Analysing Opponent Positioning Relative to the Ball: A Predictive Model for Football Possession Outcomes

Sam Casey-Popovich

A dissertation submitted to:

Auckland University of Technology

In partial fulfilment of the requirements for the degree of

Master of Sport, Exercise, and Health

2024

Faculty of Health and Environmental Sciences

Primary Supervisor: Dr Kirsten Spencer

Abstract

This dissertation investigates the factors influencing possession outcomes in football using quantitative research methodologies, specifically multinomial logistic regression. Although there is a considerable amount of literature that presents regression models to predict outcomes in football matches, none investigated opposition player positioning relative to the ball. The research aims to provide an objective measure of team performance by analysing various contextual and situational variables, in order to accurately predict possession outcomes.

The study focuses on the Wellington Phoenix Academy Men's Reserve Team during their 2023 Central League campaign. The research aims to provide an objective measure of team performance by analysing various contextual and situational variables. Routinely collected data, through foot-mounted inertial measurement units, was provided by the Wellington Phoenix Academy. A total of 923 ball possessions from a 15 game sample were analysed using a multinomial logistic regression model.

The findings suggested the positioning of the opponents relative to the ball does have an impact on the chances of creating successful possession outcomes. As the number of opponents behind the ball at the end of a possession increased, the probability of a successful outcome increases, when comparison against an unsuccessful outcome. Progressive passes and backwards passes were positively correlated and significantly predicted successful outcomes. Situational variables were analysed (adapted from literature on professional matches) however these were not effective predictors.

This study presented a novel way in which to value actions, and the findings indicated that this model is a sufficient predictor tool. Ultimately, these insights contribute to a deeper understanding of tactical decision-making and performance analysis in football, emphasizing

the utility of data-driven approaches in enhancing coaching practices. Coaching staff could implement this model to assess what actions in possession are most effective for creating goal scoring opportunities.

Table of Contents

List of Figures

List of Tables

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Signed_____

Sam Casey-Popovich

Acknowledgements

Firstly, to my family for their continuous support and encouragement throughout my tertiary studies. Thank you for always being there for me.

To my partner, Lynna, for always pushing me to do my best and for supporting me throughout. I would not have been able to manage both my work and my studies without your help and for that I am so thankful.

To the Wellington Phoenix Academy, thank you providing me the foundation and support to continue my tertiary studies. I am forever grateful to all the staff there for their warm and friendly working environment, in which I have learnt so much across the last three years.

To former Wellington Phoenix Academy Director, Steve Coleman, thank you for giving me an opportunity to work in my professional football. I could not have developed into who I am now without your guidance.

Lastly, to my primary supervisor, Dr Kirsten Spencer, thank you for endless support and encouragement during an unconventional process. Your positive approach and patience really helped me get this over the line, and for that I am incredibly grateful.

Chapter 1: Introduction

This quantitative observational study explores the contextual and situational variables associated with possession outcomes, in order to predict future possession outcomes. Goals are critical to winning games and coaches want to understand how goal scoring opportunities occur. Previous studies have successfully predicted match and possession outcomes, however this has not included the impact of the opposition. This dissertation aims to provide a potential framework that can objectively measures the value of in-possession actions that positively correspond with successful and unsuccessful in-possession outcomes.

Football or soccer, is an invasion game that is regarded as the most popular sport in the world (Singh & Lamba, 2019). Football's need for complex and dynamic actions from players, means that performance analysis has become a vital part of the coaching process and it is a key component to understanding and evaluating how teams play.

Performance analysis in football has seen a significant shift in the quality of data generated over the past 15 years, this is due to the introduction of tracking data (De Pablo et al., 2019; Low et al., 2019; Sarmento et al., 2014). Positional and spatial data availability has allowed for research to better understand tactical nuances and trends while in possession (Andrienko et al., 2019; Ric et al., 2017). Furthermore, research has now integrated this data, alongside key action variables, to predict possession and match outcomes (Cao, 2024; Bailey, 2005; Felices, 2023; McCarthy et al., 2022; Prasetio, 2016; Saravia et al., 2016). These researchers have made significant progress in helping us to understand the complexity of game states and how teams play depending on the situation. Although there is previous research that considers space and pressure in relation to opposing players, there is no research, (to the author's knowledge), that has used opposition positioning in order to value the action within a possession.

The identification of variables that correlate with goal scoring has been a key topic within relevant literature (Acar et al., 2008; Lepschy et al., 2021; Wright et al., 2011). Goals are the most important variable as this is how games are won, thus, coaches want to know the most effective way to score goals and prevent their team from conceding them. Although key variables have been identified in research, the on-going issue, as stated by Mackenzie and Cushion (2014), is that a universal agreement upon the importance of specific variables is yet to be achieved. Due to the complex nature of the sport, defining the importance and value of specific variables is subjective and is highly dependent on the intention of the coaches. Performance models presented in research are specific to the team analysed and may not always corroborate findings with another team, due to the various ways in which teams create goal scoring opportunities. This is an important limitation in relevant literature as is the influence of the opposition is a key component when applying these models into coaching practice. As football is a high pressure and high demanding environment, coaching staff within an organisation may change regularly, moreover, it is common for coaches to frequently adapt and develop their style of play or coaching philosophy overtime (Diquigiovanni & Scarpa, 2019). For a performance model to be successful long-term, it needs to consider objective measures that are not specific a coach's style of play and can be integrated into any coaching set up.

1.1 Research Aim and Purpose

The aim of this study is to investigate the possessions and possession outcomes of the Wellington Phoenix Academy Men's Reserve Team performances in the 2023 Central League season. The two key questions underpinning this research are:

1. How does the distribution of opponents behind and in front of the ball influence the prediction of a successful possession outcome?
2. Do situational variables influence possession outcome?

The key outcome of this research is to:

1. Provide a performance model that accurately predicts possession outcome using objective performance measures.

1.2 Structure of Report

Chapter two of this report provides a review of performance analysis in football, with a key focus on tactical and technical behavioural research, and the quantification of performance indicators. Furthermore, contemporary research analysing the association of performance indicators with match outcomes is a key focus of this literature review. Finally, the application of performance data, through statistical models is explored to examine the extent of implementation into coaching practice.

Chapter three is an outline of the research methodology. It is quantitative research and defines the specific method used for the data visualisation. This chapter provides information on the data collection, procedures, data analysis, participants, ethical considerations and the background of the researcher.

Chapter four presents the results of the regression analysis, the justification for the model used and absolute frequency of variables and multinomial logistic regression are presented. Key trends and patterns in the data are identified.

Chapter five is the discussion section of this dissertation. This section explains the trends identified in the data. Key ideas are discussed and related back to relevant literature.

Chapter six concludes this dissertation. It summarises the findings within the research conducted, and presents limitations and recommendations for future research.

Chapter 2: Literature Review

2.1 Introduction

Football, or soccer, is an invasion game that is regarded as the most popular sport in the world (Singh & Lamba, 2019) and with approximately 130,000 male professional football players worldwide (FIFA, 2023). This sport demands a high level of tactical, technical, physical, and psychological skill from players (Lepschy et al., 2018). Due to the complex and dynamic nature of the sport, goal scoring, which is the primary objective within the game, is not common. Therefore, teams cannot be accurately assessed on goal scoring alone, and other factors during games need to be investigated in order to understand how teams are classified as successful.

Performance analysis has become a highly influential tool in professional environments to better understand the nature of football matches (Lepschy et al., 2018; Clemente, 2012). Research within this field has been apparent since the 1960s (Carling et al., 2005), however, a key limitation of performance research in football has been the universal agreement to objectively measure performance variables (Carling et al., 2014). Due to the complexity of tactical behaviours from different coaches and teams, the learnings found in observational data research will not always accurately reflect how other teams should play. For example, many researchers have looked at how teams create chances, specifically in high-level tournaments, such as the FIFA World Cup (Acar et al., 2008; Lepschy et al., 2021; Simiyu, 2013), these articles suggested that the most successful teams created chances from wide areas with crosses. Although this does highlight a possible indicator for actions of successful teams, it does not factor the quality of the teams playing or the favoured tactical strategies of the coaches and therefore cannot conclude that crossing will be successful for all teams. Because of this, research that aims to create an objective measure for performance indicators cannot solely focus on quantifying the actions leading up to a goal, but must also consider the tactical and technical behaviours of both teams to understand how goals are created.

The purpose of this chapter is to provide a comprehensive scope of current research within the field of performance analysis in football, with a key focus on contextual variables used to measure and predict the success of team possessions. Although there is extensive research focused on objective measures to value in-possession actions, a universal agreement, in relation to the importance of the variable being measured and how variables are measured, is yet to be established (Carling et al., 2014).

This literature review provides a broad overview of performance analysis in football, with a key focus on tactical and technical behavioural research, and the quantification of performance indicators. Furthermore, contemporary research analysing the association of performance indicators with match outcomes will also be a key focus of this literature review. Finally, the application of performance data, through statistical models, will also be explored to examine the extent of implementation into coaching practice.

2.2 Performance Analysis in Football

Performance analysis in team sport is believed to have originated in basketball and American football in the 1960s, it was then adopted by football soon after (Carling et al., 2005). From this, performance analysis has become a vital part of the coaching process in football, as it is now recognised by coaches, players, and supporting staff as a key part of the preparation, reflection and feedback process (Drust, 2010). Memmert and Raabe (2017) have defined the evolution of performance analysis in football into four stages:

- Match Analysis 1.0 is the first phase of performance analysis, and this initial phase was limited to frequencies of basic statistics in football (shots, passes, etc.).
- Match Analysis 2.0 is said to have started in 1988, where analysis turned to more subjective assessment of game situations through the lens of a coach or “expert”.

- Match Analysis 3.0 and 4.0 define the transition of focus from video-specific data to positional data. Through the use of new technologies, such as wearable sensors , GPS tracking, and machine-learning algorithms, both technical, tactical, and physiological data can now be streamlined to (semi-) automatically produce more in-depth data of game situations and individuals, in real-time of post event (Memmert & Raabe, 2017).

Although professional environments now have a full-time role dedicated to performance analysis, there is still a major discrepancy between research and application (Goes et al., 2021; Mackenzie & Cushion, 2013; Wright et al., 2014). The dynamic and complex nature of a game environment makes it difficult to quantify specific actions in football, and McLean et al. (2017) states that research in this area does not consider these factors. Because of this, there is not a universal agreement in terms of how actions of quantified, as different teams will value specific actions differently in relation to their style of play. Thus, quantifying these actions has to consider the influence of technical and tactical behaviours of both teams to understand how teams create goal scoring opportunities.

Analysis of game of situations and individual performances is primarily driven through the investigation of Key Performance Indicators (KPIs). A KPI is defined as an action variable that correlates or influences aspects of successful performances (Hughes & Bartlett, 2002). The significance and type of KPIs will vary for each team environment, based on the tactical philosophy of the coach. Analysis of KPIs allows for teams to gather information on tactical and technical behaviours and the strengths and weaknesses of themselves and the opposition, greater informing the determinants of a positive match outcome and enhancing coach decision-making (Memmert et al., 2017).

Previous research in tactical behaviours in football has highlighted the effectiveness of positional data to accurately analyse and understand the dynamics of a football match (De Pablo et al., 2019; Low et al., 2019; Sarmento et al., 2014). One key focus within this area is

spatial analysis (Andrienko et al., 2019). This research identified both space and time dimensions as a possible variable to better understand the dynamics of a football match, stating that spatial configurations and player movements will change in response to different phases of match situations. This framework also supports the ideas presented by Ric et al., (2017), which suggested that successful teams maintain greater spaces between teammates which allows for greater passing sequences and progression closer to the opposition goal. Although these articles do show tactical nuances within possessions, it does not factor space relative to the opposition. Pressing styles and formations of the opposition will dictate the space afforded to a player. Therefore, positional data of opposition players must also be considered when investigating spatial analysis, in order to more accurately inform the coaches decision-making process.

Another key element of tactical behaviours that has benefitted from the introduction of positional data is passing networks (Caicedo-Parada et al., 2020). Passing networks look into the sequence of passes within a possession, providing greater context compared to when looking at passes individually (Buldú et al., 2018). The same study concluded that a greater number of passes within a network correlated with successful performance outcomes, such as goal scoring opportunities and maintaining the majority of possession throughout the game. This also corroborated findings from Mendes et al. (2018) which concluded that teams with greater passing sequences had more positive match outcomes (wins or draws) in comparison to teams with low-count passing sequences. This shows that analysing passing sequences is a fundamental KPI to effectively assess team possessions. However, more contextual variables need to be considered when analysing passing sequences, as there are variations in the way teams can create passing sequences (Caicedo-Parada et al., 2020).

The effectiveness of analysing passing networks highlights the emergence of network sciences and its ability to interpret sports data (Ma et al., 2021). Several studies have used network science to analyse passing networks through different scopes, through the variability

of players, positions on the pitch, and both factors simultaneously (Cintia et al., 2015; Grund, 2012; Narizuka et al., 2014). Studies that focused on player passing networks investigated network properties, such as centrality and clustering, and their relationship with KPIs associated with positive possession outcomes (Grund, 2012). Moreover, studies that analysed passing networks in relation to pitch location found corresponding positional patterns and positive possession outcomes (Cintia et al., 2015). Pitch-player passing network analysed both the player and their positioning during passing moments, in order to identify if positioning influences passing behaviours and their role within the network (Felices, 2023; Narizuka et al., 2014). Buldú et al., 2018 proposes a framework that analyses four dimensions of the passing network: dynamics, space, time, and layers. This variation uses the same applied methods as the pitch-player passing network, yet expands on this through the combination of opponents passing network, allowing for analysis of tactical behaviours being influenced by the opponents structure. This approach to network analysis demonstrates the importance of contextual variables, more specifically, factoring in variables associated with the opposition, to better assess the dynamic situations in a football match. By factoring the context of the opposition, the analysis can focus on how this influences a team's objective to keep the ball and execute successful outcomes from possessions.

