each true effect was assessed using non-clinical magnitude based inferences (Hopkins, 2007) expressed as probabilities that the true magnitude of the effect was substantially positive and negative (calculated from standard errors assuming a normal distribution for a non-clinical study) (Hopkins, Marshall, Batterham, & Hanin, 2009). If the probabilities that the true value is substantially positive and negative were both >5%, the effect was deemed unclear, otherwise the effect was clear (to clarify the use of unclear and clear effects in relation to p-values, the unclear results are $p>0.05$ and the clear results $p<0.05$) and reported as either trivial or substantial (whichever was larger) and the probabilities that it has that magnitude. The probabilities were interpreted with the following scale: possible >25 & ≤75%; likely >75 & ≤95%; very likely >95 & ≤99.5%; most likely >99.5%.

Recursive partitioning and decision tree analyses were performed on metrics where the difference between fixing and normal play was deemed substantial. This was used to determine to what extent each performance measure associated with match-fixing behaviour (Breiman, Friedman, Olshen, & Stone, 1984). The decision tree employed was a coarse tree with the Gini's diversity index split criterion and a maximum number of 7 splits. The accuracy of the model was determined with 5-fold cross validation (Fushiki, 2011). These measures were employed to avoid overfitting plus they produced the highest accuracy of predictions. Results were displayed using a tree visualization and predictor importance plot.

Chapter 5: Results

5.1 Physical performance metrics

Fixing play had substantial effects on eleven of the twelve speed comparisons (Figure 5-1). For all metrics of HIR and MIR, fixing play had *likely and very likely* small increases compared to normal play. For sprint, distance travelled and time spent had both *likely* small increases in fixing play compared to normal play while percentage of total distance covered was unclear. For LIR, percentage of total distance covered and time spent had both *very likely* small decreases in fixing play compared with normal play while distance travelled had a *very likely* moderate increase.

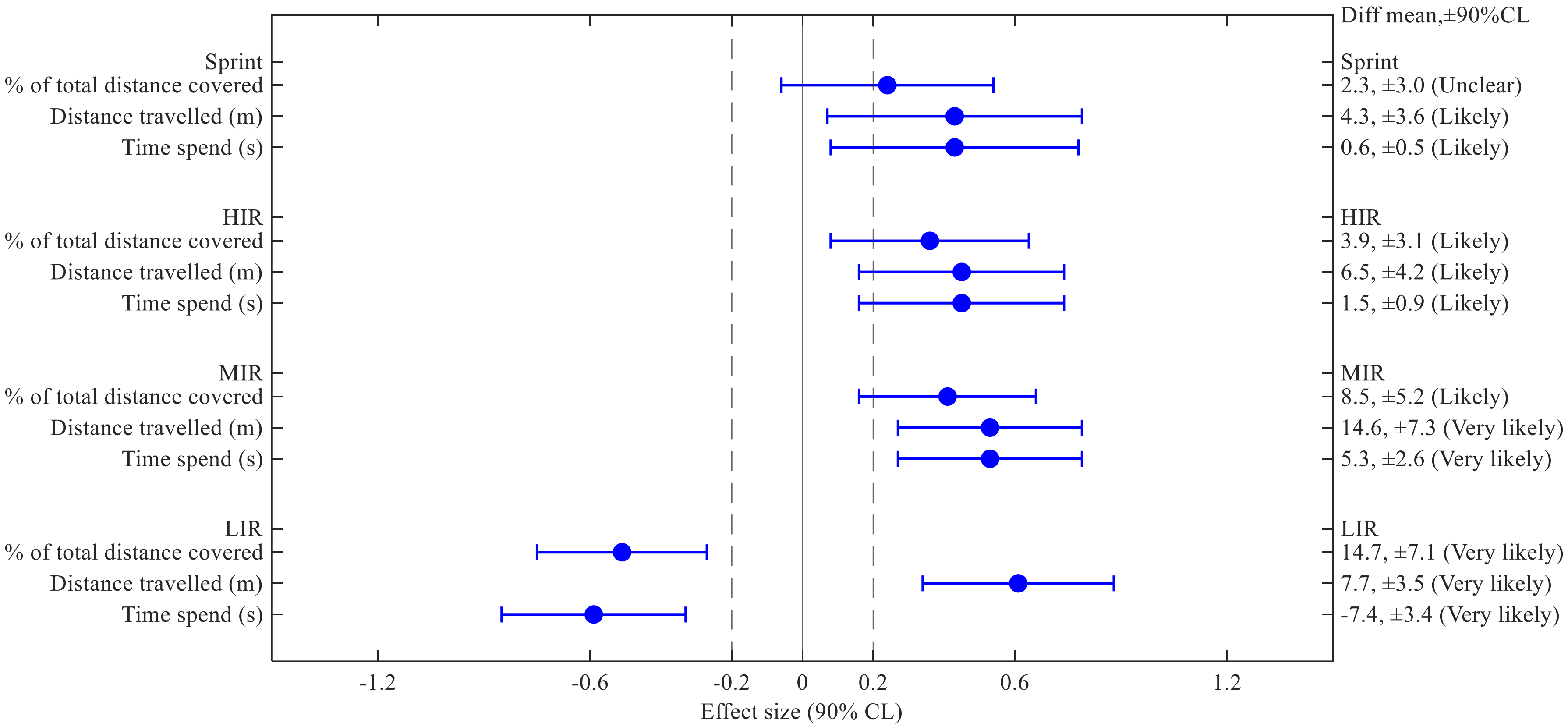

Figure 5-1: Standardized differences per minute between fixing play and normal play with 90% confidence limits for different speed bands, along with the mean difference and standard deviation for each metric. LIR = low intensity running with speed <1.9 m·s^{-1}; MIR = medium intensity running with speed 1.9-3.8 m·s^{-1}; HIR = high intensity running with speed 3.8-5.6 m·s^{-1}; Sprint with speed >5.6 m·s^{-1}. Right side of the graph means higher speed per minute for the fixing player in fixing play (lower values for normal play), left side lower speed values. Dotted lines are thresholds for the smallest important effect.

5.2 Positional performance metrics

Fixing play had substantial and trivial effects on distance to the team and position-specific centroids but no clear effect on the Voronoi area (Figure 5-2) compared to normal play. Distance to the team centroid was smaller during the fixing play compared with normal play, with decrease in distance ranging from *likely* trivial (Y-distance) to *possibly* small decrease (X-distance). Fixing play had a substantial effect on distance to position-specific centroid, it had *most likely* small increases in distance for X and Y during fixing play compared with normal play. Differences between the normal play of fixing players and normal players for the metrics distance to position-specific centroid, team centroid and Voronoi area were unclear.

Fixing play showed substantial differences in the regularity of play for distance to position-specific centroid, distance to team centroid and Voronoi area (Table 5-1; Figure 5-2). Distance to position-specific centroid was more irregular in fixing play compared with normal play, increase in irregularity were ranging from *possibly* small increase (X-distance) to *most likely* large increase (Y-distance). Distance to team centroid had a *possibly* small increase in irregularity in X-distance and *likely* small increase in Y-distance in fixing play compared with normal play. Voronoi area had a *likely* small increase in irregularity in fixing play compared with normal play. Differences between the normal play of fixing players and normal players for distance to position-specific centroid, distance to team centroid and Voronoi area were unclear.

