An Investigation into Kicking in Women's Australian Football

Emily Elizabeth Cust

Bachelor of Exercise and Sport Science

Bachelor of Business (Sport Management)

Bachelor of Applied Science – Human Movement (Honours)

Thesis submitted in fulfilment of the requirements of the degree of

Doctor of Philosophy

Supervisor: Professor Sam Robertson

Co-supervisor: Associate Professor Kevin Ball

Co-supervisor: Dr. Alice J. Sweeting

The Institute for Health and Sport (IHES)

Victoria University, Melbourne, Australia

&

Western Bulldogs Football Club,

Melbourne, Australia

March 2020

Victoria University Student Declaration

I, Emily Cust, declare that the PhD thesis by Publication entitled **An investigation into kicking in women's Australian football** is no more than 100,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work.

Signature [redacted]

Date: 27/03/2020

Abstract

In Australian Rules football (AF), kick skill performance involvements, notably the drop punt, are statistically strong contributors towards team match success. The start of a National women's AF competition (AFLW) in 2017 created opportunity for new knowledge to be established around the characteristics of AFLW athletes' skilled performances. Using developments in inertial measurement unit (IMU) technology and analytical methods, this thesis takes a multi-disciplinary approach to analysing AFLW skilled performances and subsequently proposes a concept of a semi-automated AF kick type classification system for skill monitoring in an applied environment. Specifically, the thesis: 1) evaluates the research literature on machine learning for sport-specific movement recognition, 2) determines the importance of AFLW athlete skilled performance indicator contributions during match play, 3) defines AFLW drop punt kick kinematics, and 4) evaluates AF kick type classification models using IMUs as a proof-of-concept to support further developments in the area.

Understanding analytical methods previously implemented with IMU or computer vision data and the evaluated capacity of these models in sport-specific movement recognition literature, is important in the adaptation for, and application towards new problems in sport. The first part of this thesis focuses on the experimental set-up, data pre-processing, and model development methods in the relevant literature on recognition of sport-specific movements in-field using IMU or computer vision technology. Of the 52 studies identified, 29 used IMUs, 22 used vision data and one study integrated both technologies. Supervised machine learning models were the dominant approach for developing sport specific movements recognition systems. Although nine studies implemented deep learning

algorithms which comparatively indicated superior results to machine learning models, and demonstrated the advantages and potential of these model types. This study also highlights the importance of considering the model and overall system development in relation to the targeted sports movement(s) when progressing future research in the field.

The applications of IMUs for sport skill recognition and subsequently performance analysis in-situation demonstrated in the literature may be beneficial in AF. As AF matches are technically skilled in nature, this thesis sought to investigate relationships of AFLW athlete skill performances in explaining team quarter and match success which knowledge was previously limited. Performance indicator distributions in explaining match quarter outcomes show the strongest skilled contributions from key high performing athletes, and the overall team strongest features related to kick performance indicators. Considering the importance of the kick in AF, the thesis then continued to define the kinematics of AFLW athlete's drop punt kicks across leg preferences which was unknown. Several key differences from men's AF kicks were found, also, women's kick movement patterns quantified which is beneficial for specific coaching practices. Developments in IMU use for sport-specific movement recognition through machine learning models demonstrate advantages in sporting performance analysis applications. In the final section, these technological developments are investigated for the concept of a semi-automated AF kick monitoring system using IMUs. The work is applied in an AFLW training environment as a unique study for capturing the importance kick skill performance towards team match success and differentiation from men's AF kick biomechanics. The findings indicate that kick types can be sufficiently distinguished from one another which creates scope for further applied work in AF training sessions. Overall, the work in this thesis is the first to establish the biomechanical characteristics of elite women's AF kicks and enhances the knowledge of skilled performances in the AFLW. Furthermore, it is the first to implement IMUs for on-field AF kick recognition. Increasing automation in sport-specific movement recognition can be applied in AF kick skill monitoring; particularly as a unique forefront in AFLW sport science applications towards kick performance improvement. The methods used and findings of this thesis can also be transferred to other elite women's team sporting leagues involving kicking actions such as Rugby and Gaelic football.

Acknowledgements

Well that was a three-year journey that packed a punch! Reflecting on what I have learnt, experienced, and shared has been a real eye-opener on where I first started to where and who I am now. There are most certainly many people who I am so grateful for in getting me to the finish line and beyond.

It is no secret that I had The Dream Team of supervisors. Thank you so much to Professor Sam Robertson for the opportunity to undertake my studies with a great team and alongside the inaugural AFLW competition with the Western Bulldogs. I am incredibly grateful for all your hard work, guidance, and knowledge passed on during the last four years. Secondly, to Associate Professor Kevin Ball; I'm sure your humour and calmness kept me sane during my biomechanics study when I couldn't bear to label another Vicon dot again and analysis wasn't always working out! Thank you for the wealth of knowledge you passed on in kicking biomechanics, Vicon, and conference presentations. Last but not least, thank you to Dr. Alice Sweeting for your super paper edits, study design and writing knowledge, football liaising, plus all the support and kind words. Also, thank you to Professor Rob Aughey for having me as part of the FIFA research team; it was an incredible learning experience that I will remember forever.

I feel very fortunate to have been able to complete my PhD with Victoria University and the Western Bulldogs football club. Thank you to all those as at IHeS who work in supporting all research students, furthermore to all the technical staff who helped during my research. Thanks to Suus Ellens for always being a friendly face and willing to give me a hand with my projects. Working in a professional sporting club was a fantastic experience and I'd like to particularly thank Chris Grant, Paul Groves, Debbie Lee, Jason Smith, and all the AFLW players that made my time at the Bulldogs a rewarding one. The W202 crew at Whitten Oval made the last three years awesome with many laughs, extended balcony lunch breaks, support for one another, and office baking! Thanks Elissa Denton for all the coffee chats and great cooking.

When the heartbreak of cancer hit our family during 2019, Mum was still my rock and never let me lose sight of my goals when I was so close to finishing my studies and I didn't want to leave her side. Big love to my sister Sarah for all the hugs whilst we kept ourselves grounded; congrats on getting to the Dr. before me!

Table of Contents

Victoria University Student Declaration ii

Abstract iii

Acknowledgements vi

Table of Contents vii

List of Publications and Presentations xi

Abbreviations xii

List of Tables xiv

List of Figures xvi

Chapter One: Introduction 1

1.1 Introduction 1

1.2 Aims of the Thesis 3

Chapter Two: Review of Literature 4

2.1 Sport performance analysis overview 4

2.2 Australian Rules football and kick skill performance analysis 7

2.2.1 Australian Rules football League overview 7

2.2.2 Australian Rules football performance analysis 8

2.2.3 Australian Rules football kick biomechanics research to practice 17

2.3 Inertial measurement units 25

2.3.1 Accelerometers 29

2.3.2 Gyroscopes 31

2.3.3 Magnetometers 32

2.3.4 Validity and reliability of inertial measurement units 33

2.3.5 Inertial measurement unit signal processing for movement recognition 39

2.4 Machine learning for automated movement recognition analysis 41

2.4.1. Machine learning theory overview 41

2.4.2 Model training and algorithms 44

2.4.3 Feature extraction 50

2.4.4 Feature selection...54

2.4.5 Model and algorithm performance evaluation...58

2.4.6 Deep learning..59

2.4.7 Computer vision in sport movement recognition..60

Chapter Three: Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance 61

3.1 Abstract...61

3.2 Introduction...62

3.3 Methods...67

3.3.1 Search strategy 67

3.3.2 Data extraction 70

3.4 Results 72

3.4.1 Experimental design 73

3.4.2 Inertial measurement unit specifications...74

3.4.3 Vision capture specification 80

3.4.4 Inertial measurement unit recognition model development methods.................83

3.4.5 Inertial measurement unit recognition model evaluation 97

3.4.6 Vision recognition model development methods..106

3.4.7 Vision recognition model evaluation 114

3.5 Discussion..120

3.6. Conclusions...126

Chapter Four: Inertial measurement units and machine learning for sport movement recognition: an update of the published article presented in Chapter Three 128

4.1 Sport-specific movement recognition: an update...128

Chapter Five: The relationship of team and individual athlete performances on match quarter outcome in elite women's Australian Rules football......................................147

5.1 Abstract......147

5.2 Introduction......148

5.3 Methods......151

5.4 Results......154

5.5 Discussion......159

5.6 Conclusion......162

Chapter Six: Biomechanical characteristics of elite female Australian Rules football preferred and non-preferred drop punt kicks......164

6.1 Abstract......164

6.2 Introduction......165

6.3 Methods......166

6.3.1 Participants......166

6.3.2 Research procedures......166

6.3.3 Data analysis......167

6.4 Results......169

6.5 Discussion......170

6.6 Conclusion......174

Chapter Seven: Pilot analysis of IMUs in Australian Rules football kick movement detection......176

7.1 Introduction......176

7.2 Methods......178

7.2.1 Data analysis......183

7.3 Results......184

7.4 Discussion......186

7.5 Conclusion......189

Chapter Eight: Classification of Australian football kick types in-situation via ankle-mounted inertial measurement units 190

8.1 Abstract..190

8.2 Introduction..191

8.2 Methods..193

8.2.1 Participants...193

8.2.2 Design...193

8.3 Methods..193

8.3.1 Statistical analysis...195

8.4 Results..198

8.5 Discussion...203

8.6 Conclusion..208

Chapter Nine: Discussion...209

9.1 Future directions for research 217

Chapter Ten: Conclusion 219

References..220

Appendix..252

List of Publications and Presentations

The research in this thesis has been published in peer reviewed journals or presented at peer reviewed scientific meetings as listed below:

International and National Journals

Cust, E. E., Sweeting, A. J., Ball, K., & Robertson, S. (2019). Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance review of model development and performance. *Journal of Sports Sciences*, *37*(5), 568–600. (Chapter Three)

Cust, E. E., Sweeting, A. J., Ball, K., Anderson, H., & Robertson, S. (2019). The relationship of team and individual athlete performances on match quarter outcome in elite women's Australian Rules football. *Journal of Science and Medicine in Sport*, *22*(10). https://doi.org/10.1016/j.jsams.2019.05.004. (Chapter Five)

International Conferences

Cust, E. E., Ball, K., Sweeting, A. J., & Robertson, S. (2019). Biomechanical characteristics of elite female Australian rules football preferred and non-preferred drop punt kicks. *Proceedings of the 7th International Conference on Sport Sciences Research and Technology Support - Volume 1: IcSPORTS*, 32–37. https://doi.org/10.5220/0008066300320037. (Chapter Six)

Cust, E.E., Ellens, S., Ball, K., Stephens, R., Sweeting, A.J., and Robertson, S., 2018. Analysis of Australian Rules football kick accelerometer ranges. *9th World Congress on Science and Football (WCSF 2019)*, Melbourne, Australia.

Awards

Best student presentation at The 7th International Congress on Sport Sciences Research and Technology Support, IcSPORTS 2019, Vienna, Austria.

Abbreviations

2D	Two-dimensions
3D	Three-dimensions
ADASYN	Adaptive Synthetic Sampling Method
AF	Australian football
AFL	Australian Football League
AFLW	Australian Football League Women's
AUC	Area-Under-Curve
BLSTM	Bi-directional long short-term memory network
CA	Classification accuracy
CHAID	Chi-squared automatic interaction detection
CNN	Convolutional neural network
DFT	Discrete Fourier Transform
FFT	Fast Fourier Transform
g	*g*-forces
Ga	Gauss
GBRT	Gradient boosted regression trees
GLM	Generalised linear mixed model
GPS	Global positioning system
HAR	Human activity recognition
Hz	Hertz
IFRP	Inside Football Player Ratings
IMU	Inertial measurement unit
kNN	k-Nearest Neighbour
LOO-CV	Leave-one-out cross validation
LPS	Local positioning system
LR	Linear regression
MAE	Mean absolute error
MEMS	Micro-Electro-Mechanical Systems
MRE	Mean relative error

NB	Naïve Bayesian
NN	Neural network
PCA	Principal Component Analysis
rbf	Radial basis function
RF	Random forest
RMSE	Root mean square error
ROC	Receiver operation characteristic curve
RPE	Rating of perceived exertion
SD	Standard deviation
SMOTE	Synthetic Minority Oversampling Technique
SVM	Support vector machine
μT	Micro Tesla

List of Tables

Table 3.1. Key word search term strings per database.

Table 3.2. Study inclusion and exclusion criteria.

Table 3.3. Inertial measurement unit specifications.

Table 3.4. Vision-based camera specifications.

Table 3.5. Inertial measurement unit study description and model characteristics**.**

Table 3.6. Inertial measurement unit study model performance evaluation characteristics.

Table 3.7. Vision-based study description and model characteristics.

Table 3.8. Vision-based study model performance evaluation characteristics.

Table 4.1. Inertial measurement unit specifications.

Table 4.2. Inertial measurement unit study description and model characteristics.

Table 4.3. Inertial measurement unit study model performance evaluation characteristics.

Table 5.1. Model results across data variables evaluated by mean absolute error.

Table 6.1. Definitions of measured kick parameters.

Table 6.2. Impact characteristics for preferred and non-preferred drop punt distance kicks for elite women's AF. Data reported as mean and standard deviation values and results of statistical tests comparing preferred and non-preferred leg kicks.

Table 7.1. Protocol of study part one.

Table 7.2. The mean group results of kick detection on the right and left legs.

Table 7.3. Drop punt maximum impulse accelerometer values for point of impact.

Table 8.1. Confusion matrix and performance metrics achieved by the random forest in the 4-Kick class condition.

Table 8.2. Confusion matrix and performance metrics achieved by the random forest in the 2-Kick class condition.

Table 8.3. Confusion matrix and performance metrics achieved by the random forest in the 4-Kick class condition rebalanced using ADASYN over-sampling method.

Table 8.4. Confusion matrix and performance metrics achieved by the random forest in the 2-Kick class condition rebalanced using Random over-sampling method.

Table 8.5. Confusion matrix and performance metrics achieved by the random forest for kick distance classification.

List of Figures

Figure 3.1. PRISMA flow diagram for study search, screen and selection process.

Figure 5.4-1. Regressor decision tree output of model 14, quarter score points and performance indicator relative values.

Figure 5.4-2. Regressor decision tree output of model 15, quarter score points and performance indicator feature distributions.

Figure 6.3-1. Athlete marker set-up.

Figure 6.3-2. Football reflective marker positions.

Figure 7.2-1. Part One IMU placement and axes orientation schematic.

Figure 7.2-2. Schematic of study Part One protocol set-up.

Figure 7.2-3. Schematic of study Part Two protocol set-up.

Figure 7.2-4. Study Part Two analogue accelerometer placement and axes orientation schematic.

Figure 7.3-1. IMU signal plot traces a maximal drop punt kick. The plot y-axis represents the *g*-force values.

Figure 8.3-1. Kick protocol schematic.

Figure 8.4-1. The comparison of model performances for kick classification in the 4-Kick and 2-Kick class conditions.

Figure 8.4-2. Standardised feature importance for random forest models in both kick conditions.

Figure 8.4-3. Central processing unit system user times for each model validation and testing computations in milliseconds.

Chapter One: Introduction

1.1 Introduction

The motivation for the research undertaken in this thesis came from the establishment of a National women's Australian football competition in 2017, Australian Football League Women's (AFLW). As has been observed in other sports, the research also aimed to display how different areas of sport science could be investigated to provide possible unique performance benefits to elite women's Australian Football (AF). The inaugural season of the AFLW competition occurred in the summer of 2017, and although expanded as of 2020, is still currently semi-professional and stand-alone from the long-standing men's professional competition. The competition structure means there is currently reduced whole team training practice opportunities for AFLW athletes comparative to the professional men's competition. It is hypothesised that specificity in AFLW training structures is required to suit the physical and technical AFLW match demands, which may be impacted due to the lower team contact coaching time. Hughes et al. (2007) and O'Donoghue (2014) explain the importance that performance analysis has in providing quality feedback for improvements towards an athlete's technical, tactical, and biomechanical skilled performances. Furthermore, the analytical requirements and methods to asses performances will differ based on the individual factors associated with each sport (Hughes, 2017; O'Donoghue, 2014). Therefore, research on the performance aspects of AFLW are important to progressing individual athletes, and the AFLW competition as a whole.

Continual improvements of data collection methods and higher-level information gains in sports analytics can be attributed to the increased integration of technology (Camomilla et al., 2015; Chambers et al., 2015). Inertial measurement units (IMUs) and computer vision data are key resources used for investigating semi-automated sport-specific movement recognition. Yet, a collective review of what has previously been researched and implemented across sports has not been undertaken and is required to inform for future practices. For example, as the capture of athlete skills in AF is largely manual and laborious in nature, the use of IMUs may be an approach to more efficient data collection for reporting on kick skill performance in AF (Robertson, Back, & Bartlett, 2016; Stewart, Mitchell, & Stavros, 2007).

The findings and methods of this thesis contribute notably to the research domain in two defined ways. Firstly, advancing knowledge on the current methods for semi-automated sport-specific movement recognition using IMUs and computer vision. The literature review presented in Chapter Three highlights the importance of considering adaptions to future data collection and model development methods in relation to the characteristics of the targeted sporting movement(s). Therefore, creating stimulation for further developments by evaluating previous methods to inform new research potential in specific sporting applications. Secondly, extending the body of research in elite women's AF skill performance analysis as a stand-alone from men's AF research. Specifically, generating foundational understandings of AFLW athlete match performance and kick characteristics. Quantifying the differences between women's and men's AF research for relationships in athlete performances in matches contributing to match success, and the biomechanical characteristics of drop punt kicks allows for targeted tactical and skill coaching to be formulated respectively.

1.2 Aims of the Thesis

The overarching objective of this thesis was to undertake a multi-disciplinary investigation into the use of wearable IMUs and analytical methods for sport-specific movement recognition; including specific applications in elite Women's Australian Rules football kick skill performance analysis. In detail, the aims were:

- To systematically review the research literature on machine and deep learning for sport-specific movement recognition using IMUs and, or computer vision data inputs (Chapter Three).

- To evaluate the relationship of AFLW athlete skill performance indicator distributions to explain match quarter outcomes during the 2017 and 2018 seasons, and secondly, compare quarter outcome model error rates from separate machine learning approaches based on the varied input feature sets of performance indicator variables (Chapter Five).

- To analyse the biomechanical characteristics of elite female AF drop punt kicks through 3-dimensional optoelectronic motion analysis for both the preferred and non-preferred kick legs (Chapter Six).

- To investigate IMU implementation and data detection analysis of kicks within a semi-controlled protocol in order to inform future method practices (Chapter Seven).

- To evaluate AF kick type classification models in an applied on-field environment using ankle-mounted IMUs (Chapter Eight).

Chapter Two: Review of Literature

2.1 Sport performance analysis overview

Analysis of athlete skilled performances provides important information on current athletic levels and how improvements can be made. Feedback to athletes from coaches and sport scientists can provide evaluations on technical, tactical, or biomechanical performances (Hughes et al., 2007; O'Donoghue, 2014). The type and effectiveness of feedback is also important for skill acquisition and development (Farrow & Robertson, 2017; Hodges & Williams, 2012; Phillips et al., 2013). Each sport will have different analytical requirements depending on factors such as whether it is an individual or team sport, discrete or continuous, or invasion-based (Hughes et al., 2007). For example, closed skilled sports such as golf and diving are predominately focused on the movement patterns of key skills that have a criterion set for high performance levels. Sports such as football and tennis have open skills which require focus on the execution of the skills plus the athlete's or team's decision making and tactical processes.

Sport biomechanics research provides a quantified analysis of sport-specific movements, which is advantageous in individual technical based sports such as javelin throwing and gymnastics where skill techniques can be critical to successful performances (Reily, Zhang, & Hoff, 2017; Robertson, Caldwell, Hamill, Kamen, & Whittlesey, 2013). This research domain also includes the identification of kinematic performance variables for key sporting skills such as kick foot speed in football codes (Ball et al., 2013; Barfield et al., 2002; Orloff et al., 2008) (see section 2.2.4 for AF kick biomechanics research). Methods for sport biomechanics include quantitative video feedback through computer annotation programs, force plates, electromyography, and 3D optoelectronic motion analysis (Robertson et al., 2013).

Objective performance analysis and quantified scoring of a sporting performance is important in providing unbiased information. The information can aid in decision making relative to an athlete's past performance or between teams and opposition athletes. Furthermore, athlete or team performances can be affected by contextual variables such as the opposition, game location, match phase, and field position (Alexander, Spencer, Sweeting, Mara, & Robertson, 2019; Ruano, Serna, Lupo, & Sampaio, 2016; Sampaio, Lago, Casais, & Leite, 2010; Taylor, Mellalieu, James, & Shearer, 2008).

Skilled performances of athletes and teams can be captured and analysed by a range of methods. Notational analysis was the early method of capturing and tracking athlete movements, performances, and tactical interactions (Hughes & Franks, 2004). Hughes (2017) defines a notational analysis six-step process to delivering accurate and precise feedback for sport performance improvements. These steps involve determining the specific performance indicators for a sport, ensuring the reliability of the data collected and in sufficient quantities, analysing the data, and modelling sport performance (Hughes, 2017). Examples of notational analysis include, analysing centre passes and possession chains in netball (Pulling et al., 2017), and the current application in AF matches to capture athlete performance measures and rankings (Jackson, 2016). Performance indicators (Hughes & Bartlett, 2002) can be used as a comparative method to measure performances in teams or individual athletes relative to previous performances or the opposition. Performance indicators are based on prior knowledge and are a "selection, or a combination, of action variables that aims to define some or all aspects of a performance" (Hughes, 2017). The advantage of performance indicators is that they are independent of units as they are expressed as a non-dimensional ratio (Hughes & Bartlett, 2002). Analysis of performance indicators can increase understandings of the physical, technical, and tactical characteristic of a sport in guiding training practices for the demands of competition. Relationships between athlete or team performance indicators and competition success has been researched across several sports for example AF (McIntosh, Kovalchik, & Robertson, 2018a; Robertson, Back, et al., 2016; Stewart et al., 2007), rugby (Higham et al., 2014; Jones et al., 2004; Parmar et al., 2017), basketball (Gómez et al., 2013; Moreno et al., 2013), and soccer (Harrop & Nevill, 2014; Lago-Peñas et al., 2011).

Athlete movements and skilled performances can also be captured using video and tracking technologies such as accelerometers, GPS, LPS, and semi-automated computer vision (O'Donoghue, 2014). Notational analysis using vision to provide statistics on athlete events is laborious given the manual work required to replay and annotate vision on a regular basis for training and competition (Barris & Button, 2008). Furthermore, as human input is required to categorise the movements, human error and subjectivity can affect the consistency in reproducible results (Barris & Button, 2008; Duthie et al., 2003). Integration of technology and computer analysis methods are allowing for improved efficiency in data analysis and speed of results on the quantification of athlete's skilled performances. Statistical modelling and machine learning methods are increasingly applied to find insightful patterns in performance and tracking data, for example, as shown in soccer tactical performance analysis (Herold et al., 2019) and sport-specific movement recognition (refer to section 2.4). Deep learning applications in performance analysis are also showing capability for improved results using inertial measurement units (IMUs) and vision data (Brock, 2018; Bulling et al., 2014); although implementation challenges such as the amount of quality data and hardware requirements need to be overcome. When implementing new technology in a sport performance department, several considerations exist such as recognition of the problem before the solution, practicality of data management (Torres-Ronda & Schelling, 2017), awareness of device limitations, and the need for industry standards for the generation of quality data (Cardinale & Varley, 2017). Semi-automated computer vision from either broadcast or a camera set-up has proven many applications in sport analysis, and is an unobtrusive method for the athletes (Barris & Button, 2008; Thomas et al., 2017). Another example includes the use of commercial GPS devices with in-built accelerometers in rugby to detect collisions (Hulin et al., 2017; Kelly et al., 2012), scrums (Chambers et al., 2018), and tackles (Chambers et al., 2019). Further examples of the use and methods of IMUs and computer vision for semi-automated sport skill analysis are in Chapter Three.

2.2 Australian Rules football and kick skill performance analysis

2.2.1 Australian Rules football League overview

The Australian Football League (AFL) operates a long running men's national competition that consists of a 23 home-and-away season plus a finals series, from March to September. Currently in 2020, 18 teams compete. The male players are considered professional athletes and exhibit high level physical and technical football characteristics (Gray & Jenkins, 2010). The AFL underwent an expansion to their structure in 2017 by introducing a national women's Australian football (AF) league, the Australian Football League Women's (AFLW). As of 2020, the AFLW season is held before the men's AFL regular season during the months of February and March. All current AFLW teams align with the long-established men's AFL teams. The 2017 and 2018 seasons consisted of a seven-round home-and-away competition incorporating eight teams. The 2019 season included two new teams, and a further four in the 2020 season. The AFL is aiming for a full 18 team AFLW competition aligning with the men's teams by the 2023 season. At season four of the AFLW competition, players are not considered full time professional footballers. Several match rules were modified from the men's game as the league's intention to promote improved match play and suit the conditions of the AFLW competition. Key changes include, match quarter times in the women's game are 15 minutes plus time on in comparison to 20 minutes plus time on in the men's game. A maximum of 16 players per team are allowed on the field at one time, five interchange players and uncapped interchanges in AFLW. This is in contrast to 18 players per team on ground, four interchange players and a maximum 90 interchanges during a match for the men's competition. Unlimited interchanges were permitted as the AFLW season is played during the summer months. Also, women play with a size four AF ball size, whilst men play with a size five. Findings from ball size effects in youth males footballers showed AF ball size from four to five did not influence kick performance as participants were able to sustain accuracy and quality of ball spin from the smaller to larger size (Hadlow et al., 2017). The key differences between the women's and men's AF training and match practices may affect the performance monitoring services between the cohorts. Analysing the performances of athletes in training and matches may improve understandings of how the differences in the AF competition structures affect the physical, technical and tactical characteristics of men's and women's AF.

2.2.2 Australian Rules football performance analysis

Sport performance analysis can provide measured information about technical, tactical, and physical athlete and team performances. This information is captured and interpreted by sports science analysts who report quantified data in a useful form back to coaching personnel (Hughes et al., 2007). Monitoring of training and match performances also allows for athlete inter- and intra- differences to be determined. Athlete performance analysis research and monitoring in AF training and matches include common methods of notational analysis, GPS, vision, and data statistical modelling. Video recording and manual coding analysis (O'Donoghue, 2014, Chapter Seven), typically using the software SportsCode (Sportstec Inc., Warriewood NSW), is a common method during training sessions to record skilled performances. Early manual video analysis of match vision resulted in a descriptive analysis of AF male athlete's movement patterns and activities in relation to their field position (Dawson et al., 2004). GPS data analysis and statistical methods using performance indicator data have been researched in-depth resulting in greater insights into AF performances for men and women. Skill performance measures, referred to as performance indicators, describe action variables that explain performance of an individual or team in relation to a successful outcome (Hughes & Bartlett, 2002). The relationships of team performance indicators and match outcome have been investigated in team ball sports for example, rugby codes (Higham et al., 2014; Jones et al., 2004) and soccer (Cintia et al., 2015; Lago-Peñas et al., 2011).

In men's AF the relationships between match outcome and player performance indicators have been assessed in relation to match win/loss (Robertson, Back, et al., 2016), score margin (Stewart et al., 2007; Young, Luo, Gastin, Tran, & Dwyer, 2019), distributions of individual player contributions (Robertson, Gupta, & McIntosh, 2016), player impact rating weighted according to match situation and relative to time on ground (Heasman et al., 2008), player ranking system distributions (McIntosh, Kovalchik, & Robertson, 2018b), and modelling to determine of high frequency representative events of individual athlete match kick performance data (Robertson, Spencer, Back, & Farrow, 2018). Stewart et al. (2007) assessed data from the 2002 – 2006 AFL seasons modelling the relationships between match winning margin and 47 player performance statistics to identify those most closely related to winning margins. The results of the correlation coefficients of variables contributing most to the winning margin indicated that each bounce when running with the ball in general play (r = 0.56), long kicks (r = 0.53), clearances from the centre bounce (r = 0.51), and successful kicks in general play (r = 0.46) were the strongest factors (Stewart et al., 2007). The important player statistics explain variability in the dependent variable score margin at 41% (R-squared 0.41); although the authors noted there were several other factors contributing to team winning margins, the coefficients can determine which areas of performances are contributing most to successful outcomes (Stewart et al., 2007). AFL match performance indicators from the 2013 and 2014 seasons were modelled to determine their ability to explain match outcome defined as win or loss (Robertson, Back, et al., 2016). Binary logistic regression and chi-squared automatic interaction detection (CHAID) classification trees were modelled on the 2013 season data with performance indicators in their relative form and fitted to the 2014 data to test generalisability. The logistic regression achieved a classification accuracy of 88.3%, and CHAID 89.8%. Findings indicated that higher team kick, goal conversion values, and inside 50s were the most important performance indicators contributing towards team match success in elite men's AF (Robertson, Back, et al., 2016). An extended analysis by Young et al. (2019) using a C4.5 decision tree and generalised linear mixed model (GLM) was implemented with a database including the 2001 to 2016 AFL seasons and 97 performance indicators. The decision tree models had a classification accuracy of 88.9% for win/loss and 70.3% for score margin (Young et al., 2019). The GLM showed a prediction score of 6.8 points RMSE for score margin, and classification accuracy of 95.1% for win/loss (Young et al., 2019). Overall, relative to the opposition, the highest predictors towards match outcome were turnovers forced score, inside 50s per shot, metres gained, and time in possession (Young et al., 2019). The work by Young et al. (2019) in analysing relationships between performance indicators and match outcome over different time periods also supports the changes in the tactical nature of AFL match play as found in Woods et al. (2017).

Further data-driven analysis approaches for AF match-play and training applications include the modelling of in-match kick skill performance data using rule induction to quantify the interactions of the constraints experienced on kicks and the relationships of successful kicks (Robertson et al., 2018). The modelling results from Robertson et al. (2018) indicated that a time in possession of the ball of less than two seconds negatively affected kick execution and outcome. High kick efficacy were represented by situations involving a kicker who was under no pressure and kicked either less than 40 m or to an open target (Robertson et al., 2018). Rule induction models can provide a method to quantify the level of representation within a task which can be applied in team sports such as in AF to improve training plan design for replicating match demands (Robertson et al., 2018). The framework which evaluates how accurately a training plan relates to actual competition is known as representative learning design (Pinder et al., 2011). The results from rule induction models can provide a way to quantify contextual competition characteristics for improved evaluation of athlete performances. For example, identifying how relationships of kick skill constraints translate to athlete skill strengths and weaknesses during match play (Robertson et al., 2018). A different analysis approach was taken to implement objective representative training design practices in men's AF using ecological dynamics, a skill acquisition theory (Davids et al., 2013; Woods et al., 2019). The comparison of AF categorised task, environmental, and individual constraint characteristic data derived from coded match vision showed differences in performance indicator behaviours between male elite and semi-elite athletes. Overall, elite men's AF match play was defined by having a higher percentage of total disposals under greater temporal and spatial constraints compared to semi-elite AF (Woods et al., 2019).

Data focused analysis of athlete performances through the use of statistical modelling may improve the quantified insights on team performances. The evolution of match-play in the AFL was investigated from the seasons of 2001 to 2015, overall showing the men's game to have become faster and more skillful (Woods et al., 2017). A multivariate analytical method was used on 18 team performance indicators across the 15 seasons to map whole-of-team game styles and capture temporal trends across the variables (Woods et al., 2017). From the 2005 season data, a rapid change in playing styles to a possession game with increased handballs, disposals, uncontested possession, clangers, marks, and tackles. The trend in game styles shifted to defensive team zoning tactics for repossession of the football from the 2010 season and then to a more blended style of the two trends from 2014 season onwards (Woods et al., 2017). Insights from data-driven approaches towards investigating the training and matchplay characteristics of AF can also be combined with external load measures such as GPS as highlighted below.