Transitions are another key phase of possession that coaches and teams will prioritise within their style of play. A transition is defined as a quick change in possession, whether from attack to defence or defence to attack, with the objective to rapidly capitalise on either winning the ball back or attacking the opponent's goal (Hewitt et al., 2016). Numerous studies have shown that offensive transitions are an effective means to create goal scoring opportunities and accounts for a high conversion rate of shots to goals (Eusebio et al., 2024; Yiannakos & Armatas, 2006). A study by Gonzalez-Rodenas et al. (2016) showed that during transition moments, the tactical and positional behaviour of individuals on both sides created a possession outcomes only specific to these moments during games. This shows that by

understanding tactical factors related to successful transitions, teams can understand how these moments influence opponent's positioning and the opportunities it creates for goal scoring opportunities. This also enhances the idea that understanding possession dynamics is achieved in greater detail when factoring in variables related to the opposition. Transition to attack moments have a high chance of creating goal scoring opportunities as these moments create instability in which opposing players are having to quickly transition to defend. Analysing what players do in these situations will inform the coach on how effective their team is in utilising these opportunities in games.

Set piece actions are defined as all situations that restart the game after the ball has gone out of play or a player has been fouled (e.g. corners, free kicks, throw-ins, penalty kicks). Due to the amount of control a team can have in these situations, in terms of dictating player positioning and ball delivery destination, set pieces are considered one of the most important elements of tactical strategy for some coaches (Casal et al., 2015; Maneiro et al., 2021). The importance of set pieces has been highlighted in studies analysing goals scored in Men's International tournaments (Acar et al., 2008; Vergonis et al., 2021; Yiannakos & Armatas, 2006). During the 2018 FIFA Men's World Cup, 43% of all goals came from set piece actions (Vergonis et al., 2021), and a similar percentage was found during the 2004 European Championship with 35.6% of all goals scored coming from set pieces (Yiannakos & Armatas, 2006). These statistics highlight the significance of set pieces during important games, therefore, it is important for coaches to understand what actions lead to set piece situations.

Different types of set-piece actions are considered more important than others for the attacking process in football, primarily based on its success for creating a goal scoring opportunities (Casal et al., 2015; Dizdar et al., 2016). One of which is the corner kick, due to this always being in the attacking half and frequently leading, either directly or indirectly, to a cross into the opposition penalty box. Findings by Kubayi and Larkin (2020) supported this idea, identifying a positive correlation between winning teams and the number of corner kicks they

are awarded in games. Furthermore, Casal et al. (2015) stated that a goal scored from a corner either won or drew the game on 76% of occasions. Although there is clear importance in set piece actions, and corners in particular, analysing the removal and addition of opposing players during set piece actions will most likely not provide value to the action due to the clustered nature of attacking set pieces. Primarily during corners and in-direct freekicks, both attacking and defending teams will put the majority of their outfield players in the penalty box, in doing so, almost all of the outfield players are now within a small area of the pitch and distances between players will be small. Therefore, it is more appropriate to identify the value of possession variables that result in a set piece action, rather than attempt to value the set piece action itself.

Goal kicks are another type of set piece action and are defined as a moment in which a team will restart play after the opposing team has kicked the ball over the defending teams by-line (TheFA, n.d.). Traditionally, teams would use these moments to kick the ball long to gain territorial advantage, however, a recent trend in football is for goal kicks to now be a short grounded kick to a teammate, with the aim to invite and play through opposing pressure (Casal et al., 2023; Mendes et al., 2020). A study by Casal et al. (2024) investigated the offensive participation of goalkeepers in a league setting and concluded that goalkeeper actions on the ball are more effective, in relation to creating goal scoring opportunities, when executing short passes to teammates in comparison to kicking the ball long. Contrarily, another study looked into goal kicks at the Men's 2020 European Competition and found that long goal kicks were more effective in achieving possession progression (Mendes et al., 2020). This contrast in findings may relate to the difference in competition, as one compares one team across various games and one compares multiple teams across fewer games. This, alongside the overall lack of research on football goal kicks, highlights a key area in football that needs to be investigated further. Although it is classified as a set piece, goal kicks will initiate a start in possession play and therefore should not be considered a possession outcome, rather, a possession start.

Analysing possessions starting from short goal kicks will give a clear indication of how successful teams are at playing through an opposing teams pressure deep in their defending half and the outcomes of possessions started from short goal kicks.

2.3 Predicting Match Outcomes in Football

Through the use of technical and tactical behaviours, and the integration of positional data, contemporary research has now looked into predicting match outcomes based on specific variables. A team's main objective is to score a goal and many studies conclude that teams have a greater chance of scoring goals if they create a high frequency of goal scoring opportunities, such as shots, crosses, and touches in the opposition penalty box (Kan et al., 2004; James et al., 2004; Wright et al., 2014). However, insights from these behaviours is limited without the understanding of how these chances are created (Mackenzie & Cushion, 2013). Contemporary research has investigation two categories of variables to predict how match outcome and how teams score goals: analysing actions/moments leading up to goal scoring chances (contextual variables) and events/influences outside of the two teams playing styles (situational variables) (Ballesteros et al., 2012; Mendes et al., 2018; McCarthy et al., 2022; Pollard, 2008; Prasetio, 2016; Saravia et al., 2016).

Situational variables have been studied to show a significant influence on match outcome (Almeida et al., 2014; Liu et al., 2016; Pollard, 2008). The first example of situational variables investigated was a comparison of team performance when playing at their home venue and team performances when playing in away venues. Pollard (2008) identified multiple factors relating to home advantage, including crowd influence, territoriality, and travel effects. The same study states that these factors are known to coaches and players and in turn, psychologically effects them before and during games. Other studies found that playing at home increases team cohesion during possession plays (Mendes et al., 2018). Overall, when compared to other sports, home advantage in football was deemed more influential on match

outcome (Jamieson, 2010). There are no statistics available that detail average crowd attendance for the National League – Regional Phase games in Wellington, however, pitch and venue familiarity, alongside travel affect, may influence performance and therefore needs to be taken into account.

Team quality is another situational variables that has been investigated heavily in football (Ballesteros et al., 2012: Bloomfield et al., 2015; Srgo & Lipoma, 2016). In most competitive leagues in professional football, there is significant disparity in technical proficiency between the top ranked and lower ranked teams (Sgro & Lipoma, 2016). Therefore, studies have looked into why this disparity is consistent in seasonal competitions. One example is the Bloomfield et al. (2015) study of the English Premier League 2003-2004 season, which identified that the top three teams in the league dominated possession statistics whether they were winning, losing, or drawing. A potential reason as to why top teams dominate possession was highlighted in a study by Liu et al., (2016), which found that the players in higher ranked teams in the Spanish first division had more ball touches, passes, through balls, higher passing accuracy and assists, when compared to the rest of the teams in the division. Furthermore, another study of the Spanish first division found that higher ranked teams also had a higher frequency of winning the ball back in higher zones of the pitch, whereas the most common area to win the ball back for the rest of the teams in the division was in zones closer to their own goal (Fernandez-Navarro et al., 2020). It is important to note that this study was only conducted across a 10 game sample, meaning the findings may vary if conducted across a larger sample size. Schwartz and Barsky (1977) concluded that both game location and team quality were equally important in determining match outcome in sport due to the home advantages higher ranked teams would have over lower ranked teams. Because of this, both key situational variables will most likely have an effect on performance and in doing so, must be considered simultaneously when study team performance.

Another example of situational variables are ones that can happen during the match, with the most commonly studied one being scoring first in a match. Due to the nature of football games being low scoring, with most games ending with three goals or less (Anderson & Sally, 2013), every goal scored highly influences match outcome (Liu et al., 2021). Win percentages analysed in football tournaments have supported this idea. Vergonis et al. (2019) looked into win percentage for teams that scored first at the 2018 FIFA Men's World Cup and found that 71.17% of team won their match after scoring first. Additionally, Molinuevo and Bermejo (2012) found that across a five year period (2005 to 2010) in the Spanish first division, 79.32% of teams that scored first won their match. The suggested reason for this high percentage is matches won when scoring first is due to its influence on team dynamics, as the competitive balance of the game breaks, both psychologically and tactically, players and coaches will feel a need to change strategies or styles of play to alter the score outcome (Almeida et al., 2014; Ruiz-Ruiz et al., 2013). Evidently, the score of the game may influence a team's style of play and thus must be a factor to consider when analysing a team's style of play when they have possession.

Although situational variables have a considerable effect on performance, coaches and researchers are primarily concerned with the technical and tactical variables associated with how teams play and their correlation with positive possession outcomes, such as goals, shots, and touches in the opposition penalty box. Understanding how possession play turns into these goal scoring opportunities has been looked into using the concept of perturbations (Fischer et al., 2022; Kim et al., 2019a; Link et al., 2016). Perturbations are the triggers that alter a dynamic system and changes a stable state into one that is unstable (Fischer et al., 2022). This concept promotes the idea that game performance of both teams during a game should be considered simultaneously, as the actions of one team will trigger a response in the other team (Mackenzie & Cushion, 2013; Kim et al., 2019a). Link et al. (2016) presents a quantification of possession play by considering 4 key factors of goal scoring: pitch location, time/space for the player on the

ball, pressure from an opposing player, and the density of players in proximity of the ball and the goal. These findings highlight that perturbations have significant importance when understanding the transition of game states and the quantification of said states.

Kim et al. (2019b) conducted research that identified that teams create unstable situations differently. The stability of these situations were based on the location of possession and the positioning of the opposing players to operationalise these situations. The framework of this study presented multiple contextual variables that influenced the change in game stability, with the frequencies presented to show what actions were most utilised to create goal scoring chances. Results showed that in the team studied, they predominately used the wide areas to progress possession play into goal scoring chances, crosses and touches in the opposition penalty box. These were identified as the key variables influencing the transition in game stability. This study provided an innovative approach to classifying the key moments of possession play that evolve to goal scoring opportunities, however, findings are limited by the lack of situational variables and opposition context. These factors would need to be considered to more accurately predict match outcome.

Another factor to be considered is the number of contextual variables or KPIs used when predicting match outcome. Lago-Peñas et al. (2017) measured 20 variables to identify which playing styles were utilised in possession. Whereas Kim et al. (2020) measured 23 variables to analyse the defence performance of wide and central defenders. The only study that directly discusses the appropriate amount of variables to use is Lepschy et al. (2018). This study used 25 variables to measure goal scoring, passing, and defence, concluding that 25 variables was an appropriate number to accurately analyse these sequences. Although there is no definitive agreement on the amount of variables required for sufficient analysis, these studies provide a reasonable scope for the amount of variables needed and should be considered in future studies.

With an understanding of which variables affect game context and possession moments, research in performance analysis has extensively looked into predicting possession and match outcomes through the use of regression models (Cao, 2024; Bailey, 2005; Felices, 2023; McCarthy et al., 2022; Prasetio, 2016; Saravia et al., 2016). Bailey (2005) was one of the first accounts of predicting football match outcomes. Using a linear regression model, he analysed situational variables (distance travelled, home advantage, etc.) to predict match success, and concluded a 66.7% prediction accuracy. Another study by Snyder (2011) used a similar model with more variables regarding the team performance context (matches won or loss, goal difference, etc.) and concluded a relatively similar prediction accuracy of 51.6%. These studies highlight the influence of game context on match performance, but they also show that regression models are a sufficient predictor model to apply within football research.

With the introduction of tracking data, prediction research has now looked into variables within the match to understand how teams win. Prasetio (2016) categorically analysed match performance by identifying four game phases: Home Offense, Home Defence, Away Offense, Away Defence. The findings from this study suggested that teams with high a percentage of winning the ball back during defence phases, both home and away, were more likely to win games, however, findings were limited by the lack of context within these game moments. The latest evolution of prediction research has been the investigation of passing patterns and their accuracy to predict goal scoring opportunities (Cao, 2024; Felices, 2023; McCarthy et al., 2022). Through the use of passing networks and logistic regression models, this research shows that advanced statistical models can be applied to effectively evaluate passing patterns and identify which variables are most significant for prediction. Ultimately, this provides a substantial foundation for future research and indicates key areas of focus. For example, Cao (2024) stated a limitation of their study was that it did not consider opposition positioning during passing patterns, emphasising the idea that opposition positioning, in relation to the ball, can identify what passing patterns are most effective to break down defences.

It is important to note that not all research into style of play analysis uses a quantitative methodology. Previous researchers have used a qualitative approach to understandstyle of play analysis and its application in coaching practice (Aguado-Mendez et al., 2021; Sarmento et al., 2020). For example, Sarmento et al. (2020) used a mixed methods approach to compare statistical insight and the perspective of the coaching staff. Network centrality analysis showed central midfielders as key recipients to progress play towards the opposing goal and identified the pitch zones immediately before the opposition's penalty box as key receiving areas to create goal scoring opportunities. Semi-structured interviews with the coaching staff revealed that this was the intended style of play and thus corroborated the findings of the network data to accurately determine the team's style of play. This is an insightful approach to style of play analysis, as most research does not consider if the trends they identify are the intended patterns the coaches want to implement. Although it is not included in this study, future research should consider comparing the trends identified and the intentions of the coaches.

2.4 Application of Statistical Models

Evidently, there is extensive research that investigates the variables related to teams creating goal scoring opportunities and thus show how teams win games. However, Hewitt et al. (2016) stated that these discoveries in patterns are only useful if they successfully inform the coach and effectively shape the training and preparations of the team. This idea is highlighted in the complex statistical models and programming researchers use to conduct and present their research. Although the models are effective in identifying trends in football, in most cases, the applicability of these models in a football teams coaching environment is not viable (Terblanche, 2020). Football environments are often faced-paced and dynamic, as coaches analyse opposition and team performance on a weekly basis. This means that tactics and context of games are always changing, and thus, the time between collating and delivering information to influence coaching practice is critical.

Another factor to consider is the coach. One idea highlighted by Butterworth and Woodward (2023) was that coaches are reluctant to apply statistical models due to the lack of empirical evidence. Due to footballs “unpredictable” nature, it is common for coaches to believe statistics cannot effectively capture the nature of a football match. However, statistical models have been shown to be important for pedagogical practice and may critically inform decision-making if applied efficiently (Pupo-Ge & Hernández-Ávila, 2020). Lastly, it is important to consider applicability in the long-term, as the complexity of statistical models will result in extensive work to generate and analyse data. When performance analysis research focuses on one team, it most cases it identifies which patterns of play are most effective to create goal scoring opportunities and concludes a model grounded in these identified KPIs. However, the application of this model may become redundant if a new coach fills the position and creates new possession patterns, resulting in the model generating process needing to start again. Because of this, a broader, football-wide, objective data model, that does not consider passing patterns an indicator alone, needs to be considered in order for a model to be applied sustainably.