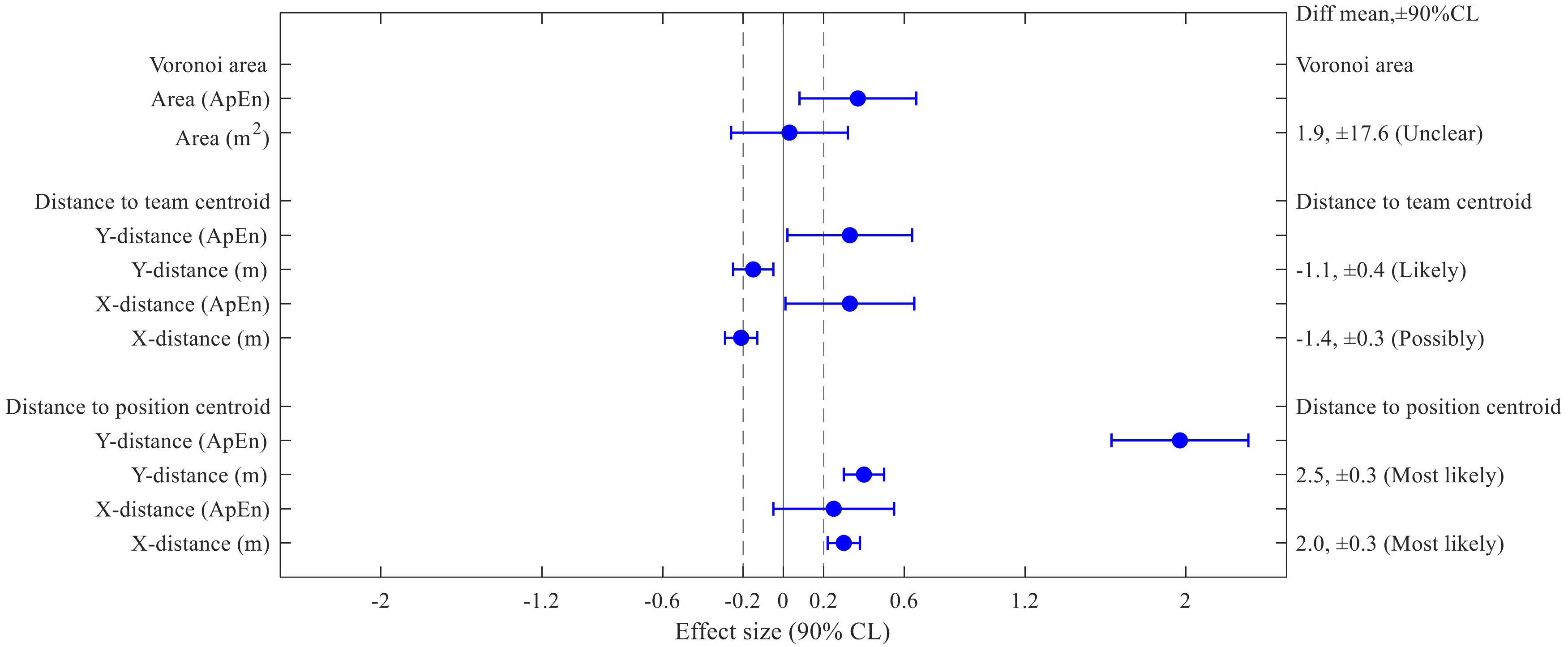

Figure 5-2: Standardized difference between fixing play and normal play with 90% confidence limits for Voronoi area, distance in x (longitudinal, goal to goal) and y (lateral) direction to team and position centroid, along with the mean difference and 90% confidence limits for each of the metrics in meters. Right side of the graph means higher values for the fixing player in fixing play (lower values for normal play), left side means lower values. Dotted lines are thresholds for the smallest important effect.

Table 5-1: Mean and standard deviation of the regularity (ApEn value) for each individual based metric along with the standardized difference between the regularity of fixing play and normal play with 90% confidence limits. Higher standardized difference means more irregularity in fixing play. Normal play of the column fixing players only includes data of the normal play of the fixing players. The column normal play includes normal players' data, the fixing players normal play data is included when compared with their fixing play.

	ApEn			
	Fixing players		**Normal play**	**Standardized difference, ± 90% CL**[a]
	Fixing play Mean ± SD	**Normal play** Mean ± SD	Mean ± SD	
Distance to position centroid				
X-distance from player to centroid	0.46 ± 0.03		0.45 ± 0.04	0.25, ±0.3$^{S\bullet}$
		0.45 ± 0.04	0.45 ± 0.06	0.0, ±0.25$^{T\circ\circ}$
Y-distance from player to centroid	0.50 ± 0.04		0.42 ± 0.04	1.97, ±0.34$^{L\bullet\bullet\bullet\bullet}$
		0.43 ± 0.06	0.42 ± 0.05	0.20, ±0.33$^{S\bullet}$
Distance to team centroid				
X-distance from player to centroid	0.73 ± 0.1		0.69 ± 0.12	0.33, ±0.32$^{S\bullet}$
		0.71 ± 0.1	0.70 ± 0.09	0.11, ±0.32$^{T\circ\circ}$
Y-distance from player to centroid	0.75 ± 0.07		0.78 ± 0.09	0.33, ±0.31$^{S\bullet\bullet}$
		0.76 ± 0.1	0.76 ± 0.12	0.0, ±0.28$^{T\circ\circ}$
Voronoi				
Voronoi area of player	0.60 ± 0.15		0.54 ± 0.16	0.37, ±0.29$^{S\bullet\bullet}$
		0.31 ± 0.16	0.31 ± 0.16	-

[a]Superscripted letters indicate effect size as follows: ^{T}Trivial, ^{S}Small, ^{M}Moderate, ^{L}Large;
Superscripted symbols indicate the probabilities of an effect being $^{\circ}$trivial or $^{\bullet}$substantial (whichever was larger).
Open circles indicate trivial effects as: $^{\circ}$Possibly, $^{\circ\circ}$Likely, $^{\circ\circ\circ}$Very likely, $^{\circ\circ\circ\circ}$Most likely
Full circles indicate substantial effects as: $^{\bullet}$Possibly, $^{\bullet\bullet}$Likely, $^{\bullet\bullet\bullet}$Very likely, $^{\bullet\bullet\bullet\bullet}$Most likely

Fixing play had a substantial effect on the team based metric position stretch index in Y-direction (Figure 5-3). Position-specific stretch index in Y-direction had a *likely* small increase in fixing play compared with normal play. Differences between fixing play and normal play for the metrics position-specifc stretch index in X-direction and team stretch index were unclear. Fixing play had substantial effects on the regularity of play for position-specific stretch index (Table 5-2; Figure 5-3). The position-specific stretch index was more irregular in fixing play compared with normal play, increases in irregularity were *most likely* moderate (X-, and Y-distance). Differences in regularity between fixing play and normal play for team stretch index were unclear.

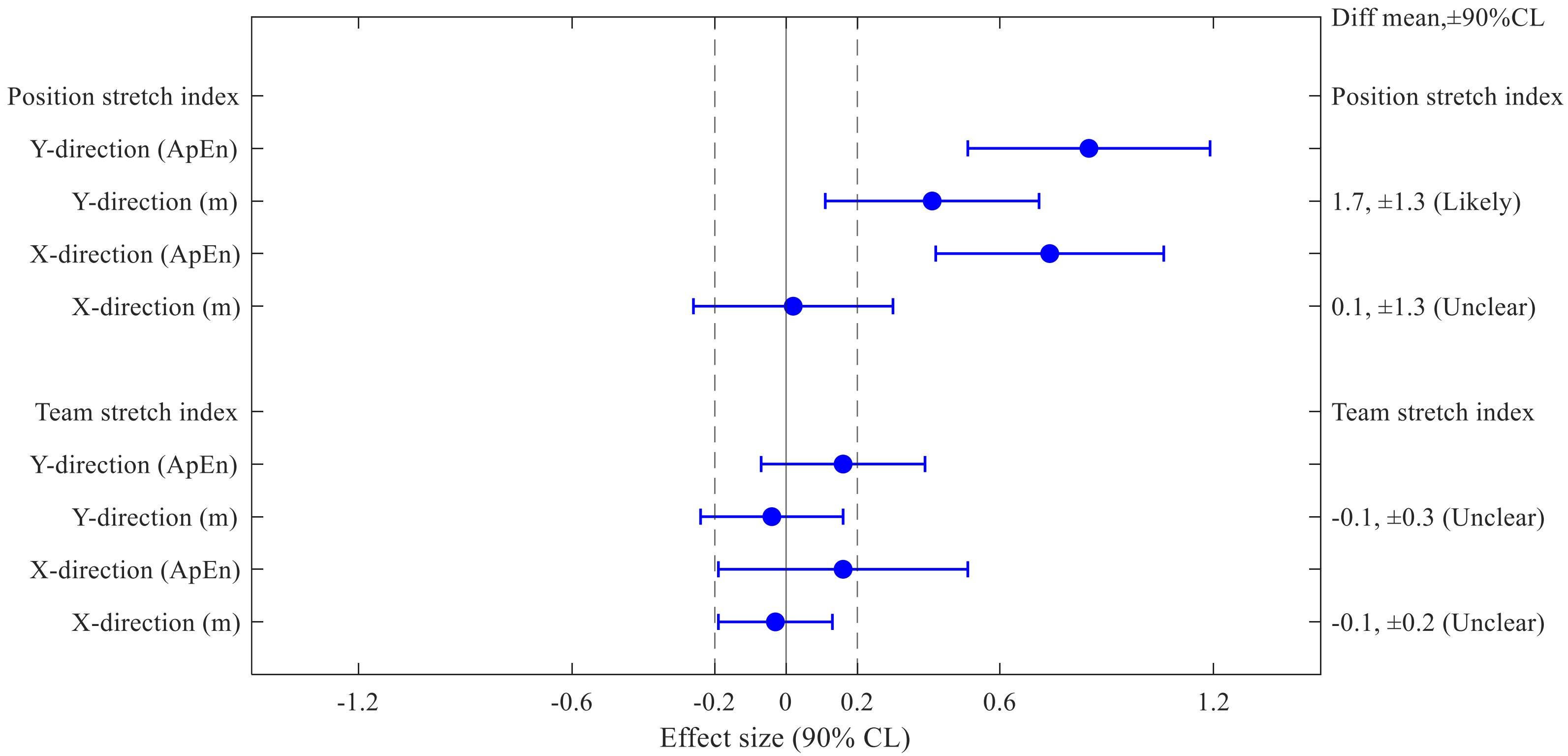

Figure 5-3 : Standardized differences between fixing play and normal play with 90% confidence limits for team/position specific stretch index in the x and y direction, along with the mean difference and 90% confidence limits for each of the metrics in meters. Right side of the graph means higher values (larger stretch index) for fixing play (smaller stretch index for normal play), left side a smaller stretch index. Dotted lines are thresholds for the smallest important effect.

Table 5-2: Mean and standard deviation of the regularity (ApEn value) for each team-based metric along with the standardized difference between the regularity of fixing play and normal play with 90% confidence limits. Higher standardized difference means more irregularity in fixing play.