Research using GPS data in men's elite AF has investigated match physical demands (Coutts et al., 2010; Gastin et al., 2013; Gray & Jenkins, 2010; Hiscock et al., 2012; Johnston et al., 2012), playing positional physical and skill variations (Kempton et al., 2015), quantified training and match load (Boyd et al., 2013), quantified the relationship between athlete time on field, skilled, and physical output (Corbett et al., 2017), and determined collective team behaviour through spatiotemporal variables (Alexander, Spencer, Mara, et al., 2019). External (GPS) and internal (rating of perceived exertion; RPE) training load measures were analysed together to quantify training and competition loads across a season in men's elite AF (Ritchie et al., 2016). These data informed AF training periodisation and athlete loading across the different season stages (Ritchie et al., 2016). In the AFLW, the physical running demands as measured by GPS data of total distance, high-speed running, and sprinting distance across playing positions were benchmarked from the inaugural season (Clarke et al., 2018), although this only included data from one AFLW team. Match running movements have also been compared between AFLW and sub-elite women's AF athletes, where AFLW athletes showed moderately greater relative high-speed running and sprint distances (Clarke et al., 2019). In men's elite AF, an athlete's skill performance measures, quantified as performance indicators, have shown to contribute more to their coaches' perceptions of performance than match activity profiles in a GPS data report (Sullivan et al., 2014a). But team and individual player match performances should still be considered as a multifactorial relationship between their physical and skill abilities for holistic understanding of the characteristics associated with a successful match outcome (Kempton et al., 2015; Sullivan et al., 2014b).

Specifically in the AFLW, the relationships between player technical involvements, match outcomes, score margins, and ladder position for season one of the AFLW were evaluated (Black et al., 2018a). The first season of the AFLW showed that winning teams per match recorded significantly greater kick numbers ($P = 0.008$), marks ($P = 0.025$), uncontested possessions ($P = 0.022$), disposal efficiency ($P = 0.002$), and greater ratio of inside 50: goals scored ($P = 0.002$) than their losing counterparts (Black et al., 2018a). When assessing score margin, a significantly negative relationship was found between marks inside 50 ($P = 0.040$), the inside 50: goals scored ratio ($P = 0.007$) and larger winning margins, also between losing margins and inside 50s ($P = 0.019$) (Black et al., 2018a). The results suggest that a key aspect to winning matches in the AFLW is being able to maintain the greatest time in possession by executing efficient disposals and winning marking contests. Interestingly, ladder position had a significant relationship with kicks ($P = 0.034$) and contested marks ($P = 0.04$) showing that as numbers of each skill decreased a team's ladder position moved further away from the top position (Black et al., 2018a). Although, modelling AF match performance indicator data by each quarter rather than as a whole match may allow for more specific information on the interplay between variables and success (Gómez et al., 2014). Due to the dynamic nature of AF matches and the ability to reconvene at quarter time breaks, quarter specific data may alter coaching strategies as the game progresses. For example, knowing the changes in skill contributions on a team and individual athlete basis at different match stages could allow coaches to alter tactical approaches during the quarter time breaks based on the game context (Gómez et al., 2013). The contribution of athletes' skilled performances towards a successful team outcome in AFL matches is relatively even across the team (Robertson, Gupta, et al., 2016). But this is yet to be investigated in AFLW, which may result in differences due to factors such as reduced AF game experience of AFLW athletes, a higher number of athletes transitioning from another sport, and lower resourced professional support structures and training opportunities in comparison to the AFL.

Efficient kicking as a distance and accuracy measure in AF has been identified as a key performance indicator contributing to successful outcomes in matches for both the men's (Robertson, Back, et al., 2016; Stewart et al., 2007) and women's leagues (Black et al., 2018a). Long kicks are highly valuable with modelling by Stewart et al. (2007) showing that each additional long kick that was successful added 0.99 of a point to a team's winning margin. Contrastingly, kicks that resulted in a turn-over to the opposition reduced a team's winning margin by 0.62 of a point. The high importance also of inside 50s and maximising metres gained could also be related to having successful teams kicks in creating faster and more efficient ball dominant movements around the field (Young et al., 2019).

Determining the importance of individual and team skill performances towards team match success allows for objective informed decision making across a range of team areas. Performance data relating to player actions has also been analysed to determine player roles in AF. The 2016 season AFL player rating data was modelled using a supervised decision tree, and unsupervised method of Euclidian distances, to classify players into seven a priori determined field match roles and also show levels of player similarity in styles of match play, respectively (McIntosh et al., 2018a). The results provide an objective way to support decision making in team selection, recruitment, injured player replacement, and list management solutions for AF clubs. The overall classification accuracy of the decision tree model was reported as 74.3% and showed that relative to other players the defenders had the majority of their rating points from intercepts, 47.1% for key defenders and 54.5% for general defenders (McIntosh et al., 2018a). The forwards gained the most rating points from mid chain possessions, 49.1% for key forwards and 32.2% for general forwards (McIntosh et al., 2018a). McIntosh et al. (2019) also investigated the relationship between subjective and objective performance rating measures for AF player performances to improve on understandings of how human subjective decision making can relate to the objective measures of AF performances. The findings support the constructive use of both subjective and objective player performance evaluations for AF decision making in clubs. A Pearson's correlations analysis showed a moderate association (r = 0.60) between the objective AFL player ratings data and the subjective Inside Football Player Ratings (IFPR) data collated from the 2013 – 2017 season matches (McIntosh et al., 2019). A linear mixed model showed that kicks and handballs contributed most to explaining performance associated with IFPR values: Beta coefficients 0.844 and 0.646, respectively (McIntosh et al., 2019).

Spatiotemporal metrics from GPS data have been used to assess differences in collective team behaviour during match play phases (Alexander, Spencer, Mara, et al., 2019), and the influence of match phase and field positions on team behaviours game styles (Alexander, Spencer, Sweeting, et al., 2019). Specifically, results from a multivariate analysis of variance to compare match phase, team, and half, demonstrated how collective team behaviour was influenced by match phase (offensive, defensive, contest), with team surface area, length, and width greater during offensive match phase passages (Alexander, Spencer, Mara, et al., 2019). Determining player position during specific match phases to infer game style and team tactics provides a more informed approach than performance indicators and match outcome alone (Alexander, Spencer, Sweeting, et al., 2019). Specifically, the phase of play had greater impact on the team length, width, and surface area in comparison to field position (defensive 50, defensive mid, forward mid, forward 50) of players. Descriptively for example, spatiotemporal analysis showed that players in men's AF tended to position closer to their team's defensive end so as to restrict space when the ball was closer to their goal (Alexander, Spencer, Sweeting, et al., 2019). The analysis of GPS data derived spatiotemporal metrics have improved understandings and formed practical visualisations of AF match play styles (Alexander, Spencer, Mara, et al., 2019). This form of analysis can also be integrated with AF match skill events and outcomes to assess different strategies that teams adopt during certain match phase situations (Alexander, Spencer, Mara, et al., 2019; Alexander, Spencer, Sweeting, et al., 2019). The combination of IMU and GPS devices may provide a more comprehensive data source for performance measures by overcoming the limitations of each device in dynamic field sports such as AF, although further validation and harmonisation of data is required (Camomilla et al., 2018).

Training and match play performance analysis on a team and individual level provides quantified information that can be applied in areas of training design, skill development, team tactics, strength and conditioning programming, and athlete recruitment (Hughes, 2015; Zambom-Ferraresi et al., 2018). Further research and applications of IMUs in AF may provide benefits in improving performance analysis practices towards semi-automated skill recognition.

2.2.3 Australian Rules football kick biomechanics research to practice

The drop punt kick in Australian Rules football is the preferred kick style on-field due to the kicker's ability to better control the accuracy, distance, and speed of execution compared to other kick styles. The characteristic flight style of the ball rotating backwards end-over-end allows for ease of catching for the receiver (Ball, 2008). At the foot and ball contact phase, the orientation of the ball is vertical on its long axis, which on foot impact creates the backspin motion of the ball in flight (Ball, 2011). This flight motion provides better stability through the air compared to other kicks (Peacock & Ball, 2017). Although is reliant on the characteristics imparted on the ball during the foot-to-ball contact phase (Peacock & Ball, 2018a, 2018b). Execution of the drop punt kick action across six defined phases (Ball, 2008) sees the ball being dropped from the hands at hip height towards the kicking foot (Ball, 2011). The ability to drop punt kick proficiently on both legs in AF matches is a tactical advantage in the dynamic unpredictable nature of match play (Ball, 2008, 2011).

Biomechanical analysis of the drop punt kick using a purposely designed laboratory mechanical kicking leg machine (Peacock & Ball, 2017, 2018d, 2018c) has identified key influences on kick execution and outcome. Mainly, the foot impact location plus the ball's orientation during the foot-ball impact (Peacock & Ball, 2018a), and the importance of the kicker's ankle rigidity during foot-ball impact both have marked influence on ball flight characteristics (Peacock & Ball, 2018d, 2018c). Increased stiffness of the ankle joint allows for a greater effective mass to be transferred from the shank during foot-ball impact, which resulted in greater ball velocities during mechanical leg trials (Peacock & Ball, 2018c). Controlled laboratory trials on a mechanical kicking machine have also shown that ball velocity, elevation angle and spin rate are influenced most by the ball orientation about the x-axis (Peacock & Ball, 2017). The strongest relationship between ball orientation and velocity was identified for maximum velocities occurring at a 43° ball orientation about the x-axis (Peacock & Ball, 2017). Recently, the determinant of impact location of a footballer's boot during a drop punt kick showed positive linear relationships between the medial-lateral impact location with the ball azimuth flight angle, and the proximal-distal impact location with ankle plantar/dorsal flexion control (Peacock & Ball, 2018a). Conclusions were drawn that there is a 'sweet spot' on the football boot for ball impact which represented the medial-proximal aspect on the boot. Impact locations around of this point caused decreased foot-to-ball speed ratios (Peacock & Ball, 2018a). Impacting the ball too distally on the boot forces the ankle into higher degrees of plantar flexion placing larger forces through the foot and ankle joints increasing injury risks and reducing ball velocity (Peacock & Ball, 2018c).

Kicking impact characteristics during the drop punt kick need to be coordinated by the player as the execution has direct influence on the ball flight characteristics and therefore success of the kick. Key ball flight characteristics studied have been ball velocity, elevation angle, azimuth angle, and ball spin rate (Peacock & Ball, 2017). On-field analysis of kick impact variables and drop punt kick accuracy showed that azimuth ball flight trajectory was the most influential ball flight characteristic for an accurate kick (Peacock & Ball, 2018b).

The concepts and research methods of biomechanical analysis in soccer kick skills are logical to review to inform of women's specific differences compared to men's for rationale behind undertaking AFLW kick skill research. Interestingly, several kinematic differences were found for elite soccer instep, inside, and curve kicks across leg preferences for both men and women (Alcock et al., 2012; Barfield et al., 2002; Sakamoto et al., 2014; Sakamoto & Asai, 2013). Significantly lower ball velocities and mean foot velocities from female athletes for instep and inside kicks were shown compared to male athletes (Barfield et al., 2002; Sakamoto et al., 2014; Sakamoto & Asai, 2013). Although female athletes generally produced lower ball velocities in maximal kicks, findings do suggest that female athletes are capable of producing similar velocities to those reported for elite males when the accuracy constraints imposed are realistic to match conditions (Alcock et al., 2012). The mean ball: foot velocity ratio of female's instep kicks (1.23 ± 0.16 m·s^{-1} females; 1.31 ± 0.18 m·s^{-1} males) and inside kicks (1.37 ± 0.14 m·s^{-1} females; 1.41 ± 0.16 m·s^{-1} males) was also significantly lower than male athletes (Sakamoto & Asai, 2013). Quantified full-body kinematics of elite female soccer athletes showed that resultant foot velocities at ball impact were not different for instep and curve kick types (Alcock et al., 2012). The results importantly identified key coaching points for how foot velocities are generated for the two kick types. In summary, instep kicks were characterised by using a faster run approach and significantly greater linear hip velocity (1.34 ± 0.48 m·s^{-1} curve; 1.81 ± 0.56 m·s^{-1} instep) and knee velocity (3.38 ± 0.49 m·s^{-1} curve; 4.09 ± 0.71 m·s^{-1} instep) during impact. Whereas, the curve kick required significantly greater knee angular velocity (31.0 ± 3.7 rad·s^{-1} curve; 28.4 ± 4.7 rad·s^{-1} instep) as a control mechanism for foot orientation and foot speed generation (Alcock et al., 2012). The progressive kick biomechanics research undertaken in soccer female athletes has been influential in highlighting technique differences under varying constraints during kicking. Notably, research specifically looking at elite women's soccer attributes of successful direct free kicks and mechanisms for expert kick skill performances (Alcock, 2010). Specialised research may allow for more tailored coaching approaches for elite female athletes. Kinematic characteristics using male participants in AF may not apply to females if significant differences have been found in kicking styles for soccer. Also physiological factors such as muscular strength and power, and striking mass between the sexes may affect foot velocity and the mechanical quality of the kick (Sakamoto & Asai, 2013). As there has been limited biomechanical research in elite women's AF, it is warranted to further investigate stand-alone women's kicking data especially given the already established numerous pre and post foot-to-ball impact characteristics that have an effect the intended kick success for distance and accuracy constraints.

Characterised foot-to-ball interactions in AF kicks as a measure of performance between the preferred and non-preferred leg provide information for defining measured kick efficiency (Ball et al., 2010; Smith et al., 2009). This information can be also practically be applied in areas of skill acquisition coaching, and strength and conditioning techniques for AF players as a method to improve kick skill. Maximising distances of the drop punt kick have been associated biomechanically with greater foot speeds and larger shank angular velocities at ball contact, a larger last step on the stance leg prior to the kick, and a ball drop contact point higher off the ground further from the support leg (Ball, 2008). Statistically significant measurement parameter differences were found between the preferred and non-preferred kick legs for experienced male AF players (Smith et al., 2009). Comparisons of these results were later made to junior male AF players for the preferred kick leg showing differences between kick impact characteristic measures (Ball et al., 2010).

The speed of the foot prior to ball contact has shown to be the key variable in the energy transfer ability from foot to the ball; in order to increase the resulting ball velocity (Ball, 2008; Peacock et al., 2017; Peacock & Ball, 2017) and in controlling kick distances (Peacock et al., 2017). A higher foot velocity resulted in a statistically significant linear increase in ball velocity and ball spin rate on a mechanical kicking leg machine (Peacock & Ball, 2017). Foot velocities for preferred and non-preferred leg submaximal drop punt kicks have been reported in men's elite AF as 22.4 ± 0.7 m·s^{-1} on the preferred kick leg and 19.2 ± 1.0 m·s^{-1} on the non-preferred kick leg (Ball, 2013). When aiming for greatest distance kicks, a foot velocity of 26.4 ± 1.2 m·s^{-1} was reported as an elite men's benchmark (Ball, 2008).

The ball: foot ratio represents a widely reported measure of kick impact efficiency in AF (Ball et al., 2010, 2013; Peacock & Ball, 2018d; Smith et al., 2009) and soccer (Nunome et al., 2018; Sakamoto & Asai, 2013; Shinkai et al., 2009). The ratio provides a single metric evaluation to describe the efficiency of energy transfer from the foot to the ball, and consequently a reflection of the summation of forces through segment chains of the legs and hips (Ball, 2011). No significant differences have been found between the ratio and kick leg preferences in elite men's AF (Smith et al., 2009). This result indicates that the faster ball velocities on preferred leg kicks are attributed to faster leg swing (higher knee angular velocity) generating faster foot velocities and therefore more force imparted on the ball (Nunome et al., 2006; Smith et al., 2009). It has been suggested that direct comparison of the ball: foot ratio value between male and female, senior and junior athletes may be confounded by the effects of body mass (Shinkai et al., 2013). Therefore other kinematic variables should be evaluated instead to explain inter-group differences.

As the kick leg swings through, the support leg provides the platform of stability to facilitate efficient movement (Ball, 2013). Different kinematic characteristics of the support leg have been suggested based on results in AFL research under changing distance and accuracy constraints imposed (Blair, Robertson, et al., 2018). Drop punt kicks requiring higher accuracy and performed at lower speeds show greater support leg knee flexion leading to a lower centre of gravity to create more stability (Dichiera et al., 2006). In contrast, distance kicks have shown a more extended knee on the support leg (Ball, 2013; Blair, Robertson, et al., 2018). During high impact kicks, athletes tend to adopt a more upright position through the torso and therefore a higher hip position in order to generate faster foot velocities during the leg swing prior to ball impact (Ball, 2013). Further research to investigate the adoption of different techniques by athletes under differing accuracy and distance constraints would be of benefit to kick skill coaching (Blair, Robertson, et al., 2018).

Kinematic assessment of movement patterns for drop punt kick performance provide insight into the muscle interplay of the kick motions and outcomes (Ball, 2011; Coventry et al., 2015; Dichiera et al., 2006). Significant differences have been found between the kinematic kicking patterns of dominant and non-dominant kicking legs in AFL (Ball, 2011) and skilled accurate and lower skilled kickers (Dichiera et al., 2006). Preferred leg kicks showed greater knee angular velocities, and foot and ball velocities. Non-preferred leg kicks produced higher hip angular velocities and hip angles in both men's (Ball, 2011). Changes in movement patterns between kick leg types may indicate imbalances in stability or less efficient use of sequential summation of momentum through the kicking motion (Ball, 2011). The velocity of the foot leading into ball contact is a function of the linear and angular velocities of the knee and shank where greater amounts of work through these segments would be originating from the thigh angular velocity (Dörge et al., 2002). Efficient transfer of velocity from the foot to the ball may be reliant on the inter-segmental patterning of hip and leg segments (Ball, 2011; Dichiera et al., 2006; Dörge et al., 2002). The technical differences found between kick legs suggest altered strategies are used by athletes during high impact kicks (Ball, 2008; Blair, Robertson, et al., 2018). Similar kick performance outcomes may be achieved with either movement strategy as performance indicators of foot velocity and kick distance were not significantly different between each approach for men's AF (Ball, 2008), although this is yet to be determined in women's AF. Kinematic kick data scaled on a thigh or hip –to-knee angular velocity continuum shows that the majority of preferred leg kicks would be classified as a knee dominant strategy with increased contribution from the knee segment and lower hip or thigh involvement (Ball, 2008, 2011). This is the opposite for non-preferred leg kicks in men's AF (Ball, 2008, 2011).

Research into muscle recruitment and activity during a drop punt kick via MRI scans showed that the gracilis, tensor fascia latae, semitendinosus and rectus femoris muscles were most active (Baczkowski et al., 2006). Throughout the kick phases the quadriceps were the most active muscle group contracting eccentrically during the wind-up phase then concentrically during the forward swing phase. Although quadricep activity does decrease markedly at the end of the forward swing phase (Orchard et al., 2002). These findings are consistent with the Australian Football injury reports citing quadricep injuries as one of the most common injury category, particularly rectus femoris muscle tears and strains (Orchard et al., 2002). This is thought to be caused by the mechanisms of the kick where by at the time of foot-ball contact the impeding torque from the ball is transferred to the extended and contracted thigh muscles (Orchard et al., 2002).

The movement control and strength of the hip and lower limb muscles for both the kicking and support leg are influential factors in the ability to accurately drop punt (Dichiera et al., 2006), and in producing faster foot speeds especially as fatigue increases (Coventry et al., 2015). A fatigue induced protocol resulted in reduced support leg control at the hip and knee which would likely reduce the energy transfer into the kick (Coventry et al., 2015). Interestingly, foot speed differences pre and post the fatiguing protocol were not significant and increases in foot speed were found during short-term fatigue (Coventry et al., 2015). Possible movement adaptions may be made by experienced athletes under short-term fatigue to maintain foot speeds such as increased range of motion through the pelvis region (Coventry et al., 2015).

When specifically looking at the technique of distance goal kicking, technical differences were found when comparing successful goal kick distances of 30 to 40 metres (Blair, Robertson, et al., 2018). The study findings indicated that 40 m goal kicks had greater support-leg knee extension during stance phase, higher kick foot speeds (19.9 $m \cdot s^{-1}$ for 40 m, and 18.0 $m \cdot s^{-1}$ for 30 m), and higher shank (1736 $deg \cdot s^{-1}$ for 40 m and 1642 $deg \cdot s^{-1}$ for

30 m) and knee angular velocities (1632 deg·s^{-1} for 40 m and 1446 deg·s^{-1} for 30 m) at ball contact (Blair, Robertson, et al., 2018). These current results improve upon the differentiating results between the technical aspects of goal kicking performance when testing for accuracy and distances measures. As athletes are faced with both constraints during AF matches, biomechanical and kinematic analysis enhances the evidence-based practices for coaching and conditioning. For example, facilitating range of motion and the ability to produce higher angular velocities through the hips and lower limbs can be developed through dedicated strength and conditioning programs and joint mobility manipulation treatments. Coaching of kicking movement cues and technique variables should also be considerate of individual player differences to ensure adaptations are optimal for that athlete (Ball, 2008).

In translating the laboratory and on-field biomechanical research of AF kicking to practical applications, there are several key strategies suggested in order to improve one's kick skill. These include: to increase ankle rigidity, increase foot velocity prior to ball contact, and refine the impact location of the football boot contacting the ball. Muscle activation and strengthening of the ankle dorsal flexors, foot muscles and ankle tendons also appear important in decreasing the passive ankle plantarflexion experienced during the ball impact phase. Improving kick technique may also have implications for prevention of injury, for example by limiting unnecessary movement generated at the ankle joint or inefficient transfer of forces through the lower limb which may creating imbalances. Coached movement cues and kick drill protocols practicing maximum kick intensity and distance targets have been linked to improving drop punt kick distance ability and foot speeds (Ball, 2008). Evidence-based kick skill assessments are an important tool for quantifying the impact of changes in technique. Specifically, the constraints of AF match kicking were taken into account to develop a representative AF kick proficiency assessment, Australian Football

Field-Based Dynamic Kicking Assessment (AFFB-DKA) (Bonney et al., 2019). Using laboratory-based research on in-context for an applied on-field kick assessment contributes towards improved evidence-based practices for talent identification and match specific kick skill development across age groups (Bonney et al., 2019).

The technical level of women's AF is rising alongside participation rates throughout the competition level pathways. In turn, this creates demand for improved skills and technique at an elite level to enable faster and more professional match play. Findings from kinematic studies using male athletes may not apply to females, and analysis of kick techniques specific to female population is required. Women's AF kick kinematic analysis would be beneficial in identifying the technique modifications required to achieve efficient kicks and assist coaches in focussing on cues required for individual athletes in their kick skill development. Knowledge of the biomechanical characteristics of AF kicking can be incorporated into methods for IMU kick performance monitoring. For example, refining the placement of the sensor in relation to the movement planes the kicking motion to record a more accurate representation of the kick over other movements performed around a kick.

2.3 Inertial measurement units

Inertial measurement units (IMUs) integrate a combination of accelerometer, gyroscope and magnetometer sensors. Each individual sensor type can be formatted as a single-, dual- or triaxial configuration. For the purposes of this thesis, only triaxial sensors, which record along the x, y, and z axes will be implied or used. The sensors measure across mutually orthogonal sensitive axes (Sabatini, 2011). When mounted in alignment on a targeted body segment will measure in correspondence to the three anatomical axes; anterior-posterior, medio-lateral, and vertical. The positioning of the sensor on a particular limb or trunk body area is important in relation to the targeted movement, and should align with the anatomical axes for accurate data representations (Fong & Chan, 2010). Further explanations of each measurement sensor type are detailed in the subsequent sub-sections 2.3.1 to 2.3.3. IMUs are named with the term "inertial' due to the in-built accelerometer and gyroscope measuring by the principle of inertia (Camomilla et al., 2018). Devices just containing accelerometers and gyroscopes will provide orientation in two-dimensions (2D), whereas the inclusion of a magnetometer allows for three-dimensional (3D) measurements (Camomilla et al., 2018). The 3D orientation of a rigid body can also be described by Euler angles using the combination of three rotations around different axes (Sabatini, 2011). The Euler angles are a way of defining spatial orientation of a frame of space in relation to a referenced fixed-frame. Developments in fabrication technology of Micro-Electro-Mechanical Systems (MEMS) has allowed for sensors to be manufactured at lower costs, with improved power consumption efficiency, and smaller in size (Sabatini, 2011). Advantages of IMUs include: can provide close real-time data both in laboratory and field settings (Chambers et al., 2015; Mayagoitia et al., 2002; Parrington et al., 2016), are relatively inexpensive comparative to common GPS technology, have a low burden on an athlete's movements, and attest a battery life substantial enough for extended recordings in true motion context (Aminian & Najafi 2004; Bulling, Blanke & Schiele 2014). Being self-contained and sourceless units, they are not reliant on external devices such as antenna receivers for positional reference (Neville et al., 2011). This key property makes IMUs accessible and portable for field-based sport data collection and monitoring practices. The sensor advantages allow for full repeated movement patterns to be captured during sports training and competition (Brodie, Walmsley & Page 2008; Fasel et al. 2015; Walker et al. 2016).

The placement or mount site of an inertial sensor on the human body should be in relation to the targeted movement but data accuracy may be affected by this choice (Fong & Chan, 2010). It is recommended that the axes of the sensor align with the anatomical axes of the body segment for which it is mounted (Fong & Chan, 2010). Soft tissue artefacts, referred to as oscillations between the sensor, skin, and underlying bone, are considered a common error source in biomechanical kinematic motion measurements (Camomilla et al., 2017). Subject individual anthropometry may also introduce different levels of muscle contraction force and skin tissue movement. This may require either standardisation or compensation of sensor site attachment and fixation method for more accurate inter- and intra-trial data (Camomilla et al., 2018; Cereatti et al., 2017; Forner-Cordero et al., 2008). Taking into account soft tissue artefacts is an important factor for accurate biomechanical bone acceleration and angular velocity measurements (Cereatti et al., 2017; Liu, Inoue, & Shibata, 2009). Therefore, consideration as a source of error may also be due in sports movement classification. For specificity in sport movement recognition, the location of the sensor in relation to the targeted movement, i.e. on the wrist for a tennis stroke, and a standardised fixation protocol, i.e. 2 cm above the lateral malleolus, may be important variables more so than physiological soft tissue artefacts. A proper sensor fixation method will aid in reducing errors from skin movement and imprecise alignment (Fong & Chan, 2010). Although, a multi-sensor alignment would be preferential in recognition of a range of motions and their hierarchical complexities (Bulling et al., 2014). This set-up is generally not feasible in a sport's training or competition space, as opposed to capturing activities of daily living or clinical assessments. Therefore, investigation into best approaches using inertial sensors within individual sports is required. Locations on the torso such as the lower back, upper back or hip are common and can capture major sporting movements (Camomilla et al., 2018; Chambers et al., 2015). But data can be prone to underestimating specific limb dominant movements (i.e. kicking or cycling) as the sensors are not proportional the kinematic joint actions (Mannini et al., 2013; Norris et al., 2014). Accelerometers embedded within a GPS device housed within an elasticised harness on the back between the shoulder blades have shown poor reliability and validity in representing thoracic acceleration, 3D centre of gravity acceleration, and peak vertical ground reaction force (Edwards et al., 2019). Also, by housing the unit in a harness rather than directly on the skin, study results confirm that this is a major contributor to irrelevant accelerometer magnitudes within the output data (Edwards et al., 2019). An accelerometer will directly measure the acceleration of the segment it is attached to; as such, caution is required when assessing whole-body accelerometry in team-sports (Nedergaard et al., 2017). It has been suggested that lower limb mounted accelerometers are more accurate than trunk mounted for whole body loading with the larger impact force exposure during foot-to-ground contact (Nedergaard et al., 2017). The use of lower limb mounted accelerometry in dance sports provided improved load metrics than the trunk mounted accelerometer and therefore was recommended as a more accurate measure of dance athlete movement and injury monitoring (Brogden et al., 2018). When considering the use of IMUs for a continual monitoring practice of athletes, the placement of the IMU must be unobtrusive and not affect the ability of the athlete to conduct their normal sport movements.

IMUs capturing in 3D provide important kinematic measurements of linear acceleration, angular velocities, and orientation of a body segment or sporting apparatus, for example a golf club (Jiao, Bie, et al., 2018). The problem domain of a specific sport will dictate the inertial sensor implementation protocol. The use and data fusion of individual sensor data types within an IMU may also impact the movement recognition problem, especially considering that signal property features for the same activity may differ largely depending on the placement of the IMU relative to the body reference (Zimmermann et al., 2018). The advantageous properties and data capture potential of IMUs allow for many potential specialised applications in sport science programs to suit the sporting context (Camomilla et al., 2018).

2.3.1 Accelerometers

Accelerometer sensors detect acceleration along one to three orthogonal axes in proportion to external forces (Yang & Hsu, 2010). Acceleration is measured by the principles of Newton's 2nd law of motion, Law of Acceleration (Equation 1), and Hooke's Law (Equation 2), to result in an equation for acceleration (Equation 3) (Kavanagh & Menz, 2008; Robertson et al., 2013).

Equation 1. Law of Acceleration

$$F = ma$$

Where m = mass, a = acceleration

Equation 2. Hooke's Law

$$F = kx$$

Where F= force, k= spring constant, x = spring displacement

Equation 3. Resulting equation for acceleration

$$a = (-kx)/m$$

(Busa & McGregor, 2008)

Measurement units are represented as gravitational acceleration g-force (g), in which 1 g = 9.81 m·s^{-1}. This represents the change of velocity with respects to time. Dynamic and static accelerations are measured and detected then converted to an electrical signal. Dynamic acceleration refers to forces other than the gravitational force applied to a rigid body. Whereas, static acceleration represents the gravitational force experienced by a rigid body. Technically, there are numerous types of accelerometers including MEMS, strain, gauge, capacitive, piezoresistive, and piezoelectric accelerometers (Kavanagh & Menz, 2008; Robertson et al., 2013). The preferred type for motion-sensing is a MEMS accelerometer. The internal mechanisms of the sensor constitutes a mass-spring system which when subjected to compression or stretching forces from a movement will cause the spring to produce a reinstating force directly proportional to the inflicted force (Kavanagh & Menz, 2008).

Accelerometers can measure properties of human movements (Godfrey et al., 2008; Pelham et al., 2006) and sporting activities (Wundersitz, Gastin, Robertson, et al., 2015). Accelerometer-based activity monitoring in everyday physical activity and sedentary behaviour research as an objective measure is prominent and moving towards incorporating advanced machine learning (Attal et al., 2015; Farrahi et al., 2019). Triaxial accelerometers output acceleration amplitude and direction in a 3D space which can be used to determine the frequency and intensity of human movements (Tamura, 2014). A direct measure of acceleration is often preferred due to increases in signal noise when distinguishing the signal from calculating the first or second derivatives of velocity or displacement respectively (Tamura, 2014). A MEMS based triaxial accelerometer allows for the direct measurements. Inferences of loads placed though the body are made based on Newton’s second law of motion (Ahmadi et al., 2015), and have been applied to quantify physical loads during training and competition (Gabbett et al., 2012). Although caution should be used when inferring whole body loads in the field using body-worn accelerometers as recent research found weak linear relationships and overestimation of whole-body mechanical loading variables from segmental-accelerometer data (Nedergaard et al., 2017). Furthermore, the use of segmental acceleration signals as a valid estimate of ground-reaction forces for measures of whole-body biomechanical loading during running-based sports has shown to be inaccurate (Verheul, Gregson, et al., 2019; Verheul, Warmenhoven, et al., 2019). On-field application examples of accelerometer data include the quantification of tackling impacts and loads in AF (Gastin et al., 2013, 2014) and rugby league collisions (Gabbett et al., 2010). The functionality of a commercial IMU designed to measure head impacts through linear and angular velocities was tested for accuracy in AF (McIntosh et al., 2019). Although the authors indicated caution in regards to the data interpretation due to measurement errors and further laboratory research is required (McIntosh et al., 2019). The combination of an accelerometer and gyroscope in an inertial sensor can improve the capture of measurement variables for evaluating movements (Tamura, 2014).

2.3.2 Gyroscopes

The gyroscopes referred to in this thesis and most commonly implemented in movement analysis measure angular velocity through an internal vibrating mechanical element reacting to Coriolis acceleration, which equates to the force of a vibrating mass relative to its velocity and the angular velocity of the rotating frame (Aminian & Najafi, 2004; Zeng & Zhao, 2011). To illustrate the internal mechanisms of a gyroscope, it consists of a spinning wheel on a moveable frame that will maintain its original orientation in space when spinning, irrespective of the central applied forces (Tamura, 2014). The rotation rate can be measured in reference to one to three rotational axes referred by aviation terminology as yaw (z-axis), pitch (y-axis), and roll (x-axis). The advantages of a gyroscope and differences from an accelerometer sensor include: there is no signal influence from gravitational and linear forces and that the signal output will generally contain less noise (Aminian & Najafi, 2004).