2.5 Conclusion

In view of these literature findings, a study on the positioning of opponents relative to the ball during ball possessions in football is justified. There are evidently numerous studies that have looked into tactical and technical variables to provide context to positive possession outcomes. Although key limitations in all these findings are the contextual variables identified. These variables are specific to one team’s or one coach’s style of play. The findings for one team or one coach will not always associate with other teams and coaches, therefore, the variables identified as key performance indicators, or the performance models suggested, will not only be limited in their use to other coaches and teams, but may become obsolete over time, once that team has brought in a new coach, or the same coach has installed a new style

of play. Because of this key limitation, the aim of this research is to create a performance model that identifies objective KPIs that are not specific to a team or coaches' style of play and thus can be integrated across team setup at any professional level.

Predicting match outcome has been a common research question in relevant literature, and the studies reviewed show that analysing situational variables can effectively predict match outcome. Because of this, it is important that this study considers the context of matches analysed and to investigate if this has an effect on playing patterns and team performance throughout a competitive season.

Chapter 3: Methodology

3.1 Introduction

The aim of this chapter is to outline the research methodology for this study, data collection, procedures and data analysis. This chapter will also provide information on the participants, ethical considerations and the background of the researcher.

3.2 Quantitative Research

The methodology used for this study was a quantitative research design. Quantitative research is a systematic investigation that focuses on quantifying variables in order to collect and analyse numerical data (Snyder & Bisch, 1989). This data is then used to find patterns, make predictions or examine relationships, ultimately answering the 'how' and 'why' to a question. When conducting research, the process should be guided by the ontological and epistemological beliefs of the researcher. Quantitative research sits within the positivist paradigm, as it seeks to uncover the objective truth of existing reality and explain phenomena through natural laws (Bayley, 2013). There are several forms of quantitative research approaches that can be used when conducting study for objective truth, these approaches can range from identifying behaviours through manipulation, to comparing previously collected data with newly collected data (Snyder & Bisch, 1989). Deciding which research method to use is predominantly dependant on its relativeness to the aim(s) of the research.

3.3 Research Design

As the aims of this research requires manipulation of independent variables to measure cause and effect on a dependant variables, its falls within the scope of an experimental design (Myers et al., 2013). The method chosen for this research was a multinomial logistic regression model (MLogit). MLogit is a statistical model that predicts the probability of multiple categorial

outcomes (dependant variables) using predictor (contextual and situational) variables (Hosmer et al., 2014), MLogit is an extension of the binary logistic regression model and can be thought of as J-1 binary logit model, where J equals the number of categories for the outcome variable (Long et al., 2014). This logit function predicts the log odds, or logits, of a case falling into a different a different category on the dependant variable when compared against a baseline (reference) category, representing a non-linear transformation of the probability of a case falling into the alternative category (reference category).

This study will regress possession outcomes (0 = successful outcome, 1 = possession continued, 2 = unsuccessful outcome) onto final result, match location, team quality, match status, pitch location of possession start, pitch location of possession outcome, possession duration, progressive passing count, progressive passing score, progressive dribbling count, progressive dribbling score, backwards passing count, backwards passing score, and opponents behind the ball at the end of possession. This MLogit model is comprised of two binary logit models where (a) successful outcomes and (b) possession continued are compared against a (c) reference (baseline) category comprised of unsuccessful outcomes.

3.4 Background of the Researcher

I am performance analyst that has worked in semi-professional and professional football environments in New Zealand for the past 4 years. I have worked at both club and international level. My role within these environments has primarily been supporting the coaching staff and players through video and data analysis. Weekly tasks include analysing upcoming opposition, review game day footage, and provide statistical insight of player performance to both coaches and players.

I wanted to continue studying and complete a post graduate degree after working in these environments as they highlighted the importance of data-driven processes in coaching practice. Focusing my study on statistical models and data analysis will improve my

understanding of how data is interpretated and applied into a sport setting. Therefore, I can take these learnings into football environments, with the aim to efficiently implement data-driven processes into the coaching practice.

3.5 Sample

Any athlete that played a competitive game for the Wellington Phoenix Academy Men's Reserve Team during their 2023 Central League campaign was considered a participant. Most participants are considered semi-professional footballers, as they compete in a league in which players are not contracted on a full-time basis. Some players were considered professional footballers as they had either signed a professional contract with the Wellington Phoenix First Team during the campaign or had been dropped from the first team to play in for the reserve team. In total, 28 players participated in this study.

3.6 Sample Size

Data was collected from the Wellington Phoenix Academy Men's Reserve Team during their 2023 Central League campaign in the New Zealand National League – Regional Phase. Data from 15 competitive games was collected. Although 18 games were played during this campaign, 3 games had to be excluded due to poor video quality.

3.7 Ethical and Cultural Considerations

Before the commencement of this research, the academy director of the Wellington Phoenix was presented all the relevant information of the proposed report and agreed to supply the routinely collected data from the Men's Reserve Team (Appendix A). The team were given a comprehensive information sheet and asked to signed consent forms if they agreed to the conditions of the research (Appendix B; Appendix C). The information sheet contained relevant information on the purpose and procedure of the research, including the detail of their in-

possession actions be valued and ranked. It was also emphasised that participation was completely voluntary and no tracking data would be used without the consent of the athlete.

To uphold anonymity, all data that was routinely collected and shared for this project was de-identified, ensuring that player identification could not be connected back to the data given. Data security was also upheld, with all data and video collected stored on an external hard drive that was not within access to anyone.

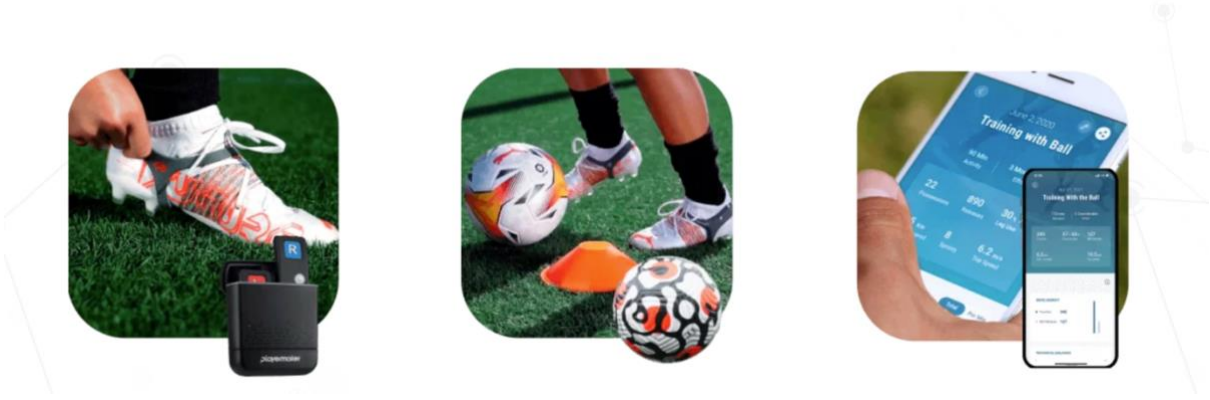
When considering possible ethical dilemmas, it is important to balance key principles of ethical research to ensure integrity is upheld (Vanclay et al., 2013). The principle of non-maleficence was a key principle I wanted to communicate to my participants, due to the evaluation and ranking of subjective actions in football. I wanted to ensure that all participants understood that this process was for the aim of creating an objective measure to quantify football actions, rather than a tool to critique and rank player ability.

This study was ethically approved by the Auckland University of Technology Ethics Committee (AUTEC) on the 31st January 2024 (Appendix D).

3.8 Data Access and Procedure

This project used pre-existing data which has been collected using an innovative foot-mounted inertial measurement unit (see fig 1; IMU; PlayerMaker™, Tel Aviv, Israel).

Figure 1. *Photo of Playermaker unit (https://www.playermaker.com/us/how-it-works/)*



Data has been routinely collected by Wellington Phoenix Academy Men's Reserve Team during the 2023 Central League competition. Each IMU includes two components from the MPU- 9150 multi-chip motion tracking module (InvenSense, California, USA), being a 16 g triaxial accelerometer and a 2000°•sec−1 triaxial gyroscope. Housed in manufacturer- supplied ergonomic silicone straps. Each player was equipped with two IMUs (one for each foot), which were located on the outside of the players boots using the silicone straps. To diminish issues related to inter-unit reliability, players used the same IMUs throughout the data collection period.

All data and footage from the 18 competitive games were collected from the Wellington Phoenix Academy via an external hard drive. All data was collected in XML format and all player information was de-identified. The next phase of data collected was a variable labelling process using Hudl Sportcode (version 12.36.0). Corresponding game XML and footage were synced into Hudl Sportcode to create a Sportcode package that had all team possessions and passing counts. Using the label mode, pitch location of possession start and end, progressive passes, progressive dribbles, backwards passes, opponents behind the ball at the end of possession and outcomes (goal, shot on target, shot blocked, shot off target, corner awarded, possession continued, turnover – ball in play, and turnover – ball out of play) were labelled onto the possession instances. All match information variables (final result, match location, team quality, match status) were attached during this process.

3.9 Data Reliability

This study used an intra-observer reliability test. Once all playermaker data was coded by the researcher, a randomised 10% (n = 92) of the total possessions analysed were selected to be reassessed. Cohen's Kappa coefficient calculation (Cohen, 1960) was used to quantify the intra-observer reliability of the data collected by the researcher. Reliability of each

contextual variable's data is presented in table 1, with the dribbling score presenting the lowest value (0.88), indicating excellent reliability according to Fleiss et al. (2003).

3.10 Research Questions

1. How does the distribution of opponents behind and in front of the ball influence the likelihood of a successful possession outcome?
2. Do situational variables influence the outcomes of possession?

3.11 Data Analysis

After completing the labelling process, the Sportscode packages are downloaded into CSV and manually arranged into rows according to the variables listed in table 1. The CSV files are then combined into one master file and exported into SPSS Statistics (version 29.0.2.0) for analysis.

Research Question 1

Statistics were generated using the multinomial logistic regression tool in SPSS Statistics. Findings were presented in beta coefficients (β), *p*-value, and odds ratios. A positive beta coefficient will identify the extent of high probability of the dependant variable (successful outcome and possession continued) when compared to directly with the reference category (unsuccessful outcomes). The odds ratio provides insight into the decrease and/or increase in likelihood based on the predictor variable measured, in relation to the reference category within the categorical predictor variables. The level of significance for each test was set at 5% (p = 0.05%) as per Di Leo and Sardanelli (2020).

For this research question, analysis specifically look into the pitch location start and end, progressive pass count and score, progressive dribble count and score, backwards pass count and score, and opposition behind the ball at the end of the possession.

Research Question 2

Using the same data generation as before, this research question looks into the variables defined as situational. This included match result, match location, team quality, match status. These variables were then directly compared to the possession outcomes to analyse if there was influence on performance.