	ApEn		
	Fixing play Mean ± SD	**Normal play** Mean ± SD	**Standardized difference, ± 90% CL**^{a}
Team stretch index			
X direction (goal to goal)	1.01 ± 0.08	1.00 ± 0.06	0.16, ±0.35$^{T\circ}$
Y direction	0.95 ± 0.05	0.81 ± 0.12	0.16, ±0.23$^{T\circ}$
Position stretch index			
X direction (goal to goal)	1.15 ± 0.09	1.09 ± 0.08	0.74, ±0.32$^{M\bullet\bullet\bullet\bullet}$
Y direction	1.18 ± 0.09	1.12 ± 0.07	0.85, ±0.34$^{M\bullet\bullet\bullet\bullet}$

^{a}Superscripted letters indicate effect size as follows: ^{T}Trivial, ^{S}Small, ^{M}Moderate, ^{L}Large; Superscripted symbols indicate the probabilities of an effect being $^{\circ}$trivial or $^{\bullet}$substantial (whichever was larger).
Open circles indicate trivial effects as: $^{\circ}$Possibly, $^{\circ\circ}$Likely, $^{\circ\circ\circ}$Very likely, $^{\circ\circ\circ\circ}$Most likely
Full circles indicate substantial effects as: $^{\bullet}$Possibly, $^{\bullet\bullet}$Likely, $^{\bullet\bullet\bullet}$Very likely, $^{\bullet\bullet\bullet\bullet}$Most likely

Fixing play showed substantial differences in player pitch positioning past the midline (Table 5-3 and Figure 5-4) compared with normal play. The fixing players spend more time pass the midline compared to their normal play, increase in time was *very likely* large. Time spend past the midline during fixing play in the upper part of the pitch (Zone W: 4:6; Zone H:4) had a *very likely* large increase compared with normal play and was unclear for the middle part of the pitch (Zone W: 4:6; Zone H:3) . Differences in time spent pass the midline when the same time periods of normal players were compared were unclear.

Table 5-3: Mean and standard deviation of time spend in percentage past the midline and of specific zones past the midline, along with the standardized difference of fixing play and normal play with 90% confidence limits.

	Average time spend (%) in zones		
	Fixing play Mean ± SD	**Normal play** Mean ± SD	**Standardized difference, ± 90% CL**^{a}
Past midline: Zone W: 4:6 Zone H: 1:4	33.8, ±6.1	26.2, ±6.3	1.17, ±0.65$^{L\bullet\bullet\bullet}$
Past midline: Zone W: 4:6 Zone H: 4	17.1, ±7.6	12.3, ±2.1	2.13, ±1.80$^{L\bullet\bullet\bullet}$
Past midline: Zone W: 4:6 Zone H: 3	17.0, ±7.5	12.9, ±10.4	0.37, ±0.57$^{S\bullet}$

^{a}Superscripted letters indicate effect size as follows: ^{T}Trivial, ^{S}Small, ^{M}Moderate, ^{L}Large; Superscripted symbols indicate the probabilities of an effect being $^{\circ}$trivial or $^{\bullet}$substantial (whichever was larger).
Open circles indicate trivial effects as: $^{\circ}$Possibly, $^{\circ\circ}$Likely, $^{\circ\circ\circ}$Very likely, $^{\circ\circ\circ\circ}$Most likely
Full circles indicate substantial effects as: $^{\bullet}$Possibly, $^{\bullet\bullet}$Likely, $^{\bullet\bullet\bullet}$Very likely, $^{\bullet\bullet\bullet\bullet}$Most likely

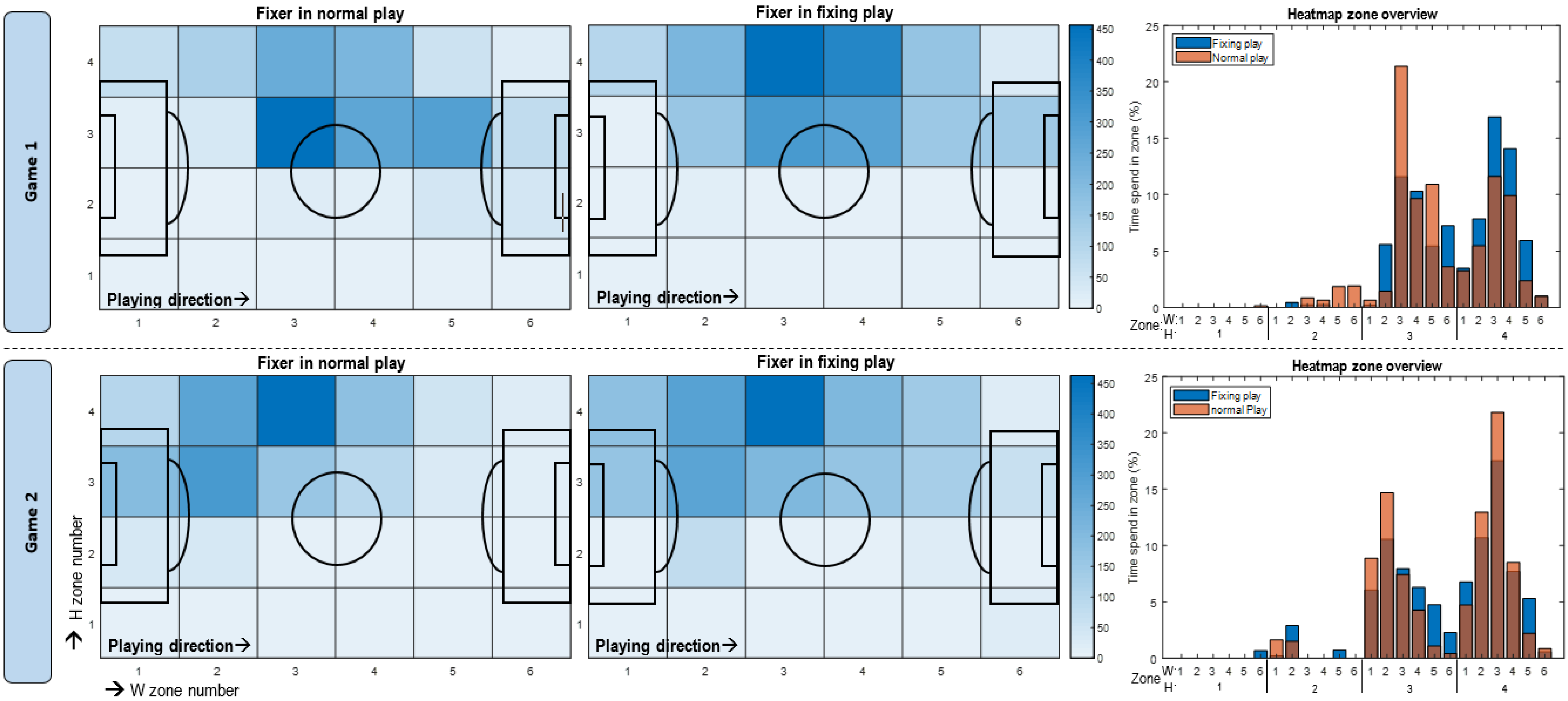

Figure 5-4: Pitch positioning analyses of the two games. Football pitches show a heatmap of the time spent in seconds in each of the zones for the fixing players in normal play (left) and fixing play (right). The bar graphs are an overview of the time in percentage spend in each of the zones, where zone numbers of W and H refer to the zone number on the heatmaps.

The full recursive partitioning and coarse tree model is presented in Figure 5-5 along with the predictor importance estimates in Figure 5-6. The used zone numbering of the heatmap for the model was different than explained in the methods section. Zones 1 to 24 have been used instead of width and height numbers (Figure 5-7), this is because the model needs a consecutive series of numbers to be able to divide the zones. Only features relating to pitch positioning contribute to the model, this can be confirmed by the predictor importance estimates of the model. The predictor importance estimates of the model showed that the most used zone of the heatmap (heatzone1) was the most important predictor for the model, followed by the Y distance to position centroid (YDtPosCtr). The model indicates that zones at upper part of the pitch are important indicators for match fixing behaviour along with a greater Y distance towards the position specific centroid in zones close to the goal. The accuracy of this model for explaining fixing and normal behaviour was 52.2%.

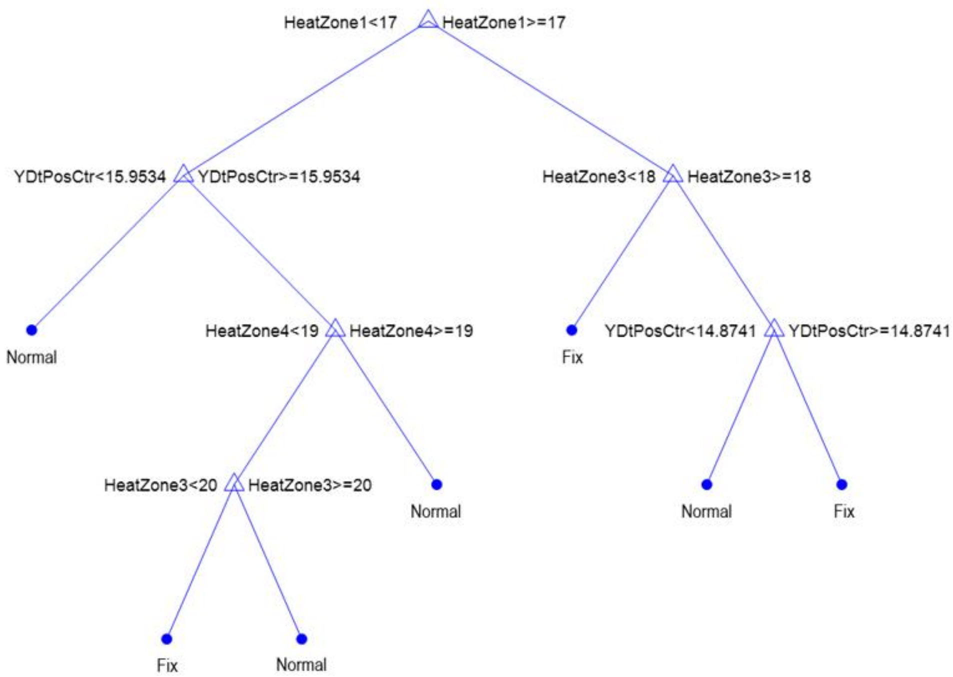

Figure 5-5: Recursive portioning and coarse tree model explaining fixing behaviour in football. The final node variables outline the model-expected behaviour (fix or normal) of a fixing player. *Heatzone1* = most used zone of heatmap; *Heatzone3* = third most used zone of heatmap; *Heatzone4* = forth most used zone of heatmap; *YDtPosCtr* = Y distance in meters to position centroid.