Although, gyroscopes are limited by their sensitivity to shock forces (Aminian & Najafi, 2004) and the orientation change calculations are open to integration drift (Luinge, 2002). This limitation can be compensated for through mathematical techniques which integrate accelerometers to limit the errors from the gyroscope integral (Tamura, 2014). As human movements are largely rotational around the body joints (Ahmadi et al., 2015), gyroscopes may provide further movement signal features for greater discriminatory analysis of dynamic in-field sensor data compared to accelerometer data alone (Tamura, 2014).

2.3.3 Magnetometers

Magnetometers are not typically used for IMU human movement data analysis. The functionality is to determine the direction of travel for the sensor relative to the Earth's magnetic north pole through the strength of the local magnetic field (del Rosario et al., 2015). These measurements contribute to the sensor orientation detection of pitch (attitude), roll (inclination), and yaw (heading) (del Rosario et al., 2015). As orientation is determined by the surrounding magnetic field, this sensor is sensitive to other electronic device outputs that can affect the magnetometer signals. Furthermore, in comparison to both the gyroscope and accelerometer, the frequency response of magnetometers is weak (del Rosario et al., 2015). Therefore, generally magnetometers are seen as additional data during IMU sensing, and less commonly used for sport movement recognition model development (Camomilla et al., 2018). Results from an investigation of the ability to compute angular velocity from 3D magnetometer data for HAR in comparison to gyroscope data (Kunze et al., 2010), made a general case for possible data integration with other sensor types, but was not conclusive in the magnetometer's reliability for HAR.

The features associated with accelerometers, gyroscopes, and, or magnetometers are considered complimentary when integrated in an IMU for providing accurate estimates of

3D motion sensing (Aminian & Najafi, 2004; Sabatini, 2011). Investigating data fusion methods from the individual sensor types or a network of IMUs may provide greater accuracy in classification of discrete skills in sport. Also, it may favour subject individualised signature signal feature extraction.

2.3.4 Validity and reliability of inertial measurement units

Validity refers to the degree in which the measure or variable reflects that of the criterion or true values that are being characterised; this principle is also denoted as concurrent validity (Hopkins, 2000). Or put simply, the capacity for the device, in this case an IMU, to output what it is intended to quantify (Atkinson & Nevill, 1998). IMUs should accurately quantify the acceleration and angular velocities for specific movements or impacts. Also, when used in these situations having the appropriate hardware capabilities to fulfil the demands required. Therefore, be able to provide clear output for further post-processing to detect and recognise sport-specific movements through semi-automated processes. Reliability testing involves the reproducibility of output from repeated trials under the same testing conditions or same measurement device (Hopkins, 2000). Or in further terms, the minimisation of measurement error to a level that is “deemed acceptable for the effective practical use as a measurement tool” (Atkinson & Nevill, 1998, p. 219). The two types of error in reliability assessment are systematic bias; the tendency for measurements to change in a positive or negative direction across repeated trials. And random error, the inherent biological or mechanical variations in a measurement device or protocol (Atkinson & Nevill, 1998). A degree of error is generally going to be present for continuous measurements so it is important to identify and attempt to minimise these where possible. Inertial sensors need to be reliable in providing consistent measurement outputs for athlete performance monitoring over time. Laboratory-based validation work on commercial triaxial IMUs against a gold standard custom made mechanical testing apparatus showed accuracy measurements within 0.6 degrees and precision within 0.1 degrees for static sensor orientation (Taylor, Miller, & Kaufman, 2017). Angular velocity was measured accurate within 4.4 $deg \cdot s^{-1}$ and precise within 0.2 $deg \cdot s^{-1}$ (Taylor et al., 2017). Clinical research using IMUs and showing valid results highlight the advantages of the devices and extends scope for their use outside a clinical environment (Ermes et al., 2008; Karantonis et al., 2006; Mayagoitia et al., 2002).

Optoelectronic or optical motion capture systems are considered the gold standard in motion analysis (Corazza et al., 2010; van der Kruk & Reijne, 2018) and commonly used in assessing the output of inertial sensor measures (Cuesta-Vargas et al., 2010). Three-dimensional optical motion capture analysis systems constitute an accepted criterion for measuring linear and non-linear human movements through a multi high-definition camera setup capturing light-reflective markers mounted on specific anatomical landmarks (Richards, 1999). Biomechanical measurements of displacement and velocity in relation to the movements and body positionings are calculated through digitising the multiple frames of the markers captured. The Vicon capture system (Vicon Nexus v2, Oxford, UK) is a common software program used in sport and clinical analysis (Blair, Duthie, et al., 2018; Cuesta-Vargas et al., 2010; Seel et al., 2014). The accuracy and precision of the Vicon system has shown to be strong given a favourable camera set-up and calibration, although performance is also dependent on several factors such as occlusion, system noise, optical biases, and the digitising process (Windolf et al., 2008). The concurrent validity of a whole-body inertial measurement system (Xsens MVN system, Xsens Technologies B.V., Enschede, The Netherlands) was tested against Vicon for measuring lower body kinematics during AF, rugby union and rugby league kicking (Blair, Duthie, et al., 2018). Concurrent validity was assessed with a linear mixed model showing trivial to small measurement errors between the inertial measurement system and Vicon across all kinematic parameters (from

0.1% to 5.8%) (Blair, Duthie, et al., 2018). The results from this research support the use of IMUs to quantify biomechanical properties of sport kicking movements in an applied field setting. Evaluation of a commercial manufacturer inertial movement analysis (IMA) software for jump detection and jump height quantification (Catapult Innovations, Melbourne, Australia) based on the inertial sensor data from the manufacturer's device (Catapult MinimaxX S4) was compared against a 36-camera (detection component) and 12-camera (quantification component) 3D motion analysis system (Spangler et al., 2018). The IMA jump detection overall accuracy was 99.3%, specificity 99.7%, and sensitivity 95.8%. The IMA software underestimated jump height quantification with a significant mean bias and moderate absolute error (Spangler et al., 2018).

Clinical laboratory-based analysis assessed the concurrent validity of multiple uniaxial accelerometer and gyroscope sensors mounted on the lower limbs (sampling at 100 Hz) against a video based motion capture system (sampling at 50 Hz, Vicon, Oxford, UK) to examine gait kinematics at various treadmill walking speeds (1.4 km/h to 4.6 km/h) (Mayagoitia, Nene & Veltink 2002). The kinematic parameters of the foot and shank in the sagittal plane were: shank angle, thigh angle, knee angle shank angular velocity, thigh angular velocity, knee linear acceleration, shank angular acceleration and thigh angular acceleration (Mayagoitia, Nene & Veltink 2002). The model in 2D included the shank and thigh rigid segments connected by the knee hinge joint in which the shank comprised of both the foot and shank parameters (Mayagoitia, Nene & Veltink 2002). Both the accelerometer and gyroscope outputs were similar to Vicon during treadmill walking, overall root mean square error 6.64% and standard deviation 4.13%, suggesting feasible evidence for the use of IMUs as an alternative to traditional lab-based motion analysis systems in gait kinematics. Future considerations did arise for accelerometer use due to several signal peak deformations at higher walking speeds, noted as possible contact or vibrations during heel strike

(Mayagoitia, Nene & Veltink 2002). Accelerometers and gyroscopes have shown the capability to provide reliable results for running gait parameter assessment (Norris et al., 2014). The ability to accurately detect running gait temporal events as a function to quantify inner-stride phase durations was evaluated using a foot-worn IMUs against a force plate reference system treadmill (T-170-FMT, Arsalis, Belgium) (Falbriard et al., 2018). Repeated 30 second run trials under a protocol starting at 8 km/h and incrementing 2 km/h until maximum speed were obtained from each participant. The study noted the effect of the kinematic signals and features in avoiding errors during stride phase detection through gait and temporal feature detection algorithms proposed against a reference threshold on the vertical ground reaction force recorded by the treadmill. Ground contact time, flight time, and step and swing time were quantified with an inter-trial median ± IQR bias < 12 ± 10 milliseconds and precision < 4 ± 3 milliseconds (Falbriard et al., 2018). Although changes in the running speeds significantly affected the biases of the estimate metrics, in which the authors suggested altering the algorithms to apply a speed-dependent correction for possible system accuracy improvement (Falbriard et al., 2018). The validity of using IMUs for measuring 3D joint kinematics in three functional movements of a bilateral squat, single-leg squat, and countermovement jump showed RMSE and range of motion error evaluations below 3 degrees for all joint measurements in these movements; mean coefficient of multiple correlation values ranging from 0.77 to 1 (Teufl et al., 2019). A sensor-fusion algorithm calculated the IMU kinematic values for evaluation against an optical motion caption system. Although, the more dynamic task of a counter movement jump did produce error measures of approximately 1 degree higher than the bilateral and single-leg squats (Teufl et al., 2019).

Specific applied field research by Parrington et al. (2016) assessed the validity of a triaxial IMU sampling at 500 Hz (IMeasureU Blue Thunder sensor, Auckland, New Zealand) against a laser speed gun (Laveg Sport Jenoptik, Germany) for 100 m running sprint repetitions. Results indicated a strong correlation with the Laveg laser criterion data for average split velocities after the first 10 m (r = 0.85 – 0.95) and peak velocity (r = 0.92). Further field concurrent validity testing against a Vicon camera system include a multi-IMU set-up (IMeasureU, Auckland, NZ) for discus throwing torso and pelvis dynamics (Brice et al., 2018) and tennis shot metrics derived from two commercial IMU sensors designed and marketed for tennis applications (Keaney & Reid, 2018), indicated applied potential but also noted areas for further investigations. Discus throw measurements showed good validity for individual segment transverse plane orientation data, although validity was low for shoulder-pelvis separation angle measurements (Brice et al., 2018). Agreement between the IMU and motion capture time-series transverse plane orientation data resulted in RMS values indicating accuracies between 2% and 3% (Brice et al., 2018). When measuring tennis shot type using the impact location and racket speed of two commercial tennis racquet sensors (Babolat Play tennis racquet and Zepp sensor), both sensors detected the same total stroke volume as Vicon, but only had a moderate agreement with actual stroke type classification; Cohen's kappa: Babolat = 0.730 and Zepp = 0.612 (Keaney & Reid, 2018). A minimal agreement was found for racquet impact location; Cohen's weighted kappa: Babolat = 0.412 and Zepp = 0.217. The Zepp sensor showed a close perfect agreement with Vicon on racket speed with an intra-class correlation coefficient of 0.983 ($p < 0.001$) (Keaney & Reid, 2018). In swimming, validity and reliability testing of accelerometers assessed the ability to detect lap time, velocity, stroke duration, stroke rate, and stroke phase (Callaway, 2015). Four triaxial accelerometers were mounted on the upper body of twelve swimmers which were assessed in-situation against video analysis from an underwater side view camera on a moveable trolley and a stationary global camera used as a global timestamp measure. Lap time results showed strong positive correlations (r = 0.98, $p < .001$) between video time and accelerometer time assessed via Pearson's r correlation (Callaway, 2015). Paired samples t-tests showed no significant differences between video time and accelerometer time (Callaway, 2015). The validity of stroke count showed strong positive correlations using Pearson's r correlation ($r = 0.95$, $p < .001$) (Callaway, 2015). Validation of stroke rate using t-test had a strong positive correlation ($r = 0.92$, $p < .001$), and a Bland-Altman plot produced a mean error of -0.25% (Callaway, 2015). Stroke duration also showed strong significant positive correlations between video and accelerometer output ($r = 0.64$, $p < .001$) (Callaway, 2015). The reliability of stroke phase detection produced low mean absolute errors for each stroke phase of entry (0.06 MAE ± 0.06 SD), pull (0.07 ± 0.08), push (0.06 ± 0.05), and recovery (0.08 ± 0.1) (Callaway, 2015). In spring board diving the validity of a triaxial gyroscope (IMeasureU, Auckland, New Zealand) against a 3D optical system (Cortex 3.3 Motion Analysis Corporation, USA) to measure angular velocity was first assessed mechanically in a lab before applied in a field setting (Walker et al., 2017). Laboratory validation results showed a Pearson's correlation of $r = 1.00$ ($p < .001$) from all individual tested gyroscopes with the optical system and a linear regression output of $R_2 = 1.00$ between the mean gyroscope and optical system angular velocity measurements (Walker et al., 2017).

The use of IMUs for sporting and everyday activity as a performance measure tool appears feasible across several movement domain. Commercially available sensors are not required to release details of their data measurement and signal processing techniques or conform to player analysis data standards outside competitions for a certain sporting governing body (Nicoella et al., 2018), therefore research on the validity and reliability of the sensors is required.

2.3.5 Inertial measurement unit signal processing for movement recognition

The characteristics of activities performed whether whole body or an isolated limb movement are assumed to translate to IMU signal pattern representations which can be identified through computer analytical methods for movement recognition (Bulling et al., 2014; Ordóñez & Roggen, 2016). Key challenges that can arise with activity recognition using IMUs are intraclass variability, where the same activity is performed differently and therefore represented uniquely in the sensor data; so models require adaptation to the variability of the same actions (Bulling et al., 2014; Nweke et al., 2018). Likewise, interclass similarity, when different actions have similar features in the sensor data and could be classed as the same action (Bulling et al., 2014; Nweke et al., 2018). In sport, data collection may involve varying amounts of individual action performance instances leading to class imbalances; this may be overcome by oversampling methods (Chapter Eight). The collection of IMU data in the sporting field must also be as unobtrusive for the athletes as possible, but undertaken in a manner that is consistent with each collection period for model accuracy by attaining clear signal information specific to the domain activities performed and their unique characteristics.

In Bulling et al. (2014), an activity recognition chain (ARC) framework is described for supervised model training and classification when working in HAR. The four stages in order are: sensor data acquisition (Section 2.3) and pre-processing, data segmentation, feature extraction and selection (Sections 2.4.3 and 2.4.4), training and classification (Section 2.4.2), and finally performance evaluations (Section 2.4.5) (Bulling et al., 2014). Signal data from inertial sensors in its raw format is often inconsistent in measures, contain missing values, noisy artefacts or outliers which can limit the efficiency of machine learning algorithms (Bux et al., 2017; Kautz, 2017). Pre-processing and data segmentation are important stages in machine learning model development requiring domain knowledge in order to improve the overall model performance results.

Inertial sensor movement capture data can also result in high-dimensional datasets based on the sensor capture settings. For example, if the IMU is sampling at 500Hz for each of the 9 signal channels, and a 2 second window is applied to create specific movement segments, each of the segmented windows would evaluate to a 500 x 9 x 2 = 9000-dimensional vector. Consequently, several pre-processing stages are required to create a suitable data form for input into the classifier algorithm (Figo et al., 2010). Pre-processing of IMU signal data for HAR and sport-specific movement recognition can involve methods such as calibration, normalization, down-sampling, data synchronisation, and filtering (Camomilla et al., 2018; Figo et al., 2010).

The data segmentation or activity detection stage aims to identify those segments of the pre-processed data that have the specific features of the activities (Bulling et al., 2014). Dividing the sensor data into smaller segments can be categorised as either activity-defined windows, event-defined windows or sliding windows (Banos et al., 2014). Activity and event-defined windows are related to splitting data based on either changes in activity or detection of specific events, respectively (Banos et al., 2014). Sliding windows are fixed lengths with no data gaps between and may overlap in some situations (Banos et al., 2014). Analysis on window sizes in general HAR have provided guidelines for how to implement windowing segmentation (Banos et al., 2014; Niazi et al., 2017). In sport-specific movement recognition for example, sliding windows have been used in tennis (Conaire et al., 2010), golf (Jensen et al., 2015), swimming (Jensen et al., 2016), and weightlifting (Adelsberger & Tröster, 2013). In the detection of skateboard tricks, firstly a sliding window of 1 s with a 0.5 s overlap was used followed by calculating the energy of the windows (Groh et al., 2015).

Energy-based data segmentation is set based on the different energy levels of the sensor signals created as the activities are performed at different intensities (Bulling et al., 2014).

Another important consideration in data acquisition and processing is the sampling rate. The minimum sampling rate theorem is called the Nyquist sampling frequency (f_n) (Jerri, 1977). It states that the frequency of the signal acquired for processing must be sampled at more than twice of the highest frequency detected within that signal (Derrick, 2004). The sampling rate can change for several reasons depending on the sport activities and also sensor power saving or operating system requirements (Bulling et al., 2014). Sampling rates are usually much lower in general human activity recognition (Janidarmian et al., 2017; Niazi et al., 2017). The theory and literature on analytics using machine learning for movement recognition, specifically focusing on sport activities will be detailed in the subsequent sections.

2.4 Machine learning for automated movement recognition analysis

2.4.1. Machine learning theory overview

Machine learning derives from statistical learning theory (Hastie et al., 2009) in which algorithms learn from data input to automated model building and perform tasks without being explicitly programmed. Learning algorithms are essentially a sequence of instructions that model a function to produce the required output from data input. The algorithm goal is to output a response function $h\sigma(\bar{x})$ that will predict a ground truth variable y from an input vector of variables $\bar{x}$ (Bzdok et al., 2017). Machine learning is data driven and forms around the design of a model to test the hypothesis function (Raschka, 2018b), generate important captured data, and extract features from the original data input. The aim overall is to develop a machine learning model that generalises best to new data (Raschka, 2018b). The model designs relate to the processing methods to generate and learn from the input data in order to improve the task and generalise well to new data inputs in the future (Deisenroth et al., 2018). Model learning looks to optimise the parameters assigned in order to discover the patterns within a dataset to make output predictions from the input data (Deisenroth et al., 2018). In general, the adaption of model learning algorithm parameters to conform to new unseen data based on the data available is known as training a system (Deisenroth et al., 2018).

The distinction between statistical modelling and machine learning has been discussed in the literature (Boulesteix & Schmid, 2014; Breiman, 2001b; Bzdok et al., 2018). No evidence was found within clinical research prediction modelling that machine learning models are routinely better in performance compared to a logistic regression algorithm (Christodoulou et al., 2019). Machine learning models will learn from data in an automated way by applying heuristics and numerical optimisation to extract data patterns, to be then characterised by the model's performance (Bzdok et al., 2017). Whereas statistical-based models derive from theory and require the user's knowledge of the system to predict values, infer relationships and determine how significant these relationships are between variables (Bzdok et al., 2018). The use of either approach depends largely on the purpose and research question at hand. Also, there are several trade-offs between each approach, for example, the level of flexibility in algorithms, interpretability of the results, and degrees of freedom (Boulesteix & Schmid, 2014; Christodoulou et al., 2019). The concept of "no-free lunch theorem" (Wolpert, 2012) explains how no single supervised learning prediction model will perform optimally across all problems and datasets (Wolpert & Macready, 1997); and in relation to search and optimisation algorithms (Wolpert & Macready, 1997).

Machine learning has three main approach types being supervised, unsupervised, and reinforcement learning (Géron, 2019; Mohammed et al., 2016). Supervised learning means each data instance has a known label assigned to it (Géron, 2019; Kotsiantis et al., 2007). Specifically, it infers a function from the labelled training data in which a training example consists of an input vector X and an output vector y label, the explanation of the X input (Hastie et al., 2009). Supervised learning is either a classification or regression problem (Géron, 2019). Classification will predict a target class by mapping onto a categorical variable which can be either a binary, multi-class, multi-labelled or a hierarchical outcome (Kotsiantis et al., 2007; Sokolova & Lapalme, 2009). Classification algorithms include support vector machine (SVM), random forests (RF), and neural networks (NN). Regression problems predict on a continuous scale, for example house prices, and algorithms include logistic regression and polynomial regression (Géron, 2019). Unsupervised learning problems are data driven as they involve a data input that is unlabelled, and therefore the aim is to find hidden structures and relationships within the data (Géron, 2019). The training data is not structured like supervised learning and can contain noise or unknown data. Unsupervised algorithms include k-means clustering, hidden Markov models, and gaussian mixture model. A learning problem does fall between these two approaches, semi-supervised learning (Chapelle et al., 2003), where a subset of the training data is labelled which with clustering can be used to label the clusters whilst the majority unlabelled data defines the cluster boundaries. Lastly, reinforcement learning (Sutton, 1992) reacts to the training environment and will train on data continually through a trial and error process. Machine learning models can also be defined as either parametric or non-parametric (Breiman, 2001b; Géron, 2019). Generally, parametric models will have a finite number of parameters or in statistical theory there is prior knowledge of the data distribution. Non-parametric models have a potential for infinite parameters which can expand model complexity in relation to the amount of training data; or a statistical approach sees no prior knowledge assumed on the distribution of the data.

2.4.2 Model training and algorithms

When training the learning algorithm, total input data is generally split into training and testing datasets. The training dataset is used to teach the algorithm and build the model. Though, it is recommended to split the training dataset into train and validation sets in which multiple models with various hyperparameters are trained on the train set and evaluated on the validation set to provide an unbiased evaluation of model fit on the training dataset; especially with smaller data sets (Lever et al., 2016; Raschka, 2018b). Briefly, hyperparameters refer to a parameter of the learning algorithm that controls details of the learning algorithm itself (Raschka, 2018b). They must be set prior to training and remain constant during training, for example the learning rate in a NN or the number of leaves and depth of a RF tree. More data should be allocated to the train set, and a strategy to avoid withholding to much data in the validation set is cross-validation. For example a *k*-fold or leave-one-out cross-validation method which are resourceful when using smaller datasets (Raschka, 2018b). Once fully trained, the best performing model is then implemented on the unseen test dataset to provide the final generalisation error evaluation. Common total data split ratios are 70% train set, 15% validation set and 15% test set, or 80% training set and 20% testing set. The complexity of the model as a function of the chosen type, input and parameter numbers will affect the bias and variance of a model which may lead to one that either overfit or underfits (Géron, 2019; Lever et al., 2016). The bias is the difference between the expected estimator value and the true value (Raschka, 2018b). The variance in a model is the variability in the predictions from a given expected value; if the model learning is highly sensitive to the small fluctuations in the training set then it has high variance (Raschka, 2018b). Models that are too basic will underfit and not learn the full characteristics of the data; this is usually characterised by not performing well on both the training and test sets (Raschka, 2018a). Whereas, overfitting indicates that the model is too complex for the data characteristics learning outside the real data, i.e. random noise, and will perform highly during training but cannot adapt as well to new a data instance therefore have low test set accuracy (Domingos, 2012; Hawkins, 2004; Lever et al., 2016). The classification algorithms used within this thesis, Chapter Eight, will be briefly explained below including the theory and relevant literature examples.

SVM algorithms are also called kernel-based methods as they use the kernel trick technique to transform or map linearly a set of low dimensional non-linear data to higher dimensional linearly separable observations (Abe, 2005). The SVM works by detecting the best separating vector in features to differentiate between them considering the margin hyperplanes separating the classes (Kotsiantis et al., 2007; Wundersitz, Josman, et al., 2015). The aim is to find the hyperplane in a defined N-dimensional feature space that maximises the margin; the distance between data points of both classes so as to clearly classify the data classes (Mohammed et al., 2016). As SVM are robust algorithms, they model and adapt well to instances of large feature numbers relative to the training data instances. Although they can have poor interpretability in comparison to other models such as decision trees (DT), and can be prone to overfitting due to high variance within the algorithm (Kotsiantis et al., 2007).

A RF classifier is an ensemble model of several DTs built until their peak with each tree dependent upon independently sampled vector values (Breiman, 2001a; Wundersitz, Josman, et al., 2015). RF will differ from a DT which divides each node by “best spilt” across all variables, as a RF splits nodes based on the “best among a subset of predictors randomly chosen at the node” (Liaw & Wiener 2002, p. 18). RF has several advantages in that it will operate efficiently on larger datasets, is more robust to missing data, and overfitting is lower and preventable by taking the mean classification performance variable of each tree (Breiman, 2001a).

A k-Nearest Neighbour (KNN) is considered an instance-based learning algorithm as the function is only approximated locally and the computation delayed until classification. It is a non-parametric method that considers occurrences in a dataset to be close with others of similar properties (Kotsiantis et al., 2007). KNN classification will output class membership, where an occurrence is classed by the majority vote located amongst the k nearest occurrences (Kotsiantis et al., 2007). A limitation of KNN models is that a larger storage space and computational allowance are required (Salman et al., 2017). Naïve Bayesian (NB) classifiers are based on the Bayes’ theorem (Mohammed et al., 2016) and simple forms of Bayesian networks which are graphical representations of probability in a set of variables (Mohammed et al., 2016). Essentially, the classifier will independently consider features contributing to the probability of the classification of instances (Yang & Hsu, 2010). A benefit of NB is the reduced computational time, although the classifier in general is considered less accurate due to the assumptions of independence (Kotsiantis et al., 2007). Finally, an adaptive boosting (AB) classifier is an ensemble-based learning algorithm which combines the output from several lower-level machine learners through a majority vote to establish an output in a stepped process training at each stage (Mansbridge et al., 2018; Zaki & Meira Jr, 2014). A multilayer perceptron (MLP) has more than one linear layer of neurons that interact through weighted connections (Pal & Mitra, 1992).

In the sport-specific movement recognition literature, SVM and RF algorithms have been most commonly implemented and achieved high results in comparison to other machine learning algorithms. For example, in classifying between expert and beginner weight lifting athletes, a SVM model achieved 94% classification accuracy (Adelsberger & Tröster, 2013), and 94% during volleyball skill assessment classification (Wang et al., 2018). Recognition of basketball activities using a RF model achieved a classification accuracy of 87.5% in comparison to using a KNN which achieved 83.6% (Holzemann & Van Laerhoven, 2018).

Selecting the right machine learning algorithm hyperparameters is important for the efficiency in the learning process and ultimately the success of the model. A model parameter is a property of the training data that is learnt during the training process, for example in Natural Language Processing the sentence lengths or word frequencies. A hyperparameter is set prior to the start of the training process representing higher level controlling properties of the actual model such as the capacity to learn and its complexity of implementation. Optimisation of model hyperparameters is a procedure that benefits from automation strategies (Snoek et al., 2012). The problem of identifying good hyperparameters for the learning model is called hyperparameter optimisation, three of the more common approaches are outlined subsequently.

Grid search operates by running every possible combination of hyperparameters specified in the learning model, in which all hyperparameter combinations are tested through a series of cross-validation passes. The number of combinations will grow exponentially with the number of hyperparameters that can leave grid search strategies to suffer from the curse of dimensionality (Bergstra & Bengio, 2012). Assessing all potential value combinations means grid search methods can become very computationally heavy, especially as the dimension space grows. Another limitation of grid search is that it tends to fail to provide appropriate coverage to the dimensions that are important to search and continues to allocate extra trials to searching dimensions that have less weight importance to the learning model

(Bergstra & Bengio, 2012). Grid search strategies for classifier optimisation have been implemented in the relevant sport-specific movement recognition literature, for example in tennis stroke type classification (Conaire et al., 2010) and skateboard trick type classification (Groh et al., 2015).

A random search strategy for hyperparameter optimisation has shown to be more efficient than grid search in use for practical reasons related to the statistical independence of every trial (Bergstra & Bengio, 2012). Also, it operates better in high-dimensional spaces and requires less computational time (Bergstra & Bengio, 2012). The approach works by selecting random combinations of values from a set-up grid of hyperparameter values. The number of search iterations performed is based on the time and resources allocated. A random search is simpler to implement and will produce a near-optimal set of parameters faster than a grid search approach.

Bayesian optimisation for hyperparameter searching is considered a highly efficient approach and powerful strategy for finding the extrema of objective functions that may otherwise be expensive to assess with other approaches (Brochu et al., 2010). A Bayesian approach has shown to outperform other state-of-art global optimisation algorithms (Snoek et al., 2012). In comparison, grid or random search approaches are uninformed strategies as they search the total space of hyperparameter combinations without considering its past results. This approach is time and computer processing consuming when in larger parameter spaces. Bayesian optimisation derives its name and functions from the Bayes' theorem (Bayes, 1991; Bernardo & Smith, 2009). Simply stated, Bayesian approaches seek to describe the probability of an event based on prior knowledge of the conditions that might be related to the event. More specifically, the Bayes' Theorem equation is stated in Brochu et al. (2010). The Bayesian optimiser is categorised in the class of sequential model-based optimisation algorithms as it uses previous observations of the loss function to determine the next more optimal sample. The optimiser works by assuming the unknown function was sampled from a Gaussian process and maintains a posterior distribution for this function from running learning algorithm experiments as the different hyperparameters are observed (Snoek et al., 2012). Past evaluation results are kept track of which are used to form probabilistic model mapping of hyperparameters to a probability score on an objective function. The major advantage of this is that less iterations are required to find the optimal hyperparameter values as areas that won't contribute highly are disregarded.

When defining data terms generally, an attribute is the type of data, for example run distance, and an instance (referred to interchangeably as a feature or variable in the literature) is the added value of the attribute for example a run distance of 10 km (Géron, 2019). The property of features can be discrete, continuous or nominal in nature and classed in terms of their influence to the model learning task as either relevant, irrelevant or redundant (Ladha & Deepa, 2011). The features used as input to the learning model should be relevant to the domain specific task to increase model output accuracy (Ladha & Deepa, 2011). Yet defining and creating feature subsets increases the dimensionality of the data input into the learning algorithm. High dimensional data is problematic for machine learning algorithms as such data creates large memory usage and high computational costs for processing (Janecek et al., 2008; Khalid et al., 2014). Data dimensionality refers to the number of features or input variables within a dataset (Liu, Motoda, Setiono, & Zhao, 2010). A reduction in data dimensionality as a pre-processing method is done so to avoid the curse of dimensionality where the numbers of feature inputs are greater relative to the number of actual dataset observations (Khalid et al., 2014; Mannini & Sabatini, 2010). Proficiency of a machine learning model is heavily dependent on the design of data representations using features derived or transformed from the original input. Devising such data representations also allows for incorporation of domain knowledge into the feature sub-sets for a more specific capture of instances (Guyon & Elisseeff, 2003). Creating feature sub-sets can reduce the dimensionality of data, but the art lies in finding the best sub-sets of the original features with the least number of dimensions that will most contribute to the learning algorithms for highest possible accuracy (Khalid et al., 2014; Kohavi et al., 2011; Ladha & Deepa, 2011). Two main methods include feature selection and feature extraction can be used individually or in combination. The choice of these two methods depends on the dataset specifics and the type of domain application (Khalid et al., 2014). Both methods will be explained in detail in the following subsections.

2.4.3 Feature extraction

Machine learning algorithms are not able to process raw sensor data effectively, and data after pre-processing can still be abstract and high-dimensional (Kautz, 2017). Signal feature extraction and selection favour increased overall processing efficacy by reducing data to critical features that can discriminate the targeted activities (Bulling et al., 2014). Feature extraction involves the generation, identification or transformation of key features from the raw data that help maximise classifier success (Mannini & Sabatini, 2010). The features can be transformed into a more suitable format in subspaces with reduced dimensions for input into the learning model, and are aimed at preserving or improving on key discriminative abilities of the feature variables (Ghojogh et al., 2019; Mannini & Sabatini, 2010). Feature extraction is considered a problem-specific processing stage in which domain knowledge is expected in order to generate quality features for training the learning algorithms and achieve maximum performance on the required task (Guyon et al., 2008). Although, two main disadvantages of feature extraction techniques are that information about the level of meaning and contribution of the original data features may overall be minimal or absent, and the linear contributions of the original features may have poor interpretability (Janecek et al., 2008). Dimensionality reduction can be categorised as either supervised or unsupervised (Ghojogh et al., 2019). Unsupervised methods are used for data reconstruction purposes and work on the variation and patterns in the data. Supervised methods use the labels and defined classes of the data to aid in algorithm prediction task performance improvements. Methods can also be linear such as principal component analysis (PCA) and linear discriminant analysis (LDA) or non-linear for example, non-linear PCA implemented by nonlinear multilayer perceptron (MLP). For example, PCA is a common feature extraction linear transformation methods and works by generating a new set of uncorrelated variables called principal components in which the first few ordered variables have the strongest patterns from all the original variables (Jolliffe, 2002). The aim is to preserve important variation information from the original dataset by capturing the linear dependencies of the variables whilst reducing the feature space dimensionality of the many original interrelated variables (Janecek et al., 2008; Jolliffe, 2002). The reduction in the feature space is done by eliminating the last principal components that do not contribute significantly to the observed variability. Using the PCA methods for feature extraction helps to avoid over-fitting learning models (Khalid et al., 2014). Although there are several limitations of PCA including that assumptions are made that the relationships between variables are linear, it does not have a probabilistic model structure, and the interpretations of the method are only reasonable if all the variables are assumed to be scaled at the numeric level.