Figure 2. *Code Window*

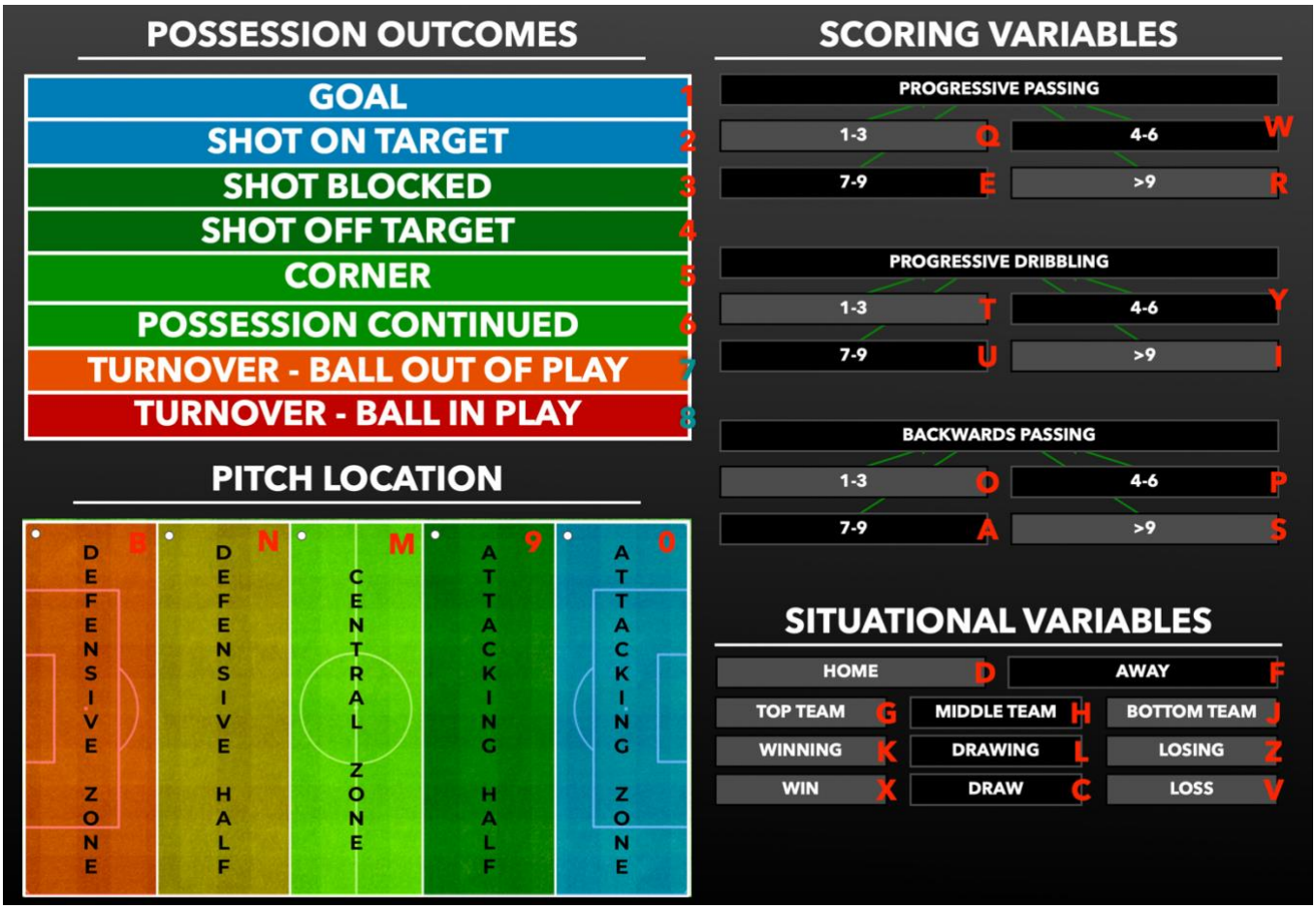


Table 1. *Operation Definitions*

Variables	Category
Match Result	**Win:** The observed team has scored more goals than the opposition after the full time whistle
	Draw: The observed team has scored the same amount goals as the opposition after the full time whistle
	Loss: The observed team has scored less goals than the opposition after the full time whistle
Location	**Home:** The observed team is playing at their home ground
	Away: The observed team is playing at the home ground of an opposing team
Team Quality	**Top Team:** The three teams that finished in the top half of the Central League 2023 competition table, after 18 matches have been played **Middle Team:** The four teams that finished in the middle of the Central League 2023 competition table, after 18 matches have been played **Bottom Team: :** The four teams that finished bottom of the Central League 2023 competition table, after 18 matches have been played
Match Status	**Winning:** The observed team has scored more goals than the opposition, at the time of this possession
	Drawing: The observed team has scored the same amount of goals as the opposition, at the time of this possession **Losing:** The observed team has scored less goals than the opposition, at the time of this possession
Pitch Location of Possession Start	**Defending Zone:** The first action of the possession is performed in the defending Zone
	Defending Half: The first action of the possession is performed in the defending Half
	Central Zone: The first action of the possession is performed in the Central Zone
	Attacking Half: The first action of the possession is performed in the Attacking Half
	Attacking Zone: The first action of the possession is performed in the Attacking Zone
Pitch Location of Possession Outcome	**Defending Zone:** The last action of the possession is performed in the defending Zone
	Defending Half: The last action of the possession is performed in the defending Half
	Central Zone: The last action of the possession is performed in the Central Zone
	Attacking Half: The last action of the possession is performed in the Attacking Half
	Attacking Zone: The last action of the possession is performed in the Attacking Zone
Possession Duration	**0-30 Seconds:** The time between the first and last action in the possession is equal to or less than 30 seconds
	31-60 Seconds: The time between the first and last action in the possession is greater than 30 seconds and equal to or less than 60 seconds **61-90 Seconds:** The time between the first and last action in the possession is greater than 61 seconds and equal to or less than 90 seconds **>90 Seconds:** The time between the first and last action in the possession is greater than 91 seconds
Progressive Passing	Player A completes a successful pass to Player B, with the ball travelling in the direct towards the opponent's goal (attacking end) Any opponents that were between Player A and B once the pass is completed are considered "removed" from defending (they can no longer defend until they are in front of the ball in relation to the goal they are defending) Anytime a pass is played forward that "removes" an opponent or opponents, a progressive pass is counted
Count	**0:** No progressive passes were completed before the last action of the possession
	1-3: A count of 1-3 progressive passes were completed during the possession duration
	4-6: A count of 4-6 progressive passes were completed during the possession duration
	7-9: A count of 7-9 progressive passes were completed during the possession duration
	>9: A count of progressive passes greater than 9 were completed during the possession duration
Score	**0:** No opponents were removed from the game, through progressive passing, during the possession duration
	1-3: 1-3 opponents were removed by progressive passing during the possession duration
	4-6: 4-6 opponents were removed by progressive passing during the possession duration

	7-9: 7-9 opponents were removed by progressive passing during the possession duration
	>9: More than 9 opponents were removed by progressive passing during the possession duration
Progressive Dribbling	Any dribbling action that "removes" opponents from defending when the dribbling action is taking place
Count	**0:** No progressive dribbles were completed during the possession duration
	1-3: A count of 1-3 progressive dribbles were completed during the possession duration
	4-6: A count of 4-6 progressive dribbles were completed during the possession duration
	7-9: A count of 7-9 progressive dribbles were completed during the possession duration
	>9: A count of progressive dribbles greater than 9 were completed during the possession duration
Score	**0:** No opponents were removed from the game, through progressive dribbles, during the possession duration
	1-3: 1-3 opponents were removed by progressive dribbling during the possession duration
	4-6: 4-6 opponents were removed by progressive dribbling during the possession duration
	7-9: 7-9 opponents were removed by progressive dribbling during the possession duration
	>9: More than 9 opponents were removed by progressive dribbling during the possession duration
Backwards Passing	Player A completes a successful pass to Player B, with the ball travelling in the direct towards the team's goal (defending end) Any opponents that were between Player A and B once the pass is completed are considered "added" back into defending (now they are in front of the ball in relation to the goal they are defending) Anytime a pass is played forward that "adds" an opponent or opponents, a backward pass is counted
Count	**0:** No backward passes were completed during the possession duration
	1-3: A count of 1-3 backward passes were completed during the possession duration
	4-6: A count of 4-6 backward passes were completed during the possession duration
	7-9: A count of 7-9 backward passes were completed during the possession duration
	>9: A count of backward passes greater than 9 were completed during the possession duration
Score	**0:** No opponents were added back into the game during the possession duration
	1-3: 1-3 opponents were added back into the game, through backward passes, during the possession duration
	4-6: 4-6 opponents were added back into the game, through backward passes, during the possession duration
	7-9: 7-9 opponents were added back into the game, through backward passes, during the possession duration
	>9: More than 9 opponents were added back into the game, through backward passes, during the possession duration
Opponents Behind The Ball At The End Of Possession	**0:** There are no opponents behind the ball, in relation to opponents defending goal, when the possession ends
	1-2: There are between 1 and 2 opponents behind the ball, in relation to opponents defending goal, when the possession ends
	3-5: There are between 3 and 5 opponents behind the ball, in relation to opponents defending goal, when the possession ends
	6-9: There are between 6 and 9 opponents behind the ball, in relation to opponents defending goal, when the possession ends
	10: There are 10 opponents behind the ball, in relation to opponents defending goal, when the possession ends

Chapter 4: Results

4.1 Introduction

This chapter outlines the results generated from interpretating the data. Firstly, a Cohen's Kappa value test and model fitting tables is presented to show the validity and justification of the coding process and the statistical model used. A total of 923 ball possessions across a 15 game sample were analysed and categorised based on the outcome. These findings will be displayed using an absolute frequencies w table, showing percentages based on possession outcome type. Finally, the data generated from the MLogic model is presented in a table, with findings represented in beta coefficients, *p*-value, and odds ratio.

4.2 Data Validity and Statistical Model Justification

A total of 92 possessions were randomly selected for re-evaluation. Table 2 shows the results of the Kappa value test, with Progressive Dribbling Count (0.87) being the lowest valued variable and 6 of 17 variables being valued exact (1.00) when compared to initial evaluation. Any value over 0.81 is considered almost perfect agreement (Fleiss et al., 2003), as all variables are above this threshold, the data generating process is deemed reliable.

Table 2. *Intra-observer Reliability Test using Cohen's Kappa Calculation*

Criteria	Intra-rater value
Pitch Location of Possession Start	1.00
Pitch Location of Possession Outcome	0.95
Progressive Passing Count	0.98
Progressive Passing Score	0.94
Progressive Dribbling Count	0.87
Progressive Dribbling Score	0.90
Backwards Passing Count	0.95
Backwards Passing Score	0.96
Opponents Behind The Ball At The End Of Possession	0.91
Goal	1.00
Shot On Target	1.00
Shot Blocked	0.98
Shot Off Target	0.98
Corner Awarded	1.00
Ball Possession - Set Piece	1.00
Turnover - Ball In Play	0.95
Turnover - Ball out of Play	1.00
Ktotal	**0.96**

The overall fit of the model was evaluated in Table 3 and Table 4. Evaluation was conducted using the likelihood ratio chi-square test and Pearson and Deviance chi-square test. The likelihood ratio chi-square test was statistically significant ($\chi^2(6) = 3321.556$, $p<.001$) indicating the model, when containing the full set of predictors, fits the data significantly better than a null or intercept-only. Conversely, the Pearson chi-square test ($\chi^2(1576)=2341.664$, $p=1.00$) and the Deviance chi-square test ($\chi^2(1576)=1990.807$, $p=1.00$) were both deemed statistically insignificant. However, this reinforces the model being a good fit to the data (Long & Freese, 2014).

Table 3. *Model Fitting Information for Multinomial Logistic Regression*

	Model Fitting Criteria			Likelihood Ratio Tests		
Model	**AIC**	**BIC**	**-2 Log Likelihood**	**Chi-Square**	**df**	**Sig.**
Intercept Only	5345.409	5362.452	5339.409			
Final	2293.853	3077.845	3077.845	3321.556	135	<.001

Table 4. *Goodness of Fit*

	Chi-Square	**df**	**Sig.**
Pearson	2341.664	3459	1.00
Deviance	1990.807	3459	1.00

4.3 Contextual and Situational Variables as Predictors of Possession Outcomes

Table 5. *Absolute Frequencies and Percentage Occurrence of Possession Outcome Variables*

	Possession Outcomes							
	Successful Possession Outcome						Unsuccessful Possession Outcome	
	Goal	Shot On Target	Shot Blocked	Shot Off Target	Corner Awarded	Ball Possession Set Piece	Turnover - Ball In Play	Turnover - Ball Out of Play
Final Result								
Win	14 (56.0%)	6 (27.3%)	12 (41.4%)	21 (40.4%)	28 (41.2%)	110 (31.9%)	96 (25.2%)	113 (35.0%)
Draw	5 (20.0%)	4 (18.3%)	8 (27.6%)	12 (23.1%)	17 (25.0%)	75 (21.7%)	69 (18.2%)	89 (27.5%)
Loss	6 (24.0%)	12 (54.4%)	9 (31.0%)	19 (36.5%)	23 (33.8%)	160 (46.4%)	215 (56.6%)	121 (37.5%)
Match Location								
Home	12 (48.0%)	7 (31.8%)	12 (41.4%)	19 (36.5%)	33 (48.5%)	157 (45.5%)	188 (49.5%)	137 (42.4%)
Away	13 (52.0%)	15 (68.2%)	17 (58.6%)	33 (63.5%)	35 (51.5%)	188 (54.5%)	192 (50.5%)	186 (57.6%)
Team Quality								
Top Team	5 (20.0%)	7 (31.8%)	8 (27.6%)	16 (30.8%)	19 (27.9%)	116 (33.6%)	156 (41.0%)	89 (27.5%)
Middle Team	9 (36.0%)	9 (40.9%)	12 (41.4%)	14 (26.9%)	26 (38.3%)	143 (41.5%)	145 (38.2%)	99 (30.7%)
Bottom Team	11 (44.0%)	6 (27.3%)	9 (31.0%)	22 (42.3%)	23 (33.8%)	86 (24.9%)	79 (20.8%)	135 (41.8%)
Match Status								
Winning	11 (44.0%)	6 (27.3%)	13 (44.8%)	11 (21.2%)	19 (27.9%)	106 (30.7%)	69 (18.1%)	94 (29.1%)
Drawing	11 (44.0%)	6 (27.2%)	12 (41.4%)	21 (40.4%)	38 (55.9%)	116 (33.6%)	139 (36.6%)	125 (38.7%)
Losing	3 (12.0%)	10 (45.5%)	4 (13.8%)	20 (38.4%)	11 (16.2%)	123 (35.7%)	172 (45.3%)	104 (32.2%)
Pitch Location of Possession Start								
Defending Zone	6 (24.0%)	4 (18.2%)	11 (21.1%)	11 (21.2%)	6 (24.0%)	45 (13.0%)	65 (17.1%)	57 (17.6%)
Defending Half	11 (44.0%)	9 (40.9%)	15 (51.5%)	18 (34.6%)	11 (44.0%)	140 (40.6%)	142 (37.4%)	116 (35.9%)
Central Zone	4 (16.0%)	4 (18.2%)	4 (13.7%)	12 (23.1%)	22 (32.4%)	102 (29.6%)	130 (34.2%)	92 (28.5%)
Attacking Half	3 (12.0%)	4 (18.2%)	4 (13.7%)	9 (17.3%)	9 (13.2%)	36 (10.4%)	41 (10.8%)	38 (11.8%)
Attacking Zone	1 (4.0%)	1 (4.5%)	0 (0.0%)	2 (3.8%)	5 (7.4%)	22 (6.4%)	2 (0.5%)	20 (6.2%)
Pitch Location of Possession Outcome								
Defending Zone	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (1.5%)	32 (9.3%)	17 (4.6%)	29 (9.0%)
Defending Half	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	7 (10.2%)	116 (33.6%)	63 (16.6%)	99 (30.7%)
Central Zone	0 (0.0%)	0 (0.0%)	1 (3.5%)	2 (3.9%)	25 (36.8%)	93 (27.0%)	120 (31.6%)	91 (28.1%)
Attacking Half	4 (16.0%)	8 (36.4%)	13 (44.8%)	10 (19.2%)	18 (26.5%)	58 (16.8%)	101 (26.6%)	65 (20.1%)
Attacking Zone	21 (84.0%)	14 (63.6%)	15 (51.7%)	40 (76.9%)	17 (25.0%)	46 (13.3%)	79 (20.8%)	39 (12.1%)
Possession Duration								
0-30	17 (68.0%)	10 (45.6%)	20 (69.0%)	40 (76.9%)	48 (70.5%)	281 (81.4%)	258 (67.9%)	263 (81.4%)
31-60	5 (20.0%)	11 (50.0%)	5 (17.2%)	10 (19.3%)	15 (22.1%)	57 (18.5%)	104 (27.4%)	52 (16.1%)
61-90	3 (12.0%)	1 (1.9%)	4 (13.8%)	1 (1.9%)	4 (5.9%)	6 (1.7%0	15 (3.9%)	8 (2.5%)
>90	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (1.9%)	1 (1.5%)	1 (0.3%)	3 (0.8%)	0 (0.0%)
Progressive Passing Count								
0	5 (20.0%)	5 (22.8%)	8 (27.6%)	16 (30.8%)	10 (14.7%)	99 (28.7%)	63 (16.6%)	77 (23.8%)
1-3	18 (72.0%)	11 (50.0%)	17 (58.6%)	31 (59.6%)	49 (72.0%)	203 (58.8%)	262 (68.9%)	219 (67.8%)
4-6	2 (8.0%)	5 (22.8%)	3 (10.4%)	4 (7.7%)	7 (10.3%)	41 (11.9%)	43 (11.3%)	20 (6.2%)
7-9	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (1.9%)	1 (1.5%)	0 (0.0%)	9 (2.4%)	5 (1.5%)
>9	0 (0.0%)	1 (3.4%)	1 (3.4%)	0 (0.0%)	1 (1.5%)	2 (0.6%)	3 (0.8%)	2 (0.7%)