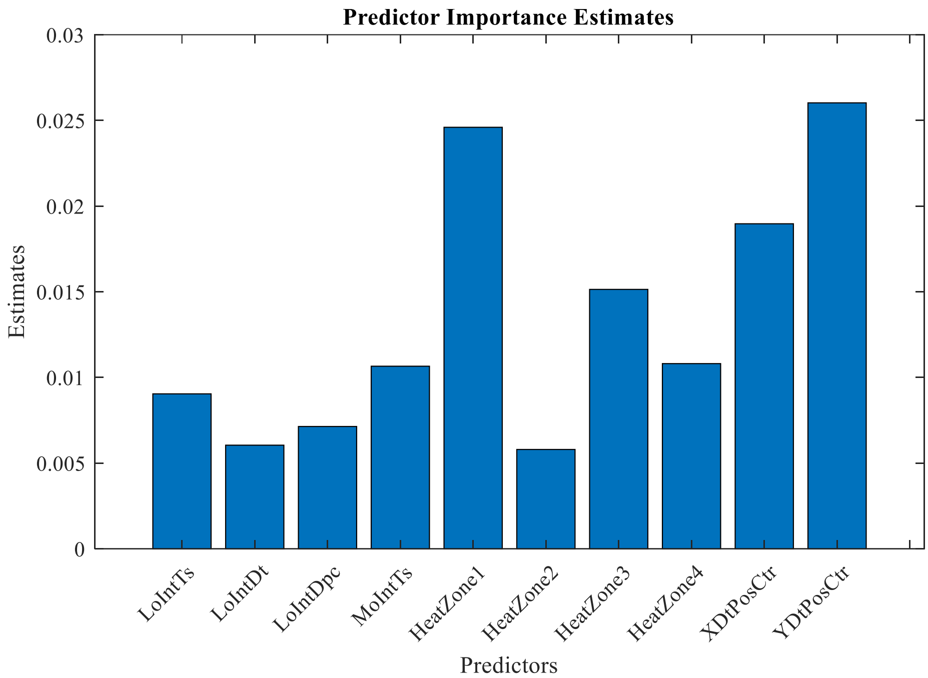

Figure 5-6: Predictor importance estimates of the used metrics for the model. Higher estimate values indicate a higher importance for the predictor. *LoIntTs* = time spend in seconds in LIR; *LoIntDt* = distance covered in meters in LIR; *LoIntDpc* = percentage of total distance covered in LIR; *MoIntTs* = time spend in seconds in MIR; *Heatzone1* = most used zone of heatmap; *Heatzone2* = second most used zone of heatmap; *Heatzone3* = third most used zone of heatmap; *Heatzone4* = forth most used zone of heatmap; X*DtPosCtr* = X distance to position centroid; *YDtPosCtr* = Y distance to position centroid.

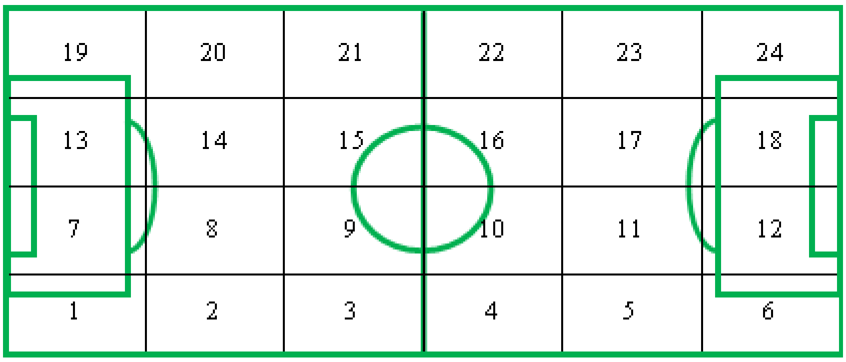

Figure 5-7: Overview of the used zone numbers of the heatmap for the recursive partitioning and coarse tree model

Chapter 6: Discussion

This thesis aimed to explore how player performance metrics derived from players' positional x and y coordinates are related to match-fixing behaviour in football.

This research may be the first to investigate the effect of match-fixing on football players' performance, analysing their performance metrics during a game. The key finding of this thesis was that fixing players changed their playing behaviour during a game. Specifically, fixing play had substantial effects on distance travelled in all speed zones where more distance was covered in fixing play compared with normal play. Furthermore, fixing players spent less time and had a lower percentage of total distance covered in the low intensity running zone. Fixing players had more irregularity in all performance metrics except team-based metrics, which were not associated with match-fixing behaviour. Finally, fixing play had substantial effects on player pitch positioning as fixing players moved forward on the pitch and kept more distance towards the position-specific centroid resulting in more spread of play in the lateral direction. These findings provide insights to player performance metrics underpinning match-fixing behaviour for defence players which can possibly assist in providing supporting evidence to prosecute match-fixing players.

It is suggested that higher approximate entropy values in player performance metrics and greater distance towards the centroid may characterize match-fixing behaviour for defence players.

6.1 Physical performance metrics

The fixing players did alter their movement, as evidenced by the findings of differences in distance travelled, percentage of total distance covered and time spent per minute in all speed zones. Fixing players covered more distance per minute in all speed zones in fixing play and distance showed a higher percentage of the total distance covered for medium intensity running and higher (>1.9 m·s-1). The findings indicate the fixing player spent more time in higher speed intensities, which may be linked to creating open space and being unpredictable (Gonçalves, Figueira, Maçãs, & Sampaio, 2013). Fixing players in fixing play covered more distance in low intensity running while spending less time in the zone, indicating the fixing player moved more often at low intensities in fixing play while standing still in normal play. The higher approximate entropy values of fixing players suggest that fixing players move more randomly and were not coordinated with the collective team behaviour. The higher irregularity in movement and not being coordinated with the team behaviour could have resulted in the fixing players covering more distance. Speed zones are valuable metrics to determine if a player behaves differently according to their physical capacity. Speed profiles of players will likely be a good indicator to discriminate fixing and normal play based on their physical capacity.

6.2 Positional performance metrics

6.2.1 Voronoi area

The fixing player did not affect the Voronoi area, indicated by no difference between normal play and fixing play. This might be explained by taking into account that the analysis approach averages the Voronoi area across fixing and normal play. The Voronoi area showed a high degree of variability, indicated by a coefficient of variation of 72% and high ApEn values, similar to findings in previous research (Gonçalves, et al., 2017; Fonseca, Milho, Travassos, & Araújo, 2012). This trait in itself might make this measure less likely to be able to detect differences in play due to this inherent variability. The spatial organization of a team, defined by the Voronoi area, is influenced by the offensive and defensive phases during a game (Fonseca, Milho, Travassos, & Araújo, 2012). The Voronoi area appear to be larger for teams in the offensive phase and smaller in the defensive phase. As such, the average Voronoi area might not possess sufficient sensitivity to discriminate fixing and normal behaviour without taking into account the defensive and offensive phases. Post-hoc analyses of Voronoi area across time indicated high peaks of Voronoi area during fixing play with peaks ranging from 40 m^2 to 80 m^2 above the Voronoi area of a normal player with the same positional role in the same time period (Figure 6-1). The high peaks of Voronoi area were identified by using time analysis window of 10 seconds, the high peaks were clouded by the inherent variability for longer time windows. Given the Voronoi area is influenced and calculated based on the positions of all players on the pitch, the high peaks during fixing play indicates the fixing player is creating space from his teammates and opponent players. Thus, the Voronoi area across time could possibly be used to discriminate fixing and normal behaviour. However, further research into Voronoi area across time should be done to identify if the peaks in fixing play are indeed related to fixing play and not to the variability of the metric.

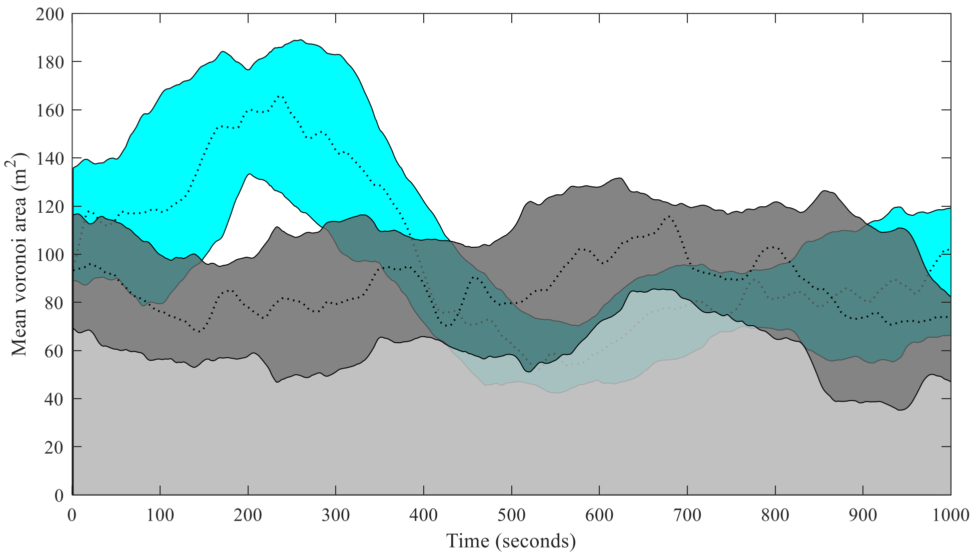

Figure 6-1: Example of a high Voronoi area peak across time for one of the fixing players in fixing play (*blue area*) and of a normal player with the same positional role in the same time period (*grey area*). Dotted line represents the mean value and solid lines the standard deviation.