Features extracted from the input signal data can be categorised into several types and transformed into different domains of representation. Feature extraction method approaches include: data descriptive statistics, time domain features, frequency domain features, and time-frequency domain or wavelet analysis. Data descriptive statistics for IMU data include variance, kurtosis, correlation coefficient, and standard deviations, for example as applied in the classification of lunge biomechanics (Reilly et al., 2017a) and team sport movements (Wundersitz, Josman, et al., 2015) using inertial sensor data.

Time domain features are statistical measures generated from a segmented window of data (Preece, Goulermas, Kenney, Howard, et al., 2009). Examples include the calculated mean of windows in a data sample which can be used as both a feature in activity recognition but also applied to the raw data to remove noise and for smoothing the dataset (Figo et al., 2010). The variance metric represents the average of squared differences from the mean and informs about the variability in the dataset, stability of the signal, and probability distribution (Figo et al., 2010). The signal cross correlation measures the relationship between waveforms and can be used to find patterns in prolonged signal data (Figo et al., 2010). The signal magnitude area can be used to differentiate between static and dynamic activity states and is defined as the sum of area encompassed by the magnitude of each the three axis signals (Figo et al., 2010).

Frequency domain features are also derived from a segmented window but the data is first transformed into the frequency domain usually through a discrete Fourier transform (Preece, Goulermas, Kenney, Howard, et al., 2009) and used to capture the repetitive nature of sensor signal data (Figo et al., 2010). A Fourier transform expresses a signal in terms of its basic sinusoidal components being sines and cosines, and gives a quantitative measure of how fast the signal moves (Prandoni & Vetterli, 2005). Common Fourier transformation algorithms in human activity recognition are the fast Fourier transform (FFT) and discrete Fourier transform (DFT). Representation of the signal in the frequency domain allows for further important features of a time-series signal to be extracted such as the direct current component and can be more robust in discriminating movements based on different distinct signal frequency complexities (Mannini & Sabatini, 2010). Key frequency domain features for human activity recognition include the power spectral density which represents the frequency corresponding to the highest computed power spectrum density over when using a sliding window approach. The direct current signal component will provide the power spectral density at the frequency f = 0 Hz. The signal entropy allows for discrimination between movements that have the same power spectral density but different patterns of movement (Figo et al., 2010). Lastly, energy features can anlayse a movement's unique properties, which can be important when studying the energy expenditure of a subject for the targeted movement (Mannini & Sabatini, 2010). Evaluation of time-domain and frequency-domain features on data from two (walking, running) and three (walking, running, jumping) activity scenarios showed that frequency domain features were more robust in distinguishing activities in the three activity dataset whereas time-domain features performed better on the two activity dataset (Figo et al., 2010). It was expected that frequency domain metrics performed better on the higher dimensional data due to the computational efficiencies of the analysis (Figo et al., 2010). Although as a higher degree of processing for frequency domain features is required as they have longer window times and greater computational demands (Khan et al., 2013). This may be a disadvantage in real-time activity recognition, whereas time domain features are less computationally demanding during real-time processing but domain knowledge around the understandings of how each unique movement signal pattern are differentiated is required to increase application specificity (Khan et al., 2013; Preece, Goulermas, Kenney, Howard, et al., 2009).

The time-frequency domain through signal analysis involves wavelet analysis and can provide both time and frequency feature characteristics which can be decomposed into individual coefficients containing data on a specific frequency band (Preece, Goulermas, Kenney, Howard, et al., 2009). Wavelet transforms can be more efficiently computed over the entire spectrum, without requiring a dominant frequency band compared to a Fourier transform wavelet analysis (Bajric et al., 2016; Figo et al., 2010). Furthermore, can anlayse non-stationary movement signals where the frequency environment changes over time (Preece, Goulermas, Kenney, Howard, et al., 2009), and are more effective for transient features extraction (Bajric et al., 2016).

2.4.4 Feature selection

Feature selection, also referred to as subset selection, is a method of feature reduction used to select domain relevant and explanatory data features (Guyon et al., 2008). It is aimed at defining the smallest optimal feature subset according to a define criterion that is representative of the original features within the defined problem domain (Ladha & Deepa, 2011; Liu & Motoda, 2012). Relevant features are those that if removed from the subset will cause the performance measure of the remaining features to decrease (Liu & Motoda, 2012). Whereas, a redundant feature does not bring new insightful information about the target variable which would lead to measured model improvements. The advantages of using feature selection within data processing include:

- A reduction in dimensionality of the feature space which limits storage requirements therefore increasing model speeds,
- A reduced chance of overfitting the learning model which will enhance the generalisability of the model,
- Greater understandings of the data features and their relationships to the target variable and about the process that generated the data,
- A method to improve data visualisations,

- Improved efficiency and speed of the model training process,

- Improved learning accuracy due to domain specific input features,

(Beniwal & Arora, 2014; Guyon et al., 2008; Khalid et al., 2014; Ladha & Deepa, 2011; Liu et al., 2010).

Three key methods for devising subsets of the feature space are filter, wrapper and embedded methods (Ghojogh et al., 2019; Liu et al., 2010). Filter methods use statistical measures to anlayse and evaluate the general features of the data without the use of a learning algorithm as a pre-processing step (Liu et al., 2010). This means the filter method is independent of the later implemented machine learning algorithm and does not have any bias associated with a learning algorithm (Janecek et al., 2008; Liu et al., 2010). Filter methods are a one-shot process that output a ranked set of the most influential features in a subset for input to training the model learning algorithm. Because of the independence from a learning algorithm, it allows filter methods to have a simpler design structure making them more computationally efficient and easier to implement and understand (Guyon & Elisseeff, 2003; Liu et al., 2010). A limitation of filter methods are that most are univariate in operation, meaning that features are evaluated separately which causes feature dependencies to be ignored and a possible reduction in the final learning model performance (Beniwal & Arora, 2014).

In contrast, wrapper methods (Kohavi et al., 2011) use the predetermined learning algorithm intended for the later predictive task as a black box to evaluate feature subsets by their inclusion effect on model performance (Guyon & Elisseeff, 2003). Wrappers are iterative methods as at each iteration through the search, subsets of features are generated and evaluated using the learning algorithm to determine which will be carried over to the next iteration. To implement a wrapper method, the search strategy, model type, machine learning algorithm, and accuracy evaluation measure need to be specified. Wrapper feature selection is often viewed as a search problem as it searches through the space of the feature subset to measure the performance of the learning algorithm on the features which can be added or removed based on criterion from the feature subset (Janecek et al., 2008; Liu & Motoda, 2012). Several search directions and strategies can be implemented each with varying speeds, computational costs, and performance accuracy outputs depending on the current dataset domain and size of the search space (Liu & Motoda, 2012). Search directions include, sequential forward selection which starts with an empty feature set and sequentially adds features one at a time based on a criterion (Kohavi et al., 2011; Liu & Motoda, 2012). Sequential backwards elimination begins with a full feature set and removes the least important features at that search space one at a time based on a criterion (Kohavi et al., 2011; Liu & Motoda, 2012). Random generation starts at a random direction and adds or deletes features also in an un-systematic way (Liu & Motoda, 2012). Lastly, bi-directional generation conducts two searches concurrently which will each start at different directions. The searches will stop when either both searches meet in the middle of the search space or when one search detects the best feature subset before reaching the mid-point (Liu & Motoda, 2012). Search strategies (Liu & Motoda, 2012) include, exhaustive or complete searches where all possible subsets will be explored to find the optimal set. Heuristics searches use heuristics to collate the subsets; this is faster and less computationally expensive than an exhaustive search but the optimal subset is not guaranteed to be reached. A nondeterministic search will look for the next subset at random and the search doesn't need to be completed for the best subset to be found, but it is unclear when in the processing it will be determined.

By using the pre-determined machine learning algorithm as a black box, the wrapper feature selection process is considered part of model training, therefore separate validation and testing datasets are required to evaluate the final model error. The final error will also reflect the performance of the wrapper feature selection method. The value of the chosen feature subset is directly measured by the performance of the machine learning algorithm applied in the wrapper feature selection process (Beniwal & Arora, 2014). Performance measures across both wrapper and filter methods can be evaluated in terms of accuracy, consistency, information gain, distance or dependence (Liu & Motoda, 2012). Cross-validation methods for evaluation of wrapper generated feature subsets are commonly chosen (Liu et al., 2010).

Finally, embedded feature selection methods incorporate the selection process built into the learning model fitting and training process itself. It is considered a hybrid approach that includes the advantages of both filter and wrapper methods (Guyon & Elisseeff, 2003; Khalid et al., 2014). The resulting output will either be a selected feature subset or measured weights representing the features' relevance in relation to the target variable (Liu et al., 2010). Embedded feature selection methods do not need to split the training data into training and validation sets thereby making the training process more efficient by not retraining a classifier for each feature subset to reach the results quicker than a wrapper method (Guyon & Elisseeff, 2003). Although filter methods are computationally faster and more efficient due to their simple structure and independence from a learning algorithm; wrapper and embedded methods will generally perform to a higher degree as the process is optimised for the classifier taking into account the dependencies between the features (Guyon & Elisseeff, 2003; Janecek et al., 2008). Wrapper methods can be computationally expensive when faced with a large feature space as each feature subset must be evaluated with the training classifier (Khalid et al., 2014). Yet search strategies can be devised to reduce the computational demands for greater model efficiency (Guyon & Elisseeff, 2003).

2.4.5 Model and algorithm performance evaluation

To determine the algorithm for the problem task and therefore the model for the algorithm's hypothesis space, several comparison evaluation methods are used depending on the characteristics of the data and problem to solve (Raschka, 2018b; Sokolova & Lapalme, 2009). For example, multiple independent training and test sets would be more appropriate for large datasets when comparing algorithms, whereas nested cross-validation would suit small datasets (Raschka, 2018b). There are several measures for classification tasks including the accuracy, precision, recall, and F1-score (or F-measure, or F-score) (Sokolova & Lapalme, 2009). Clinical prediction models most commonly reported Area-Under-Curve (AUC) then sensitivity and specificity as performance measures (Christodoulou et al., 2019). As a single number evaluation of classifiers the AUC or F1-score tend to be better alternatives over classification accuracy which underestimates the ability of the classifier on the smaller classes under class imbalance (Bulling et al., 2014; Forman & Scholz, 2010; Ling et al., 2003). The F1-score provides the harmonic mean, this gives increased weight to lower values, but to achieve a good result the recall and precision must both be high (Géron, 2019). A precision-recall trade-off exists that means increasing either precision or recall will cause a decrease in the other measure (Géron, 2019). Although, accuracy is generally inherently unbiased being "expressed in terms of a binomial distribution"; and an F1-score can be open to estimate biases (Forman & Scholz, 2010). Evaluations can be by the weighted F1-score as a single evaluation metric for the balance between precision and recall on each class (Forman & Scholz, 2010). The weighted average calculates the metrics of each class and finds the average weight by the support factor to account for any class imbalance in comparison to a macro average or treating all classes equally with a micro average. The AUC score is better for comparing the overall performances of multiple classifiers (Ling et al., 2003), and tells how well the model is capable of distinguishing between classes (Hanley

& McNeil, 1982). An AUC score of 0.5 shows model predictions are almost random, and an AUC of 1.0 represents a perfect classifier.

2.4.6 Deep learning

Deep learning is a division of machine learning based on the underlying architectural concepts of neural networks in that they are similar in mathematical properties but deep learning algorithms are expressed through a vast network of layers (Bengio, 2013; Lecun et al., 2015). The formed deeper hierarchical models of multiple hidden layers are based on representative learning, and several processing and abstraction layers (Bux et al., 2017; Yang, Nguyen, San, Li, & Krishnaswamy, 2015). Progressing through the layers sees linear and non-linear transforms applied to the output from the previous layer, therefore model learning is undertaken at each layer (Chen & Xue, 2015; Sugomori, Kaluža, Soares, & Souza, 2017; Yang et al., 2015). Direct raw data input can be handled with features automatically extracted and transformed to be represented in a hierarchy from low-to-high level feature representations of the data (Nweke et al., 2018). This feature extraction process is advantageous in comparison to the handcrafted "shallow" features of conventional machine leaning systems (Ravi et al., 2017), which present challenges as detailed in Nweke et al. (2018). For example, problems in capturing the detailed spatial and temporal variations in activities. Feature extraction through deep learning means the features become task-dependent making them more robust to overfitting (Alsheikh et al., 2016). Comparisons of feature learning methods for activity recognition using IMUs showed the differences and benefits of varying deep learning architectures for obtaining the characteristic features of both short- and long-term movement data time dependencies (Li, Shirahama, Nisar, Köping, & Grzegorzek, 2018). Also, deep learning avoids several pre-processing stages required in machine learning model setup, reducing the overall computational times and designer domain specific knowledge required. In training the deep neural networks, a key algorithm is backpropagation which is a computationally efficient and fast process that uses gradient descent and allows the neural network training to be tractable (Sze et al., 2017). Deep learning models have shown improved results in computing and model performances for human activity recognition tasks (Ignatov, 2018; Zebin et al., 2016; Zimmermann et al., 2018). Also, novel methods for IMU data feature representation extraction (Eyobu & Han, 2018; Li et al., 2018). For sport-specific recognition movement tasks, studies that have implemented deep learning algorithms are detailed in Chapter Three. Since the manuscript publication of Chapter Three, further research using deep learning methods have been published in the topic area which are summarised in Chapter Four.

2.4.7 Computer vision in sport movement recognition

Computer vision in sport has been widely adopted for a range of applications in training, coaching, in-competition referee systems, and commercial broadcast of competitions (Thomas et al., 2017). Extracting relevant information as a semi-autonomous process from video data can provide an efficient and important resource for coaches and athletes to review performances. Action recognition in sport presents challenges including larger variations in pose, viewpoints, and distances of the actions from the camera (Barris & Button, 2008). Also, particularly in team sports, occlusions from numerous athletes in the frame and the dynamic, unpredictable nature of many sports in general (Thomas et al., 2017). Action recognition in sport often requires both problems of object tracking and action detection then classification. Specifically, temporally cropping the action from a continuous video and then tracking the athlete in order to classify the action required (Nibali et al., 2017). Further details of the use of computer vision in sport for action recognition are in Chapter Three, as computer vision is not a core research component for this thesis.

Chapter Three: Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance

---

Manuscript copied and adapted from the original published article: **Cust, E. E.**, Sweeting, A. J., Ball, K., & Robertson, S. (2019). Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance review of model development and performance. Journal of Sports Sciences, 37(5), 568–600. https://doi.org/10.1080/02640414.2018.1521769 (Appendix B).

3.1 Abstract

Objective assessment of an athlete's performance is of importance in elite sports to facilitate detailed analysis. The implementation of automated detection and recognition of sport-specific movements overcomes the limitations associated with manual performance analysis methods. The object of this study was to systematically review the literature on machine and deep learning for sport-specific movement recognition using Inertial measurement unit (IMU) and, or computer vision data inputs. A search of multiple databases was undertaken. Included studies must have investigated a sport-specific movement and analysed via machine or deep learning methods for model development. A total of 52 studies met the inclusion and exclusion criteria. Data pre-processing, processing, model development and evaluation methods varied across the studies. Model development for movement recognition were predominantly undertaken using supervised classification approaches. A kernel form of the Support Vector Machine algorithm was used in 53% of IMU and 50% of vision-based studies. Twelve studies used a deep learning method as a form of Convolutional Neural

Network algorithm and one study also adopted a Long Short-Term Memory architecture in their model. The adaptation of experimental set-up, data pre-processing, and model development methods are best considered in relation to the characteristics of the targeted sports movement(s).

3.2 Introduction

Performance analysis in sport science has experienced considerable recent changes, due largely to access to improved technology and increased applications from computer science. Manual notational analysis or coding in sports, even when performed by trained analysts, has limitations. Such methods are typically time intensive, subjective in nature, and prone to human error and bias. Automating sport movement recognition and its application towards coding has the potential to enhance both the efficiency and accuracy of sport performance analysis. The potential automation of recognising human movements, commonly referred to as human activity recognition (HAR), can be achieved through machine or deep learning model approaches. Common data inputs are obtained from inertial measurement units (IMUs) or vision. Detection refers to the identification of a targeted instance, i.e., tennis strokes within a continuous data input signal (Bulling et al., 2014). Recognition or classification of movements involves further interpretations and labelled predictions of the identified instance (Bulling et al., 2014; Bux et al., 2017), i.e., differentiating tennis strokes as a forehand or backhand. In machine and deep learning, a model represents the statistical operations involved in the development of an automated prediction task (Lecun et al., 2015; Shalev-Shwartz & Ben-David, 2014).

Human activities detected by inertial sensing devices and computer vision are represented as wave signal features corresponding to specific actions, which can be logged and extracted.

Human movement activities are considered hierarchically structured and can be broken down to basic movements. Therefore, the context of signal use, intra-class variability, and inter-class similarity between activities require consideration during experimental set-up and model development. Wearable IMUs contain a combination of accelerometer, gyroscope, and magnetometer sensors measuring along one to three axes. These sensors quantify acceleration, angular velocity, and the direction and orientation of travel respectively (Gastin et al., 2014). These sensors can capture repeated movement patterns during sport training and competitions (Camomilla et al., 2018; Chambers et al., 2015; Wagner, 2018). Advantages include being wireless, lightweight and self-contained in operation. IMUs have been utilised in quantifying physical output and tackling impacts in Australian Rules football (Gastin et al., 2013, 2014) and rugby (Gabbett et al., 2011, 2012; Howe et al., 2017; Hulin et al., 2017). Other applications include swimming analysis (Mooney et al., 2015), golf swing kinematics (Lai et al., 2011), over-ground running speeds (Wixted et al., 2010), full motions in alpine skiing (Yu et al., 2016); and the detection and evaluation of cricket bowling (McNamara et al., 2015, 2018; Wixted et al., 2011).

Computer vision has applications for performance analysis including player tracking, semantic analysis, and movement analysis (Stein et al., 2018; Thomas et al., 2017). Automated movement recognition approaches require several pre-processing steps including athlete detection and tracking, temporal cropping and targeted action recognition, which are dependent upon the sport and footage type (Barris & Button, 2008; Saba & Altameem, 2013; Thomas et al., 2017). Several challenges including occlusion, viewpoint variations, and environmental conditions may impact results, depending on the camera set-up (Poppe, 2010; S. Zhang et al., 2017). Developing models to automate sports-vision coding may improve resource efficiency and reduce feedback times. For example, coaches and athletes involved in time-intensive notational tasks, including post-swim race analysis, may benefit from rapid objective feedback before the next race in the event program (Liao et al., 2003; Victor et al., 2017). For detecting and recognising movements, body worn sensor signals do not suffer from the same environmental constraints and stationary set-up of video cameras. Furthermore, multiple sensors located on different body segments have been argued to provide more specific signal representations of targeted movements (Yang et al., 2015). But it is not clear if this is solely conclusive, and the use of body worn sensors in some sport competitions may be impractical or not possible.

Machine learning algorithms learn from data input for automated model building and perform tasks without being explicitly programmed. The algorithm goal is to output a response function $h\sigma(\bar{x})$ that will predict a ground truth variable y from an input vector of variables $\bar{x}$. Models are run for classification techniques to predict a target class (Kotsiantis et al., 2007), or regression to predict discrete or continuous values. Models are aimed at finding an optimal set of parameters σ to describe the response function, and then make predictions on unseen unlabelled data input. Within these, model training approaches can generally run as supervised learning, unsupervised learning or semi-supervised learning (Mohammed et al., 2016; Sze et al., 2017).

Processing raw data is limited for conventional machine learning algorithms, as they are unable to effectively be trained on abstract and high-dimensional data that is inconsistent, contains missing values or noisy artefacts (Bux et al., 2017; Kautz, 2017). Consequently, several pre-processing stages are required to create a suitable data form for input into the classifier algorithm (Figo et al., 2010). Filtering (Figo et al., 2010; Wundersitz, Gastin, Robertson, et al., 2015), window capture durations (Mitchell et al., 2013; Preece, Goulermas, Kenney, & Howard, 2009; Wundersitz, Josman, et al., 2015), and signal frequency cut-offs (Wundersitz, Gastin, Richter, et al., 2015; Wundersitz, Gastin, Robertson, et al., 2015) are common techniques applied prior to data prior to dynamic human movement recognition. Well-established filters for processing motion signal data include the Kalman filter (Kautz, 2017; Titterton & Weston, 2009; D. Wagner et al., 2017) and a Fourier transform filter (Preece, Goulermas, Kenney, Howard, et al., 2009) such as a fast Fourier transform (Kapela et al., 2014; Preece, Goulermas, Kenney, & Howard, 2009). Near real-time processing benefits from reducing memory requirements, computational demands, and essential bandwidth during whole model implementation. Signal feature extraction and selection favours faster processing by reducing the signals to the critical features that can discriminate the targeted activities (Bulling et al., 2014). Feature extraction involves identifying the key features that help maximise classifier success, and removing features that have minimal impact in the model (Mannini & Sabatini, 2010). Thus, feature selection involves constructing data representations in subspaces with reduced dimensions. These identified variables are represented in a compact feature variable (Mannini & Sabatini, 2010). Common methods include principal component analysis (PCA) (Gløersen, Myklebust, Hallén, & Federolf, 2018; Young & Reinkensmeyer, 2014), vector coding techniques (Hafer & Boyer, 2017) and empirical cumulative distribution functions (ECDF) (Plötz et al., 2011). An ECDF approach has been shown to be advantageous over PCA as it derives representations of raw input independent of the absolute data ranges, whereas PCA is known to have reduced performance when the input data is not properly normalised (Plötz et al., 2011). For further detailed information on the acquisition, filtering and analysis of IMU data for sports application and vision-based human activity recognition, see (Kautz, 2017) and (Bux et al., 2017), respectively.

Deep learning is a division of machine learning, characterised by deeper neural network model architectures and are inspired by the biological neural networks of the human brain (Bengio, 2013; Lecun et al., 2015; Sze et al., 2017). The deeper hierarchical models create a profound architecture of multiple hidden layers based on representative learning with several processing and abstraction layers (Bux et al., 2017; Yang et al., 2015). These computational models allow data input features to be automatically extracted from raw data and transformed to handle unstructured data, including vision (Lecun et al., 2015; Ravi et al., 2017). This direct input avoids several processing steps required in machine learning during training and testing, therefore reducing overall computational times. A current key element within deep learning is backpropagation (Hecht-Nielsen, 1989; Yann LeCun et al., 1998). Backpropagation is a fast and computationally efficient algorithm, using gradient descent, that allows training deep neural networks to be tractable (Sze et al., 2017). Human activity recognition has mainly been performed using conventional machine learning classifiers. Recently, deep learning techniques have enhanced the bench mark and applications for IMUs (Kautz et al., 2017; Ravi et al., 2017; Ronao & Cho, 2016; J. B. Yang et al., 2015; Zebin et al., 2016; M. Zeng et al., 2014) and vision (Ji et al., 2012; Karpathy et al., 2014; Krizhevsky et al., 2012; Nibali et al., 2017) in human movement recognition producing more superior model performance accuracy.

The objective of this study was to systematically review the literature investigating sport-specific automated movement detection and recognition. The review focusses on the various technologies, analysis techniques and performance outcome measures utilised. There are several reviews within this field that are sensor-based including wearable IMUs for lower limb biomechanics and exercises (Fong & Chan, 2010; O'Reilly et al., 2018), swimming analysis (Magalhaes et al., 2015; Mooney et al., 2015), quantifying sporting movements (Chambers et al., 2015) and physical activity monitoring (Yang & Hsu, 2010). A recent systematic review has provided an evaluation on the in-field use of inertial-based sensors for various performance evaluation applications (Camomilla et al., 2018). Vision-based methods for human activity recognition (Aggarwal & Xia, 2014; Bux et al., 2017; Ke et al., 2013; Zhang et al., 2017), semantic human activity recognition (Ziaeefard & Bergevin, 2015) and motion analysis in sport (Barris & Button, 2008) have also been reviewed. However, to date, there is no systematic review across sport-specific movement detection and recognition via machine or deep learning. Specifically, incorporating IMUs and vision-based data input, focussing on in-field applications as opposed to laboratory-based protocols and detailing the analysis and machine learning methods used.

Considering the growth in research and potential field applications, such a review is required to understand the research area. This review aims to characterise the evolving techniques and inform researchers of possible improvements in sports analysis applications. Specifically: 1) What is the current scope for IMUs and computer vision in sport movement detection and recognition? 2) Which methodologies, inclusive of signal processing and model learning techniques, have been used to achieve sport movement recognition? 3) Which evaluation methods have been used in assessing the performance of these developed models?

3.3 Methods

3.3.1 Search strategy

The preferred PRISMA recommendations (Moher et al., 2009) for systematic reviews were used. A literature search was undertaken by the first author on the following databases; IEEE Xplore, PubMed, ScienceDirect, Scopus, Academic Search Premier, and Computer and Applied Science Complete. The searched terms were categorised in order to define the specific participants, methodology and evaluated outcome measure in-line with the review aims. Searches used a combination of key words with AND/OR phrases which are detailed in Table 3.1. Searches were filtered for studies from January 2000 to May 2018 as no relevant studies were identified prior to this. Further studies were manually identified from the bibliographies of database-searched studies identified from the abstract screen phase, known as snowballing. Table 3.2 provides the inclusion and exclusion criteria of this review.

Table 3.1. Key word search term strings per database.

Database key word searches
IEEE Xplore:
((((inertial sensor OR accelerometer OR gyroscope OR IMU OR microsensor)) AND (sport OR athlete* OR match OR game OR training)) AND (detection OR recognition OR classification)) AND (movement OR skill)
((((sport OR athlete* OR player*)) AND (video OR vision)) AND movement classification)
PubMed:
((((inertial sensor OR accelerometer OR gyroscope OR IMU OR microsensor)) AND (sport OR athlete* OR match OR game OR training)) AND (detection OR recognition OR classification)) AND (movement OR skill)
((((((((Vision OR video OR camera OR footage OR computer vision)) AND (sport OR athlete* OR match OR game OR training)) AND (detection OR recognition OR classification)) AND (movement OR skill))) AND human) NOT clinical)) NOT review
ScienceDirect:
((sport OR athlete* OR player*)) and ((inertial sensor OR accelerometer)
((sport OR athlete* OR player*)) and TITLE-ABSTR-KEY((vision OR video OR camera) AND (detection OR classification)).
Scopus:
((((inertial sensor OR accelerometer OR gyroscope OR IMU OR microsensor)) AND (sport OR athlete* OR match OR game OR training)) AND (detection OR recognition OR classification)) AND (movement OR skill)
((((sport OR athlete* OR player*)) AND (video OR vision)) AND movement classification)
Academic Search Premier:
((((inertial sensor OR accelerometer OR gyroscope OR IMU OR microsensor)) AND (sport OR athlete* OR match OR game OR training)) AND (detection OR recognition OR classification)) AND (movement OR skill)
((((sport OR athlete* OR player*)) AND (video OR vision)) AND movement classification)
Computer and Applied Science Complete:
((((inertial sensor OR accelerometer OR gyroscope OR IMU OR microsensor)) AND (sport OR athlete* OR match OR game OR training)) AND (detection OR recognition OR classification)) AND (movement OR skill)
((((Vision OR video OR camera OR footage OR computer vision)) AND (sport OR athlete* OR match OR game OR training)) AND (detection OR recognition OR classification)) AND (movement OR skill)

* Entails truncation, i.e., finding all terms that begin with the string of text written before it.

Table 3.2. Study inclusion and exclusion criteria.

Inclusion criteria	Exclusion criteria
• Original peer reviewed published manuscripts • Aimed at a sport-specific movement or skill, • Used IMUs and/or computer vision input datasets for model development • Investigated as an in-field application of the technology to the sporting movement • Defined clear data processing and model development methods inclusive of machine or deep learning algorithms for semi-automated or automated movement recognition • Published as full-length studies written in English	• Solely investigated gait analysis for clinical purposes • Solely investigated every day or non-sport-specific locomotion i.e., walking downstairs • Solely investigated player field positional tracking methods using data such as X, Y coordinates or displacement without any form of sport-specific skill detection and classification associated to it • Used ball trajectory and audio cue data as the major determinant for event detection • Data collection conducted within a laboratory setting under controlled protocol • Data processing pipelines or recognition model development methodology not clearly defined • Review studies

3.3.2 Data extraction

The first author extracted and collated the relevant information from the full manuscripts identified for final review. A total of 18 parameters were extracted from the 52 research studies, including the title, author, year of publication, sport, participant details, sport movement target(s), device specifications, device sample frequency, pre-processing methods, processing methods, feature selected, feature extraction, machine learning model used, model evaluation, model performance accuracy, validation method, samples collected, and computational information. A customised Microsoft Excel™ spreadsheet was developed to categorise the relevant extracted information from each study. Participant characteristics of number of participants, gender, and competition level, then if applicable a further descriptor specific to a sport, for example, 'medium-paced cricket bowler'. Athlete and participant experience level was categorised as written in the corresponding study to avoid misrepresentations. The age of participants was not considered an important characteristic required for model development. The individual ability in which the movement is performed accounts for the discriminative signal features associated with the movements. For the purposes of this review, a sport-specific movement was defined from a team or individual sport, and training activities associated with a particular sport. For example, weight-lifting as strength training, recognised under the Global Association of International Sports Federations. The targeted sports and specific movements were defined for either detection or recognition. Model development techniques used included pre-processing methods to transform data to a more suitable form for analysis, processing stages to segment data for identified target activities, feature extraction and selections techniques, and the learning algorithm(s). Model evaluation measures extracted were the model performance assessment techniques used, ground-truth validation comparison, number of data samples collected, and the model performance outcomes results reported. If studies ran multiple experiments using several algorithms, only the superior algorithm and relevant results were reported as the best method. This was done so in the interest of concise reporting to highlight favourable method approaches (Sprager & Juric, 2015). Any further relevant results or information identified from the studies was included as a special remark (Sprager & Juric, 2015). Hardware and specification information extracted included the IMU or video equipment used, number of units, attachment of sensors (IMUs), sample frequency, and sensor data types used in analysis (IMUs). Studies identified and full data extracted were reviewed by a second author.

3.4 Results

An outline of the search results and study exclusions has been provided in Figure 3.1. Of the initial database search which identified 4885 results, a final 52 studies met criteria for inclusion in this review. Of these, 29 used IMUs and 22 were vision-based. One study (Conaire et al., 2010) used both sensors and vision for model development separately then together via data fusion. Tables 3.3 to 3.8 provide a description of the characteristics of the reviewed studies, detailed in the following sections.