Progressive Passing Score								
0	5 (20.0%)	5 (22.7%)	8 (27.6%)	8 (15.4%)	10 (14.7%)	99 (28.7%)	63 (16.6%)	77 (23.8%)
1-3	7 (28.0%)	1 (4.5%)	4 (13.8%)	10 (19.2%)	22 (32.4%)	83 (24.1%)	107 (28.2%)	97 (30.0%)
4-6	6 (24.0%)	8 (36.4%)	8 (27.6%)	7 (13.4%)	14 (20.6%)	81 (23.5%)	115 (30.3%)	80 (24.8%)
7-9	3 (12.0%)	4 (18.2%)	4 (13.8%)	11 (21.2%)	12 (17.6%)	38 (11.0%)	43 (11.3%)	35 (10.8%)
>9	4 (16.0%)	4 (18.2%)	5 (17.2%)	16 (30.8%)	10 (14.7%)	44 (12.8%)	52 (13.7%)	34 (10.5%)
Progressive Dribbling Count								
0	14 (56.0%)	10 (45.5%)	21 (72.4%)	33 (63.5%)	42 (61.8%)	231 (67.0%)	248 (65.3%)	203 (62.8%)
1-3	11 (44.0%)	12 (54.5%)	8 (27.6%)	18 (34.6%)	26 (38.2%)	111 (32.1%)	132 (34.7%)	117 (36.3%)
4-6	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (1.9%)	0 (0.0%)	3 (0.9%)	0 (0.0%)	3 (0.9%)
7-9	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
>9	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Progressive Dribbling Score								
0	13 (52.0%)	10 (45.5%)	21 (72.4%)	29 (55.8%)	38 (55.9%)	201 (58.4%)	212 (55.8%)	168 (52.0%)
1-3	10 (40.0%)	11 (50.0%)	7 (24.1%)	21 (40.4%)	30 (44.1%)	126 (36.6%)	141 (37.1%)	133 (41.2%)
4-6	1 (4.0%)	1 (4.5%)	1 (3.4%)	1 (1.9%)	0 (0.0%)	16 (4.7%)	27 (7.1%)	19 (5.9%)
7-9	1 (4.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.3%)	0 (0.0%)	3 (0.9%)
>9	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Backwards Passing Count								
0	20 (80.0%)	10 (45.5%)	21 (72.4%)	40 (76.9%)	36 (52.9%)	221 (64.0%)	250 (65.8%)	198 (61.3%)
1-3	3 (12.0%)	11 (50.0%)	7 (24.2%)	10 (19.3%)	28 (41.3%)	106 (30.7%)	119 (31.3%)	109 (33.7%)
4-6	1 (4.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (2.9%)	14 (4.1%)	5 (1.3%)	7 (2.2%)
7-9	1 (4.0%)	1 (4.5%)	1 (3.4%)	2 (3.8%)	2 (2.9%)	3 (0.9%)	5 (1.3%)	6 (1.9%)
>9	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.3%)	1 (0.3%)	3 (0.9%)
Backwards Passing Score								
0	18 (72.0%)	9 (40.9%)	21 (72.4%)	38 (73.1%)	34 (50.0%)	219 (63.5%)	249 (65.5%)	196 (60.7%)
1-3	2 (8.0%)	6 (27.3%)	3 (10.4%)	5 (9.6%)	24 (35.2%)	70 (20.2%)	77 (20.3%)	67 (20.7%)
4-6	2 (8.0%)	5 (22.7%)	2 (6.9%)	5 (9.6%)	5 (7.4%)	32 (9.3%)	33 (8.7%)	31 (9.6%)
7-9	2 (8.0%)	0 (0.0%)	2 (6.9%)	1 (1.9%)	1 (1.5%)	13 (3.8%)	11 (2.9%)	13 (4.0%)
>9	1 (4.0%)	2 (9.1%)	1 (3.4%)	3 (5.9%)	4 (5.9%)	11 (3.2%)	10 (2.6%)	16 (5.0%)
Opponents Behind The Ball At The End Of Possession								
0	3 (12.0%)	8 (36.4%)	10 (34.6%)	18 (34.6%)	22 (32.4%)	134 (38.8%)	140 (36.8%)	120 (37.2%)
1-3	6 (24.0%)	4 (18.2%)	8 (27.6%)	10 (19.2%)	23 (33.8%)	93 (27.0%)	112 (29.5%)	77 (23.8%)
4-6	4 (16.0%)	2 (9.1%)	7 (24.1%)	8 (15.4%)	10 (14.7%)	45 (13.0%)	47 (12.4%)	48 (14.9%)
7-9	4 (16.0%)	3 (13.6%)	3 (10.3%)	9 (17.3%)	6 (8.8%)	45 (13.0%)	51 (13.4%)	53 (16.4%)
9-11	8 (32.0%)	5 (22.7%)	1 (3.4%)	7 (13.5%)	7 (10.3%)	28 (8.2%)	30 (7.9%)	25 (7.7%)

Table 6. *Multinomial Logistic Regression to Predict Successful Outcomes, Possession Continued, vs Unsuccessful Outcomes (adapted from Casal et al. 2024)*

	Possession Outcomes							
	Successful Outcomes				Possession Continued			
Predictor	**β**	***P***	Odds Ratio	CI (95%)	**β**	***P***	Odds Ratio	CI (95%)
Final Result								
Win#								
Draw	-0.393	0.185	0.675	0.377 - 1.208	-0.348	0.141	0.706	0.444-1.122
Loss	-1.197	0.027	0.302	0.105 - 0.87	-0.487	0.201	0.614	0.291-1.296
Match Location								
Home#								
Away	0.563	0.015	1.755	1.115-2.764	0.285	0.1	1.33	0.947-1.868
Team Quality								
Top Team#								
Middle Team	-0.767	0.057	0.464	0.211-1.024	-0.131	0.609	0.877	0.53-1.45
Bottom Team	-1.23	0.026	0.292	0.099-0.865	-0.918	0.016	0.399	0.19-0.841
Match Status								
Winning#								
Drawing	-0.156	0.522	0.855	0.53-1.381	-0.312	0.11	0.732	0.499-1.073
Losing	-0.502	0.132	0.605	0.315-1.164	-0.439	0.08	0.645	0.395-1.054
Pitch Location of Possession Start								
Defending Zone#								
Defending Half	-0.15	0.593	0.861	0.497-1.49	0.568	0.013	1.765	1.129-2.76
Central Zone	-0.707	0.018	0.493	0.274-0.886	0.649	0.008	1.914	1.182-3.098
Attacking Half	-0.323	0.348	0.724	0.368-1.421	0.743	0.014	2.103	1.162-3.803
Attacking Zone	-0.712	0.144	0.491	0.189-1.275	1.642	<.001	5.165	2.389-11.17
Pitch Location of Possession Outcome								
Defending Zone#								
Defending Half	0.778	0.477	2.178	0.254-18.652	-0.293	0.323	0.746	0.418-1.333
Central Zone	2.056	0.05	7.815	1.003-60.905	-0.831	0.007	0.436	0.239-0.796
Attacking Half	2.958	0.005	19.268	2.491-149.035	-1.139	<.001	0.32	0.169-0.608
Attacking Zone	4.122	<.001	61.662	8.049-472.413	-1.134	<.001	0.322	0.164-0.631
Possession Duration								
0-30	-0.529	0.314	0.589	0.21-1.651	0.773	0.153	2.166	0.75-6.257
31-60	-0.691	0.181	0.501	0.182-1.378	0.453	0.402	1.573	0.545-4.542
61-90#								
>90	0.648	0.622	1.912	0.146-25.095	0.629	0.638	1.875	0.136-25.846

Progressive Passing Count								
0#								
1-3	1.351	0.171	3.862	0.559-26.671	-0.015	0.987	0.985	0.149-6.504
4-6	0.457	0.625	1.58	0.253-9.868	0.257	0.782	1.294	0.208-8.029
7-9	-0.833	0.479	0.435	0.043-4.361	14.605	0.981	4.54E-07	*
>9								
Progressive Passing Score								
0#								
1-3	-1.909	0.062	0.148	0.02-1.103	-0.409	0.678	0.664	0.096-4.579
4-6	-1.711	0.095	0.181	0.024-1.35	-0.336	0.733	0.715	0.104-4.933
7-9	-1.293	0.211	0.274	0.036-2.079	-0.024	0.981	0.977	0.14-6.798
>9	-0.724	0.453	0.485	0.073-3.21	0.097	0.918	1.102	0.173-6.997
Progressive Dribbling Count								
0#								
1-3	1.061	0.019	2.889	1.188-7.028	0.182	0.515	1.199	0.694-2.071
4-6	-6.407	0.981	0.002	6.124E-238-4.442E+231	13.103	0.959	490361	1.308E-212-1.838E+223
7-9								
>9								
Progressive Dribbling Score								
0#								
1-3	-0.849	0.049	0.428	0.184-0.998	-0.342	0.17	0.71	0.436-1.157
4-6	-2.974	<.001	0.051	0.011-0.237	-0.648	0.13	0.523	0.226-1.209
7-9	6.932	0.98	1024.4	3.824E-232-2.744E+237	-13.352	0.958	1.59E-06	4.219E-224-5.988E+211
>9	3.874	*	48.149	48.149-48.149	4.198	0.999	66.584	*
Backwards Passing Count								
0#								
1-3	-2.373	0.012	0.093	0.015-0.596	-0.628	0.531	0.534	0.075-3.805
4-6	-2.345	0.053	0.096	0.009-1.035	0.718	0.521	2.051	0.228-18.424
7-9	-1.271	0.301	0.281	0.025-3.12	-0.695	0.623	0.499	0.031-7.969
>9	-15.593	0.982	1.69E-07	*	0.351	0.836	1.421	0.051-39.955
Backwards Passing Score								
0#								
1-3	2.453	0.011	11.619	1.761-76.66	0.59	0.558	1.805	0.25-13.038
4-6	2.447	0.014	11.555	1.631-81.859	0.497	0.627	1.643	0.222-12.156
7-9	2.047	0.052	7.746	0.978-61.332	0.351	0.738	1.42	0.182-11.063
>9	2.523	0.023	12.471	1.417-109.782	-0.373	0.754	0.688	0.067-7.107
Opponents Behind The Ball At The End Of Possession								
0#								
1-3	0.056	0.826	1.057	0.644-1.735	0.066	0.724	1.068	0.742-1.536
4-6	0.178	0.549	1.195	0.667-2.139	0.157	0.499	1.17	0.742-1.844
7-9	-0.482	0.126	0.617	0.333-1.145	0.027	0.909	1.027	0.65-1.622
9-11	0.037	0.92	1.038	0.499-2.159	0.274	0.373	1.315	0.72-2.401

Unsuccessful outcome is the outcome reference category

#, Variable reference category; B, Beta Coefficient; CI, Confidence Interval; $p < 0.05$; *, insufficient data

Final Result

As expected, 56% of the goals scored in this study (n = 14) were scored in games that resulted in a match win. Moreover, there was a significant increase in turnover – ball in play (56.6%, n = 215) in matched that resulted in a loss. The multinomial analysis shows that there is statical significance between loss and unsuccessful outcomes (p = 0.027), with the odds of possessions continuing are 0.302 time lower than unsuccessful outcomes (CI: 0.105-0.87).

Match Location

Match location does not show any major discrepancy in frequency data when comparing home and away matches. The only significant difference in frequencies was that the team produced more shots on target (68.2%, n = 15) and shots off target (63.5%, n = 35) during away matches. Away matches did significantly increase the odds of possession continuing, when compared to unsuccessful outcomes, by 1.755 times (CI: 1.115-2.764) and this was deemed a statistically significant result (p = 0.027). Away matches also had a high odds ratio for possession continued and unsuccessful outcomes at 1.33 times (CI: 0.947-1.868). However, this result was deemed not significant (p = 0.1)

Team Quality

The quality of the teams played against did have an impact on the frequency of variables measured. The majority of goals were scored against the bottom (44%, n = 11) and middle (36%, n = 9) teams. Furthermore, turnover – ball in play had a much higher frequency against the top (41%, n = 156) and middle (38.5%, n = 145) teams . Table 4 shows possession continuing is 0.464 times lower in comparison to top teams.

Match Status

The Wellington Phoenix Men's Reserve Team only scored 3 goals (12%) while in losing positions. It is also interesting to note that 55.9% of all crosses awarded (n = 38) when there was an equal score line and turnovers – ball in play were significantly lower when winning (18.1%, n = 69). Regression analysis shows that losing decreases the odds of a successful outcome by 0.605 times (CI: 0.315-1.164), however, it is not statistically significant (p = 0.132). Similarly, drawing decreased the odds of keeping the ball at the end of a possession (possession continued) by 0.732 times (CI: 0.499-1.073) and also was statistically insignificant with a p value of 0.609.

Pitch Location of Possession Start

Possessions starting in the defensive half increased the odds of possession continuing by 1.765 when compared to possessions started in the defending zone, with a p-value of 0.013, indicating a positive and significant predictor of possession continuing. The most significant predictor for possessions continuing was the attacking zone with a highly significant p-value of <0.001 and an odds ratio of 5.165 (CI: 2.389-11.17) when compared to the defending zone. Lastly, the central zone was the most significant (p = 0.018) when comparing against the defending zone in successful outcomes. An odds ratio shows that a successful possession outcome is 0.492 times more likely when starting from the central zone, in comparison to starting in the defending zone.

Pitch Location of Possession Outcome

The beta coefficient for attacking zone location in successful possession outcomes (β = 4.122) shows that there is a high chance that a possession ending in this pitch location will end in a successful outcome. Furthermore, the p-value (<0.001) indicates that this is a high significant result. Attacking half was also highly significant (p = 0.005), with an odds ratio of

19.28 (CI: 2.491-149.035) when compared to the defending zone. *P*-value was also highly significant for possessions continued that ended in the attacking half (p = <0.001) and attacking zone (p = <0.001).

Possession Duration

The regression analysis shows that there were no significant results when investigating the time duration of possessions. Odds ratios indicated that a possession lasting longer than 90 seconds were 1.912 (CI: 0.146-25.095) times more likely to continue possession when compared to possessions lasting between 61 and 90 seconds, however, this test was deemed insignificant (p = 0.622). Similarly, possessions lasting less than 30 seconds were 2.166 times (CI: 0.75-6.257) more likely to continue possession when compared to possessions lasting between 61 and 90 seconds. This test was also not significant (p = 0.153).