6.2.2 Team metrics

The fixing player did not affect the positional performance metric team stretch-index, indicated by no difference between normal play and fixing play. This might be due to the fact that the fixing player represented only one of the ten players in the team stretch index calculations. Given the team stretch index is based on the average distance towards the team centroid of all players in a team, the change due to only one player may be clouded by the calculation process and weighted towards the remaining nine players. The finding is also possibly due to the emergent behaviour of the team, players might have adjusted their position due to the fixing player's different movement pattern to maintain a certain 'shape'. When the fixing player in defence plays differently than usual, players around him will constantly adjust their position and distance towards the team centroid, conditioned by the fixing player (and their teammates) positioning (Poplu, Ripoll, Mavromatis, & Baratgin, 2008; Sampaio & Maçãs, Measuring tactical behaviour in football, 2012). As the collected data was of players that train and compete together, a high level of inter-player coordination is expected (Gonçalves, Figueira, Maçãs, &

Sampaio, 2013). Similar research has also shown that inter-player coordination of nearest teammates in defence and midfield are highly regular (Gonçalves, et al., 2017) and is utilised to reduce open spaces to prevent scoring opportunities by the opponent (Duarte, et al., 2012).

Findings in this thesis indicate the fixing player is staying closer to the team centroid in fixing play compared to normal play. To determine if the nearest midfield players to the fixing player altered their position patterns accordingly, post-hoc analyses of their distance towards the team centroid in fixing play and normal play were performed. This analysis indicated the nearest midfield players might have altered their relative positioning towards the team centroid in fixing play as indicated by a *possibly* small increase (1.3 ±0.3m) and therefore also mask the possible change of the team stretch-index. As such, the team stretch-index in fixing play and normal play are possibly similar because of the inter-player coordination and position adjustments of the normal players. Certainly, team-based metrics were not sensitive enough to detect fixing in this instance and will likely be a poor indicator where altered positioning is a part of the fixing strategy.

6.2.3 Position-specific metrics

The fixing player did affect the positional performance metric position-specific centroid metrics, indicated by differences between the regularity in normal play and fixing play along with differences in distance to centroid and stretch index. Specifically, the fixing player created more irregularity (higher ApEn) and spread of play to the lateral side (Y-direction) of the pitch. The higher irregularity and distance is in contrast with the principles of the defending play where high coordination and regularity between players are expected, mostly due to the similar roles and smaller distances between players, and needed to reduce open spaces and opportunities for the opponent to score (Headrick, et al., 2012; Duarte, et al., 2012; Gonçalves, Figueira, Maçãs, & Sampaio, 2013). The fixing player is possibly stretching the defence to create space and cause gaps between the defending players to create opportunities for the opposition players to pass or run through to the goal (Quellette, 2004). Lateral changes in distance are associated with critical game periods (Frencken, Poel, Visscher, & Lemmink, 2012), indicating that the fixing player is destabilizing his team with larger distance/spread and irregularity in the lateral (Y) direction to create opportunities for the opponent to score. The fixing player created more irregularity and space in the longitudinal (X) direction of the pitch compared with his normal play based on an increase in distance towards the position-specific centroid and higher ApEn in the X position stretch index. However, teammates with the same positional role are covering up this distance, shown by no difference in the distance of position stretch index in X. The increase in irregularity indicates players were not stable in their positioning and kept changing their position in the longitudinal direction. This could be due to teammates covering up the created distance by the fixing player (which is more irregular and less predictable) to maintain and keep the defence shape or to keep the longitudinal distance short to use the offside rule to keep attackers away from the goal (Wade, 1996). The higher irregularity in both X and Y direction for the position based metrics during fixing play are in line with the fact that the fixing player is moving around more often as discussed in section 6.1.

The longitudinal based position metrics were not sensitive enough to detect fixing in this instance, possibly due to teammates being more coordinated in the longitudinal than lateral direction (Sampaio & Maçãs, Measuring tactical behaviour in football, 2012). However, regularity and distance measures of position-specific metrics in the lateral direction were strong indicators to discriminate fixing and normal play and will likely be good indicators where altered positioning is a part of the fixing strategies.

6.2.4 Heatmap

Fixing players positioned themselves more forward past the midline and more laterally on the pitch in fixing play compared with normal play. As the fixing players were defensive players, the findings of the fixing players staying closer to the team centroid (which is more forward than the position centroid) in longitudinal direction during fixing play also confirms fixing players staying more forward on the heatmap.

The fixing player positioned himself more forward on the heatmap, when viewed in conjunction with an increased distance in the longitudinal direction to the position-specific centroid in combination with no clear difference in the position stretch index it indicates the other defensive players also moved forward. By doing so defenders reduced the stretch index in the longitudinal direction and compressed the defence (Quellette, 2004) to reduce open spaces and opportunities for the opponent to score (Headrick, et al., 2012). In fact, these findings help confirm the behaviour of covering space and maintaining the defensive shape in the fixing play as previously discussed in section 6.2.3.

The fixing player stayed more lateral on the heatmap in fixing play and stayed closer to the team centroid in the lateral direction, indicating the average overall lateral position of teammates was closer to the fixing player. However, the fixing player in fixing play stayed further away from the position centroid in lateral direction suggesting that the players with the same positional role had more spread to the opposite lateral direction. This finding helps confirm that fixing players stretch the defence in the lateral direction. Thus, heatmaps will likely be a good indicator to discriminate fixing and normal play where altered positioning is a part of the fixing strategies as heatmaps provide an overview of players' most frequently used positions on the pitch.

6.3 Recursive portioning and coarse tree

The recursive partitioning and coarse tree model provided a visual representation of which performance metrics associated more with fixing or normal play. The fixing behaviour was explained by position-based features Y-distance to position centroid and the most, third most and fourth most used zone of the heatmap. The model confirms the fixing player positioned himself differently by moving forward on the pitch and keeping more distance towards his position centroid as previously discussed in sections 6.2.3 and 6.2.4. Applications of the model have the potential to be beneficial for identifying metrics associated with fixing behaviour and eliminate the need of monitoring other metrics which do not associate with fixing behaviour. Further, the visual representation of the model provides a clear overview of performance metrics associated with fixing behaviour. The visual representation makes the interpretability of the model easier for anyone interested in the variables.

6.4 Limitations and recommendations for future research

Some limitations of this study should be noted. The performance profile for a player's normal play was based on one game, so this does not provide a broad perspective of player typical values and variability within each performance metric. To get a broader perspective of a player's normal playing performance profile, using more games can provide a more representative view of a normal profile. The use of more games might provide a more substantial evaluation of the typical values and variability of the individual performance metrics of a player which in turn would strengthen any substantial deviation from this baseline. However, it should not discount the present work. Substantial differences existed in this study for normal and fixing play and this was against the same opponent and under the same conditions. These are factors that might confound the normal playing performance profile created with use of more games, as games would include different opponents playing different styles and under different weather conditions and ground sizes. A combination of both single game and numerous game analyses is worth further evaluation.

Another limitation of this study was that no match event data was available for the analyses and performance metrics derived from position data only were used. However, player behaviour is constrained by several other factors such as environmental factors, goals scored, receiving cards and the ball location (McGarry, Anderson, Wallace, Hughes, & Franks, 2002; Gonçalves, Figueira, Maçãs, & Sampaio, 2013). For example, if a player received a yellow card during the game, this could influence the way he is playing, he can be more cautious with his movements to prevent him from receiving a red card. Ball position is another factor which can add valuable information about team tactics on the pitch. The ball location can reveal offensive and defensive phases with accompanying team tactics during the game, which are currently unknow. Offensive and defensive phases during the game could clarify positioning of players, e.g. players can be more forward-positioned because the offensive phase is higher. However, the current analysis consisted of big enough time samples between fixing and normal play to be able to avoid bias of offensive and defensive phases in the analysis. Ball location could also give insight into player tactics. A common match fixing technique for forwards, explained in chapter 2.3, is to move the ball straight to the opposing team to allow them to take possession of the ball. This is part of their tactics where there is currently no insight to. Having access to the position data of the ball provides more information of team tactics and can possibly give insight to if fixing may influence defensive or offensive behaviour. Ball position would also provide more information of players' relation to the ball which can possibly give insight to players' intention for moving over the pitch. Match event data and ball location should be included in future analyses to get a complete overview of the game.