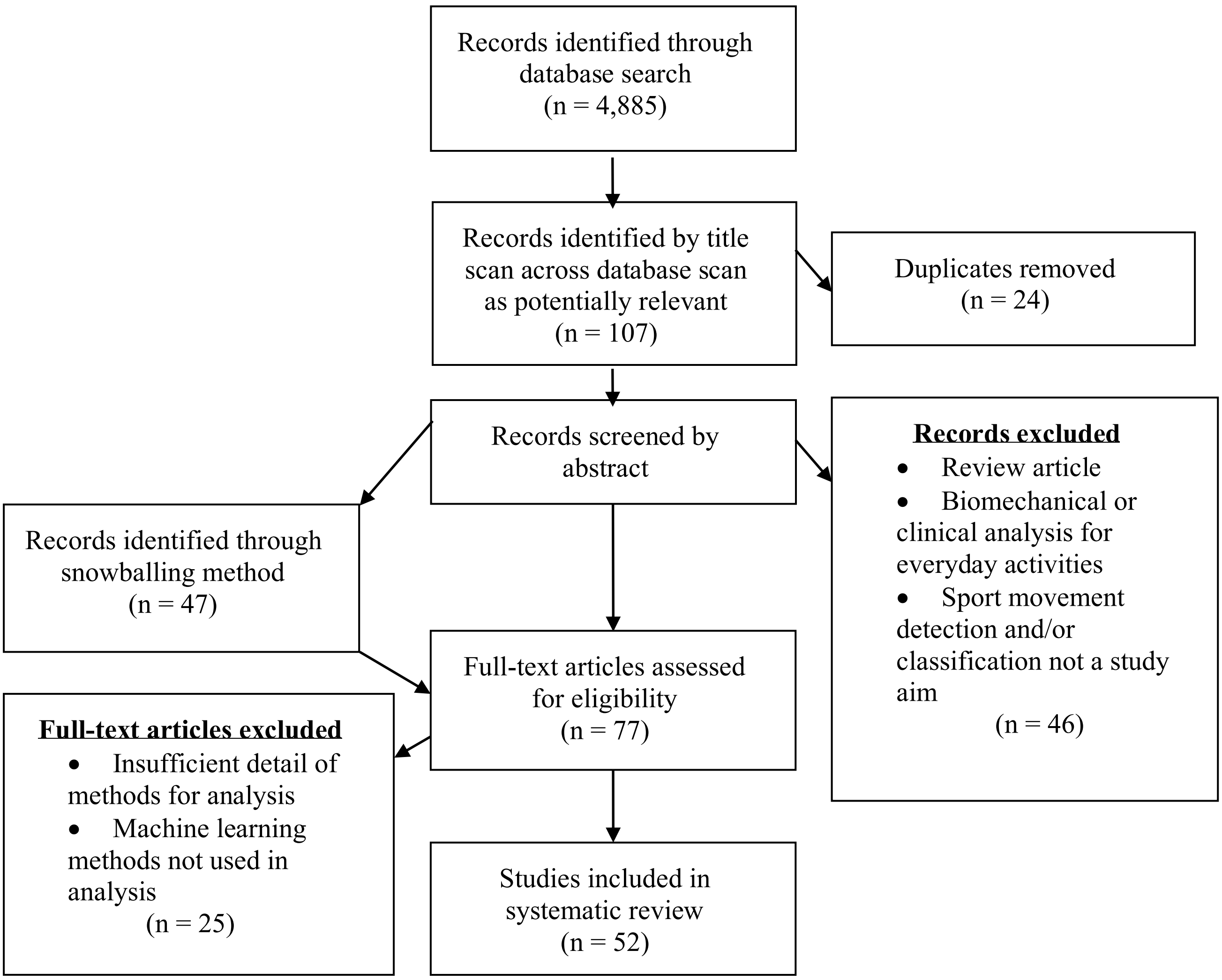

Figure 3.1. PRISMA flow diagram for study search, screen and selection process.

3.4.1 Experimental design

A variety of sports and their associated sport-specific movements were investigated, implementing various experimental designs as presented in Tables 3.5 and 3.7. Across the studies, sports reported were tennis (n = 10), cricket (n = 3), weightlifting or strength training (n = 6), swimming (n = 4), skateboarding (n = 2), ski jumping (n = 2), snowboarding (n = 1), golf (n = 4), volleyball (n = 2), rugby (n = 2), ice hockey (n = 2), gymnastics (n = 2), karate (n = 1), basketball (n = 3), Gaelic football (n = 1), hurling (n = 1), boxing (n = 2), running (n = 2), diving (n = 1), squash (n = 1), badminton (n = 1), cross-country skiing (n = 2) and soccer (n = 4). The Sports 1-M dataset (Karpathy et al., 2014) was also reported, which consists of 1,133,158 video URLs annotated automatically with 487 sport labels using the YouTube Topic API. A dominant approach was the classification of main characterising actions for each sport. For example, serve, forehand, backhand strokes in tennis (Conaire et al., 2010; Connaghan et al., 2011; Kos & Kramberger, 2017; Shah et al., 2007; Srivastava et al., 2015), and the four competition strokes in swimming (Jensen et al., 2013, 2016; Liao et al., 2003; Victor et al., 2017). Several studies further classified sub-categories of actions. For example, three further classes of the two main classified snowboarding trick types Grinds and Airs (Groh et al., 2016), and further classifying the main tennis stroke types as either flat, topspin or slice (Srivastava et al., 2015). Semantic descriptors were reported for classification models that predicted athlete training background, experience and fatigue level. These included running (Buckley et al., 2017; Kobsar et al., 2014), rating of gymnastic routines (Reily et al., 2017), soccer pass classification based on its quality (Horton et al., 2014), cricket bowling legality (Qaisar et al., 2013; Salman et al., 2017), ski jump error analysis (Brock et al., 2017; Brock & Ohgi, 2017) and strength training technique deviations (O’Reilly et al., 2015; Reilly et al., 2017b, 2017a). One approach (Yao & Fei-Fei, 2010), encoded the mutual context of human pose and sporting equipment using semantics, to facilitate the detection and classification of movements including a cricket bat and batsman coupled movements.

Total participant numbers for IMU-based studies ranged from one (Qaisar et al., 2013) to 30 (Kautz et al., 2017). Reported data individual instance sample sizes for sensor studies ranged from 150 (Salman et al., 2017) to 416, 737 (Rassem et al., 2017). Vision-based studies that explicitly reported total participant details ranged from five (Conaire et al., 2010) to 40 (Victor et al., 2017). Vision dataset sample sizes varied across studies, from 50 individual action clips (Liao et al., 2003) to 15, 000 (Victor et al., 2017). One study (Karpathy et al., 2014) used the publicly available Sports-1M, as previously described. Vision-based studies also reported datasets in total time, 10.3 hours (Bertasius et al., 2017), 3 hours (Montoliu et al., 2015), 1, 500 minutes (Shah et al., 2007), and 50 hours (Kapela et al., 2014), and by frame numbers, 6, 035 frames (Zhu, Xu, Huang, & Gao, 2006) and 10, 115 frames (Reily et al., 2017).

3.4.2 Inertial measurement unit specifications

A range of commercially available and custom-built IMUs were used in the IMU-based studies (n= 30), as presented in Table 3.3. Of these, 23% reported using a custom-built sensor. Of the IMU-based studies, the number of sensors mounted or attached to each participant or sporting equipment piece ranged from one to nine. The majority of studies (n= 22) provided adequate details of sensor specifications including sensor type, axes, measurement range, and sample rate used. At least one characteristic of sensor measurement range or sample rate used in data collection was missing from eight studies. All studies used triaxial sensors and collected accelerometer data. For analysis and model development, individual sensor data consisted of only accelerometer data (n = 8), both accelerometer and gyroscope data (n = 15), and accelerometer, gyroscope and magnetometer data (n = 7). The individual sensor measurement ranges reported for accelerometer were ± 1.5 *g* to ± 16 *g*, gyroscope ± 500 ∘/s to ± 2000 ∘/s, magnetometer ± 1200 µT or 1.2 to 4 Ga. Individual sensor sample rates ranged from 10 Hz to 1000 Hz for accelerometers, 10 Hz to 500 Hz for gyroscopes and 50 Hz to 500 Hz for magnetometers.

Table 3.3. Inertial measurement unit specifications.

Reference	Sensor model	Sensor No.	Sensor placement	Accelerometer			Gyroscope			Magnetometer		
				Axes	Range	Sample rate	Axes	Range	Sample rate	Axes	Range (1 Ga = 100 µT)	Sample rate
(Adelsberger & Tröster, 2013)	Ethos	3	Left ankle, wrist, lower back	3	± 6 *g*	NR	3	± 2000 ◦/s	NR	3	4 Ga	NR
(Anand et al., 2017)	Samsun Gear 2 smart watch	1	Wrist of hitting hand	3	± 8 *g*	100 Hz	3	± 2000 ◦/s	100 Hz			
(Brock & Ohgi, 2017)	Logical Product SS-WS1215/SS-WS1216, Fukuoka, Japan	9	Pelvis, right and left thighs, right and left shanks, right and left upper arms, ski blades	3	± 5 *g* (body) ± 16 *g* (ski)	500 Hz	3	± 1500 ◦/s	500 Hz	3	± 1.2 Gauss full-scale	500 Hz
(Brock et al., 2017)	Logical Product SS-WS1215/SS-WS1216, Fukuoka, Japan	9	Pelvis, right and left thighs, right and left shanks, right and left ski anterior to ski binding, right and left upper arm	3	± 5 *g* (body) ± 16 *g* (ski)	500 Hz	3	± 1500 ◦/s	500 Hz	3	± 1.2 Gauss full-scale	500 Hz
(Buckley et al., 2017)	Shimmer3 (Realtime Technologies Lab. Dublin, Ireland)	3	Right and left shanks 2cm above lateral malleolus, 5th lumbar spinous process	3	± 8 *g*	256 Hz	3	± 1000 ◦/s	256 Hz	3	± 4 Gauss full-scale	256 Hz
(Buthe et al., 2016)	EXLs33 IMU	3	Tennis racquet, on each shoe	3	± 16 *g*	200 Hz	3	± 500 ◦/s	200 Hz	3	NR	200 Hz
(Connaghan et al., 2011)	Custom Tyndall developed TennisSense WIMU system	1	Forearm of racquet arm	3	NR	NR	3	NR	NR	3	NR	NR

Table 3.3. continued.

Reference	Sensor model	Sensor No.	Sensor placement	Accelerometer			Gyroscope			Magnetometer		
				Axes	Range	Sample rate	Axes	Range	Sample rate	Axes	Range (1 Ga = 100 µT)	Sample rate
(Groh et al., 2015)	miPod sensor system	1	Underside of skateboard on the right side of front axis.	3	± 16*g*	200 Hz	3	± 2000 °/s	200 Hz	3	± 1200 µT	200 Hz
(Groh et al., 2016)	miPod sensor system	1	Top of snowboard behind the front binding	3	± 16 *g*	200 Hz	3	± 2000 °/s	200 Hz	3	± 1200 µT	200 Hz
(Groh et al., 2017)	miPod sensor system	1	Underside of skateboard on the right side of front axis.	3	± 16 *g*	200 Hz	3	± 2000 °/s	200 Hz	3	± 1200 µT	200 Hz
(Jiao, Wu, et al., 2018)	NR	2	Golf club (location not specified)	3	NR	NR	3	NR	NR			
(Jensen et al., 2015)	Shimmer™ 2R sensor nodes (Realtime	1	Golf club head	3	± 1.5 *g*	256 Hz	3	± 500 °/s	256 Hz	NR	NR	NR
(Jensen et al., 2016)	Shimmer™ 2R sensor nodes (Realtime Technologies Lab. Dublin, Ireland)	1	Back of head under a swim cap	3	± 1.5 *g*	10.24 Hz to 204.8 Hz	3	± 500 °/s	10.24 Hz to 204.8 Hz	NR	NR	NR
(Jensen et al., 2013)	Shimmer™ (Realtime Technologies Lab. Dublin, Ireland)	1	Back of head above swim cap	3	± 1.5 *g*	200 Hz	3	± 500 °/s	200 Hz	NR	NR	NR
(Kautz et al., 2017)	Bosch BMA280	1	Wrist of dominant hand	3	± 16 *g*	39 Hz	NR	NR	NR	NR	NR	NR
(Kelly et al., 2012)	SPI Pro	1	Between the shoulder blades	3	NR	39 Hz	NR	NR	NR	NR	NR	NR

Table 3.3. continued.

Reference	Sensor model	Sensor No.	Sensor placement	Accelerometer			Gyroscope			Magnetometer		
				Axes	Range	Sample rate	Axes	Range	Sample rate	Axes	Range (1 Ga = 100 μT)	Sample rate
(Kobsar et al., 2014)	G-Link wireless accelerometer node (Microstrain Inc., VT)	1	Lower back on the L3 vertebra region	3	± 10 *g*	617 Hz	NR	NR	NR	NR	NR	NR
(Kos & Kramberger, 2017)	Custom sensor	1	Wrist of racquet arm	3	± 16 *g*	NR	3	± 2000 ◦/s	NR	NR	NR	NR
(Conaire et al., 2010)	Custom sensor	6	Left and right wrists, left and right ankles, chest, lower back	3	± 12 *g*	120 Hz	NR	NR	NR	NR	NR	NR
(O'Reilly et al., 2015)	Shimmer™ sensor (Realtime Technologies Lab. Dublin, Ireland)	1	5th lumbar vertebra	3	± 16 *g*	51.2 Hz	3	± 500 ◦/s	51.2 Hz	3	± 1 Ga	51.2 Hz
(Reilly et al., 2017a)	Shimmer™ sensor (Realtime Technologies Lab. Dublin, Ireland)	5	5th lumbar vertebra, mid-point on right and left thighs, right and left shanks 2cm above lateral malleolus	3	± 2 g	51.2 Hz	3	± 500 ◦/s	51.2 Hz	3	± 1.9 Ga	51.2 Hz
(Reilly et al., 2017b)	Shimmer™ sensor (Realtime Technologies Lab. Dublin, Ireland)	5	Spinous process of the fifth lumbar vertebra, mid-point of both femurs, right and left shanks 2 cm above the lateral malleolus	3	± 2 *g*	51.2 Hz	3	± 500 ◦/s	51.2 Hz	3	± 1.9 Ga	51.2 Hz

Table 3.3. continued.

Reference	Sensor model	Sensor No.	Sensor placement	Accelerometer			Gyroscope			Magnetometer		
				Axes	Range	Sample rate	Axes	Range	Sample rate	Axes	Range (1 Ga = 100 µT)	Sample rate
(Pernek et al., 2015)	Custom sensor	5	Chest, left and right wrists, left and right upper arms	3	NR	30 Hz	NR	NR	NR	NR	NR	NR
(Qaisar et al., 2013)	Custom sensor	3	upper arm, elbow, wrist	3	NR	150 Hz	3	NR	150 Hz	NR	NR	NR
(Rassem et al., 2017)	NR	1	NR	3	NR	50 Hz						
(Rindal et al., 2018)	IsenseU Move+	2	Chest, Lower arm	3	NR	20 Hz	3	NR	20 Hz			
(Salman et al., 2017)	Custom sensor	3	Bowling arm: upper arm, forearm, wrist	3	NR	150 Hz	3	NR	150 Hz	NR	NR	NR
(Schuldhaus et al., 2015)	Custom sensor	2	Cavity of each shoe	3	± 16 *g*	1000 Hz	NR	NR	NR	NR	NR	NR
(Srivastava et al., 2015)	Samsung Gear S smart watch	1	Wrist of racquet arm	3	± 8 *g*	25 Hz	3	± 2000 ◦/s	25 Hz	NR	NR	NR
(Whiteside et al., 2017)	IMeasureU IMU (Auckland, New Zealand)	1	Wrist of racquet arm	3	± 16 *g*	500 Hz	3	± 2000 ◦/s	500 Hz	3	± 1200 µT	500 Hz

g G-forces, *Ga* gauss, *Hz* Hertz, *IMU* inertial measurement unit, *µT* micro Tesla

NR not reported: study either did not directly report the specification or the device did not include the sensor type

3.4.3 Vision capture specification

Several experimental set-ups and specifications were reported in the total 23 vision-based studies (Table 3.4). Modality was predominately red, green, blue (RGB) cameras. Depth cameras were utilised (Kasiri-Bidhendi et al., 2015; Kasiri et al., 2017; Reily et al., 2017), which add depth perception for 3-dimensional image mapping. Seven studies clearly reported the use of a single camera set-up (Couceiro et al., 2013; Díaz-Pereira et al., 2014; Hachaj et al., 2015; Kasiri-Bidhendi et al., 2015; Kasiri et al., 2017; Nibali et al., 2017; Reily et al., 2017). One study reported 16 stationary positioned cameras at a 'bird's eye view' (Montoliu et al., 2015), and Ó Conaire et al. (2010) reported the use of one overhead and eight stationary cameras around a tennis court baseline, although data from two cameras were only used in final analysis due to occlusion issues. Sample frequency and, or pixel resolution were reported in seven of the studies (Couceiro et al., 2013; Hachaj et al., 2015; Kasiri-Bidhendi et al., 2015; Kasiri et al., 2017; Montoliu et al., 2015; Victor et al., 2017; Zhu, Xu, Huang, & Gao, 2006), with sample frequencies ranging from 30 Hz to 210 Hz.

Table 3.4 Vision-based camera specifications.

Reference	Camera model	Modality	Camera No.	Data collection setting
(Bertasius et al., 2017)	GoPro Hero 3 Black Edition	RGB	1	100 fps 1280 x 960 pixels
(Couceiro et al., 2013)	Casio Exilim - High Speed EX-FH25. Focal length lens of 26 mm	RGB	1	Resolution 480 x 360 pixels 210 Hz
(Díaz-Pereira et al., 2014)	Sony Handycam DCR-SR78	RGB	1	
(Hachaj et al., 2015)	Kinetic 2 SDK system	3 Dimensional	1	30 Hz
(Horton et al., 2014)	NR	NR	NR	NR
(Ibrahim et al., 1971)	NR	NR	NR	NR
(Kapela et al., 2015)	NR	NR	NR	NR
(Karpathy et al., 2014)	NR	NR	NR	NR
(Kasiri-Bidhendi et al., 2015)	Swisse-range SR4000 time-of-flight (MESA Imaging AG, Switzerland)	Depth Camera at 5 m overhead height	1	25 fps 176 x 144 pixels
(Kasiri et al., 2017)	Swisse-range SR4000 time-of-flight (MESA Imaging AG, Switzerland)	Depth Camera at 5 m overhead height	1	25 fps 176 x 144 pixels
(Li et al., 2018)	iPhone5s, 6, 6plus, 6s, 7	RGB	1	30 fps
(Liao et al., 2003)	NR	RGB	NR	NR
(Lu et al., 2009)	NR	RGB	NR	NR

Table 3.4 continued.

Reference	Camera model	Modality	Camera No.	Data collection setting
(Montoliu et al., 2015)	NR	NR	16 synchronised and stationary with a 'bird's eye view' positioned along a soccer pitch	25 fps
(Nibali et al., 2017)	NR	RGB	One fixed	NR
(Conaire et al., 2010)	IP camera	RGB	One overhead and eight around court baseline positioned	NR
(Ramanathan et al., 2016)	NR	NR	NR	NR
(Reily et al., 2017)	Kinetic 2	Depth Camera	1	NR
(Shah et al., 2007)	NR	RGB	NR	NR
(Tora et al., 2017)	NR	NR	NR	NR
(Victor et al., 2017)	NR	RGB	NR	Swimming: 50 fps Tennis: 30 fps
(Yao & Fei-Fei, 2010)	NR	RGB	NR	NR
(Zhu, Xu, Huang, & Gao, 2006)	Live Broadcast vision	RGB	NR	Video compressed in MPEG-2 standard with a frame resolution 352 x 288 pixels

fps frames per second, *Hz* hertz, *MPEG* Moving Picture Experts Group, *RGB* red green blue

NR not reported: study either did not directly report the specification or the device did not include the sensor type

3.4.4 Inertial measurement unit recognition model development methods

Key stages of model development from data pre-processing to recognition techniques for IMU-based studies are presented in Table 3.5. Data pre-processing filters were reported as either a low-pass filter (n = 7) (Adelsberger & Tröster, 2013; Buckley et al., 2017; Kelly et al., 2012; O'Reilly et al., 2015; Reilly et al., 2017b, 2017a; Rindal et al., 2018), high-pass filter (n = 2) (Kautz et al., 2017; Schuldhaus et al., 2015), or calibration with a filter (Salman et al., 2017). Processing methods were reported in 67% of the IMU-based studies (Adelsberger & Tröster, 2013; Anand et al., 2017; Brock et al., 2017; Buckley et al., 2017; Buthe et al., 2016; Conaire et al., 2010; Groh et al., 2015, 2017, 2016; Jensen et al., 2016, 2015; Jiao, Wu, et al., 2018; Kautz et al., 2017; Kobsar et al., 2014; Pernek et al., 2015; Qaisar et al., 2013; Reilly et al., 2017a, 2017b; Salman et al., 2017; Schuldhaus et al., 2015). Methods included, calibration of data (Groh et al., 2017, 2016; Jensen et al., 2015; Qaisar et al., 2013), a one-second window centred around identified activity peaks in the signal (Adelsberger & Tröster, 2013; Schuldhaus et al., 2015), temporal alignment (Pernek et al., 2015), normalisation (Conaire et al., 2010), outlier adjustment (Kobsar et al., 2014) or removal (Salman et al., 2017), and sliding windows ranging from one to 3.5 seconds across the data (Jensen et al., 2016). The three studies that investigated trick classification in skateboarding (Groh et al., 2015, 2017) and snowboarding (Groh et al., 2016) corrected data for different rider board stance styles, termed Regular or Goofy, by inverting signal axes.

Movement detection methods were specifically reported in 16 studies (Adelsberger & Tröster, 2013; Anand et al., 2017; Conaire et al., 2010; Connaghan et al., 2011; Groh et al., 2015, 2017, 2016; Jensen et al., 2013, 2015; Kautz et al., 2017; Kelly et al., 2012; Kos & Kramberger, 2017; Rindal et al., 2018; Salman et al., 2017; Schuldhaus et al., 2015;

Whiteside et al., 2017). Detection methods included thresholding (n = 5), windowing segmenting (n = 4), and a combination of threshold and windowing techniques (n = 5).

Signal feature extraction techniques were reported in 80% of the studies, with the number of feature parameters in a vector ranging from a vector of normalised X, Y, Z accelerometer signals (Conaire et al., 2010) to 240 features (Reilly et al., 2017a). Further feature selection to reduce the dimensionality of the feature vector was used in 11 studies. Both feature extraction and selection methods varied considerably across the literature (Table 3.5).

Algorithms trialled for movement recognition were diverse across the literature (Table 3.5). Supervised classification using a kernel form of Support Vector Machine (SVM) was most prevalent (n = 16) (Adelsberger & Tröster, 2013; Brock et al., 2017; Brock & Ohgi, 2017; Buckley et al., 2017; Buthe et al., 2016; Conaire et al., 2010; Groh et al., 2015, 2017, 2016; Jensen et al., 2016; Kautz et al., 2017; Kelly et al., 2012; Pernek et al., 2015; Salman et al., 2017; Schuldhaus et al., 2015; Whiteside et al., 2017). The next highest tested were Naïve Bayesian (NB) (n = 8) (Buckley et al., 2017; Connaghan et al., 2011; Groh et al., 2015, 2017, 2016; Kautz et al., 2017; Salman et al., 2017; Schuldhaus et al., 2015) and k-Nearest Neighbour (kNN) (n = 8) (Buckley et al., 2017; Conaire et al., 2010; Groh et al., 2015, 2017, 2016; Kautz et al., 2017; Salman et al., 2017; Whiteside et al., 2017), followed by Random Forests (RF) (n = 7) (Buckley et al., 2017; Groh et al., 2017; Kautz et al., 2017; Reilly et al., 2017b, 2017a; Salman et al., 2017; Whiteside et al., 2017). Supervised learning algorithms were the most common (n = 29). One study used an unsupervised discriminative analysis approach for detection and classification of tennis strokes (Kos & Kramberger, 2017). Five IMU-based study investigated a deep learning approach including using Convolutional Neural Networks (CNN) (Anand et al., 2017; Brock et al., 2017; Jiao, Wu, et al., 2018; Kautz et al., 2017; Rassem et al., 2017) and Long Short Term Memory (LSTM) (Hochreiter &

Schmidhuber, 1997) architectures (Rassem et al., 2017; Sharma et al., 2017). In order to assess the effectiveness of the various classifiers from each study, model performance measures quantify and visualise the predictive performance as reported in the following section.

Table 3.5 Inertial measurement unit study description and model characteristics.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Adelsberger & Tröster, 2013)	Weight-lifting: thruster (squat press)	16: four females and 12 males, beginner to expert		Low-pass filter	1 s window	Heuristically found threshold value to derive start and end indices of each thruster episode	Accelerometer magnitude modelled on sum of six Gaussian functions with four parameters each: scale αi, amplitude offset βi, standard deviation σi, and mean value μi	1.5 s window around detected signal peaks. Nelder Mead simplex direct search	SVM
(Anand et al., 2017)	Tennis: 7 stroke types Badminton: 4 stroke types Squash: 3 stroke types	31 tennis players, 34 badminton players, 5 squash players	Total training set: ~8500. Total testing set: ~ 7100			Detection shot: 3 cues to identify shot regions across the three sports: 1) threshold, 2) jerk based detection, 3) shot shape-based detection. Fixed number or sample before and after impact point assigned as shot region	Seven shot windows developed for each stage of a shot. Three feature set types generated from all shot windows resulting in ~2000 features including: 1) statistical features, 2) pairwise correlation coefficients between elements of the window set, 3) shape-based features	Pearson correlation coefficient minimum redundancy maximum relevance (MRMR) technique	LR, bi-directional LSTM
(Brock & Ohgi, 2017)	Ski Jumping: error jump, non-error jump	Four: male, junior athletes					Set 1: discrete feature values based on one-dimensional data points built from the raw and processed data of every sensor Set 2: different time-series features based on the estimated positions and orientations of every sensor		SVM, DTW

Table 3.5 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Brock et al., 2017)	Ski jumping: nine motion style errors in flight and landing (5 errors during aerial phase/ 4 error during landing phase)	Three: ski jump athletes	85 measured jump motions		1) removal of internal noise 2) sensor alignment to bone direction of mounted segment using standardised calibration measurement 3) neutralisation 4) segmentation of motion streams into jump phases 5) all sensor streams down-sampled by factor of 2 along temporal domain		CNN model - transformed every pre-processed data segment into a multi-channel motion image of size [R, C, D] with D = 3		CNN, SVM
(Buckley et al., 2017)	Running: classification of running form as a non-fatigued or fatigued state	21: 11 females, 10 males, recreationally active	584 extracted stride repetitions labelled as 292 non-fatigued and 292 fatigued	Low-pass Butterworth filter with a frequency cut-off of 5 Hz od order n = 5	Additional signals computed: Euler, pitch, roll, yaw and Quaternion W, X, Y, Z using algorithms on board the Shimmer IMUs. Stride segmentation by an adaptive algorithm		16 time-domain and frequency-domain features computed to describe the 16 IMU signals over each stride repetition.	Wilcoxon Rank Sum Test, the top 20 signal features extracted	RF, SVM, kNN, NB
(Buthe et al., 2016)	Tennis: 6 shot types and movements	Four: male athletes, three intermediate and 1 advanced	Shots n = 200 Steps n = 640		Shots: discretise data using k-Means algorithm Steps: deadreckoning technique				Shots: LCS Steps: SVM

Table 3.5 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Connaghan et al., 2011)	Tennis: serve, forehand, backhand	Eight: two novices, three intermediate, three advanced athletes	2543			Compute length 3D acceleration vector with a W s window around largest absolute magnitude			NB
(Groh et al., 2015)	Skateboarding: ollie, Nollie, kickflip, heelflip, pop shove-it, 360-flip	Seven: male, advanced skateboarders as three regular and four goofy stance directions	210		Rider stance correction: x-axes and z-axes for all goofy rider stance data inverted	Accelerometer signal segmented into window lengths 1 s with 0.5 s overlap. Energy of window calculated as sum of squares of all axes. Threshold-based detection defined	Total 54 features calculated: mean, variance, skewness, kurtosis, dominant frequency, bandwidth, x-y-correlation, x-z-correlation, y-z-correlation	Embedded Classification Software Toolbox using the best-first forward selection method	NB, PART, SVM (radial bases kernel), kNN
(Groh et al., 2016)	Snowboarding: two trick categories (Grinds and Airs) with three trick classes each category	Part A Four: male snowboarders Part B Seven: male snowboarders	275 tricks total (119 Grinds and 156 Airs)		Calibration of accelerometer and gyroscope data using static measurements and rotations about all axes. Rider stance correction: x-axes and z-axes of all goofy rider stance data inverted	Peak detected in accelerometer signal landing after trick. L^1-norm $S\alpha, t$ computed for all times t. Window-based threshold of length 50 samples (0.25s), overlap 49 samples. Threshold determined by LOOCV	Trick category: defined threshold approaches from magnetometer signals Trick class: nine gyroscope signal features of total rotation, rotation for first half of trick, and rotation from s half of trick for each axis		Trick category: NB Trick class: NB, kNN, SVM, C4.5

Table 3.5 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Groh et al., 2017)	Skateboarding: 11 trick types, trick fail, resting period	11: skateboard athletes	905 trick events		Calibration. Signal y-axes and z-axes inverted	Accelerometer peaks and gyroscope landing impact signals	Accelerometer: x–z-axes correlation after a landing impact Gyroscope: correlation of the x–y-, x–z- and y–z-axes, and specified rotation features	Trick event interval defined as 1 s before and 0.5 s after landing impact	NB, RF, LSVM, SVM (radial-basis kernel), kNN
(Jensen et al., 2015)	Golf: putt phases, putt event, no-putt event	15: inexperienced golfers	272		Sensor data calibration using the 9DOF Calibration Software (version 2.3). Sensor data transformation using a Direction Cosine Matrix	HMM with sliding windows (500 samples, 1.95 s) with a 50% overlap	31 kinematic parameters from 6D IMU data: (1) phase length and ratios of phase lengths (2) angles and ratios of angles (3) velocity at impact (4) summed acceleration around impact (5) velocity and acceleration profiles in fore-swing		AB
(Jensen et al., 2016)	Swimming: rest period, turn, butterfly, backstroke, breaststroke, freestyle	11: high level junior swimmers			Sliding windows between 1 s to 3.5 s with 0.5 s increments. Feature normalization		48D feature vectors per window, computed on each axis: signal energy, min, max, mean, STD, kurtosis, skewness, variance	Best First Search wrapper algorithm	AB, LR, PART, SVM

Table 3.5 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Jensen et al., 2013)	Swimming: butterfly, backstroke, breaststroke, freestyle, turns	12: five females and 7 males, high-level swimmers				Spatial energy and head position	48 features total (8 features x 6 axes): mean, STD, variance, energy, kurtosis, skewness, min, max		DT
(Jiao, Wu, et al., 2018)	Golf: nine swing types	Four: amateur to professional ranked golfers	213 raw samples, 917 samples after augmentation		Dataset augmented to balance swing counts in each class				Vanilla CNN
(Kautz et al., 2017) Machine learning approach	Volleyball: nine shot skill types, one null class	30: 11 females and 19 males, novice to professional	4284	High-pass Butterworth filter with an 8 Hz cut-off frequency	$L1$-norm of the high-passed signal was computed. Signal was smoothed using a low-pass Butterworth filter with a 3 Hz cut-off frequency	Threshold based approach with calculated indicators. C4.5 with LOOCV	39 features: median, mean, STD, skewness, kurtosis, dominant frequency, amplitude of spectrum at dominant frequency, max, min, position of the max, position of the minimum, energy. Pearson correlation coefficients for the correlations between axes	Filter based on the Adjusted Rand Index	SVM, (radial basis kernel function), kNN, Gaussian NB, CART, RF, VOTE

Table 3.5 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Kautz et al., 2017) Deep learning approach	Volleyball: nine shot skill types, one null class	30: 11 females and 19 males, novice to professional	4284		Resampling of raw data				Deep CNN defined as two conv layers with ReLUs and max-pooling, followed by two FC layers with soft-max
(Kelly et al., 2012)	Rugby Union: tackle and non-tackle impacts	Nine: professional athletes		Low-pass filter on magnitude signals		Local maxima with an amplitude cut-off of 0.25 Hz	Static window features: max, min, mean, variance, kurtosis, skewness Impact region features: calculated from a window with dynamically calculated start and end points. Impact region signal features: temporal changes in each accelerometer raw data signals		SVM, HCRF, Learning Grid approach with model fusion by AB

Table 3.5 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Kobsar et al., 2014)	Running: motion patterns to predict training background and experience level	14, soccer athletes. 16, first time marathon runners. 12, experienced marathon runners	Per participant: 15 s accelerometer data equating to ~20 – 25 footfalls		RMS of accelerations in the vertical, medio-lateral, anteroposterior, and resultant direction calculated. The economy of accelerations determined as the RMS in each axis divided by the gait speed. Outliers adjusted using a Winsorizing technique. All variables standardised to a mean of 0 and a STD of 1		DWT procedure of 5-level wavelet decomposition using Daubechies 5-mother wavelet	PCA	LDA (binary classification)
(Kos & Kramberger, 2017)	Tennis: forehand, backhand, serve	Seven: junior to senior athletes	446			Defined threshold based on two-point derivative of acceleration curves			Unsupervised discriminative analysis
(Conaire et al., 2010)	Tennis: serve, backhand, forehand	Five: elite nationally ranked	300		Normalisation of stroke data by rescaling for variance to equal 1	1 s window over accelerometer peaks detected from a threshold approach	Normalised signal x, y, z vectors		SVM (radial basis function kernel), kNN
(O'Reilly et al., 2015)	Squat: correct or incorrect technique and specific technique deviations	22: 4 females and 18 males, with prior experience and regular squat training in regime	682	Low-pass Butterworth filter with a frequency cut-off of 20 Hz			30 feature set		Back-propagation NN