Progressive Passing Count

The beta coefficient for passes count 1-3 (β = 1.351) shows that there is a high probability of progressive passing count 1-3 resulting in a successful outcomes as opposed to an unsuccessful outcome. The odds ratio shows that progressive passing count 1-3 increases the odds of a successful outcome by 3.862 times (0.559-26.671), when compared to a possession that had 0 progressive passes. Ultimately, this result was not significant (p = 0.171).

Progressive Passing Score

A progressive passing score of 1-3 had a significantly lower probability of possession ending in a successful outcome when compared to an unsuccessful outcome (β = -1.909) and is considered a marginally significant result (p = 0.062). A progressive passing score of 4-6 also showed a low probability of possession ending in successful outcome (β = -1.711) and had marginal significance (p = 0.095). An odd ratio of 1.102 times (CI: 0.173-6.997) shows a positive

correlation progressive passing scores exceeding 9 and possession continuing as opposed to a possession that does not receive a progressive score (0).

Progressive Dribbling Count

A possession that has a count of progressive dribbles between 4-6 is shown to have a very low probability of achieving a successful outcome when compared against an unsuccessful outcome (β = -6.407). The *p*-value of this result was highly significant (*p* = 0.002). Contrarily, a possession that has a count of progressive dribbles between 4-6 is shown to have a very high probability of a continuing possession in comparison to an unsuccessful outcome (β = 13.103). However, the *p*-value of this test (*p* = 0.958) showed it was not significant.

Progressive Dribbling Score

A progressive dribbling score of 4-6 was deemed highly significant (*p* = < 0.001) and showed a considerably low probability of achieving a successful outcome when compared to unsuccessful outcomes (β = -2.974). There was significance (*p* = 0.049) and low probability ((β = -0.849) for progressive dribbling scores between 1-3, when comparing successful and unsuccessful outcomes. A progressive dribbling score of 1-3 also showed a low probability when comparing possession continued and unsuccessful outcome (β = -0.628), however, this was not d significant (*p* = 0.534).

Backwards Passing Count

Backwards passing counts exceeding 9 showed a substantially low probability for successful outcomes as opposed to unsuccessful (β = -15.593), however, this test was not significant (*p* = 0.982). The odds ratio for 4-6 backwards passing count ending in possession continued shows this type of outcome is 2.051 times more likely as opposed to 0 backwards passes. This result was also not significant (*p* = 0.521).

Backwards Passing Score

All backwards passing scores ending in a successful outcome showed a high probability when compared to unsuccessful outcomes (1-3, β = 2.453; 4-6, β =2.447; 7-9, β = 2.047; >9, β = 2.523). All of these results were significant (1-3, p = 0.011; 4-6, p = 0.014; 7-9, p = 0.052; >9, p = 0.023). Odds ratios support these trends, a backwards passing score exceeding 9 increases the odds of a successful outcome by 12.471 (CI: 1.417-109.782) as opposed to a backwards score of 0.

Opponents Behind the Ball at the End of Possession

All the odds ratios for possession continued showed a positive correlation when compared to 0 (1-3, OR = 1.068, CI: 0.742-1.536; 4-6, OR = 1.17, CI: 0.742-1.844; 7-9, OR = 1.027, CI: 0.65-1.622; 9-11, OR = 1.315, CI: 0.72-2.401). All p-values for possession continued were not significant ($p < 0.05$). Possessions ending with 7-9 opponents behind the ball has a lower probability (β = -0.482) in ending in an successful outcome as opposed to a unsuccessful outcome. However, this result was also not significant (p = 0.126).

Chapter 5: Discussion

5.1 Opposition Positioning in Relation to the Ball at the End of Possessions

The aim of this study was to investigate the possessions and possession outcomes of the Wellington Phoenix Academy Men's Reserve Team performances in the 2023 Central League season in order to create an objective measure of team performance that does not factor coach playing styles. The regression analysis broke down all components of a possession directly relating to adding and removing opponents, while also showing a total number of opponents removed at the end of a possession. The main findings from this study is that teams are more likely to score when all opponents are behind the ball. It is also identified a team is more likely to gain a successful outcome of continue possession after one or more opponents has been beaten.

There has been extensive research on passing patterns in football in order to determine the most effective way to maintain the ball and create goal scoring opportunities (Cintia et al., 2015; Grund, 2012; Narizuka et al., 2014). This study looked into the amount of times during a possession a progressive pass was made to better understand how they break down a defensive setup. The odds ratios of progressive pass count 1-3 (OR = 3.862, CI: 0.559-26.671) shows that likelihood of a successful outcome greatly increased when 1-3 progressive passes have been made in comparison to 0. This is to be expected, as it would be an extremely rare event for goals to be achieved with all opponents in front of the ball. However, when compared with the odds ratios of a 4-6 and 7-9 progressive pass count (4-6, OR = 1.58, CI: 0.253-9.868; 7-9, OR = 0.435, CI: 0.043-4.361), it highlights the possibility that shorter, more effective possession play will create more goal scoring chances. Previous studies have identified short passing sequences as more effective means to score goals (De Pablo et al., 2019; Low et al., 2019; Sarmento et al., 2014). Therefore, reinforcing the justification of this trend.

After analysing the count of progressive passes, the value of each pass can be determined to understand how many opponents are being removed. Valuing passes has had differing degrees on success in the past (Goes et al., 2021; Wright et al., 2014) and Mackenzie and Cushion (2014) highlight the issue of a universal objective agreement for measures due to the subjectivity of what is considered a good pass or bad pass. A possession that had a progressive passing score of 1-3 had a significantly lower probability of possession ending in a successful outcome when compared to an unsuccessful outcome (β = -1.909) and is considered not significant (p = 0.062). A progressive passing score of 4-6 also showed a low probability of possession ending in successful outcome (β = -1.711) and no significance (p = 0.095). These results are surprising, and potentially indicate that playing the ball forward is not always the best option. Playing a ball forward may remove multiple opponents from the game, however, if the recipient does not have enough space or does not have a connecting pass available, then these moments will often lead to a turnover, hence the value of the beta coefficient. This could be understood in greater detail if spatial data was considered for this study, however, this resource is not available at this level of competition.

To consider all possible actions that can remove an opponent, dribbling was accounted for in this study. There was very limited research in the field of quantifying dribbling actions, with only Brink et al. (2023) quantifying the value of a 1v1 dribble, however, this does not factor all dribbling moments as it is dependent on a defender within close proximity. The results the regression analysis showed that a possession that has a count of progressive dribbles between 4-6 is shown to have a very low probability of achieving a successful outcome when compared against an unsuccessful outcome (β = -6.407). The p-value of this result was highly significant (p = 0.002). Whereas a possession that has a count of progressive dribbles between 4-6 is shown to have a very high probability of a continuing possession in comparison to an unsuccessful outcome (β = 13.103). However, the p-value of this test (p = 0.958) is not significance. Although this may suggest that progressive dribbling counts are likely to lead to

turnovers, these contrasting findings possibly indicate that dribbling is not a strong predictor factor for possession outcomes.

Additionally, progressive dribbling scores in successful outcomes had high significance ($p < 0.001$) when scoring between 4-6 and showed a low probability when compared to unsuccessful outcomes (β = -2.974). This was also highlighted in continued possession vs unsuccessful outcomes, as probability was low (β = -0.628) and results not significant ($p = 0.534$) when scoring between 1-3. To an extent, it can be assumed that high dribbling scores and counts can be associated with turning over the ball, however, due to the inconsistency of test significant, this cannot be definitive. Despite this, dribbling must be considered to accurately assess all moments in which an opponents can be removed from a possession.

Back passes need to be considered for this study as teams can complete multiple progressive passes during a possession. As mentioned prior, there is extensive research on passing and passing networks, however, this is no research on the orientation of a pass in relation to opposition. This research found that >9 passes that added opponents back in the game had a substantially low probability for successful outcomes as opposed to unsuccessful (β = -15.593). Furthermore the odds ratio for 4-6 backwards passing count ending in possession continued shows this type of outcome is 2.051 times more likely as opposed to 0 backwards passes. These findings suggest that teams are most susceptible to turning over the ball during long spells of possession. This would corroborate the ideas presented earlier, which stated that a pass forward is not always the correct decision as spacing is not considered. Receiving the ball in an area where you cannot progress the play and therefore must play backwards to the deeper teammates. This would provide a sufficient explanation as to why high back pass counts occur.

Backwards passing scores collectively showed a high probability (1-3, β = 2.453; 4-6, β =2.447; 7-9, β = 2.047; >9, β = 2.523) and significance 1-3, $p = 0.011$; 4-6, $p = 0.014$; 7-9, $p = 0.052$; >9, $p = 0.023$) when comparing successful and unsuccessful outcomes. The odds ratios

also support this trend within successful outcomes, with an increase in odds by 12.471 (CI: 1.417-109.782) if the backwards passing score exceeds 9, in comparison to 0. The backward passing scores show that the direction of passing is significant to understand how teams create chances. Although passing backwards it typically considered negative, these findings indicate that passing backwards may engage the opposition and create gaps for progressive passes to be made. More insight into this can be achieve if the data is continuous and sequential, although this would be highly complex and has not been attempted in relevant literature thus far.

Prior, to data collected, it was assumed that investigating the number of opponents behind the ball at end of possession would provide the most insight into this studies applicability, with the assumption that teams will create the most goal scoring chances when all or the majority of opponents (9-11) are behind the ball. The findings from this variable shows positive correlation in possession continued possessions (1-3, OR = 1.068, CI: 0.742-1.536; 4-6, OR = 1.17, CI: 0.742-1.844; 7-9, OR = 1.027, CI: 0.65-1.622; 9-11, OR = 1.315, CI: 0.72-2.401) when compared to 0. Furthermore, Possessions ending with 7-9 opponents behind the ball has a lower probability (β = -0.482) in ending in an successful outcome as opposed to a unsuccessful outcome, but this result was also not significant (p = 0.126). Finally, 9-11 opponents behind the ball has a higher probability of ending in possession continued when compared to unsuccessful outcomes (β = 0.274), where the same opponent count when comparing successful and unsuccessful had a lower probability of β = 0.037. The findings show that opponents behind the ball at the end of possession are a relatively effective variable to predict possession outcomes. Although the p-values did not indicate significance, the beta coefficients show that in almost all cases, removing opponents from the game will lead to a successful outcome, in comparison to an unsuccessful outcome. This is most likely due to the effect on game stability, as mentioned in Kim et al. (2019b), as teams defensively organise into a shape that can prevent the ball being received in space. Once the first line of defensive

organisation is broken, spaces become more available for the team in possession and opposing players may have to break their structure to tackle or put pressure on this ball. This would also explain the increase in odds ratios as the categorical number of opponents behind the ball at the end of possession increases.

5.2 Situational Variable Influence on Predicting Possession Outcomes

Variables that are not direct consequences of actions on the field have been shown to effect performance (Pollard, 2008). This research indicated that situational variables must be considered when creating an effective model in predicting team performance (Casal et al., 2023). When comparing the absolute frequencies of variables, goals (n = 14), shots blocked (n = 12), shots off target (n = 21) and corners awarded (n = 28) were all higher in frequency when compared to draws and losses. In terms of predictive value, the negative slop for loss (β = -1.197) indicates that the probability of a successful possession outcome was less than a successful outcome in a win. Evidently, this portion of the research corroborates some of the early work done in performance analysis in football. Most teams who win will create more goals and more goal-scoring chances (Wright et al., 2011).

Both match location, team quality, and match status did not show any substantial likelihood or probability trends in prediction for successful outcomes (away, β = 0.563, OR = 1.755; middle team, β = -0.767, OR = 0.464; bottom team, β = -1.23, OR = 0.292; drawing, β = -0.156, OR = 0.522; losing, β = -0.502, OR = 0.605). This was also the trend in the same variables when comparing possession continued and unsuccessful outcomes (away, β = 0.285, OR = 1.33; middle team, β = -0.131, OR = 0.877; bottom team, β = -0.918, OR = 0.399; drawing, β = -0.312, OR = 0.732; losing, β = -0.439, OR = 0.645). There are a few assumptions as to why these predictor variables did not show considerable trends. Firstly, there a no stadium fields in the Central League competition, so psychological impact would not be the same as study that have investigated this (Almeida et al., 2014). Fan attendance and crowd noise was apparent at all

games. However, according to the data, this did not affect team performance. Lastly, match status did not show any noticeable changes in possession outcome trends. The most likely explanation for this is the context of the Wellington Phoenix Academy objectives. The objective of the academy is to develop young players to then progress into the first team environment, because of this, the reserve teams will play according to the playing style of the first team coaches, in order to ease the transition. This is why game results are not considered important and are not factored when assessing players. This would also explain no predictor difference in match status, as the team would continue using their passing patterns to keep improving. If this was to be used in a professional environment, this would most likely have more predicting value as teams will change their tactics according to the state of the game.

Another variable that did not show substantial trends in prediction value was possession duration. This was initially integrated into this predictor model as contrasting research has shown that both longer spells and short spells of possession are most effective for scoring goals and creating goal scoring opportunities (Wright et al., 2011). It can be argued that this team favoured longer spells of possession (>90, $\beta = 0.648$; 0-30, $\beta = -0.529$) when comparing successful and unsuccessful outcomes. However, these tests were not found to be significant (>90, $p = 0.622$; 0-30, $p = 0.314$). Although conversely, frequency data showed that 68% (n = 17) of all goals scored in this data sample were scored in possession lasting less than 30 seconds. These results indicated an inconclusive prediction of possession duration affecting possession outcomes. Both the frequency data and the regression data highlight different ways possession duration was used, and therefore, this value would need to be looked into more closely to provide conclusive findings.

The final situational variable was the location of possession starts and end. The immergence of tracking data in recent literature has shown the significance of where teams win the ball and where the give it away (Caicedo-Parada et al., 2020). For possession starts, the attacking half ($p = 0.005$) and attacking zone ($p = < 0.001$) had high significance and high

probability of ending in a successful outcome when compared to unsuccessful. Whereas possessions ending in the attacking half (β = -1.139) and attacking zone (β = -1.134) show a lower probability of continuing possession when comparing to unsuccessful outcomes. Both results were considered highly significant (p = < 0.001). This can provide valuable insight to a coach, as this indicates that more possessions without goal scoring opportunities (shots, corners awarded) are being turned over to the opposition, instead of continuing possession through set pieces (free kick, throw in). Although insight is relatively limited due to the low number of pitch areas, this analysis indicated that pitch locations for both the start and outcome of possession are significant variables to predict possession outcomes.