Examining specific metrics, a limitation of the Voronoi area is that it is based on Euclidean distances between the players (Fonseca, Milho, Travassos, & Araújo, 2012). This assumes 'ideal' conditions of same speed and reaction time of players. Further, the area calculated around any player does not take into account the way the players are facing or their current speed nor does it factor in player visual response and their reaction time. These factors can affect how quickly a player might get to the edge of their area and means that a player in an adjacent area might, in fact, be able to get to the same position faster. The distance between ball and player might also have different influences on nearby Voronoi areas. For example, when players are far away from the ball, they have a larger area and subsequently time to reach the ball when it moves towards them, possibly resulting in a wider influence on nearby areas. Alternatively, for players close to the ball, they have a smaller area and less time to reach the ball when it moves towards them, possibly resulting in less influence on nearby areas. The Voronoi area was calculated without the goalkeeper as their movement behaviour is different from all other players and their positioning is restricted to a specific area. However, the position of a goalkeeper on the pitch could possibly influence Voronoi areas of players. Future research should investigate the influence of match-fixing behaviour on players' Voronoi area using models which take into account the goalkeeper and these factors of ball location, direction of movement and player speed (Fernandez & Bornn, 2018). This may provide more insight into the intentions of the players which in turn might assist in providing evidence for match-fixing behaviour.

Chapter 7: Conclusions

Performance metrics derived from players' positional x and y coordinates can be used to detect match fixing scenarios where players alter their game movement patterns. Both physical and positional performance metrics differed between normal and fixing play indicating that both metrics might be used.

Positional performance metrics related to pitch positioning were most associated with fixing behaviour and showed the biggest differences compared to normal behaviour. Specifically, centroid-based metrics and heatmap substantially differed between fixing play and normal play with effect sizes ranging from small to large. Voronoi area was not affected by fixing play, possibly due to the high degree of variability in the metric. Further, fixing players had more irregularity in all positional performance metrics except team-based metrics. The inter-player coordination and position adjustments of the normal players is likely to be the main cause of similar results for team metrics along with the used mathematical approach.

Physical performance metrics were less associated with fixing behaviour but showed differences compared to normal play. Specifically, fixing players covered more distance in all speed intensities and spent more time in higher speed intensities compared to their normal play. Furthermore, fixing players moved around more often with low intensities in fixing play while standing still in normal play.

A rule based recursive partitioning and coarse tree model was developed based on the metrics that differ between normal and fixing play. The model could detect match-fixing activity with an accuracy of 52.2%.

Future work should focus on the use of a wider range of fixing scenarios of numerous games to further develop the match-fixing detection framework.

The finding of this thesis can be beneficial, not only for integrity purposes of the football related society, but also for a wider spectrum of team sports using electronic performance tracking systems to measure player performance. The findings can possibly assist in providing supporting evidence to prosecute match-fixing players and provide scientific knowledge to create a match-fixing detection approach which covers both betting and non-betting related match-fixing.

References

Bartlett, R., Button, C., Robins, M., Dutt-Mazumder, A., & Kennedy, G. (2012). Analysing Team Coordination Patterns from Player Movement Trajectories in Soccer: Methodological Considerations. *International Journal of Performance Analysis in Sport, 12*(2), 398-424.

Boeri, T., & Severgnini, B. (2013). Match Rigging in Italian Professional Soccer: The Economic Determinants of Corruption. In M. Haberfeld, & D. Sheehan, *Match-Fixing in International Sports: Existing Processes, Law Enforcement, and Prevention Strategies* (pp. 101-112). Springer.

Boles, J., & Garbin, A. P. (1984). The Strip Club and Customer-Stripper Patterns of Interaction. *Sociology and Social Research, 58*, 136-144.

Boniface, P., Lacarriere, S., Verschuuren, P., Tuaillon, A., Forrest, D., Icard, J., . . . Wang, X. (2012). Retrieved from Spordiinfo: http://www.spordiinfo.ee/g22s355

Borristow, M., Bernard, J., Pendlebury, H., & Greaves, R. (1960, October 10). This is the way matches are fixed. *Daily Mail (London), 1*(11).

Bourbousson, J., Sève, C., & McGarry, T. (2010). Space-time coordination dynamics in basketball: Part 2. The interaction between the two teams. *Journal of Sports Sciences, 28*(3), 349-358.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees.* Monterey: Wadsworth.

Bricknell, S. (2015). Corruption in Australian sport. Trends & issues in crime and criminal justice. *Australian Institute of Criminology, 490*.

Brooks, J., Kerr, M., & Guttag, J. (2016). Using Machine Learning to Draw Inferences from Pass Location Data in Soccer. *Statistical Analysis and Data Mining: The ASA Data Sci Journal, 9*(5), 338-349.

Brooks, J., Kerr, M., & Guttag, J. (2016). Using Machine Learning to Draw Inferences from Pass Location Data in Soccer. *Stat. Anal. Data Min. ASA Data Sci., 9*(5), 338-349.

Buchheit, M., Allen, A., Poon, T., Modonutti, M., Gregson, W., & Di Salvo, V. (2014). Integrating different tracking systems in football: multiple camera semi-automatic system, local position measurement and GPS technologies. *Journal of Sports Sciences, 32*(20), 1844-1857.

Cairns, W. (2013). IPL spot-fixing allegations result in criminal charges. *Corruption Watch, 21*(2).

Casamichana, D., Castellano, J., & Castagna, C. (2012). Comparing the physical demands of friendly matches and small-sided games in semiprofessional soccer players. *Journal of Strength and Conditioning, 26*(3), 837-843.

Castellano, J., Álvarez, D., Figueira, B., Coutinho, D., & Sampaio, J. (2013). Identifying the effects from the quality of opposition in a Football team positioning strategy. *International Journal of Performance Analysis in Sport, 13*, 822-832.

Chater, M., Arrondel, L., Gayant, J., & Laslier, J. (2018). Fixing match-fixing. *arXiv*.

Clemente, F., Couceiro, M., Martins, F., Ivanova, M., & Mendes, R. (2013). Activity profiles of soccer players during the 2010 world cup. *Journal of Human Kinetics, 38*, 201-211.

Decker, W. (1995). *Sport in der griechischen Antike. Vom minoischen Wettkampf bis zu den Olympischen Spielen. [Sport in Greek antiquity. From the Minoan tournament to the Olympic Games].* Munchen: Beck.

Di Salvo, V., Baron, R., Tschan, H., Calderon Mont, F. J., Bachl, N., & Pigozzi, F. (2007). Performance Characteristics According to Playing Position in Elite Soccer. *International Journal of Sports Medicine, 28*, 222-227.

Di Salvo, V., Baron, R., Tschan, H., Calderon Mont, F. J., Bachl, N., & Pigozzi, F. (2007). Performance Characteristics According to Playing Position in Elite Soccer. *Int J Sports Med, 28*, 222-227.

Dixon, M., & Coles, S. (1997). Modelling association football scores and inefficiencies in the football betting market. *Applied Statistics, 46*, 265-280.

Duarte, R., Araújo, D., Freire, L., Folgado, H., Fernandes, O., & Davids, K. (2012). Intra- and inter-group coordination patterns reveal collective behaviors of football players near the scoring zone. *Human Movement Science, 31*(6), 1639-1651.

EUROPOL. (2013, February 6). *UPDATE - RESULTS FROM THE LARGEST FOOTBALL MATCH-FIXING INVESTIGATION IN EUROPE*. Retrieved from EUROPOL: https://www.europol.europa.eu/newsroom/news/update-results-largest-football-match-fixing-investigation-in-europe

EUROPOL. (2014, may 27). *UEFA and EUROPOL step up cooperation against match-fixing*. Retrieved from europol.europa: https://www.europol.europa.eu/newsroom/news/uefa-and-europol-step-cooperation-against-match-fixing

FATF. (2009, July). *Money Laundering through the Football Sector.* Retrieved from http://www.fatf-gafi.org/media/fatf/documents/reports/ML%20through%20the%20Football%20Sector.pdf

Feltes, T. (2013). Match fixing in Western Europe. In D. Sheehan, & M. Heberfeld, *Match fixing in international sports: Existing processes, law enforcement, andprevention strategies* (pp. 15-30). Springer.

Fernandez, J. (1993). *"A Worldwide Plague".* Kuala Lampur: Malay Mail.

Fernandez, J., & Bornn, L. (2018). Wide Open Spaces: A statistical technique for measuring space creation in professional soccer. *MIT Sloan Sports Analytics Conference 2018.* Boston.

FIFA. (2014). *Goal-Line Technology.* Retrieved from FIFA: https://football-technology.fifa.com/media/1011/goal-line_technology_testing_manual_2014.pdf

FIFA. (2017a). *Innovations/VAR*. Retrieved from FIFA: https://football-technology.fifa.com/en/media-tiles/video-assistant-referee-var-explained/

FIFA. (2017b, February). *FIFA strengthens global football integrity programme with Sportradar agreement*. Retrieved from FIFA: https://www.fifa.com/governance/news/y=2017/m=2/news=fifa-strengthens-global-football-integrity-programme-with-sportradar-a-2866086.html

FIFA. (2018b, June 22). *World Cup teams support FIFA's integrity initiative.* Retrieved from FIFA: https://img.fifa.com/image/upload/sflftaz9asa0wajwlcio.pdf

FIFA Communications Division. (2007, May 31). Retrieved from FIFA: https://www.fifa.com/mm/document/fifafacts/bcoffsurv/bigcount.statspackage_7024.pdf

Folgado, H., Duarte, R., Fernandes, O., & Sampaio, J. (2014). Competing with lower level opponents decreases intra-team movement synchronization and time-motion demands during pre-season soccer matches. *PLOS ONE, 10*(3).