Table 3.5 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Reilly et al., 2017a)	Lunge: discriminate between different levels of lunge performance and identify aberrant techniques	80: 23 females, 57 males, with prior experience and regular lunge training in regime	3440	Low-pass Butterworth filter with frequency cut-off of 20 Hz of order n = 8	3D orientation of IMU computed from all axes using a gradient descent algorithm. Acceleration and gyroscope magnitude calculated. Each exercise repetition resampled to length of 250 samples.		240 features per IMU calculated and extracted including:		RF
(Reilly et al., 2017b)	Deadlifting: technique deviations	135: 41 females and 94 males, with prior lifting experience	2245	Low-pass Butterworth filter with a frequency cut-off of 20 Hz	Rotation quaternions were converted to pitch, roll and yaw signals. Magnitude of acceleration and rotational velocity computed. Time-normalization by exercise repetitions resampled to a length of 250 samples		17 time and frequency domain feature each signal: mean, RMS, STD, kurtosis, median, skewness, range, variance, max, min, energy, 25th percentile, 75th percentile, fractal dimension, level crossing-rate, variance of approximate and detailed wavelet coefficients		RF

Table 3.5 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Pernek et al., 2015)	Weightlifting: six dumbbell lifting exercises	11: three females and 8 males	~ 2904		Temporal alignment. Uniform resampling of sample rate to 25 Hz		Min, max, range, arithmetic mean, STD, RMS, correlation	Sliding window approach	SVM (Gaussian radial basis function kernel)
(Qaisar et al., 2013)	Cricket: correct and incorrect medium paced bowls	One: medium paced cricket bowler	40		Calibration by filter using signal processing techniques and interpolated to smooth out the filtered data		Mean, mode, STD, peak to peak value, min, max, first deviation, second deviation	K-means clustering	K-means clustering, Markov Model, HMM.
(Rassem et al., 2017)	Cross-country skiing: gears variations	NR	416,737		Data segmented into training, validation, testing set applied with a window size 1 sec with 50% overlap				Recurrent LSTM, CNN, MLP
(Rindal et al., 2018)	Cross-country skiing: eight technique sub-classes	10: 9 male, 1 female, trained amateurs to professional world-cup skiers	8616	Chest accelerometer data filtered with Gaussian low-pass filter 0.0875 s (1.75 samples) standard deviation in the time domain			Samples were decimated or interpolated into 30 samples per cycle and then appended into one feature vector of 94 samples		NN with three hidden layers of 50, 10, 20 neurons in each layer respectively

Table 3.5 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Salman et al., 2017)	Cricket: detect legal or illegal bowls	14: male cricketers, medium and fast paced bowlers	150	Calibration and filter	Outliers removed using IQR method. Missing values in each attribute replaced with corresponding mean values of attribute, conditional of 10% limit of missing values per attribute before discarded	Data divided into tagged windows corresponding to phases of bowling action. Ball release point was the maxima to denote start process of windowing and tagging	Seven features per axis of accelerometer and gyroscope signals: mean, median, STD, skewness, kurtosis, min, max	Correlation-based feature selection with Greedy search method resulting in the top 21 features	SVM (redial basis function kernel), kNN, NB, RF, NN (three-layer feed-forward)
(Schuldhaus et al., 2015)	Soccer: shot, pass, event leg, support leg, other soccer events	23: male athletes	64 passes, 12 shots	High-pass Butterworth filter		Accelerometer peak detection using a Signal Magnitude Vector. Segmented windows of 1 s around peaks	Four features from each accelerometer axis: mean, variance, skewness, kurtosis		SVM (linear kernel), CART, NB

Table 3.5 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Srivastava et al., 2015)	Tennis: forehand, backhand, serve, sub-shot types (flat, topspin, slice)	14: five professional and nine novices	~1000 shots from professional athletes, ~1800 shots from novice athletes			Pan Tomkin's algorithm to isolate shot signal from noise. Accelerometer x-axis differentiated and squared. Moving window integration with window size 3* the sampling rate. Identified potential shot impact region using thresholding			Two Level hierarchical classifier: (1) DTW, (2) QDTW
(Whiteside et al., 2017)	Tennis: serve, forehand (rally, slice, volley), backhand (rally, slice, volley), smash, false shot	19: 8 females and 11 males, junior national development athletes	Per athlete: mean 1504 ± 971		Saturated signals reconstructed using a linear interpolation method. Signals smoothed with 50-point (0.1 sec) moving average.	Threshold algorithm with a window size 0.5 s either side of the detected shot. Shot instances temporally aligned with exported coded vision file.	40 features (5 features across 8 waveforms): min, med, integral, discrete value at time of impact		SVM (linear, quadratic, cubic, Gaussian kernels), CT (10, 25, 50 splits), kNN (k of 1, 3, 5), NN, RF, DA (linear and quadratic)

3D three dimensions, AB Adaptive Boosting, C4.5 decision tree analysis type, CART classification and regression tree, CNN convolutional neural network, CT classification tree, DA discriminative analysis, DOF degrees of freedom, DT decision tree, DWT dynamic time warp, FC fully-connected, HCRF hidden conditional random field, HMM Hidden Markov Model, HZ hertz, IMU inertial measurement unit, IQR interquartile range, kNN k-Nearest Neighbour, LCS Longest Common Subsequence algorithm, LDA linear discriminative analysis, LOOCV leave-one-out-cross-validation, LR logistic regression, LSTM long short term memory, LSVM linear support vector machine, MLPs multi-layer perceptrons, NB Naïve Bayesian, NN neural network, NR not reported, PART partial decision tree, QDTW Quaternions based Dynamic Time Warping, ReLUs rectifier linear unit, RF random forests, RMS root mean square, STD standard deviation, SVM Support Vector Machine, VOTE vote classifier.

3.4.5 Inertial measurement unit recognition model evaluation

Reported performance evaluations of developed models across the IMU-based studies are shown in Table 3.6. Classification accuracy, as a percentage score for the number of correct predictions by total number of predictions made, was the main model evaluation measure (n = 24). Classification accuracies across studies ranged between 52% (Brock & Ohgi, 2017) to 100% (Buckley et al., 2017). Generally, the reported highest accuracy for a specific movement was ≥ 90% (n = 17) (Adelsberger & Tröster, 2013; Anand et al., 2017; Buckley et al., 2017; Conaire et al., 2010; Connaghan et al., 2011; Groh et al., 2015; Jensen et al., 2013; Jiao, Wu, et al., 2018; Kobsar et al., 2014; Kos & Kramberger, 2017; Pernek et al., 2015; Qaisar et al., 2013; Reilly et al., 2017a; Rindal et al., 2018; Schuldhaus et al., 2015; Srivastava et al., 2015; Whiteside et al., 2017) and ≥ 80% to 90% (n = 7) (Brock et al., 2017; Brock & Ohgi, 2017; Groh et al., 2017; Jensen et al., 2016; O'Reilly et al., 2015; Reilly et al., 2017b; Salman et al., 2017). As an estimate of the generalised performance of a trained model on $n - x$ samples, a form of leave-one-out cross validation (LOO-CV) was used in 47% of studies (Buthe et al., 2016; Conaire et al., 2010; Groh et al., 2015, 2017, 2016; Jensen et al., 2013, 2016; Kobsar et al., 2014; O'Reilly et al., 2015; Pernek et al., 2015; Reilly et al., 2017b; Salman et al., 2017; Schuldhaus et al., 2015). Precision, specificity and sensitivity (also referred to as recall) evaluations were derived for detection (n = 6) and classification models (n = 10). Visualisation of prediction results in the form of a confusion matrix featured in six studies (Buthe et al., 2016; Groh et al., 2017; Kautz et al., 2017; Pernek et al., 2015; Rindal et al., 2018; Whiteside et al., 2017).

Table 3.6 Inertial measurement unit study model performance evaluation characteristics.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Anand et al., 2017)	Detection: precision, recall, F1-score Classification: CA		Detection of squash: • Precision 0.95 • Recall 0.96 • F1- score 0.96 CA: • Tennis: CNN 93.8% • Badminton: BLSTM 78.9% • Squash: BLSTM 94.6%	In-house developed tool to align recorded vision and sensor data to tag shot types in which tagged data serves as ground truth for analysis	
(Adelsberger & Tröster, 2013)	Detection accuracy, CA	75% / 25% train-test dataset split	Detection accuracy: • 100% (when athletes did not move between reps) Classification: • CA 94.117% (between expert and beginner level) Classification: • CA 93.395% (individual thruster instances)	Video footage with performances labelled by a certified coaching expert	Dataset split details: Tennis: training set ~4500 shots by 15 players testing set ~5000 shots by 16 players Badminton: training set ~3500 shots by 20 players testing set ~2000 shots by 14 players Squash: training set ~500 shots by 3 players testing set ~100 shots by 2 players
(Brock & Ohgi, 2017)	Precision, recall, CA, error rate		SVM: CA 52% - 82%	Video control data	For each classifier algorithm, 72 experiments were conducted varying in factor sampling rate (4 variations), windows size (6 variations) and feature selection strategy (3 variations). Error rate defined as the difference between classification accuracy and 1.0

Table 3.6 continued.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Brock et al., 2017)	CA, cross-entropy loss	8-fold cross validation	CNN 1 layer: CA 93 ± 0.08%	Jump style annotated by qualified judge under the judging guidelines of the International Skiing Federation	
(Buckley et al., 2017)	CA, sensitivity, specificity, F1-score,	LOO-CV 10-K-fold cross validation	Global Classifier: • LIMU lumbar spine CA 75% • IMU right shank CA 70% • IMU left shank CA 67% Personalised classifier: • IMU lumbar spine CA 89% • IMU right shank CA 99% • IMU left shank CA 100%	Manual labelling	Personalised classifiers appear more computationally efficient than global classifiers as they require less training data and memory storage.
(Buthe et al., 2016)	Detection accuracy, confusion matrix, recall, precision, user-specific dataset comparison for train and test	LOO-CV	Step detection accuracy: • Overall 76% • Side steps 96% • Shot steps 63% LOOCV: • Precision 0.49 ± 0.04% • Recall 0.49 ± 0.22% User-specific: • Precision 98% • Recall 87%		Gyroscope signals showed to be more suitable than accelerometer signals to separate shot movements and identify fast foot movements
(Connaghan et al., 2011)	Detection accuracy, CA	10-fold cross validation	Detection accuracy: • Candidate strokes 85% • Non-candidate strokes 85% Classification accuracy: • 3 sensor fusion overall accuracy 90% • Accelerometer 7 player model 97% • Gyroscope 7 player model 76% • Magnetometer 7 player model 76%		Accelerometer signals were the most effective at classifying different skill levels

Table 3.6 continued.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Groh et al., 2015)	Detection: sensitivity, specificity Classification: CA, computational effort	LOSO-CV	Detection: • Sensitivity 94.2% • Specificity 99.9% Classification: • CA 97.8% (NB and SVM) Computation effort (lowest): • NB (operations 360, time 6.2 s) • PART (operations 41, time 10.6 s)	Video footage and expert analysis of trick quality	Computational effort defined as the time and required operations for one model run without grid search
(Groh et al., 2016)	Precision, recall, CA	LOSO-CV	Event detection: • Recall 0.99 • Precision 0.368 Trick category classification: • Grind recall 0.966 • Grind precision 0.885 • Airs recall 0.974 • Airs precision 0.910 Trick class CA: • Grind 90.3% (SVM) • Airs 93.3% (kNN)	Video footage	
(Groh et al., 2017)	Detection: precision, recall Classification: CA, confusion matrix	Classification: LOSO-CV	Detection: • Precision 0.669 • Recall 0.964 Classification: Correct trick execution • CA 89.1% (SVM) All tricks modelled 79.8% CA (RF)	Video footage with manual annotation	

Table 3.6 continued.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Jensen et al., 2015)	Detection accuracy, false positive rate		Overall detection rate 68.2%. False positive rate 2.4%	Video footage	Detection rate: $DR = \frac{N_d}{N_p}$ False positive rate: $FPR = \frac{N_m}{N_m + N_p}$ N_d number of detected putts N_p number of performed putts N_m number of misdetected putts
(Jensen et al., 2016)	CA	LOSO-CV	Maximum CA 86.5% (SVM) Average CA 82.4% (SVM)	Video footage manually labelled	72 methodological experiments were conducted. A sampling rate of 10.25 Hz and increased window sizes produced higher classification accuracy.
(Jensen et al., 2013)	CA	LOSO-CV	Turn CA 99.8%. Swim stroke CA 95%		
(Jiao, Wu, et al., 2018)	CA, precision, recall	10-fold cross validation	CA 95% Precision 0.95 average Recall 0.95 average F1-score 0.95 average		
(Kautz et al., 2017) Machine learning approach	Confusion matrix, sample accuracy, balanced accuracy, computational time	Detection: LOSO-CV Classification: leave-three-subjects-out cross validation	Sample accuracy 67.2% (VOTE) Balanced accuracy 60.3% (VOTE) Training computational time: • 18.1 ms (NB with feature selection) Class prediction computational time: • 0.53 µs (CART)	Video footage manually labelled	Sample accuracy: $\lambda_s = \frac{\sum_{c=1}^{M} r_c}{\sum_{c=1}^{M} N_c}$ Balanced accuracy: $\lambda_b = \frac{1}{M}\sum_{c=1}^{M}\frac{r_c}{N_c}$ N_c number of samples from class *c* r_c number of sample from class *c* classified correctly M number of classes

Table 3.6 continued.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Kautz et al., 2017) Deep learning approach	Sample accuracy, balanced accuracy	Leave-two-out cross-validation	Sample accuracy 83.2% Balanced accuracy 79.5%	Video footage manually labelled	
(Kelly et al., 2012)	Recall, precision, TP, TN, FP, FN		Learning Grid approach: • Recall 0.933 • Precision 0.958	Video footage manually labelled by the medical staff of the elite rugby union team involved	
(Kobsar et al., 2014)	CA	LOO-CV	Training background CA 96.2% Experience level CA 96.4%		
(Kos & Kramberger, 2017)	CA		Serve CA 98.8%, forehand CA93.5%, backhand CA 98.6%	Video footage	Gyroscope signals were found to be more discriminative between stroke types
(Conaire et al., 2010)	Detection accuracy, CA	LOO-CV	Detection accuracy: 100% Classification: • Right arm data CA 89.41% (kNN) • Full-body data CA 93.44% (kNN)		Data fusion of accelerometer and vision data improved CA: • Vision back viewpoint with full body accelerometer 100% CA (kNN) Data fusion overcame viewpoint sensitivity • Vision trained on side viewpoint and tested on back viewpoint fused with full body accelerometer data 96.71% CA (kNN)
(O'Reilly et al., 2015)	CA, sensitivity, specificity	LOSO-CV	Binary classification: • Sensitivity 64.41% • Specificity 88.01% • CA 80.45% Multi-label classification; • Sensitivity 59.65% • Specificity 94.84% • CA 56.55%	Chartered Physiotherapist evaluation based on the National Strength and Conditioning Association guidelines	

Table 3.6 continued.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Reilly et al., 2017a)	CA, sensitivity, specificity, out-of-bag error	LOSO-CV	Classify acceptable and aberrant technique five lower limb IMU set-up: • CA 90% • Sensitivity 80% • Specificity 92% Classify specific technique deviations five lower limb IMU set-up: • CA 70% • Sensitivity 70% • Specificity 97%	Chartered physiotherapist and strength and conditioning trained practitioner. Correct technique described by the National Strength and Conditioning Association (NSCA) guidelines.	
(Reilly et al., 2017b)	CA, sensitivity, specificity	LOSO-CV	Natural technique deviations binary CA: • Global classifier 73% (RF) • Personalised classifier 84% (RF) Natural technique deviations multi-class CA: • Global classifier 54% (RF) • Personalised classifier 78% (RF)	Video footage labelled by a Chartered Physiotherapist	Personalised classifiers outperformed the global classifiers and were more computationally efficient. kNN, SVM, NB tested during analysis against RF, but did not improve results and some caused increased computational times in some cases.
(Pernek et al., 2015)	CA, prediction error, confusion matrix	LOSO-CV, 10-fold cross-validation, 75%/ 25%train-test dataset split	Methodology experiments: • CA range 84.2 ± 11.3% to 93.6 ± 0.5% Intensity error: • range 1.2% to 6.6 ± 2.5%	Video footage with manual annotation	A 2 s window size with 50% overlap data processing yielded the best performance results.
(Qaisar et al., 2013)	CA		Overall CA: 90.2% (HMM) • Wrist sensor data 100% • Elbow sensor data 88.24% • Upper arm sensor data 82.35%	Video footage	

Table 3.6 continued.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Rassem et al., 2017)	Average testing classification error over the model run. MLP model used as performance benchmark for DL models		Standard LSTM: 1.6% class error value CNN: 2.4% class error value		Data was divided into training, validation and testing sets with a segmentation process applied of window size one second with a 50% overlap.
(Rindal et al., 2018)	CA, sensitivity, precision, confusion matrix	Validation dataset was used to evaluate which of the 20 trained neural networks to use for final model. Test set created from six different athlete data	CA 99.8% on training dataset CA 96.5% on validation dataset CA 93.9% on combined tests sets	Manual video labelling	Artificially expanded training dataset by taking every cycle in the original training data and created a new cycle by keeping the x-axis and z-axis, whereas the y-axis was flipped resulting in 8616 cycles from the original 4308 training cycles.
(Salman et al., 2017)	Detection accuracy, CA, recall, precision, F1-score	LOSO-CV	Detection of ball release point 100% accuracy. CA 81 ± 3.12% (SVM) Recall 0.80 (SVM) Precision 0.82 (SVM) F1-score 0.81 (SVM)	Video footage evaluated by an expert cricketer	
(Schuldhaus et al., 2015)	CA	LOSO-CV	Set protocol conditions CA (SVM): • Leg type 99.9% • Other events 96.7% • Pass or shot 88.6% Match conditions CA (SVM): • Shot 86.7% • Pass 81.7%	Video footage manually labelled	

Table 3.6 continued.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Srivastava et al., 2015)	Detection accuracy, CA		Shot detection accuracy: • Professional 99.58% • Novice 98.96% • Total 99.41% Shot CA: • Class professional player 99.6% • Class novice player 99.3% • Sub-shot types professional player 90.7% • Sub-shot types novice player 86.2%		
(Whiteside et al., 2017)	CA, confusion matrix, precision, recall	10-fold cross-validation	Mean CA (SVM – cubic kernel): • Condition one 97.43 ± 0.24% • Condition two 93.21 ± 0.45%	Video footage manually labelled by a performance analyst	SVM algorithms were constructed using linear, quadratic, cubic and Gaussian kernels, and a one-versus-one approach. kNN classifiers were built using a k of 1,3 and 5. CT were constructed using a maximum of 10, 25 and 50 splits. NN included a conventional single-layer model and multi-layer deep network

CA classification accuracy, CART classification and regression tree, CT classification tree, FN false negative, FP false positive, Hz hertz, kNN k-Nearest Neighbour, LOO-CV leave-one-out cross validation, LOSO-CV leave-one-subject-out cross validation, MLP multi-layer perceptrons, NB Naïve Bayesian, PART partial decision tree, RF random forests, SVM Support Vector Machine, TN true negative, TP true positive, VOTE vote classifier.

3.4.6 Vision recognition model development methods

Numerous processing and recognition methods featured across the vision-based studies to transform and isolated relevant input data (Table 3.7). Pre-processing stages were reported in 14 of studies, and another varied 13 studies also provided details of processing techniques. Signal feature extraction and feature selection methods used were reported in 78% of studies.

Both machine (n = 16) and deep learning (n = 7) algorithms were used to recognise movements from vision data. Of these, a kernel form of the SVM algorithm was most common in the studies (n = 10) (Conaire et al., 2010; Couceiro et al., 2013; Horton et al., 2014; Kasiri-Bidhendi et al., 2015; Kasiri et al., 2017; J. Li et al., 2018; Montoliu et al., 2015; O'Reilly et al., 2017; Reily et al., 2017; Shah et al., 2007; Zhu, Xu, Huang, & Gao, 2006). Other algorithms included kNN (n = 3) (Conaire et al., 2010; Díaz-Pereira et al., 2014; Montoliu et al., 2015), decision tree (DT) (n = 2) (Kapela et al., 2014; Liao et al., 2003), RF (n = 2) (Kasiri-Bidhendi et al., 2015; Kasiri et al., 2017), and Multilayer Perceptron (MLP) (n = 2) (Kapela et al., 2014; Montoliu et al., 2015). Deep learning was investigated in seven studies (Bertasius et al., 2017; Ibrahim et al., 1971; Karpathy et al., 2014; Nibali et al., 2017; Ramanathan et al., 2016; Tora et al., 2017; Victor et al., 2017) of which used CNNs or LSTM RNNs as the core model structure.

Table 3.7 Vision-based study description and model characteristics.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset samples	Pre-processing	Processing	Feature extraction and selection	Recognition
(Bertasius et al., 2017)	Basketball: some-body shooting a ball, camera wearer possessing the ball, camera wearer shooting the ball	48: male US College players	10.3 hours of recorded vision			Gaussian mixture function	CNN, Multi-path convolutional LSTM
(Couceiro et al., 2013)	Golf Putting: athlete signature features	Six: male, expert level	180 trial shots (30 trials per athlete)		Darwinian particle swarm optimization method		LDA, QDA, NB with Gaussian distribution, NB with kernel smoothing density estimate, LS-SVM with RBF kernel
(Díaz-Pereira et al., 2014)	Gymnastics: 10 actions grouped into three categories of jumps, rotations, pre-acrobatics	Eight: junior gymnasts	560 video shots (5 - 7 actions per gymnast)	Motion Vector Flow Instance		PCA and LDA	kNN
(Hachaj et al., 2015)	Oyama Karate: 10 classes of actions grouped into 4 defence types, 3 kick types, 3 stands	Six: advanced Oyama karate martial artists	1236	Pre-classification: data pre-processed based on z-scores calculations for each feature value	Segmentation: GDL classifier approach training with an unsupervised R-GDL algorithm. A Baum-Welch algorithm to estimate HMM parameters	Angle-based features	Continuous Gaussian density forward-only HMM classifiers

Table 3.7 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset samples	Pre-processing	Processing	Feature extraction and selection	Recognition
(Horton et al., 2014)	Soccer: Pass quality	Dataset: English Premiership 2007/2008 season games	2932 passes across four matches			Features: basic geometric prediction variables, sequential predictor variables, physiological predictor variables, strategic predictor variables	Multinomial logistic regression, SVM, RUSBoost algorithm
(Ibrahim et al., 1971)	Volleyball: six team activity classes, seven individual athlete actions	Dataset: 15 YouTube volleyball videos	1525 annotated frames			CNN	CNN, LSTM
(Kapela et al., 2015)	Rugby, Basketball, Soccer, Cricket, Gaelic football, Hurling: 8 scene types	Dataset	50 hours	Video de-coding: storage of every 5th frame in the buffer		FFT	DT, Feed-forward MLP NN, Elman NN
(Karpathy et al., 2014)	Sports-1M dataset	Dataset	1 million YouTube videos containing 487 classes with 1000 - 3000 videos per class	Optimization: Downspur Stochastic Gradient Descent	Data augmentation: (1) crop centre region and resize to 200 x 200 pixels, randomly sampling 170 x 170 region, and randomly flipping images horizontally with 50% probability. (2) subtract constant value of 117 from raw pixel values		CNN (several approaches to fusing data across temporal domains)

Table 3.7 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset samples	Pre-processing	Processing	Feature extraction and selection	Recognition
(Kasiri-Bidhendi et al., 2015)	Boxing: 6 punch types of straight, hook, uppercut from both rear and lead hand	Eight: elite orthodox boxers	192 punches (32 for each type)		Detection of body parts: fuzzy inference method based on 2D chamfer distance and geodesic distances	Spatial-temporal features of each punch	RF, Linear SVM, Hierarchical SVM
(Kasiri et al., 2017)	Boxing: 6 punch types of straight, hook, uppercut from both rear and lead hand	14: elite orthodox and southpaw boxers across different weight classes	605 punches		Detection of body parts: fuzzy inference method based on 2D chamfer distance, depth values and geodesic distances	Transition-invariant trajectory features of hand and arm descriptors extracted. Feature ranking for feature reduction experimented using PCA, RF, SVM-reclusive feature eliminator	Multi-class SVM, RF
(Liao et al., 2003)	Swimming: backstroke, breaststroke, butterfly, freestyle	Dataset	50 clips	Associated limb region detection: RGB images converted to HSV space. Associated skin colour detection: pixels labelled between 0.3 to 1.5 hue values.	Upper body sections isolated using heuristic, threshold approach	LR analysis	DT

Table 3.7 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset samples	Pre-processing	Processing	Feature extraction and selection	Recognition
(Li et al., 2018)	Golf: key swing gesture detection		Golf front angle swing vision from 553 players, Golf side angle swing vision from 790 players, Baseball swing vision from 3363 players			Multi-scale aggregate channel feature method	AD-DWTAdaBoost Linear SVM
(Lu et al., 2009)	Ice Hockey: skating movement directions of down, up, left, right	Male unspecified athletes	5609 images of 32 x 32 grayscale images	Tracking: HSV, HOG combined with SVM. Template updating: SPPCA	Multi-target tracking by incorporated SPPCA with an action recogniser using an AB algorithm		SMLR
(Montoliu et al., 2015)	Soccer: team activities of ball possessions, quick attack, set pieces	Private dataset: professional Spanish soccer team	Two matches of 90 min each	All camera images combined via algorithmic approach for a unique image covering field length		Bag-of-Words Optical Flow	kNN, SVM, MLP
(Nibali et al., 2017)	Diving: 5 dive properties of rotation type, pose type, number of somersaults, number of twists, handstand beginning inclusion	Dataset: high-level divers from the Australian Institute of Sport	Training set: 25 hours with 4716 non-overlapping dives. Test set: day's footage of 612 dives	Temporal action localisation: TALNN - built from volumetric Convolutional layers. Smoothing: Hann Window Function	Spatial Localisation: full regression, partial regression, segmentation, and Global constraints (RANSAC algorithm).		C3D volumetric convolutional network (3x3x3 kernels, ReLUs, dropouts)

Table 3.7 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset samples	Pre-processing	Processing	Feature extraction and selection	Recognition
(Conaire et al., 2010)	Tennis: serve, forehand, backhand	Five: elite nationally ranked			Contour features: back-ground subtraction and image morphology	Player foreground region divided into 16 pie segments centred on player centroid and normalization	SVM with RBF kernel, kNN
(Ramanathan et al., 2016)	Basketball: 11 match activity classes and frame key player detection	Dataset: 257 NCAA games from YouTube	1143 training clips, 856 validation clips, 2256 testing clips	Each clip subsampled to six fps at four seconds in length		Each video-frame represented by a 1024-dimensional feature vector. Appearance features extracted using the Inception7 (Szegedy & Ibarz, 2015) network and spatially pooling the response from the lower layer. Features corresponded to a 32x32 spatial histogram combined with a spatial pyramid	LSTM and BLSTM RNNs
(Reily et al., 2017)	Gymnastics: Pommel horse routine spinning	Unspecified male gymnasts	10115 frames recorded as 16-bit PNG images, organised into 39 routines	DOI segmentation: (1) Parzen window (2) Identified signal peaks padded with neighbourhood 10% max depth		SAD3D: The gymnast in each frame is described by features: (1) width of their silhouette, (2) height of their silhouette, (3–4) depth values at the leftmost and rightmost ends of the silhouette, (5– 8) shift in the left-most x, right-most x, upper y, and lower y coordinates compared to the previous frame.	SVM with radial basis function kernel. Smoothing techniques after classification
(Shah et al., 2007)	Tennis: forehand, backhand, other	Dataset: male and female unspecified athletes	150 games each clipped to 10 min segments	Optical flow calculated between consecutive frames	Image segmentation and weight calculation by global adaptive thresholding. Player appearance modelling by Expectation Maximization algorithm	Oriented histogram of skeletonised binary images of athletes	SVM with RBF kernel

Table 3.7 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset samples	Pre-processing	Processing	Feature extraction and selection	Recognition
(Tora et al., 2017)	Ice Hockey: dump in, dump out, pass, shot, loose puck recovery	Dataset: National Hockey League videos	2507 training events, 250 testing events			Features extracted by the fc7 layers of AlexNet (Krizhevsky et al., 2012). Max-pooling of features of individual players in frames to incorporate player interactions	LSTM
(Victor et al., 2017)	Swimming: backstroke, breaststroke, butterfly, freestyle Tennis: stroke detection	Datasets: Swimming: 40 athletes Tennis: 4 athletes	15k swimming strokes labelled in 650k fames. 1.3k tennis strokes labelled in 270 frames	Swimming: pre-processed using Hough transform as in (Sha et al., 2013) to extract the lanes from colour information. Tennis: excluded unlabelled tennis strokes from input dataset. Input data frames down sampled to 192 x 128 pixels	Model parameters initialised. Adedelta optimiser. MSE loss function. All frame's pixels encoded in YUV colour-space and down sampled to 128 x 48		Regression: CNN with a base architecture based off the VGG-B CNN (Simonyan & Zisserman, 2015)
(Yao & Fei-Fei, 2010)	Human-object interaction sport activities: cricket defensive shot, cricket bowling, croquet shot, tennis forehand, tennis serve, volleyball smash	Dataset	350 images (50 images per 6 classes)	Gaussian over the number of edges and randomization of initialization connectivity to different starting points	Hill-climbing approach with a Tabu list	Parameter estimation with a max-margin learning method	Composition inference method

Table 3.7 continued.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset samples	Pre-processing	Processing	Feature extraction and selection	Recognition
(Zhu, Xu, Huang, & Gao, 2006)	Tennis: left and right swings	Professional tennis athletes	6035 frames of 1099 left swing strokes and 1071 right swing strokes		Player tracking: SVR particle filter and background subtraction.	Motion descriptor extraction: optical flow computed using Horn-Sckunck algorithm with half-wave rectification and Gaussian smoothing. Feature discrimination: slice-based optical flow histograms	SVM

2D two dimensional, BLSTM bidirectional LSTM, CNN convolutional neural network, DOI Depth of interest segmentation, DT decision tree, ELU Exponential Linear Units, FFT Fast Fourier Transform, GDL Gesture Description Language, HMM Hidden Markov Model, HOG Histogram of Oriented Gradients, HSV Hue-Saturation-Value-Colour-Histogram, kNN k-Nearest Neighbour, LDA linear discriminative analysis, LR logistic regression, LS-SVM least squares support vector machine, MLP multi-layer perceptron, NB Naïve Bayesian, NN neural network, PCA principal component analysis, PNG Portable Network Graphics, QDA quadratic discriminative analysis, RBF radial basis function, RF random forests, RUSBoost Random Under Sampling Boosting, SAD3D Silhouette Activity Descriptor in 3 Dimensions, SPPCA Switching Probabilistic Principal Component Analysis, SVM Support Vector Machine, SVR Support Vector Regression.

3.4.7 Vision recognition model evaluation

Performance evaluation methods and results for vision-based studies are reported in Table 3.8. As with IMU-based studies, classification accuracy was the common method for model evaluations, featured in 61%. Classification accuracies were reported between 60.9% (Karpathy et al., 2014) and 100% (Hachaj et al., 2015; Nibali et al., 2017). In grouping the reported highest accuracies for a specific movement that were ≥ 90% (n = 9) (Conaire et al., 2010; Hachaj et al., 2015; Karpathy et al., 2014; Kasiri-Bidhendi et al., 2015; Kasiri et al., 2017; J. Li et al., 2018; Montoliu et al., 2015; Nibali et al., 2017; Reily et al., 2017; Shah et al., 2007), and ≥ 80% to 90% (n = 2) (Horton et al., 2014; Yao & Fei-Fei, 2010). A confusion matrix as a visualisation of model prediction results was used in nine studies (Couceiro et al., 2013; Hachaj et al., 2015; Ibrahim et al., 1971; Karpathy et al., 2014; Kasiri-Bidhendi et al., 2015; Kasiri et al., 2017; Lu et al., 2009; Shah et al., 2007; Tora et al., 2017). Two studies assessed and reported their model computational average speed (Lu et al., 2009) and time (Reily et al., 2017).