Chapter 6: Conclusion

In view of these findings, there is sufficient evidence that quantifying possession actions, in relation to opposition players orientation to the ball, is an effective and objective measure to assess and predict possession outcomes. As the number of opponents behind the ball at the end of a possession increased, the probability of a successful outcome increases in comparison to unsuccessful outcomes. Furthermore, these possessions were broken down into specific actions to see what actions were most frequent and most effect when removing opponents. Although the predicting value of these actions showed mixed results, quantifying all actions involved with removing an opponent is vital to understand what is the ideal moment for a possession play to transition into a goal scoring opportunity.

Relevant literature emphasised the importance of considering situational variables when predicting possession outcomes (Ballesteros et al., 2012: Bloomfield et al., 2015; Srgo & Lipoma, 2016). The final result show an accurate prediction, in relation to goal scoring opportunities and creating positive possession outcomes. Some of the situational variables were limited with their insight, however, this was a result of the Wellington Phoenix Academy objectives and the quality of the competition. Location of possession start and end did have significance as a predictor variable, and provide even greater insight if pitch zones were more refined. Although in this case situational variables as a possession outcome predictor was limited, apply this model to a team in a professional environment will garner more insightful findings.

6.1 Limitations and Future Research Recommendations

This study has shown that a MLogit model is an effective tool to predict possession outcomes in football. Most importantly, findings indicate that this model can be implemented into any coaching set up, as it did not factor any playing patterns or coaching playing

philosophy. However, there are some key limitations to be considered. The first limitation is the sample size. This study could only analyse 15 games, due to the competition duration and poor video conditions. MLogit is most insightful with a larger sample size, therefore, a recommendation for future research is to analyse a team over course of 30-45 games.

Another key limitation is that only one team was analysed. The objective of this study was to produce a possession predictor model that could be integrated into a coaching practice without the need to factor playing patterns or coaching playing philosophy. Although this study can justify the effect of the model, more insight would be made if two teams with different styles of play were analysed and compared. This was not a possibility with the sample data provided by the Wellington Phoenix Academy as organisation is the only football organisation to have playermaker integrated into the academy system. Comparison between multiple teams in the Wellington Phoenix Academy would not provide this level of insight as all teams play the same style of play. Furthermore, evaluating the prediction value was limited by the context of the competition and the venues that the games were played in. Applying this model to a professional environment would provide greater insight into how these variables affect performance

Lastly, a future recommendation is applying this prediction values to passing network data visualisation. As mentioned in chapter 2, application of complex statistical models is not realistic in professional coaching practice due to the strict turnaround time between game completion and statistical insight visualisation to then influence the preparation leading up to the next game. Future research should look to investigate how this predictor data can be integrated into passing networks, as these models are some of the more widely accept forms of data output by coaching staff.

Reference List

Acar, M. F., Yapicioglu, B., Arikan, N., Yalcin, S., Ates, N., & Ergun, M. (2008). Analysis of goals scored in the 2006 World Cup. In *Science and football VI* (pp. 261-268). Routledge.

Aguado-Méndez, R. D., González-Jurado, J. A., Callejas-Jerónimo, J. E., & Otero-Saborido, F. M. (2021). Analysis of the goal-scoring opportunities conceded in football: a study case in the Spanish La Liga. *Quality & Quantity*, *55*(4), 1477-1496.

Almeida, C. H., Ferreira, A. P., & Volossovitch, A. (2014). Effects of match location, match status and quality of opposition on regaining possession in UEFA Champions League. *Journal of human kinetics*, *41*(1), 203-214.

Anderson, C., & Sally, D. (2013). *The numbers game: Why everything you know about soccer is wrong*. Penguin.

Andrienko, G., Andrienko, N., Budziak, G., Dykes, J., Fuchs, G., Von Landesberger, T., & Weber, H. (2017). Visual analysis of pressure in football. *Data Mining and Knowledge Discovery*, *31*, 1793-1839.

Ballesteros, J. L., Peñas, C. L., Eiras, E. R., Martinez, L. C., & Lago, E. D. (2012). Success of offensive actions in elite football. Influence of the tactical models used and situational variables. *European Journal of Human Movement*, *28*, 145-170.

Bailey, M.J. (2005). Predicting Sporting Outcomes: A Statistical Approach. Swinburne University of Technology: Faculty of Life and Social Sciences.

Bayley, R. (2013). The quantitative paradigm. *The handbook of language variation and change*, 83-107.

Buldú, J. M., Busquets, J., Martínez, J. H., Herrera-Diestra, J. L., Echegoyen, I., Galeano, J., & Luque, J. (2018). Using network science to analyse football passing networks: Dynamics, space, time, and the multilayer nature of the game. *Frontiers in Psychology*, 9. https://doi.org/10.3389/fpsyg.2018.01900

Brink, L., Ha, S. K., Snowdon, J., Vidal-Codina, F., Rauch, B., Wang, F., ... & Hosoi, A. E. (2023). Measuring skill via player dynamics in football dribbling. *Scientific Reports*, *13*(1), 19004.

Butterworth, A., & Woodward, J. (2023). Performance Analysis in the Coaching Process. In *Professional Practice in Sport Performance Analysis* (pp. 1-18). Routledge.

Caicedo-Parada, S., Lago-Peñas, C., & Ortega-Toro, E. (2020). Passing networks and tactical action in football: A systematic review. *International Journal of Environmental Research and Public Health*, *17*(18), 6649. https://doi.org/10.3390/ijerph17186649

Carling, C., Williams, A., & Reilly, T. (2005). The handbook of soccer match analysis: A systematic approach to improving performance. Abingdon: Routledge.

Carling, C., Wright, C., Nelson, L. J., & Bradley, P. S. (2014). Comment on 'Performance analysis in football: A critical review and implications for future research'. *Journal of sports sciences*, *32*(1), 2-7.

Casal, C., Stone, J., Iván-Baragaño, I., & Losada, J. (2023). Effect of goalkeepers' offensive participation on team performance in the women Spanish La Liga: a multinomial logistic regression analysis. *Biology of Sport*, *41*(1), 29-39.

Cao, S. (2024). Passing path predicts shooting outcome in football. *Scientific Reports*, *14*(1), 9572.

Cintia, P., Rinzivillo, S., & Pappalardo, L. (2015, September). A network-based approach to evaluate the performance of football teams. In *Machine learning and data mining for sports analytics workshop, Porto, Portugal*.

Clemente, F. (2012). Study of successful teams on FIFA World Cup 2010 through notational analysis. *Pamukkale Journal of Sport Sciences*, *3*(3), 90-103.

Dawson, B., Appleby, B., & Stewart, G. (2005, May). Analysis of a 16-game winning streak in Australian rules football. In *Science and Football V: The Proceedings of the Fifth World Congress on Sports Science and Football, Routledge* (p. 203).

De Pablo, M., Silva, A. T., & Ramírez, A. G. (2019). Patterns of play to score goals in uruguayan professional football. *Revista iberoamericana de psicología del ejercicio y el deporte*, *14*(2), 179-183.

Drust, B. (2010). Performance analysis research: Meeting the challenge. *Journal of Sports Science, 28*(9), 921-922.

Di Leo, G., & Sardanelli, F. (2020). Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach. *European radiology experimental*, *4*, 1-8.

Diquigiovanni, J., & Scarpa, B. (2019). Analysis of association football playing styles: An innovative method to cluster networks. *Statistical modelling*, *19*(1), 28-54.

Diznar, D., Jozak, R., & Basic, D. (2016). The importance of offense corner kicks in football with regard to final outcome of the match and league system of competition. *Journal of Sports Sciences*, *10*(2), 56-59.

Eusebio, P., Prieto-González, P., & Marcelino, R. (2024). Decoding the complexities of transitions in football: A comprehensive narrative review. *German Journal of Exercise and Sport Research*, 1-11.

Felices, A. M. (2023, April 5). Analyzing team passing networks depending on the possession outcome. *Medium*. https://medium.com/@marin11amf11/analyzing-team-passing-networks-depending-on-the-possession-outcome-5ad778788507

Fernandez-Navarro, J., Ruiz-Ruiz, C., Zubillaga, A., & Fradua, L. (2020). Tactical variables related to gaining the ball in advanced zones of the soccer pitch: analysis of differences among elite teams and the effect of contextual variables. *Frontiers in Psychology*, *10*, 3040.

FIFA. (2023). Professional Football Report 2023 [PDF]. https://digitalhub.fifa.com/m/2a5dc95026d9cf8a/original/FIFA-Professional-Football-Report-2023.pdf

Fischer, J., Fischer, D., & Keiner, M. (2022). Perturbation profile of elite football–a cross-sectional analysis of the goals and goal scoring opportunities immediately before and after goal scoring of the 1st German Bundesliga. *International Journal of Performance Analysis in Sport*, *22*(4), 491-504.

Goes, F. R., Meerhoff, L. A., Bueno, M. J. O., Rodrigues, D. M., Moura, F. A., Brink, M. S., ... & Lemmink, K. A. P. M. (2021). Unlocking the potential of big data to support tactical performance analysis in professional soccer: A systematic review. *European Journal of Sport Science*, *21*(4), 481-496.

Gonzalez-Rodenas, J., Lopez-Bondia, I., Calabuig, F., Pérez-Turpin, J. A., & Aranda, R. (2016). Association between playing tactics and creating scoring opportunities in counterattacks from United States Major League Soccer games. *International Journal of Performance Analysis in Sport*, *16*(2), 737-752.

Grund, T. U. (2012). Network structure and team performance: The case of English Premier League soccer teams. Social Networks, 34, 682-690.

Hewitt, A., Greenham, G., & Norton, K. (2016). Game style in soccer: What is it and can we quantify it? *International Journal of Performance Analysis in Sport*, *16*(1), 355–372.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression (3 rd ed.).

Wiley.

Hughes, M. D., & Bartlett, R. M. (2002). The use of performance indicators in performance analysis. *Journal of sports sciences*, *20*(10), 739-754.

James, N., Jones, P. D., & Mellalieu, S. D. (2004). Possession as a performance indicator in soccer as a function of successful and unsuccessful teams. *Journal of sports sciences*, *22*(6), 507-508.

Jamieson, J. P. (2010). The home field advantage in athletics: A meta-analysis. *Journal of Applied Social Psychology*, *40*(7), 1819-1848.

Kan, A., Shiokawa, M., Okihara, K., Soon Choi, C., Usui, S., & Yanagihara, T. D. E. (2004). The movement of players and the team: comparing two games, Japan versus UAE and J-Leaguegame [Abstract] Part II: Game activity and analysis. *J. Sports Sci*, *22*(6), 500-520.

Kim, D., Jang, J., Kim, H. J., Lim, S., Ryoo, H., Jung, T. Y., & Suh, S. H. (2020). Discovering primary indicators for evaluating defender's technical performance using multivariate statistics in football games. *International Journal of Applied Sports Sciences*, *32*(1), 27-36.

Kim, J., James, N., Parmar, N., Ali, B., & Vučković, G. (2019a). Determining unstable game states to aid the identification of perturbations in football. *International Journal of Performance Analysis in Sport*, *19*(3), 302-312.

Kim, J., James, N., Parmar, N., Ali, B., & Vučković, G. (2019b). The attacking process in football: a taxonomy for classifying how teams create goal scoring opportunities using a case study of crystal Palace FC. *Frontiers in psychology*, *10*, 2202.

Kubayi, A., & Larkin, P. (2019). Analysis of teams' corner kicks defensive strategies at the FIFA World Cup 2018. *International Journal of Performance Analysis in Sport*, *19*(5), 809-819.

Lago-Peñas, C., Gómez-Ruano, M., & Yang, G. (2017). Styles of play in professional soccer: an approach of the Chinese Soccer Super League. *International Journal of Performance Analysis in Sport*, *17*(6), 1073-1084.

Lepschy, H., Wäsche, H., & Woll, A. (2018). How to be successful in football: a systematic review. *The open sports sciences journal*, *11*(1).

Lepschy, H., Woll, A., & Wäsche, H. (2021). Success factors in the FIFA 2018 world cup in Russia and FIFA 2014 world cup in Brazil. *Frontiers in psychology*, *12*, 638690.

Link, D., Lang, S., & Seidenschwarz, P. (2016). Real time quantification of dangerousity in football using spatiotemporal tracking data. *PloS one*, *11*(12), e0168768. Liu, H., Gómez, M. A., Gonçalves, B., & Sampaio, J. (2016). Technical performance and match-to-match variation in elite football teams. *Journal of sports sciences*, *34*(6), 509-518.

Liu, T., García-de-Alcaraz, A., Wang, H., Hu, P., & Chen, Q. (2021). Impact of scoring first on match outcome in the Chinese Football Super League. *Frontiers in Psychology*, *12*, 662708.

Long, J. S., & Freese, J. (2014). Regression models for categorical dependent variables using

stata. (3rd edition ed.) Stata Press.

Low, B., Boas, G. V., Meyer, L., Lizaso, E., Hoitz, F., Leite, N., & Gonçalves, B. (2018). Exploring the effects of deep-defending vs high-press on footballers' tactical behaviour, physical and physiological performance: A pilot study. *Motriz: Revista de Educação Física*, *24*(02), e1018171.

Ma, L., Wang, Y., Wang, Y., Li, N., Fung, S. F., Zhang, L., & Zheng, Q. (2021). The Hotspots of Sports Science and the Effects of Knowledge Network on Scientific Performance Based on Bibliometrics and Social Network Analysis. *Complexity*, *2021*(1), 9981202.

Mackenzie, R., & Cushion, C. (2013). Performance analysis in football: A critical review and implications for future research. *Journal of sports sciences*, *31*(6), 639-676.

McCarthy, C., Tampakis, P., Chiarandini, M., Randers, M. B., Jänicke, S., & Zimek, A. (2022, September). Analyzing passing sequences for the prediction of goal-scoring opportunities. In *International Workshop on Machine Learning and Data Mining for Sports Analytics* (pp. 27-40). Cham: Springer Nature Switzerland.

McLean, S., Salmon, P. M., Gorman, A. D., Read, G. J., & Solomon, C. (2017). What's in a game? A systems approach to enhancing performance analysis in football. *PloS one*, *12*(2), e0172565.