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

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

football-technology FIFA. (2018). *EPTS Electronic performance and tracking systems.* Retrieved from football-technology fifa: https://football-technology.fifa.com/media/1031/epts_english.pdf

Forrest, D. (2012). The Threat to Football from Betting-Related Corruption. *International Journal of sport Finance, 7*, 99-116.

Forrest, D. (2013). Match Fixing: and economic perspective. In M. Haberfeld, & D. Sheehan, *Match-Fixing in international sports* (pp. 177-197). Switserland: Springer International Publishing.

Forrest, D., & McHale, I. (2015). *an evaluation of Sportradar's fraud detection system.* Sportradar.

Frencken, W., Lemmink, K., & Delleman, N. (2010). Soccer-specific accuracy and validity of the local position measurement (LPM) system. *Journal of Science and Medicine in Sport, 13*, 641-645.

Frencken, 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 Sport Science, 11*(4), 215-223.

Frencken, W., Poel, H., Visscher, C., & Lemmink, K. (2012). Variability of inter-team distances associated with match events in elite-standard soccer. *Journal of Sports Sciences, 30*(12), 1207–1213.

Fujimura, A., & Sugihara, K. (2005). Geometric Analysis and Quantitative Evaluation of Sport Teamwork. *Systems and Computers in Japan, 36*(6), 49-58.

Fushiki, T. (2011). Estimation of prediction error by usuing K-fold cross-validation. *Statistics and Computing*, 137-146.

Goddard, J. (2005). Regression models for forecasting goals and match results in association football. *International Journal of Forecasting, 21*, 331-340.

Goddard, J., & Asimakopoulos,, I. (2004). Forecasting football results and the efficiency of fixed-odds betting. *Journal of Forecasting, 23*, 51-66.

Gonçalves, B., Coutinho, D., Santos, S., Lago-Penas, C., Jiménez, S., & Sampaio, J. (2017). Exploring Team Passing Networks and Player Movement Dynamics in Youth Association Football. *PLoS ONE, 12*(1).

Gonçalves, B., Figueira, B., Maçãs, 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), 191-199.

Gorse, S., & Chadwick, S. (2010). Conceptualising corruption in sport: Implications for sponsorship programmes. *The European Business Review, July/August*. Retrieved April 18, 2018, from http://www.europeanbusinessreview.com/?p=1973

Gorse, S., & Chadwick, S. (2011). *The Prevalence of Corruption in International Sport: A statistical analysis.* Report prepared for the Remote Gambling Association and their Partners, The European Gaming and Betting Association and the European Sports Security Association. Retrieved from http://www.eu-ssa.org/essa_wp/wp-content/uploads/2013/08/The-Prevalence-of-Corruption-in-International-Sport.pdf

Gorse, S., Chadwick, S., & Byers, T. (2014). Emerging Trends in Corruption in Sport: Implications for Sport Management. *North American Society for Sport Management Conference* (pp. 143-144). Pittsburgh: nassm.

Gozse, I. (2015). Optical Indoor Positioning System Based on TFT Technology. *Sensors, 16*(1).

Graycar, A. (2015). Corruption: Classification and analysis. *Policy and Society, 34*(3).

Haberfeld, M., & Sheehan, D. (2013). *Match-Fixing in International Sports.* Switzerland: Springer International Publishing.

Hanson, R., & Oprea, R. (2009). A Manipulator can Aid Prediction Market Accuracy. *Economica, 76*(302), 304-314.

Harley, J. A., Lovell, R. J., Barnes, C. A., Portas, M. D., & Weston, M. (2011). The interchangeability of global positioningsystem and semiautomated video-based performance dataduring elite soccer match play. *Journal of Strength andConditioning Research, 25*, 2334–2336.

Headrick, J., Davids, K., Renshaw, I., Araujo, I., Passos, P., & Fernandes, O. (2012). Proximity-to-goal as a constraint on patterns of behaviour in attacker-defender dyads in team games. *Journal of Sports Sciences, 30*(3), 247-253.

Heilemann, T. (2014). Bestechlichkeit und Bestechung im sportlichen Wettbewerb als eigenständiges Strafdelikt – de lege lata, de lege ferenda. *Schriftenreihe Causa Sport, 8*.

Hill, D. (2010). *The Fix: soccer and organised crime.* Toronto: McClelland & Stewart.

Hiscock, D., Dawson, B., Heasman, J., & Peeling, P. (2012). Game movements and player performance in the Australian Football League. *International Journal of Performance Analysis in Sport, 12*(3), 531-545.

Hopkins, W. G. (2007). *Spreadsheet to Compare the Means of Two Groups.* Retrieved from Sportscience: sportsci.org/2007/inbrief.htm - xcl2

Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive Statistics for Studies in Sports Medicine and Exercise Science. *Medicine and Science in Sports and Exercise, 41*(1), 3-13.

Hvattum, L. M., & Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. *International Journal of Forecasting*, *26*(3), 460-470.

Hvattum, L. M., & Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. *International Journal of Forecasting*, *26*, 460-470.

INTERPOL. (2016, July 18). *More than 4,100 arrests in INTERPOL-led operation targeting Asian illegal gambling networks*. Retrieved from Interpol: https://www.interpol.int/News-and-Events/News/2016/More-than-4-100-arrests-in-INTERPOL-led-operation-targeting-Asian-illegal-gambling-networks

INTERPOL. (2019a, July). *Corruption in sport*. Retrieved from Interpol: https://www.interpol.int/en/Crimes/Corruption/Corruption-in-sport

INTERPOL. (2019b, July 9). *INTERPOL and IOC boost sports corruption investigations in the Nordic-Baltic region*. Retrieved from interpol: https://www.interpol.int/en/News-and-Events/News/2019/INTERPOL-and-IOC-boost-sports-corruption-investigations-in-the-Nordic-Baltic-region

Kannekens, R., Elferink-Gemser, M. T., & Visscher, C. (2011). Positioning and deciding: key factors for talent development in soccer. *Scandinavian Journal of Medicine & Science in Sports, 21*, 846-852.

Kapidžić, A., Mejremić, E., Bilalić, J., & Bečirović, E. (2010). Differences ins some parameters of situation efficiemcy between winning and losing teams at two levels of competition. *International journal of kinesiology, 7*(2).

Karlis, D., & Ntzoufras, I. (2003). Analysis of sports data by using bivariate Poisson models. *Statistician, 52,* 381-393.

Kerr, J. (2016). Big data, sport and match fixing [Recorded by J. Kerr, M. Pride, N. Bucci, H. Blake, & K. Rossmo].

Linke, D., Link, D., & Lames, M. (2018). Validation of electronic performance and tracking systems EPTS under field conditions. *PLoS ONE, 13*(7).

Low, B., Coutinho, D., Gonçalve, B., Rein, R., Memmert, D., & Sampaio, J. (2019). A systematic review of collective tactical behaviours in football using positional data. *Sports Medicine, 50*, 343–385.

Machado, V., Leite, R., Moura, F., Cunha, S., Sadlo, F., & Comba, J. L. (2017). Visual soccer match analysis using spatiotemporal positions of players. *Computers & Graphics, 68*, 84-95.

Macutkiewicz, D., & Sunderland, C. (2011, June). The use of GPS to evaluate activity profiles of elite women hockey players during match-play. *Journal of Sports Science, 29*(9), 967-973.

Maennig, W. (2002). On the economics of doping and corruption in international sports. *Journal of Sports Economics, 4*, 61-89.

Maennig, W. (2005). Corruption in International Sports and Sport Management: Forms, Tendencies, Extent and Countermeasures. *European Sport Management Quartely, 5*(2), 187-225.

Maher, M. (1982). Modelling association football scores. *Statistica Neerlandica, 36*, 109-118.

Masters, A. (2015). Corruption in sport: From the playing field to the field of policy. *Policy and Society, 34*(2), 111-123.

McGarry, T., Anderson, D. I., Wallace, S. A., Hughes, M. D., & Franks, I. M. (2002). Sport competition as a dynamical selforganizing system. *Journal of port Science, 20*(10), 771-781.

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

Mohr, M., Krustrup, P., & Bangsbo, J. (2003). Match performance of high-standard soccer players with special reference to development of fatigue. *Journal of Sports Sciences, 21*, 519–528.

Moroney, M. (1956). *Facts from Figures* (Vol. 3). London: Penguin.

My Football. (2013, September 16). *Victoria Police make arrests into alleged match-fixing*. Retrieved from myfootball: https://www.myfootball.com.au/news/victoria-police-make-arrests-alleged-match-fixing

Nettleton, J., & Chong, K. (2013, December 20). *Australia: Match-Fixing Criminal Offences – Guilty Pleas to Football Match-Fixing Charges.* Retrieved from addisonslawyers: http://www.addisonslawyers.com.au/knowledge/assetdoc/9692c98d9e8819ab/Australia%20Match%20Fixing%20Criminal%20Offences.pdf

Nowy, T., & Breuer, C. (2017). Match-fixing in European grassroots football. *European Sports Management Quarterly, 17*(1), 24-44. doi:https://doi.org/10.1080/16184742.2016.1193212

Numerato, D. (2016). Corruption and public secrecy: An ethnography of football match-fixing. *Current Sociology, 64*(5), 699-717.

Nye, J. (1967). Corruption and political development: A cost benefit analysis. *The American Political Science Review, 61*(2), 417-427.