Table 3.8 Vision-based study model performance evaluation characteristics.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Bertasius et al., 2017)	F1-score	24 videos for training dataset, 24 videos for testing dataset	Basketball event detection mean F1-score 0.625. Basketball athlete performance evaluation model F1-score 0.793.	Manual labelling and athlete performance assessment by a former professional basketball player	Compared model's performance to first-person activity recognition baselines and a video activity recognition baseline C3D
(Couceiro et al., 2013)	Confusion matrix, ROC		LS-SVM overall best performance		1) five classifiers evaluated for detecting signature patterns 2) best classifier method applied to extract individual golf putt signatures
(Díaz-Pereira et al., 2014)	True/ false recognition rates for binary classification, sensitivity, specificity	10-fold cross validation	Specificity 85% overall Sensitivity 90% overall		
(Hachaj et al., 2015)	CA, confusion matrix	LOO-CV	Overall CA range across classes 93 ± 7% to 100% (four-state HMM)		Five HMM classifiers tested with number of hidden states ranging from 1 (GMM) to 5
(Horton et al., 2014)	CA, precision, recall, F1-score	80%/ 20% train-test dataset split. Tests set was stratified so per class frequency was consistent with the distribution in training examples	Three-class model 85.5% (SVM)	Labelled data of pass events. Rating of pass quality by observers (6-point Likert Scale) Cohen's Kappa for heuristic measure of agreement between ratings	Experiments conducted using two labelling schemes: 1) six-class labels assigned by observers. 2) three-class scheme (aggregation of six-classes) Test dataset was stratified so per-class frequency consistent with distribution in training dataset.
(Ibrahim et al., 1971)	CA, confusion matrix	2/3rd of total data as training set, 1/3rd as testing set	51.1% CA		Compared model performance to several baseline models

Table 3.8 continued.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Kapela et al., 2015)	Modified accuracy (focused around detection performance), precision, modified precision		Overall precision 0.96	Manual annotation	$\text{Modified accuracy} = \frac{(\text{DE} - \text{DTE})}{\text{NE}}$ $\text{Precision} = \frac{\text{DTE}}{\text{DE}}$ $\text{Modified precision} = \frac{\text{DTE}}{\text{NE}}$
(Karpathy et al., 2014)	Prediction classification accuracy %, per-class average precision, confusion matrix	Dataset split: 70% training set, 10% validation set, 20% test set	CNN model average CA 63.9% Slow fusion model CA 60.9%	Labelled data classes	
(Kasiri-Bidhendi et al., 2015)	CA, confusion matrix	LOO-CV Model trained on data from seven participants and tested on withheld data from one participant	Hierarchal SVM CA 92 – 96%	Start and end frames of each punch labelled by expert analysts	
(Kasiri et al., 2017)	CA, feature numbers, confusion matrix		Hierarchical SVM CA 97.3%	Start and end frames of each punch labelled by expert analysts	
(Liao et al., 2003)	Developed scoring system based on measure of proximity to the prominent feature of a specific style				
(Li et al., 2018)	CA, precision, recall, computational time	Cross-validation (not specified). Dataset split: 80% train/ 10% validation/ 10% test set	CA 97% Average recognition time of 2.38 ms		

Table 3.8 continued.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Lu et al., 2009)	CA, average computing speed, confusion matrix		SMLR and HOG approach CA 76.37% Computing speed: average total time classification image 0.206s (SMLR and HOG approach)	Manual image retrieval and division into the four classes	Compared developed model against benchmark action recognisers.
(Montoliu et al., 2015)	CA	5-fold cross-validation, LOO-CV	RF CA 92.89 ± 0.2%	Manual vision annotation by expert	
(Nibali et al., 2017)	CA, precision, recall, F1-score		Dive property CA from 86.89 - 100%	Labelled training data	Segmentation works best (spatial localisation). Dilated convolutions boosted CA.
(Conaire et al., 2010)	CA	LOO-CV	Back viewpoint CA 98.67% (kNN) Side viewpoint CA 95% (kNN)		Data fusion of accelerometer and vision data improved CA: • Vision back viewpoint with full body accelerometer CA 100% (kNN) Data fusion overcame viewpoint sensitivity • Vision trained on side viewpoint and tested on back viewpoint fused with full body accelerometer data CA 96.71% (kNN)
(Ramanathan et al., 2016)	Mean average precision	Hyperparameters chosen by cross-validating on the validation dataset	Event classification 0.516 mean average precision Event detection 0.435 mean average precision Key player attention 0.618 mean average precision	Manually labelled videos through an Amazon Mechanical Turk task	Event classification from isolated video clips was compared against different control setting and baseline models

Table 3.8 continued.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Reily et al., 2017)	CA, computational time, error rates (RMSE, average absolute), approach tested on CAD60 dataset benchmark		ID depth interest CA 97.8% Spin detection CA 93.81% Smoothing processing improved spin CA to 94.83%. Spin consistency performance analysis in comparison to ground truth RMSE 12.9942 ms from ground truth timestamp.	Manually labelled dataset	Study model reduces late stage data amount processing to perform calculations on 37.8% of the original data.
(Shah et al., 2007)	CA, confusion matrix		Forehand CA 97.24% Backhand CA 96.42% No stroke CA 98.02%	Manually labelled segment frames	Model computational performance speed was 20 fps
(Tora et al., 2017)	CA, Confusion matrix		Overall 49.2% CA		Model compared to several baseline models
(Victor et al., 2017)	F1-score, average frame distance, average distance to smoothed	80%/ 20% train-test dataset split	Swimming F1-score 0.922 Tennis F1-score 0.977	Manually labelled dataset by expert analysts	Experimented with how temporal information incorporated into the model, data input style, and three smoothing functions. Developed model tested and validated on tennis stroke dataset
(Yao & Fei-Fei, 2010)	CA, compared developed model to previous published benchmarks and a baseline measure (bag-of-words with a linear SVM)	60%/ 40% train-test dataset split	Activity CA 83.3%	Labelled training dataset	

Table 3.8 continued.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Zhu, Xu, Huang, & Gao, 2006)	Precision, recall		Tennis stroke classification using video frames: • Left recall 84.08%, • Left precision 89.80% • Right recall 90.20%, • Right precision 84.66%. Tennis stroke classification using action clips: • Left recall 87.50%, • Left precision 90.74% • Right recall 89.80%, • Right precision 86.27%		

CA classification accuracy, CNN convolutional neural network, DE detected events, DTE detected true events, GMM Gaussian mixture model, HMM Hidden Markov Model, kNN k-Nearest Neighbour, LOO-CV leave-one-out cross validation, LOSO-CV leave-one-subject-out cross validation, LS-SVM least squares support vector machine, NE number of events, RF random forests, ROC receiver operation characteristic curve, SVM Support Vector Machine.

3.5 Discussion

The aim of this systematic review was to evaluate the use of machine and deep learning for sport-specific movement recognition from IMUs and, or computer vision data inputs. Overall, the search yielded 52 studies, categorised as 29 which used IMUs, 22 vision-based and one study using both IMUs and vision. Automation or semi-automated sport movement recognition models working in near-real time is of particular interest to avoid the error, cost and time associated with manual methods. Evident in the literature, models are trending towards the potential to provide optimised objective assessments of athletic movement for technical and tactical evaluations. The majority of studies achieved favourable movement recognition results for the main characterising actions of a sport, with several studies exploring further applications such as an automated skill quality evaluation or judgement scoring, for example automated ski jump error evaluation (Brock et al., 2017).

Experimental set-up of IMU placement and numbers assigned per participant varied between sporting actions. The sensor attachment locations set by researchers appeared dependent upon the specific sporting conditions and movements, presumably to gain optimal signal data. Proper fixation and alignment of the sensor axes with limb anatomical axes is important in reducing signal error (Fong & Chan, 2010). The attachment site hence requires a biomechanical basis for accuracy of the movement being targeted to obtain reliable data. Single or multiple sensor use per person also impacts model development trade-off between accuracy, analysis complexity, and computational speed or demands. In tennis studies, specificity whilst using a single sensor was demonstrated by mounting the IMU on the wrist or forearm of the racquet arm (Connaghan et al., 2011; Kos & Kramberger, 2017; Srivastava et al., 2015; Whiteside et al., 2017). A single sensor may also be mounted in a low-profile manner on sporting equipment (Groh et al., 2015, 2017, 2016; Jensen et al., 2015).

Unobtrusive use of a single IMU to capture generalised movements across the whole body was demonstrated, with an IMU mounted on the posterior head in swimming (Jensen et al., 2013, 2016), lower back during running (Kobsar et al., 2014), and between the shoulder blades in rugby union (Kelly et al., 2012).

The majority of vision-based studies opted for a single camera set-up of RGB modality. Data output from a single camera as opposed to multiple minimises the volume of data to process, therefore reducing computational effort. However, detailed features may go uncaptured, particularly in team sport competition which consists of multiple individuals participating in the capture space at one time. In contrast, a multiple camera set-up reduces limitations including occlusion and viewpoint variations. However, this may also increase the complexity of the processing and model computational stages. Therefore, a trade-off between computational demands and movement recording accuracy often needs to be made. As stated earlier, the placement of cameras needs to suit the biomechanical nature of the targeted movement and the environment situated in. Common camera capture systems used in sports science research such as Vicon Nexus (Oxford, UK) and OptiTrack (Oregon, USA) were not present in this review. As this review targeted studies investigating during on-field or in-situation sporting contexts, efficiency in data collection is key for routine applications in training and competition. A simple portable RGB camera is easy to set-up in a dynamic and changing environment, such as different soccer pitches, rather than a multiple capture system such as Vicon that requires calibrated precision and are substantially more expensive.

Data acquisition and type from an IMU during analysis appears to influence model trade-off between accuracy and computational effort of performance. The use of accelerometer, gyroscope or magnetometer data may depend upon the movement properties analysed. Within tennis studies, gyroscope signals were the most efficient at discriminating between stroke types (Buthe et al., 2016; Kos & Kramberger, 2017) and detecting an athlete's fast feet court actions (Buthe et al., 2016). In contrast, accelerometer signals produced higher classification accuracies in classifying tennis stroke skills levels (Connaghan et al., 2011). The authors expected lower gyroscope classification accuracies as temporal orientation measures between skill levels of tennis strokes will differ (Connaghan et al., 2011). Conversely, data fusion from all three individual sensors resulted in a more superior model for classifying advanced, intermediate and novices tennis player strokes (Connaghan et al., 2011). Fusion of accelerometer and vision data also resulted in a higher classification accuracy for tennis stroke recognition (Conaire et al., 2010).

Supervised learning approaches were dominant across IMU and vision-based studies. This is a method which involves a labelled ground truth training dataset typically manually annotated by sport analysts. Labelled data instances were recorded as up to 15, 000 for vision-based (Victor et al., 2017) and 416, 737 for sensor-based (Rassem et al., 2017) studies. Generation of a training data set for supervised learning can be a tedious and labour-intensive task. It is further complicated if multiple sensors or cameras are incorporated for several targeted movements. A semi-supervised or unsupervised learning approach may be advantageous as data labelling is minimal or not required, potentially reducing human errors in annotation. An unsupervised approach could suit specific problems to explain key data features, via clustering (Mohammed et al., 2016; Sze et al., 2017). Results computed by an unsupervised model (Kos et al., 2016) for tennis serve, forehand and backhand stroke classification compared favourbaly well against a proposed supervised approach (Connaghan et al., 2011).

Recognition of sport-specific movements was primarily achieved using conventional machine learning approaches, however nine studies implemented deep learning algorithms.

It is expected that future model developments will progressively feature deep learning approaches due to development of better hardware, and the advantages of more efficient model learning on large data inputs (Sze et al., 2017). Convolutional Neural networks (CNN) (Y LeCun et al., 1998) were the core structure of five of the seven deep learning study models. Briefly, convolution applies several filters, known as kernels, to automatically extract features from raw data inputs. This process works under four key ideas to achieve optimised results: local connection, shared weights, pooling and applying several layers (Lecun et al., 2015; J. B. Yang et al., 2015). Machine learning classifiers modelled with generic hand-crafted features, were compared against a CNN for classifying nine beach volleyball actions using IMUs (Kautz et al., 2017). Unsatisfactory results were obtained from the machine learning model, and the CNN markedly achieved higher classification accuracies (Kautz et al., 2017). The CNN model produced the shortest overall computation times, requiring less computational effort on the same hardware (Kautz et al., 2017). Vision-based CNN models have also shown favourable results when compared to a machine learning study baseline (Karpathy et al., 2014; Nibali et al., 2017; Victor et al., 2017). Specifically, consistency between a swim stroke detection model for continuous videos in swimming which was then applied to tennis strokes with no domain-specific settings introduced (Victor et al., 2017). The authors of this training approach (Victor et al., 2017) anticipate that this could be applied to train separate models for other sports movement detection as the CNN model demonstrated the ability to learn to process continuous videos into a 1D signal with the signal peaks corresponding to arbitrary events. General human activity recognition using CNN have shown to be a superior approach over conventional machine learning algorithms using both IMUs (Ravi et al., 2017; J. B. Yang et al., 2015; Zebin et al., 2016; M. Zeng et al., 2014; Zheng et al., 2014) and computer vision (Ji et al., 2012; Krizhevsky et al., 2012; Lecun et al., 2015). As machine learning algorithms extract heuristic features requiring domain knowledge, this creates shallower features which can make it harder to infer high-level and context aware activities (Yang et al., 2015). Given the previously described advantages of deep learning algorithms which apply to CNN, and the recent results of deep learning, future model developments may benefit from exploring these methods in comparison to current bench mark models.

Model performance outcome metrics quantify and visualise the error rate between the predicted outcome and true measure. Comparatively, a kernel form of an SVM was the most common classifier implemented and produced the strongest machine learning approach model prediction accuracies across both IMU (Adelsberger & Tröster, 2013; Brock & Ohgi, 2017; Buthe et al., 2016; Groh et al., 2015, 2017, 2016; Jensen et al., 2016; Pernek et al., 2015; Salman et al., 2017; Schuldhaus et al., 2015; Whiteside et al., 2017) and vision-based study designs (Horton et al., 2014; Kasiri-Bidhendi et al., 2015; Kasiri et al., 2017; J. Li et al., 2018; Reily et al., 2017; Shah et al., 2007; Zhu, Xu, Huang, & Gao, 2006). Classification accuracy was the most common reported measure followed by confusion matrices, as ways to clearly present prediction results and derive further measures of performance. Further measures included sensitivity (also called recall), specificity and precision, whereby results closer to 1.0 indicate superior model performance, compared to 0.0 or poor model performance. The F1-score (also called a F-measure or F-score) conveys the balances between the precision and sensitivity of a model. An in-depth analysis performance metrics specific to human activity recognition is located elsewhere (Minnen et al., 2006; Ward et al., 2010). Use of specific evaluation methods depends upon the data type. Conventional performance measures of error rate are generally unsuitable for models developed from skewed training data (Provost & Fawcett, 2001). Using conventional performance measures in this context will only take the default decision threshold on a model trained, if there is an uneven class distribution this may lead to imprecision (Provost & Fawcett, 2001; Seiffert et al., 2008). Alternative evaluators including Receiver Operating Characteristics (ROC) curves and its single numeric measure, Area Under ROC Curve (AUC), report model performances across all decision thresholds (Seiffert et al., 2008). Making evaluations between study methodology have inherent complications due to each formulating their own experimental parameter settings, feature vectors and training algorithms for movement recognition. The No-Free-Lunch theorems are important deductions in the formation of models for supervised machine learning (Wolpert, 1996), and search and optimisation algorithms (Wolpert & Macready, 1997). The theorems broadly reference that there is no 'one model' that will perform optimally across all recognition problems. Therefore, experiments with multiple model development methods for a particular problem is recommended. The use of prior knowledge about the task should be implemented to adapt the model input and model parameters in order to improve overall model success (Shalev-Shwartz & Ben-David, 2014).

Acquisition of athlete specific information, including statistics on number, type and intensity of actions, may be of use in the monitoring of athlete load. Other potential applications include personalised movement technique analysis (Reilly et al., 2017b), automated performance evaluation scoring (Reily et al., 2017) and team ball sports pass quality rating (Horton et al., 2014). However, one challenge lies in delivering consistent, individualised models across team field sports that are dynamic in nature. For example, classification of soccer shots and passes showed a decline in model performance accuracy from a closed environment to a dynamic match setting (Schuldhaus et al., 2015). A method to overcome accuracy limitations in dynamic team field sports associated with solely using IMUs or vision may be to implement data fusion (Conaire et al., 2010). Furthermore, vision and deep learning approaches have demonstrated the ability to track and classify team sport collective court activities and individual player specific movements in volleyball (Ibrahim et al., 1971), basketball (Ramanathan et al., 2016) and ice hockey (Tora et al., 2017). Accounting for methods from experimental set-up to model evaluation, previous reported models should be considered and adapted based on the current problem. Furthermore, the balance between model computational efficiency, results accuracy and complexity trade-offs calculations are an important factor.

In the present study, meta-analysis was considered however variability across developed model parameter reporting and evaluation methods did not allow for this to be undertaken. As this field expands and further methodological approaches are investigated, it would be practical to review analysis approaches both within and between sports. This review was delimited to machine and deep learning approaches to sport movement detection and recognition. However, statistical or parametric approaches not considered here such as discriminative functional analysis may also show efficacy for sport-specific movement recognition. However, as the field of machine learning is a rapidly developing area shown to produce superior results, a review encompassing all possible other methods may have complicated the reporting. Since sport-specific movements and their environments alter the data acquisition and analysis, the sports and movements reported in the present study provide an overview of the current field implementations.

3.6. Conclusions

This systematic review reported on the literature using machine and deep learning methods to automate sport-specific movement recognition. In addressing the research questions, both IMUs and computer vision have demonstrated capacity in improving the information gained from sport movement and skill recognition for performance analysis. A range of methods for model development were used across the reviewed studies producing varying results.

Conventional machine learning algorithms such as Support Vector Machines and Neural Networks were most commonly implemented. Yet in those studies which applied deep learning algorithms such as Convolutional Neural Networks, these methods outperformed the machine learning algorithms in comparison. Typically, the models were evaluated using a leave-one-out cross validation method and reported model performances as a classification accuracy score. Intuitively, the adaptation of experimental set-up, data processing, and recognition methods used are best considered in relation to the characteristics of the sport and targeted movement(s). Consulting current models within or similar to the targeted sport and movement is of benefit to address benchmark model performances and identify areas for improvement. The application within the sporting domain of machine learning and automated sport analysis coding for consistent uniform usage appears currently a challenging prospect, considering the dynamic nature, equipment restrictions and varying environments arising in different sports.

Future work may look to adopt, adapt and expand on current models associated with a specific sports movement to work towards flexible models for mainstream analysis implementation. Investigation of deep learning methods in comparison to conventional machine learning algorithms would be of particular interest to evaluate if the trend of superior performances is beneficial for sport-specific movement recognition. Analysis as to whether IMUs and vision alone or together yield enhanced results in relation to a specific sport and its implementation efficiency would also be of value. In consideration of the reported study information, this review can assist future researchers in broadening investigative approaches for sports performance analysis as a potential to enhancing upon current methods.

Chapter Four: Inertial measurement units and machine learning for sport movement recognition: an update of the published article presented in Chapter Three

Literature review studies on the applications of IMUs for sport performance analysis (Camomilla et al., 2018; Chambers et al., 2015; Santos-Gago et al., 2019) provide evidence for the current capabilities, potential limitations and areas for future development. Sport-specific reviews on the uses of IMUs for performance analysis and movement recognition include combat sports (Worsey et al., 2019b), rowing (Worsey et al., 2019a), swimming (Mooney et al., 2015), and wheelchair court sports (Shepherd et al., 2018). These reviews inform of the current trends, specific guidelines, and future improvements relevant to the targeted sport. A detailed review of the use of IMUs in sport movement detection and recognition is presented in Chapter Three of this thesis. Since the original manuscript publication date, 17 relevant studies have been published. This chapter provides an update to the publication in Chapter Three using the same search strategy for the dates between June 2019 to December 2019. The collated information is summarised in Tables 4.1, 4.2 and 4.3, as per the same format used in Chapter Three.

4.1 Sport-specific movement recognition: an update

IMUs have been applied across several sporting domains and demonstrated capability as a wearable technology for sport science athlete monitoring practices. The IMU sensor placement and number of sensors have varied for research methodologies in sport applications, as detailed in Chapter Three and Table 4.1. As an example, in cross-country skiing and jump technique assessment, several approaches have been used. Cross-country skiing technique variations and sub-techniques have been classified using IMU set-up configurations of 17 sensors (Jang et al., 2018), two sensors (Rindal et al., 2018), and a single sensor (Rassem et al., 2017). Different sensor data input combinations from the 17 IMU set-up in Jang et al. (2018) were fused and fed into a CNN-LSTM model. The findings indicated that the highest mean classification accuracy (95%) was achieved using the five sensor data fusion from both hands, both feet, and the pelvis (Jang et al., 2018). The two-sensor configuration in Rindal et al (2018) achieved a classification accuracy of 94% using a NN model. Lastly, a classification error value of 1.6% was achieved in Rassem et al. (2017) using a standard LSTM model with data input from one IMU sensor. In detecting and classifying ski jump errors, a nine IMU set-up was implemented (Brock et al., 2017; Brock & Ohgi, 2017). Binary classification of either a jump with error or no error using a SVM model achieved a classification accuracy of 82% (Brock & Ohgi, 2017). Whereas the recognition of nine ski jump motion errors using a CNN model achieved a classification accuracy of 93% (Brock et al., 2017). Relating back to the IMU device considerations from the literature in section 2.3, investigation is required to devise an IMU set-up that suits the sporting constraints and analysis application purpose. Just as there is no one learning algorithm model to suit all problems and datasets (section 2.4.1), IMU specifications and the data collection set-up domain used in one sporting context may not work well in another.

Four of the updated literature search studies used deep learning algorithms for sport-specific movement recognition. Two implemented CNN based models (Jiao, Bie, et al., 2018; Soro et al., 2019), two used LSTM based model (Anand, Sharma, Srivastava, Kaligounder, & Prakash, 2017; Zhang et al., 2018), and one study ran a CNN-LSTM model (Jang et al., 2018). Although none of these studies compared their results against machine learning based models as undertaken in ski jump error recognition (Brock et al., 2017) and volleyball action classification (Kautz et al., 2017). The model comparisons in these studies resulted in the deep learning models showing higher classification accuracies. Differentiating between sub-techniques of a motion may benefit from the higher-level feature extraction methods used in deep learning models for improved feature representations of the sub-technique intricates (Nweke et al., 2018). For example, differentiating with data obtained from ski jumping which can have marginal differences between what would be considered a quality jump versus a substandard performance (Brock et al., 2017). Deep learning models may also benefit analysis with data is collected in an uncontrolled sport field setting that will likely contain noise or missing data (Nweke et al., 2018). Here, the deep learning models have the capability to learn from the underlying data features rather than the shallow hand-crafted feature extraction in machine learning models that could be confounded by any imperfections in the data (Nweke et al., 2018; Ravi et al., 2017).

When looking at AF specifically, the use of IMUs and computer vision in the context of skill specific movement recognition is very limited. Given the current favourable findings from similar field team ball sports in Chapters Three and Four, there may be applications for IMUs in AF skill performance analysis. The introduction of computer vision in Chapter Two section 2.4.7 and subsequent literature analysis in Chapter Three provides important context for vision data as either an alternative or integration method to enhance sport movement recognition where IMUs may have disadvantages for the intended application. As highlighted in Chapter Two, computer vision was previously tested in AF for athlete tracking in matches (Faulkner & Dick, 2015). Although the results were not favourable at the time and indicated the limitations of using static captured vision alone in AF. Investigating the use of IMUs for AF skill recognition may provide an improved alternative or present scope for an integrated system of IMU and computer vision technologies.

With the recent formation of the AFLW competition, the knowledge of the key team and individual player skills involved in match play are relatively restricted at the current time with lack of historical data. Determining the relationships and importance of skills towards team match success in the AFLW will increase the understandings of the women's league athlete performances and provide useful evidence for how IMUs could be focussed towards priority areas in skill analysis.

Table 4.1. Inertial measurement unit specifications.

Reference	Sensor model	Sensor No.	Sensor placement	Accelerometer			Gyroscope			Magnetometer		
				Axes	Range	Sample rate	Axes	Range	Sample rate	Axes	Range (1 Ga = 100 µT)	Sample rate
(Abdullah et al., 2020)	Custom sensor	1	Underneath of skateboard fixed behind front truck	3	NR	20 Hz	3	NR	20 Hz	NR	NR	NR
(Chambers et al., 2018)	Catapult S5 OptimEye (Melbourne, Australia)	1	Between shoulder blades in manufacture made vest	3	± 16 *g*	100 Hz	3	± 2000 °/s	100 Hz	3	± 4900 µT	100 Hz
(Chambers et al., 2019)	Catapult S5 OptimEye (Melbourne, Australia)	1	Between shoulder blades in manufacture made vest	3	± 16 *g*	100 Hz	3	± 2000 °/s	100 Hz	3	± 4900 µT	100 Hz
(Derungs et al., 2018)	MVN Link IMS (Xsens Technologies B.V., Enschede, Netherlands)	14	Right and left: wrists, upper arms, bridge of feet, shins, thighs. Upper back, lower back. Below each Nordic pole handles.	3	± 16 *g*	50 Hz	3	± 2000 °/s	50 Hz	3	± 1.9 Gauss full-scale	50 Hz
(Ebner & Findling, 2019)	Xsens MTw-38A70G20 sensors	2	Throat of racket in a custom 3D printed holder and on the wrist of the racket hand	3	± 16 *g*	100 Hz	3	± 2000 °/s	100 Hz	3	± 1.9 Gauss full-scale	100 Hz
(Harding et al., 2008)	IMU (not specified)	1	Lower back at 5 cm left of the spine	3	NR	NR	NR	NR	NR	NR	NR	NR

Table 4.1. (Continued).

Reference	Sensor model	Sensor No.	Sensor placement	Accelerometer			Gyroscope			Magnetometer		
				Axes	Range	Sample rate	Axes	Range	Sample rate	Axes	Range (1 Ga = 100 µT)	Sample rate
(Holzemann & Van Laerhoven, 2018)	Custom sensor built with a MPU-9250 by Invensense	1	Wrist of dominant hand	3	± 16 *g*	25 Hz	3	± 2000 °/s	25 Hz	3	± 0.6 Gauss full-scale	25 Hz
(Jang et al., 2018)	MVN Link IMS (Xsens Technologies B.V., Enschede, Netherlands)	17	Manufacture's specified locations for body suit sensor placements	3	± 16 *g*	240 Hz	3	± 2000 °/s	240 Hz	3	± 1.9 Gauss full-scale	240 Hz
(Jiao, Bie, et al., 2018)	A micro- electro-mechanical (MEMS) accelerometer sensor and a MEMS gyroscope sensor	2	Gold club shaft	3	NR	500 Hz	3	NR	500 Hz	NR	NR	NR
(Ma et al., 2018)	IMU (not specified)	1	Wrist	NR	NR	NR	NR	NR	NR	NR	NR	NR
(McGrath et al., 2019)	OneSABEL Sense IMU (SABEL Labs, Australia)	1	Between shoulder blades in manufacture made vest	3	± 16 *g*	250 Hz	3	± 2000 °/s	250 Hz	3	± 1200 µT	250 Hz

Table 4.1. (Continued).

Reference	Sensor model	Sensor No.	Sensor placement	Accelerometer			Gyroscope			Magnetometer		
				Axes	Range	Sample rate	Axes	Range	Sample rate	Axes	Range (1 Ga = 100 μT)	Sample rate
(Op De Beéck et al., 2018)	Shimmer3 (Realtime Technologies Lab. Dublin, Ireland)	6	Both left and right: shin bone (anteromedial aspect for the distal tibia), wrist (dorsal carpal ligament) and arm (mid-point between the acromial and the radial, on the mid-line of the lateral surface of the arm)	3	NR	1024 Hz	3	NR	1024 Hz	NR	NR	NR
(Peng et al., 2018)	Custom sensor	1	Wrist	3	NR	NR	NR	NR	NR	NR	NR	NR
(Soro et al., 2019)	Huawei Watch 2 smartwatch with Android Wear 2.9 (Huawei, Shenzhen, China)	2	Right ankle and right wrist	3	NR	100 Hz	3	NR	100 Hz	3	NR	100 Hz
(Uddin Ahamed et al., 2019)	Lumo Run® (Lumo Bodytech Inc., Mountain View, CA, USA)	1	Posterior aspect of either the runner’s waistband or a running belt	3	NR	100 Hz	3	NR	100 Hz	3	NR	100Hz
(Wang et al., 2018)	Custom sensor with MPU9250 (TDK InvenSense, USA) chip.	1	Wrist	3	± 16 *g*	NR	3	NR	NR	NR	NR	NR

Table 4.1. (Continued).

Reference	Sensor model	Sensor No.	Sensor placement	Accelerometer			Gyroscope			Magnetometer		
				Axes	Range	Sample rate	Axes	Range	Sample rate	Axes	Range (1 Ga = 100 µT)	Sample rate
(Zhang et al., 2018)	mbientlab wearable motion tracking device (MBIENTLAB INC, San Francisco, CA)	1	Wrist	3	NR	100 Hz	3	NR	100 Hz	3	NR	100 Hz

g g-forces, Ga gauss, Hz Hertz, IMU inertial measurement unit, µT micro Tesla.

NR not reported: study either did not directly report the specification or the device did not include the sensor type

Table 4.2. Inertial measurement unit study description and model characteristics.

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Abdullah et al., 2020)	Skateboarding: five trick types – Ollie, Nollie FS, Shuvit, Frontside 180, Pop shove-it, Kickflip	One: experienced skateboarder	20 successful tricks				36 features (6 features, 6 axes): mean, skewness, kurtosis, peak to peak, root mean square, STD		SVM kNN ANN LR NB RF
(Chambers et al., 2018)	Rugby Union: scrum events	30: elite forwards	1057 scrum events			Criterion based measures about the sensor orientation. A Window over identified events using the mid-point as the event timestamp.		11 signal features selected through variable selection using RF.	RF

Table 4.2. (Continued).

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Chambers et al., 2019)	Rugby Union: ruck and tackle events	One team of elite male player	120 tackle events and 125 ruck events		Resultant magnitude of accelerometer data was identified. Smoothing using a low-pass 4th order Butterworth filter with a 25 Hz cut-off frequency. Movement profiles were clustered using GMM over one-second windows and classified using DTW methods.	Synchronised video and sensor data	2-second sliding window for all files and calculate relevant variables and descriptive feature sets to characterise rucks, tackles and other movements. Features: max, min, mean, variance, kurtosis, skewness, spectral bandwidth, spectral centroid, magnitude.		RF

Table 4.2. (Continued).

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Derungs et al., 2018)	Nordic walking: technique mistakes in beginners	10: beginner Nordic walkers	11247 Nordic walk strides	Low-pass filter with cut-off frequency calculated using a Fast Fourier Transformation for each participant	Hill-climb algorithm used to segment individual strides.		Feature vector of 351 time and frequency domain features from one IMU for each stride	PCA and GDB Decision Fusion used to maximise skill estimation accuracy by combining skill grades derived via PCA and GDB.	Bayesian Ridge Regression OLS-LR SVR AdaBoostR
(Ebner & Findling, 2019)	Tennis: eleven tennis stroke types	Six: semi-professional tennis players.	250 instances			Stroke segmentation with peak detection on derivative averaged accelerometer data. 1 sec window applied (0.5 sec each peak side).	FS1) same features as (Whiteside et al., 2017). FS2) extended FS1 including STD, skewness, kurtosis, interquartile range, frequency bands, zero and mean crossing rate. FS3) raw accelerometer and gyroscope values	FS1) no selection method. FS2) mutual info classifier. FS3) PCA	SVM linear SVM rbf KNN CART

Table 4.2. (Continued).