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

Memmert, D., & Raabe, D. (2017). *Revolution im Profifußball*. Springer Berlin Heidelberg.

Mendes, B., Clemente, F. M., & Maurício, N. (2018). Variance in prominence levels and in patterns of passing sequences in elite and youth soccer players: a network approach. *Journal of human kinetics*, *61*(1), 141-153.

Molinuevo, J. S., & Prieto Bermejo, J. (2012). The effect of scoring first and home advantage in professional Spanish football and indoor soccer leagues. *Revista de psicología del deporte*, *21*(2), 301-308.

Myers, J. L., Well, A. D., & Lorch Jr, R. F. (2013). *Research design and statistical analysis*. Routledge.

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

Pollard, R., Silva, C. D., & Medeiros, N. C. (2008). Home advantage in football in Brazil: differences between teams and the effects of distance traveled. *Revista Brasileira de Futebol (The Brazilian Journal of Soccer Science)*, *1*(1), 3-10.

Prasetio, D. (2016, August). Predicting football match results with logistic regression. In *2016 International Conference On Advanced Informatics: Concepts, Theory And Application (ICAICTA)* (pp. 1-5). IEEE.

Pupo-Ge, C. N., & Hernández-Ávila, M. S. Y. (2020). Statistics for decision making in sports: a theoretical contribution. Scientific journal specialized in Physical Culture and Sports Sciences, 17(45).

Ric, A., Torrents, C., Gonçalves, B., Torres-Ronda, L., Sampaio, J., & Hristovski, R. (2017). Dynamics of tactical behaviour in association football when manipulating players' space of interaction. *PloS one*, *12*(7), e0180773.

Ruiz-Ruiz, C., Fradua, L., Fernández-GarcÍa, Á., & Zubillaga, A. (2013). Analysis of entries into the penalty area as a performance indicator in soccer. *European Journal of Sport Science*, *13*(3), 241-248.

Sarmento, H., Clemente, F. M., Gonçalves, E., Harper, L. D., Dias, D., & Figueiredo, A. (2020). Analysis of the offensive process of AS Monaco professional soccer team: A mixed-method approach. *Chaos, Solitons & Fractals*, *133*, 109676.

Sarmento, H., Marcelino, R., Anguera, M. T., CampaniÇo, J., Matos, N., & LeitÃo, J. C. (2014). Match analysis in football: a systematic review. *Journal of sports sciences*, *32*(20), 1831-1843.

Schwartz, B., & Barsky, S. F. (1977). The home advantage. *Social forces*, *55*(3), 641-661.

Sgro, F., & Lipoma, M. (2016). Technical performance profiles in the European Football Championship 2016. *Journal of Physical Education and Sport*, *16*(4), 1304-1309.

Simiyu, W. W. N. (2013). Analysis of goals scored in the 2010 world cup soccer tournament held in South Africa.

Singh, P., & Lamba, P. S. (2019). Influence of crowdsourcing, popularity and previous year statistics in market value estimation of football players. *Journal of Discrete Mathematical Sciences and Cryptography*, *22*(2), 113-126.

Snyder, R. L., & Bish, D. L. (1989). Quantitative analysis. *Modern powder diffraction*, *20*, 101-144.

Terblanche, N. H. D. (2020). The coaching model derivation process: combining grounded theory and canonical action research for developing coaching models. *Coaching: An International Journal of Theory, Research and Practice*, *13*(1), 45-60.

TheFA. (n.d.). Law 16: The Goal Kick. https://www.thefa.com/football-rules-governance/lawsandrules/laws/football-11-11/law-16---the-goal-kick

Vanclay, F., Baines, J. T., & Taylor, C. N. (2013). Principles for ethical research involving humans: ethical professional practice in impact assessment Part I. *Impact assessment and project appraisal*, *31*(4), 243-253.

Vergonis, A., Michailidis, Y., & Metaxas, T. (2021). The significant role of scoring from set plays in the 2018 FIFA World Cup. *Faculty of Education*, *47*.

Vergonis, A., Michailidis, Y., Mikikis, D., Semaltianou, E., Mavrommatis, G., Christoulas, K., & Metaxas, T. (2019). Technical and tactical analysis of goal scoring patterns in the 2018 FIFA World Cup in Russia. *Facta Universitatis, Series: Physical Education and Sport*, 181-193.

Wright, C., Atkins, S., Polman, R., Jones, B., & Sargeson, L. (2011). Factors associated with goals and goal scoring opportunities in professional soccer. *International Journal of Performance Analysis in Sport*, *11*(3), 438-449.

Wright, C., Carling, C., & Collins, D. (2014). The wider context of performance analysis and it application in the football coaching process. *International Journal of Performance Analysis in Sport*, *14*(3), 709-733.

Yiannakos, A., & Armatas, V. (2006). Evaluation of the goal scoring patterns in European Championship in Portugal 2004. *International Journal of Performance Analysis in Sport*, *6*(1), 178-188.

Appendices

Appendix A: Permission to access data form

Permission for researchers to access organisation data

Project title: **Quantifying physical and technical actions in New Zealand professional soccer using foot-mounted inertial measurement units: An age group comparison**

Project Supervisor: Associate Professor Kirsten Spencer

- ○ I have read and understood the information provided about this research project in the Information Sheet dated 1st November 2023.
- ○ I give permission for the researcher to access to organisation permanently deidentified data comprising routinely collected Playermaker data.
- ○ I wish to receive a summary of results upon completion of the study

Director of Wellington Phoenix Academy:

Name: .....Steve Coleman.....

Signature: ....

Date : ...01/11/23.....

Note: The head of the organisation should retain a copy of this form.

Appendix B: Player information sheet

Participant Information Sheet

Date Information Sheet Produced:

06 November 2023

Project Title

Predicting the outcome of ball possessions in football using game stability and game effect score metrics

An Invitation:

Kia Ora, my name is Sam Casey-Popovich and I am an analyst for the Wellington Phoenix and I am a post-graduate student completing a master's in Sport, Exercise, and Health.

You, the Wellington Phoenix reserve team (uner 23s), have been invited to participate in an integral research project that will contribute to the development of Football through giving permission to access permanently deidentified data collected routinely by Wellington Phoenix in the 2023 football year. The data that has been routinely collected, via the Player Maker IMU units worn during training and matches, will enable the comparison physical, technical and tactical actions.

This project is being undertaken by Auckland University of Technology (AUT) and the Wellington Phoenix Football Club. All data collected and analysed by AUT will be permanently deidentified. It is important to note, participation in this project would be greatly appreciated however, it is voluntary, and no consequences will be given if you chose not to give permission to access the data.

What is the purpose of this research?

Reviews on performance analysis in football suggest that identifying performance variables does enhance the discovery of performance patterns, however, there is a challenge in a universal agreement of the measurement of values. Recent studies suggest an approach that categorises team possession into game states. This is an innovative approach to investigating patterns in possession, but there is still an absence of contextual information that would be of use to a coach/football team. The Wellington Phoenix Academy are a professional football team located in Wellington, New Zealand. The academy has a performance model that provides objective scoring to common actions in football games. I believe that this objective scoring model can be integrated into the game state model to enhance discovery of performance patterns. The scoring of actions in a possession will fluctuate due to the possibility of positive and negative actions, therefore, this fluctuation of score should correlate with the determinants of each game state. The purpose of this study is to investigate the relationship between score trends and the outcome of possessions, and to predict the outcome of a possession based on trends in scores and game states.

The findings of this research may be used for academic publications and presentations.

How was I identified and why am I being invited to participate in this research?

This research will be conducted upon the Wellington Phoenix Reserve Men's team. All identified participants are current players within this team.

How do I agree to participate in this research?

A consent form will be attached to this information sheet. Signing this consent form will confirm participation in this research. These forms will be presented to you in a team meeting prior to a scheduled training session.

Your participation in this research is voluntary (it is your choice) and whether or not you choose to participate will neither advantage nor disadvantage you. You are able to withdraw from the study at any time. If you choose to withdraw from the study, then you will be offered the choice between having any data that is identifiable as belonging to you removed or allowing it to continue to be used. However, once the findings have been produced, removal of your data may not be possible.

What will happen in this research?

This research project will look into the routinely collected playermaker data from the Wellington Phoenix Reserve Men's team. Therefore, there will be no direct interaction with participants. All data investigated will be de-identified, meaning no data will be attached to player names; all data will be compiled into team actions.

What are the discomforts and risks?

As data will be de-identified, there will be no collection of data that will be attached/associated to the names of the participants. If potential participants do have an concerns or queries, they are advised to use this meeting time to voice these concerns/queries, either collectively or privately.

What are the benefits?

The aim of this research is provide a more in-depth analysis of the attacking process in football. Participating in this research will results in more accurate and descriptive data being presented to the coaching staff. Aiding the coaching process will enhance decision-making and developing tactically competent football players.

How will my privacy be protected?

As data will be de-identified, there will be no collection of data that will be attached/associated to the names of the participants.

What opportunity do I have to consider this invitation?

Two weeks is the allocated amount of time to consider partaking in this research project. If you chose to partake, please return a signed copy of a consent form to the researcher.

Will I receive feedback on the results of this research?

A summary report will be offered to the team upon completion of the research project.

What do I do if I have concerns about this research?

Any concerns about this research are advised to be voiced before signing the consent form, to either the researcher or any of the coaching staff. If concerns do arise after signing the consent form, participants are encouraged to discuss these concerns with the researcher or the coaching staff. If participants feel their concerns cannot be resolved, withdrawing or declining participation will have no consequences on team selection or coach decisions.

Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor, *Sam Casey-Popovich, sam.casey-popovich@aut.ac.nz*

Concerns regarding the conduct of the research should be notified to the Executive Secretary of AUTEC, *ethics@aut.ac.nz* , (+649) 921 9999 ext 6038.

Whom do I contact for further information about this research?

Please keep this Information Sheet and a copy of the Consent Form for your future reference. You are also able to contact the research team as follows:

Researcher Contact Details:

Sam Casey-Popovich, sam.casey-popovich@aut.ac.nz

Project Supervisor Contact Details:

Kirsten Spencer, kspencer@aut.ac.nz

Approved by the Auckland University of Technology Ethics Committee on *type the date final ethics approval was granted*, AUTEC Reference number *type the reference number*.

Appendix C: Player consent and assent form

Consent and Release Form

Project title: Predicting the outcome of ball possessions in football using game stability and game effect score metrics

Project Supervisor: ***Kirsten Spencer***

Researcher: ***Sam Casey-Popovich***

- ○ I have read and understood the information provided about this research project in the Information Sheet dated 06.11.23.
- ○ I have had an opportunity to ask questions and to have them answered.
- ○ I understand that taking part in this study is voluntary (my choice) and that I may withdraw from the study at any time without being disadvantaged in any way.
- ○ I understand that if I withdraw from the study then I will be offered the choice between having any data that is identifiable as belonging to me removed or allowing it to continue to be used. However, once the findings have been produced, removal of my data may not be possible.
- ○ I understand that the videos will be used for academic purposes.
- ○ I understand that any copyright material created by the filming sessions is deemed to be owned by the Wellington Phoenix Academy and that I do not own copyright of any of the videos.
- ○ I agree to take part in this research.

Participant's signature: .....

Participant's name: .....

Participant's Contact Details (if appropriate):

.....

.....

.....

.....

Date:

Approved by the Auckland University of Technology Ethics Committee on type the date on which the final approval was granted AUTEC Reference number type the AUTEC reference number

Note: The Participant should retain a copy of this fo

Assent Form

Project title: ***xxx***

Project Supervisor: ***Kirsten Spencer***

Researcher: ***Sam Casey-Popovich***

- ○ I have read and understood the sheet telling me what will happen in this study and why it is important.
- ○ I have been able to ask questions and to have them answered.
- ○ I understand that notes will be taken during the interviews and that they will also be audio-taped and transcribed.
- ○ I understand that I can stop being part of this study whenever I want and that it is perfectly ok for me to do this.
- ○ If I stop being part of the study, I understand that then I will be offered the choice between having any information that that other people can know is about me removed or letting the researcher keep using it. I also understand that sometimes, if the results of the research have been written, some information about me may not be able to be removed.
- ○ I agree to take part in this research.

Participant's signature: .....

Participant's name: .....

Participant Contact Details (if appropriate):

.....

.....

.....

.....

Date:

Approved by the Auckland University of Technology Ethics Committee on *type the date on which the final approval was granted* ***AUTEC Reference number*** *type the AUTEC reference number*

Note: The Participant should retain a copy of this form.

Appendix D: AUTEC approval for study

Auckland University of Technology Ethics Committee (AUTEC)

31 January 2024

Kirsten Spencer
Faculty of Health and Environmental Sciences

Dear Kirsten

Re Ethics Application: **23/341 Predicting the outcome of ball possessions in football using game stability and game effect score metrics.**

Thank you for your responses to AUTEC's conditions.

Your ethics application has been approved for three years until 31 January 2027.

Non-Standard Conditions of Approval

1. Please include in the Information Sheet that it is only data from the Wellington Phoenix Academy reserve team (under 23s) that is being sought for the research.

Non-standard conditions do not need to be submitted to or reviewed by AUTEC unless requested but must be completed before commencing your study.

Standard Conditions of Approval

1. The research is to be undertaken in accordance with the Auckland University of Technology Code of Conduct for Research and as approved by AUTEC.
2. All public facing documents must have the AUTEC approval number and be of a high standard of spelling and grammar. Dates on the Information Sheet(s) and Consent Form(s) must be consistent.
3. Any amendments to the project must be approved by AUTEC prior to being implemented.
4. A progress report is due annually on the anniversary of the approval date.
5. A final report is due at the expiration of the approval period, or, upon completion of project.
6. Any serious or adverse events must be reported to AUTEC, this includes unforeseen issues that might affect continued ethical acceptability of the project.
7. AUTEC grants ethical approval only. You are responsible for obtaining management permission for access from any institution or organisation at which your research is being conducted and you need to meet all ethical, legal, public health, and locality obligations or requirements for the jurisdictions in which the research is being undertaken.

The application number and title need to be referenced on all correspondence related to this project.

All forms are available online http://www.aut.ac.nz/research/researchethics

For any enquiries, please contact ethics@aut.ac.nz
(This is a computer-generated letter for which no signature is required)

The AUTEC Secretariat
Auckland University of Technology Ethics Committee

Cc: samkcasey21@gmail.com