Okabe, A., Boots, B., Sugihara, K., & Chiu, S. N. (2000). *Spatial tesselations: Concepts and applications of Voronoi diagrams.* New York: John Wiley & Sons, Inc.

Paoli, L., & Donati, A. (2013). *The sports doping market: Understanding supply and demand, and the challenges of their control.* New York: Springer.

Perumal, W. R., Righi, A., & Piano, E. (2014). *Kelong Kings.* Invisible Dog.

Pincus, S. M. (1991). Approximate entropy as a measure of system-complexity. *Proceedings of the National Academy of Sciences of the United States of America, 88*(6), 2297-2301.

Pincus, S. M., & Goldberger, A. L. (1994). Physiological timeseries analysis – what does regularity Quantify. *American Journal of Physiology, 266*(4), 1643–1656.

Poplu, G., Ripoll, H., Mavromatis, S., & Baratgin, J. (2008). How do expert soccer players encode visual information to make decisions in simulated game situations? *Research Quarterly for Exercise and Sport, 79*, 392 – 398.

Preston, I., & Szymanski, S. (2000). Racial discrimination in English football. *Scottish Journal of Political Economy, 47*(4), 342–363.

Quellette, J. (2004). Principles of play for soccer. *Strategies: A Journal for Physical and Sport Educators, 17*(3), 26.

Ralph, J. (2011, August 3). *How Melbourne tanked in 2009*. Retrieved from Herald Sun: https://www.heraldsun.com.au/sport/afl/how-melbourne-tanked-in-2009/news-story/5b0efacdf9da532f8daea1c1313d461c

Randers, M., Mujika, I., Hewitt, A., Santisteban, J., Bischoff, R., Solano, R., . . . Mohr, M. (2010). Application of four different football match analysis systems: a comparative study. *Journal of Sports Sciences, 28*(2), 171-182.

Reade, J. (2014). Detecting corruption in football. In J. Goddard, & P. Sloane, *Handbook on the economics of professional football* (pp. 419-447). Cheltenham: Edward Elgar.

Reade, J., & Akie, S. (2013). Using Forecasting to Detect Corruption in International Football. *Working Papers 2013-005, The George Washington University, Department of Economics, Research Program on Forecasting.* Retrieved from https://ideas.repec.org/p/gwc/wpaper/2013-005.html

Reep, C., Pollard, R., & Benjamin, B. (1971). Skill and chance in ball games. *Journal of the Royal Statistical Society*, 623-629.

Reilly, T., & Thomas, V. A. (1976). A motion analysis of work rate in different positional roles in professional football match-play. *Journal Human Movement Studies, 2*, 87-97.

Richardson, J. (2012, May 29). *Latest Italian match-fixing scandal met with both fury and weary resignation*. Retrieved from The Guardian: https://www.theguardian.com/football/2012/may/28/italian-match-fixing-fury-resignation

Richman, J., & Moorman, J. (2000). Physiological time-series analysis using. *American Journal of Physiology, 278*(6), 2039-2049.

Rodenberg, R. M., & Feustel, E. D. (2014). Forensic sports analytics: detecting and predicting match-fixing in tennis. *The Journal of Prediction Markets, 8*(1), 77-95.

Rogers, M. (2012). *Martin Rogers: Japan's women's soccer team plays to intentional draw.* Yahoo Sports.

Romero Clavijo, F., Corrêa, U., & Trindade Pinheiro Menuchi, M. (2017). Network of football players interactions according to the match period: a case study of the Bayern Munich vs. Real Madrid. *Complex Systems in Sport*, (pp. 127-129). Barcelona.

Sampaio, J., & Maçãs, V. (2012). Measuring tactical behaviour in football. *International Journal of Sports Medicine, 33,* 395–401.

Sampaio, J., & Maçãs, V. (2012). Measuring tactical behaviour in football. *Int J Sports Med, 33,* 395–401.

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

Schnitzer, K. (2017, May 23). *This is how a tennis match-fixing scandal works*. Retrieved from New York Post: https://nypost.com/2017/05/23/this-is-how-a-tennis-match-fixing-scandal-works/

Siegle, M., Stevens, T., & Lames, M. (2013). Design of an accuracy study for position detection in football. *Journal of Sports Sciences*, *31*(2), 166-172.

Sport and Recreation Ministers. (2011, June 10). *National Policy on Match-Fixing in Sport.* Retrieved from Australian Government Department of Health: http://www.health.gov.au/internet/main/publishing.nsf/Content/F6DB8637F05C9643CA257C310021CCE9/$File/National%20Policy%20on%20Match-Fixing%20in%20Sport%20(FINAL).pdf

Stagno, K. M., Thatcher, R., & Van Someren, K. A. (2007). A modified TRIMP to quantify the in-season training load of team sport players. *Journal of Sports Sciences, 25*(6), 629–634.

Stergiou, N., Buzzi, U., Kurz, M., & Heidel, J. (2004). Nonlinear tools in human movement. In N. Stergiou, *Innovative analyses of human movement: analytical tools for human movement research* (1 ed., pp. 63-90). Champaign: Human Kinetics.

Suarez-Arrones, L., Torreño, N., Requena B, B., Sáez De Villarreal, E., Casamichana, D., Munguía-Izquierdo, D., & Barbero-Alvarez, J. (2014). Match-play activity profile in professional soccer players during official games and the relationship between external and internal load. *The Journal of sports medicine and physical fitness, 55*(12), 1417-1422.

Svensson, J. (2005). Eight Questions about Corruption. *Journal of Economic Perspectives, 19*(3), 19-42.

Taki, T., & Hasegawa, J.-i. (2000). Visualization of dominant region in team games and its application to teamwork analysis. *Computer Graphics International 2000.* Geneva: IEEE.

Transparency International. (2018). *ANTI-CORRUPTION GLOSSARY*. Retrieved from transparency: https://www.transparency.org/glossary/term/corruption

U.S. Government. (2008, September). *Global Positioning System standard positioning service performance standard.* Retrieved from gps.gov: https://www.gps.gov/technical/ps/2008-SPS-performance-standard.pdf

UEFA. (2018, March 26). *UEFA and ESSA join forces to combat match-fixing*. Retrieved from UEFA: https://www.uefa.com/insideuefa/protecting-the-game/integrity/news/newsid=2546483.html

UEFA. (2019, June 13). *Integrity*. Retrieved from UEFA: https://www.uefa.com/insideuefa/protecting-the-game/integrity/

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), 121-127,.

Voronoi, G. (1907). Nouvelles applications des paramètres continus à la théorie des formes quadratiques [New applications of continuous parameters to the theory of quadratic forms]. *Journal für die Reine, 133*, 97-178.

Wade, A. (1996). *Principles of Team Play.* Spring City PA: Reedswain Inc.

Weeber, K. (1991). *Die unheiligen Spiele. Das antike Olympia zwischen Legende und Wirklichkeit. [The unholy games. The antique Olympics between legend and reality].* Munchen: Artemis & Winkler.

Wen-Bin Lin, & Mei-Yen Chen. (2015). A study of efficiency monitoring systems for match-fixing players in the Chninese Professional Baseball League. *European Sport Management Quarterly, 15*(3), 301-322.

Williamson, L. (2011, December 9). *Ajax demand probe into Lyon's incredible 7-1 Champions League victory*. Retrieved from Daily Mail: https://www.dailymail.co.uk/sport/football/article-2071803/Ajax-want-probe-Lyons-7-1-win-Dinamo-Zagreb.html

Wolfers, J. (2006). Point Shaving: Corruption in NCAA Basketball. *American Economic Review, 96*(2), 279-283.

Yentes, J., Hunt, N., Schmid, K., Kaipust, J., McGrath, D., & Stergiou, N. (2013). The appropriate use of approximate entropy and sample entropy with short data sets. *Annals of biomedical engineering, 41*(2), 349-365.

Appendices

Appendix A: Table of forecasting models for goal scoring and match results

	Model	**Forecasting**
(Moroney, 1956)	Poisson distribution	Goal scoring. One of the earliest works on modelling scores, suggesting improvements by the use of negative binominal distribution.
(Reep, Pollard, & Benjamin, 1971)	Negative binominal distribution	Goal scoring. Concluded that it was not possible to predict outcomes with their model because of the noise cause by chance in the game.
(Maher, 1982)	Univariate and bivariate Poisson distributions	Goal scoring. The first model that accounted for the strengths of the different teams involved.
(Dixon & Coles, 1997)	Univariate Poisson distributions	Goal scoring and match results. This model is built on the basic model of Maher (1982).
(Karlis & Ntzoufras, 2003)	Bivariate Poisson models	Difference between the scores of the two teams. This method is somewhere in between the goal scoring and match result approaches.
(Goddard & Asimakopoulos,, 2004)	Ordered probit regression model	Match results. The model was used to forecast match results directly without deriving it of goals scored.
(Goddard, 2005)	Bivariate Poisson regression & Ordered probit regression	Goal scoring and match results. Models for forecasting goals scored and match results have been compared and yield similar performance
(Hvattum & Arntzen, 2010)	Ordered logit regression models	Match results. Using ELO ratings to derive covariates that are used in the model.
(Reade & Akie, Using Forecasting to Detect Corruption in International Football, 2013)	Ordered probit regression model	Goal scoring and match results. This method further develops methods contained in Olmo et al. (2011) and Reade (2013).
(Reade & Akie, Using Forecasting to Detect Corruption in International Football, 2013)	Regression model	Residual distribution from models. In determining potential corrupt activity, they use residuals larger than three standard deviations.