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Harding et al., 2008)	Snowboarding : movement performance assessment variables	Four: Australian half-pipe snowboarding team members	92 aerial manoeuvres		Data segmented into individual athletes runs	Sliding FFT window, power analysis, average power levels using a threshold-based algorithm.			
(Holzemann & Van Laerhoven, 2018)	Basketball: low dribble, crossover dribble, high dribble, jump shot	Three: male experienced players	235000 data points in total			Sliding window size of second over data	Arithmetic mean and the standard deviation for every axis of the acceleration data		KNN RF
(Jang et al., 2018)	Cross-country Skiing: eight ski techniques	Four: professional XC skiing athletes	24 train set files and 9 test set files.	Low-pass Butterworth filter Ski cycle detection filter data by resampling of cycles in original data.	Arrangement of cycles into tensors.				CNN-LSTM

Table 4.2. (Continued).

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Jiao, Bie, et al., 2018)	Golf: swing type	Four: professional and amateur golf players	213 golf swings		Data augmentation, shuffling and standardisation (z-score normalisation).				Four CNN-based models
(Ma et al., 2018)	Basketball: nine movement types	One player	10 repetitions of the nine movements		Derive the y-axis acceleration and z-axis angular velocity.		50 features per basketball move: STD, absolute values, max, min, integral		Two-layer FF-NN
(McGrath et al., 2019)	Cricket: fast bowling	17 male fast bowlers from Auckland cricket premier competition	522 bowls, 102 non-bowling throws	Fourth order 1Hz Butterworth low pass filter for baseline removal.	Data centred and normalised for all models except RF.	Magnitude of gyroscope axes calculated for peaks greater than 500 ◦/s to determine bowl events in a 10 sec window.	282 time and frequency domain features: mean, STD, max, skewness, kurtosis, frequency amplitude, frequency, energy, correlation between axes, position of max and min, magnitude vectors of each axis	Highly correlated features (r > 0.95) removed leaving 223 features	RF, Linear SVM, Polynomial SVM, NN, GBA

Table 4.2. (Continued).

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Op De Beéck et al., 2018)	Running: fatigue	29 runners	98 trials		Construct examples by dividing collected sensor signals into non-overlapping 10 sec windows. Standardised every signal in example to account for different pacing strategies and varied run speeds.		• 15 statistical features, • Sport science features: sample entropy, detrended fluctuation analysis, stride regularity • One symmetry feature		GBRT, NN, LR with EN, LR with LASSO
(Peng et al., 2018)	Volleyball: spike movement components	One: elite volleyball player	NR		Augmented training sample size from 12 to 96 trials.		Eight features extracted from accelerometer z-axis.	RF and python panda correlation matrix: two features removed due to minimal impact.	NN
(Soro et al., 2019)	CrossFit: ten common movements	54: beginner, intermediate, and advanced CrossFit athletes	5461 total reps across all movements						CNN

Table 4.2. (Continued).

Reference	Sport: target movement(s)	Participants Number: gender, level	Dataset sample No.	Data pre-processing			Feature extraction	Feature selection	Recognition algorithm
				Filter	Processing	Detection			
(Wang et al., 2018)	Volleyball: straight-ahead spike	Ten male players: three amateurs, three sub-elite, four elite.	100 repetitions		Three-point moving average across data		12 statistical and three morphological	PCA	SVM
(Uddin Ahamed et al., 2019)	Running: six gait variables	11: recreational runners training for a half-marathon, 10 female and 1 male.	Test day 1: average 4.1 km (1996 strides) each runner. Test day 2: average 3.7 km (1832 strides) each runner.						RF
(Zhang et al., 2018)	Tennis: forehand topspin, forehand slice, backhand topspin, backhand slice, serve	Ten players: two professional and eight amateurs	250: 50 repetitions of each action		Action segmentation through highlight sensing using sliding windows.	27-dimensional statistical feature vector			LSTM

AdaBoostR AdaBoost.R2 algorithm, BLSTM Bi-directional Long Short-Term Memory network, CNN-LSTM convolutional neural network with long short-term memory, DTW dynamic time warping, EN Elastic Net, FS feature set, FF-NN feed-forward neural network, FFT fast Fourier transform, GBA gradient boosting algorithm, GDB gradient descent boosting, GBRT gradient boosted regression trees, GMM gaussian mixture models, kNN k-Nearest Neighbour, LASSO least absolute shrinkage and selection operator regularisation, LR linear regression, LSTM long short-term memory, NB Naïve Bayesian, NN neural network, OLS-LR ordinary least square linear regression, PCA principal component analysis, RBF radial basis function, SVM support vector machine, SVR support vector regression.

Table 4.3. Inertial measurement unit study model performance evaluation characteristics.

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Abdullah et al., 2020)	CA, AUC, precision, recall, F1-score	LOO-CV	LR CA 95% NB CA 95% LR F1-score 0.95 NB F1-score 0.95		
(Chambers et al., 2018)	CA, mean CA, precision, specificity, sensitivity, ROC	10-fold cross-validation	All data model CA 91% Training session data model CA 87.6% Match data model CA 93.6%	Manual timestamp labelling of instances from video data	The orientation of the inertial sensor was estimated using a proprietary sensor fusion algorithm that included accelerometer and gyroscope data
(Chambers et al., 2019)		10-fold cross validation	RF classification results indicated that all rucks and tackles were correctly identified during match-play when 79.4 ± 9.2% and 81.0 ± 9.3% of the RF decision trees agreed with the video-based determination for these events.	Separate 177 data files with synchronised video data from the same cohort during eight international matches.	
(Derungs et al., 2018)	RMSE, MAE, Pearson correlation coefficient	LOO-CV	All three Nordic ski mistakes can be estimated with a normalised RMSE of 24.15% across all participants.	Video recording. Two professional ski cross athletes analysing scoring movements.	

Table 4.3. (Continued).

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Ebner & Findling, 2019)	CA, confusion matrix	Training partition: 10-fold cross validation with a hyperparameter grid search for each model type. Dataset split: 80% train/ 20% test	Detection across all stroke types • 98.47% wrist • 99.44% racket Classification: • Mean across all wrist models 93.4% and racket models 92.5% • Linear SVM using raw data PCA feature selection: 99.2% wrist model and 98.8% racket model	Video recording	
(Harding et al., 2008)	Detection accuracy		Two pass signal processing technique presented within this paper was able to detect 100 % of the aerial acrobatic manoeuvres performed	Video recording	
(Holzemann & Van Laerhoven, 2018)	CA, precision, recall	LOO-CV	Overall CA performance: • KNN 83.6% • RF 87.5%	Video recording labelling	
(Jang et al., 2018)	CA	LOO-CV	Test set 1 overall mean CA 87.2% Test set 2 overall mean CA 91.2% Test set 1 KNN CA 68.8% Test set 2 KNN CA 78.0%	Video frame labelling by professional cross-country skiing athletes	Five combinations of sensor data inputs tested.
(Jiao, Bie, et al., 2018)	F1-score, precision-recall curves, CA, baseline SVM model	10-fold cross validation on train set to select hyperparameters and models. Dataset split: 2/3 train set 1/3 test set	All CNN models achieved CA of > 90% GolfInception model: • Precision 0.98 • Recall 0.97 • F1-score 0.97		Data standardisation was employed to alleviate the domination of sequences from one individual sensor in the training phase.

Table 4.3. (Continued).

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Ma et al., 2018)	CA, confusion matrix	Dataset split: 70% test/15% validation/ 15% test	CA 98.9%	Video recording	
(McGrath et al., 2019)	Sensitivity, specificity, CA, F1-score, ROC, confusion matrix	10-fold cross validation used to determine optimal model tuning parameters	Overall CA across models => 95%	Ball velocities measured with Stalker radar gun (Radar Scales, Minneapolis, US) at 250 Hz	
(Op De Beéck et al., 2018)	MAE	LOO-CV	GBRT model with wrist data MAE 1.89. GBRT model with sensor data fusion wrist, left and right tibia MAE 1.74.		Four regression algorithms, sensor location, sensor data fusion, data processing, and feature extraction methods investigated.
(Peng et al., 2018)		Dataset split: 75% train/ 15% validation/ 15% test	Total recognition correct rate 89.6%	High-speed camera	
(Soro et al., 2019)	Accuracy %, MAE, MRE	5-fold cross-validation for recognition of performed exercises. LOO-CV for repetition counting evaluations.	Exercise recognition in constrained workout 99.96%. Exercise repetition counting in constrained workout MAE 0.7 reps per set and MRE 6.1%	Participant worn watch programmed to indicate start of repetitions and type; used as data labelling method.	
(Wang et al., 2018)	CA	5-fold cross-validation Dataset split: 70% train/ 30% test.	SVM and PCA CA 94% SVM CA 90% NN CA 90% NB CA 84%	High-speed camera	

Table 4.3. (Continued).

Reference	Evaluation	Cross validation or dataset split approach	Performance	Ground truth	Special remarks
(Uddin Ahamed et al., 2019)	CA, variable importance %	Subject-specific approach: one-against-another-subject CV Group-based approach: LOO-CV	Mean classification accuracy: • Subject-specific model 86.29% • Group-based model 76.17%		
(Zhang et al., 2018)	CA, confusion matrix, processing time		MV-Sports achieves 99.6% action segmentation CA. Average CA 98.01% on player action recognition. Average CA 91.20% for actions performed by users the model had not been trained on.	Video cameras calibrated to tennis area	

AUC area under curve, BLSTM bi-directional long short-term memory network, CA classification accuracy, CNN convolutional neural network, GBRT gradient boosted regression trees, kNN k-Nearest Neighbour, LOO-CV leave-one-out cross validation, LR linear regression, MAE mean absolute error, MRE mean relative error, NB Naïve Bayesian, NN neural network, PCA principal component analysis, RF random forest, RMSE root mean square error, ROC receiver operation characteristic curve, SVM support vector machine.

Chapter Five: The relationship of team and individual athlete performances on match quarter outcome in elite women's Australian Rules football

Manuscript copied and adapted from the original published article**: Cust, E. E.,** Sweeting, A. J., Ball, K., Anderson, H., & Robertson, S. (2019). The relationship of team and individual athlete performances on match quarter outcome in elite women's Australian Rules football. Journal of Science and Medicine in Sport, 22(10). https://doi.org/10.1016/j.jsams.2019.05.004 (Appendix C).

5.1 Abstract

To evaluate the relationships between the athlete distribution of team performance indicators and quarter outcome in elite women's Australian Rules football matches. Thirteen performance indicators were obtained from 56 matches across the 2017 and 2018 Australian Football League Women's (AFLW) seasons. Absolute and relative values of 13 performance indicators were obtained for each athlete, in each quarter of all matches. Eleven features were further extracted for each performance indicator, resulting in a total of 169 features. Generalised estimating equations (GEE) and regression decision trees were run across the different feature sets and dependent variables, resulting in 22 separate models. The GEE algorithm produced slightly lower mean absolute errors across all dependent variables and feature sets comparative to the regression decision tree models. Quarter outcome was more accurately explained when considered as total points scored comparative to quarter score margin. Team differential and the 75th percentile of individual athlete inside 50s were the strongest features included in the models. Modelling performance statistics by quarter outcomes provides specific practical information for in-game tactics and coaching in relation

to athlete performances each quarter. Within the current elite women's Australian Rules football competition, key high performing individual athletes' skilled performances within matches contribute more to success rather than a collective team effort.

5.2 Introduction

Match performance analysis in team sports can provide a greater understanding of the physical, technical and tactical characteristics athletes require to produce a successful competition outcome (Hughes & Bartlett, 2002). Analysis may help guide coaching staff on training practices that replicate and prepare athletes for the demands of competition (Pinder et al., 2011). Determining the form and function of events within the specifics of a sport for teams and individual athletes should inform the variables for quantification of performance and therefore the sport analytics approaches used to facilitate future coaching practise (Hughes, 2015). The relationship between match athlete performance indicators (Hughes & Bartlett, 2002) in Australian Rules football (AF) have been investigated heavily in the literature across elite male teams (McIntosh et al., 2018a; Robertson, Back, et al., 2016; Stewart et al., 2007), individual athlete contributions (Robertson, Gupta, et al., 2016), and recently, elite women's teams (Black et al., 2018a).

In 2017, AF established a national elite women's competition, the Australian Football League Women's (AFLW) in addition to the long running elite men's Australian Football League (AFL). For the purposes of this article the two competitions will be referred to as AFLW for elite women's and AFL for elite men's. The opening two seasons consisted of a seven-round home-and-away competition, incorporating eight teams. As the depth of talent and resources develop, the league has set plans for expansion to the competition. This in turn will provide further opportunities to investigate elite women's football training and match physical, technical and tactical areas. For example, information on athlete match demands may improve club training practices, assess the effectiveness of the rule changes implemented differently to the AFL competition, and inform league directors on the quality of development in the competition.

Research in women's AF is currently limited (Black et al., 2018a, 2018b; Clarke et al., 2018). Recent research on the physical demands, technical performances and activity profiling across field playing positions of match-play in AFLW (Clarke et al., 2018) has provided initial insights into match activity. There were no absolute differences between physical variables, based on match playing position, in the AFLW (Clarke et al., 2018). Furthermore, no positional group differences were noted for skill measures such as total kicks, handballs, contested possessions, uncontested possessions, and marks (Clarke et al., 2018). This is in contrast to several physical demands characteristic differences that have been observed across athlete match positions in the AFL (Boyd et al., 2013; Gray & Jenkins, 2010). The specificity of AFLW positional roles may not yet be established and consequently, athletes may be more homogenous in playing tactics and physical abilities comparative to AFL players (Clarke et al., 2018). Although there are inherent differences between the AFL and AFLW games such as amount of time and players on ground creating independent constraints between each competition. Currently focussing on the AFLW as an independent competition and quantifying match variables as the league matures may be more beneficial over a direct sport analytics comparison of the AFL and AFLW given the current game constraint differences. Match performance indicator analysis assessed the relationship between team skill involvements and match outcome in the first season of AFLW (Black et al., 2018a). Match outcome, defined as win/ loss and score margin, indicated that higher uncontested possessions and inside 50: goal score ratio were the strongest predictors for winning. Increased kick numbers and contested marks resulted in a higher team ladder position (Black et al., 2018a).

Match success in the AFL has been linked to individual athlete skill efficiency rather than their physical activity profile (Sullivan et al., 2014b). Specifically, physical activity profiles may increase, yet skill involvements efficiency may decrease when teams lose a quarter (Sullivan et al., 2014a). An analysis inclusive of athlete skilled match performances, by individual match quarter and across feature derived performance distributions, is yet to be investigated in AF. A quarter by quarter approach could provide differentiated information about specific technical and tactical foci for coaches. Situational variables such as starting quarter score, quality of opposition, and whether the team is playing at a home or away ground have shown influence on elite women's team sport quarter performances (Gómez et al., 2014; Moreno et al., 2013). Analysing by quarter could improve relevancy of results, given output may fluctuate across quarters for several reasons (Sampaio et al., 2010). During quarter time breaks, coaches can address athletes directly. Knowledge or information transfer from the coach to the playing group should be of purpose, work in context of the current events and tie in with previously delivered knowledge the coach has provided prior to the match to maximise group understandings of the information (Joshi et al., 2007). Factors may affect the extent of knowledge transfer to the playing group between the restricted quarter time frame such as the coach's communication style, clarity of information, and a player's prior involvement in the match strategy system development (Joshi et al., 2007).

Quantifiable information about skill performances, in context of the match, could further justify changes to team playing strategies based on the current situation. With respect to influence on the team match outcome, quantification of individual athlete distributions have been linked to successful match outcome (Robertson, Gupta, et al., 2016). Specifically, lower 75th, 90th and 95th percentile values for team goals and higher 25th and 50th percentile values for disposals (Robertson, Gupta, et al., 2016). Measured athlete performance distribution information calculated by individuals rather than a team data as a whole could determine the influential basis for match success in the AFLW. Information may also convey whether success in the current AFLW game constitutes a more collective team-based effort or skewed to a few stronger individual athletes. Findings may inform match team selection to suit the current game style influence or opposition at play. This may be important as several new teams are introduced to the competition over the next few years making key athlete retention or attainment a challenge.

The primary aim of this study was to evaluate the relationship of AFLW athlete skill performance indicator distributions, to explain match quarter outcomes during the 2017 and 2018 seasons. Secondly, this study aimed to compare quarter outcome model error rates from separate machine learning approaches, based on the varied input feature set variables.

5.3 Methods

All match performance indicators were obtained from the AFL match statistics provider, Champion Data Pty Ltd. (Melbourne, Australia) online portal, Coaches Information Analysis (CIA). Data collection by Champion Data involves human recordings of the statistics by working at each match, as such the inter- and intra-reliability of the data is currently unknown (McIntosh et al., 2018b; Robertson, Gupta, et al., 2016). Reliability and validity of the data has been assessed independently to determine the agreement between the Champion Data and author-coded values (Robertson, Gupta, et al., 2016). Reliability assessment showed very high agreement levels, intra-class correlation coefficient range

0.947 – 1.000. The validity of author's coding showed low absolute error in regards to the Champion Data, RMSE range 0.0 – 4.5, (Robertson, Gupta, et al., 2016) indicating the expected absolute error points between each performance indicator for each game. A total of 56 matches across the 2017 and 2018 AFLW season were obtained and 13 discrete performance indicators were selected (Black et al., 2018a; Robertson, Back, et al., 2016; Robertson, Gupta, et al., 2016). The definitions for each indicator are provided in Appendix A. Absolute values from every quarter (n = 224), match (n = 56), athlete (n = 154), and all teams (n = 7), across performance indicators, were extracted into a custom ExcelTM spreadsheet. Quarter outcome (as win = 1 or loss = 0 or draw = 2), quarter score margin (points), match outcome (win/ loss) and match score margin (points) were recorded. Score points were recorded as both their absolute values and relative values to the opposition at play. The University's Human Research Ethics Committee approved the study (application number 0000025654).

Each athlete's contribution to their team's total were converted to a relative form, as a percentage of their team total for each match (Robertson, Gupta, et al., 2016). Features extracted for each performance indicator were the minimum, maximum, mean, standard deviation and percentiles, at 0.05, 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95, resulting in a total of 143 features (11 features x 13 performance indicators) (Robertson, Gupta, et al., 2016). Features were collated with team name, round number (1 – 7), season (2017 or 2018), quarter number (1 – 4), quarter outcomes (loss, win or draw), and match outcome (loss, win or draw). The stability of the data performance profiles (Hughes, 2017) was plotted and assessed by visual inspection, and deemed acceptable to model for comparison of analysis methods and reporting of results for practical feedback.

A total of 22 models were developed. Modelling of statistics by machine learning was performed for quarter points scored (absolute), and quarter point margin relative to the opposition (relative). Four features sets were used in separate models: total performance indicator values (n = 13), performance indicator values relative (n = 13) to the opposition, derived feature distribution values for each performance indicator (n = 143), combined performance indicator total, relative and feature distribution values (n = 169).

Regression decision trees were computed with Python version 3.6.6 (Python Software Foundation, 2018), using the package Scikit-learn (Pedregosa et al., 2011). Data was split into a 70% training set and 30% testing set. Each regressor tree was computed with a minimum sample split of 30 (> 13.4% of total sample) and a maximum depth of five. Several model parameter combinations were tested to reduce the risk of overfitting whilst minimising error (Hawkins, 2004). Regression trees were also computed using the whole training set for the four feature sets as a comparison. Generalised Estimating Equations (GEE) were also constructed separately in R (R Core Team, 2018) for each dependent variable and feature sets. Team (n = 7) was considered a fixed repeated measure and a greedy feature selection was implemented for feature selection in model construction. Model evaluation was based on the mean absolute error (MAE) computed from the withheld testing set, unless otherwise indicated.

5.4 Results

The MAE results for each model are presented in Table 5.1. The GEE produced lower MAE's than the decision trees (Table 5.1). Across both analysis approaches, the influence of performance indicators was more accurately explained by quarter score points, as opposed to quarter score margin, for all input feature set variables. The mean average difference between score margin and score points MAE results was 2.32 points (Table 5.1). Modelling performance statistics by quarter score points using the relative values feature set (n = 13) resulted in one of the lower MAE scores for both the GEE (3.83) and the decision tree (5.59). The lowest prediction errors for both models were on larger feature sets. The GEE MAE was 3.60 on the 169-feature set comprised of the combined total, relative and feature distribution values. The decision tree MAE was 5.45 on the 143-feature set comprised of the derived feature distribution values.

Rule outputs from the two regressor decision tree models, with the lowest MAE, are shown in Figures 5.4-1 and 5.4-2. The relative performance indicator of team differential of inside 50 values (Figure 5.4-1) and feature distribution inside 50s in the 75th percentile (Figure 5.4-2) contributed most strongly to the models. Interpretation involves following the branches down, from the root node representing the outcomes for each test, to the final terminal node to define the regression decision rules for the model. For example, in Figure 5.4-1 following down the right side, teams with relative Inside 50s greater than -3.5, relative kicks long greater than 5.5 scored more points per quarter, model prediction of 18.5 points based on 22 samples. Teams with higher contributions from more athletes to their inside 50 count, short and long kicks, and lower contributions from more athletes to their ineffective kick counts are more successful per quarter (Figure 5.4-2). See Figure 5.4-1 and Figure 5.4-

2 for further examples for rule sets. The defined rules represent performance skill fulfilment requirements for teams to achieve a successful quarter score or score margin outcome.

Table 5.1. Model results across data variables evaluated by mean absolute error.

Algorithm	Parameters	Model	Variables (number of features)	MAE (points)
Regressor Decision Tree	Total training data	1	Quarter margin and PI totals (13)	6.65
		2	Quarter margin and PI relative (13)	5.93
		3	Quarter margin and PI features distributions (143)	6.81
		4	Quarter margin and combined PI totals, relatives and feature distributions (169)	5.98
		5	Quarter score and PI totals (13)	4.47
		6	Quarter score and PI relative (13)	4.31
		7	Quarter score and PI feature distributions (143)	4.32
		8	Quarter score and combined PI totals, relatives and feature distributions (169)	4.10
Regressor Decision Tree	70% train set and 30% test set	9	Quarter margin and PI totals (13)	8.56
		10	Quarter margin and PI relative (13)	7.63
		11	Quarter margin and PI features distributions (143)	9.57
		12	Quarter margin and combined PI totals, relatives and feature distributions (169)	8.38
		13	Quarter score and PI totals (13)	5.60
		14	Quarter score and PI relative (13)	5.59
		15	Quarter score and PI feature distributions (143)	5.45
		16	Quarter score and combined PI totals, relatives and feature distributions (169)	5.74
Generalised Estimating Equations	Total training data	17	Quarter margin and PI totals (13)	7.13
		18	Quarter margin and PI relative (13)	6.18
		19	Quarter margin and PI features distributions (143)	6.03
		20	Quarter margin and combined PI totals, relatives and feature distributions (169)	5.12
		21	Quarter score and PI totals (13)	4.48
		22	Quarter score and PI relative (13)	4.64
		23	Quarter score and PI feature distributions (143)	3.83
		24	Quarter score and combined PI totals, relatives and feature distributions (169)	3.60

MAE, mean absolute error, *PI*, performance indicator

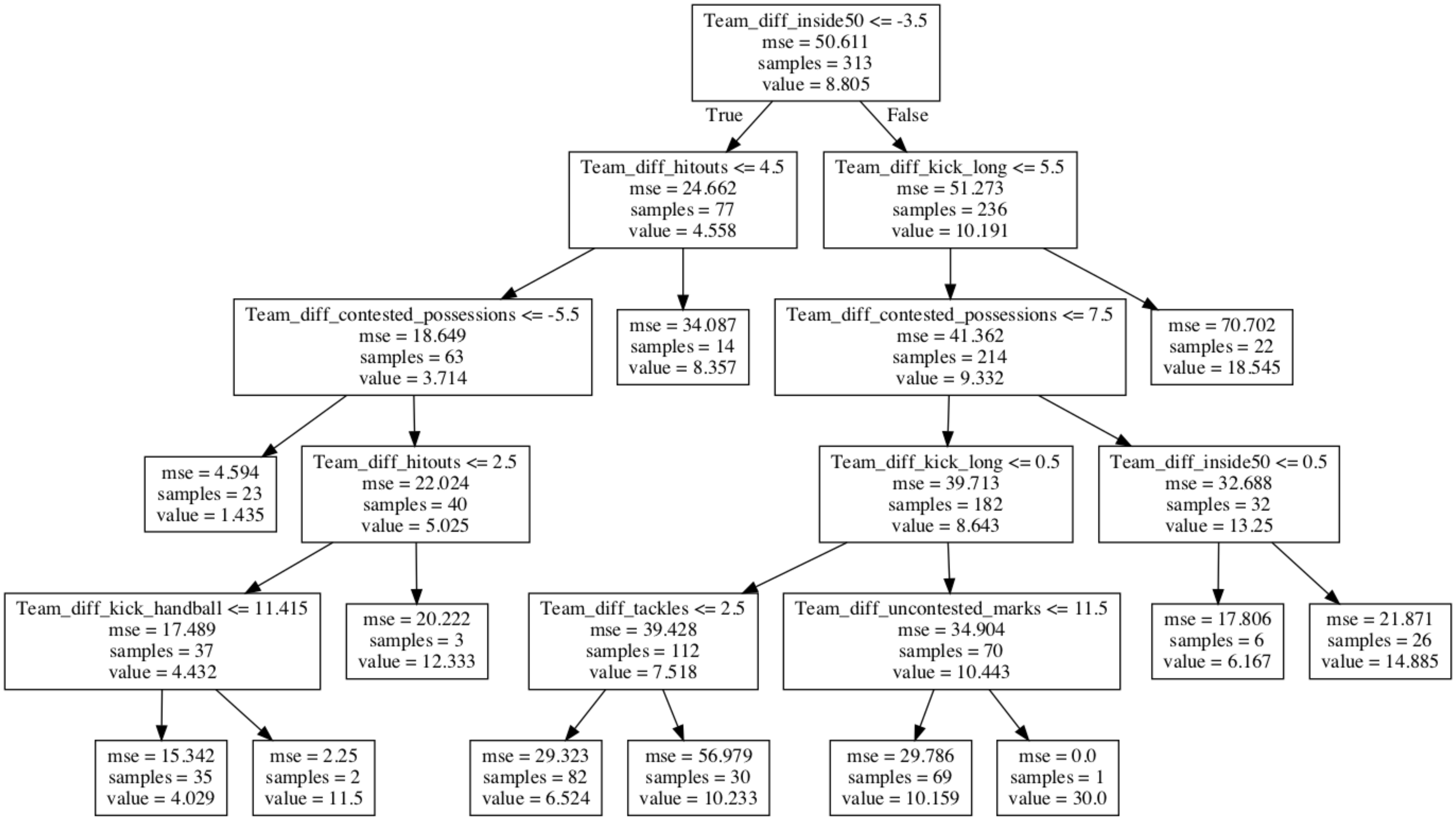

Figure 5.4-1. Regressor decision tree output of model 14, quarter score points and performance indicator relative values.

diff, differential, *mse*, mean sample error.

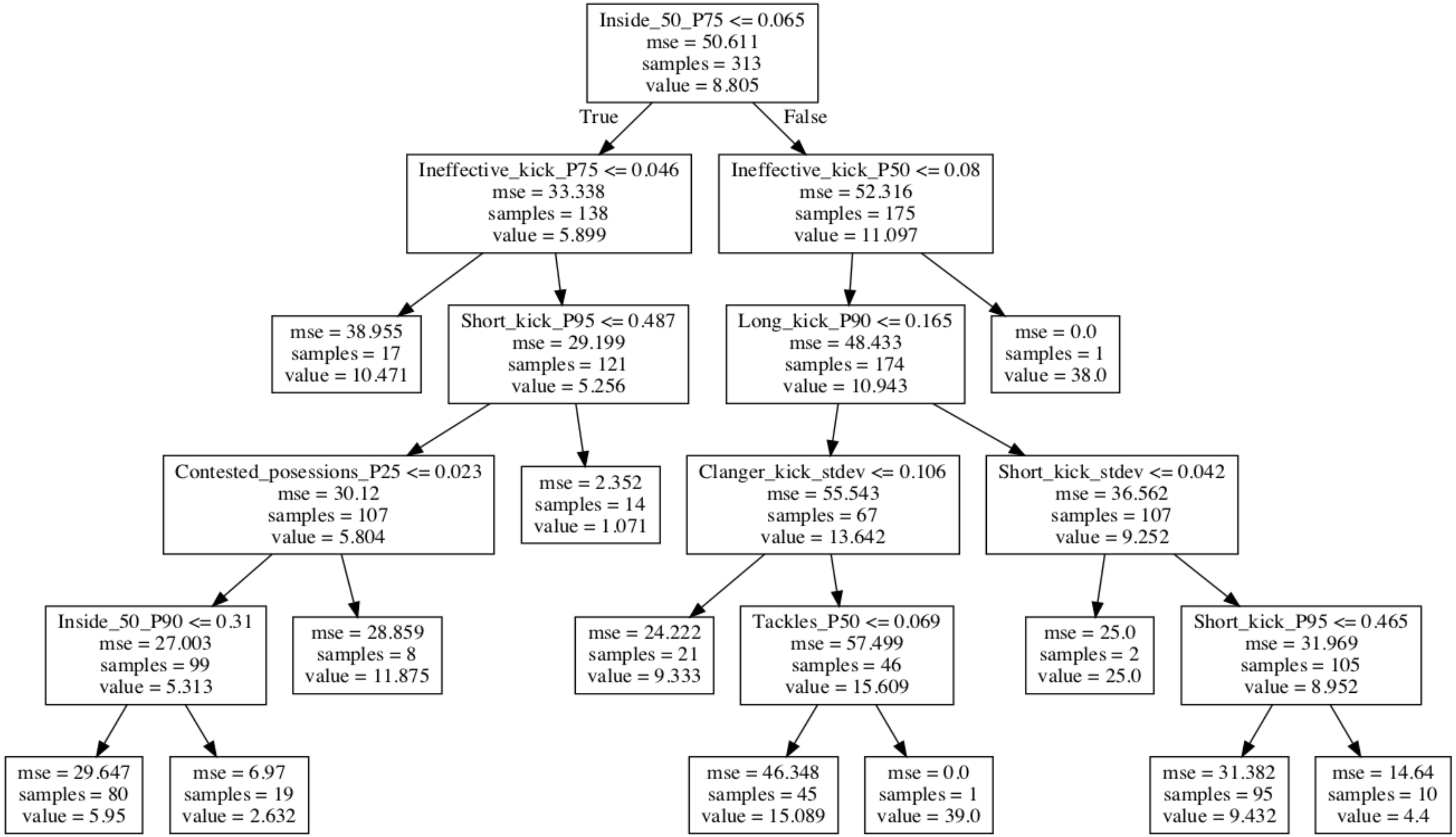

Figure 5.4-2. Regressor decision tree output of model 15, quarter score points and performance indicator feature distributions.

mse, mean sample error; *P25, P50, P75, P90, P95,* percentile level; *stdev*, standard deviation.

5.5 Discussion

This study assessed the extent to which AFLW athlete skill performance distributions explain match quarter outcome across the first two seasons of the inaugural AFLW national competition. Key results indicate that modelling data by quarter score points total was more accurate compared to quarter score margin. Teams with more successful inside 50 entries than their opposition likely scored more points in the quarter.

Modelling performance indicator data by quarter and not an entire match may allow for specific information and clearer relationships between the variables and success within different periods of a match (Gómez et al., 2014). During matches, coaches have the chance to address the playing group and reset tactics at quarter time breaks. Specific quarter-based skill influence information may aid in modifying individual athlete and team tactics, in comparison to the opposition as shown in elite women's basketball (Gómez et al., 2013). Therefore, breaking performance indicator data into the influence by quarter may provide targeted information for coaches during matches. As the league expands and more data becomes available, longitudinal comparisons would be of interest. In comparison of the two approaches, the GEE produced lower prediction errors across all data input variables. This may indicate that a simpler model approach is more appropriate for the current smaller dataset with relatively low feature dimensionality. However, to provide a practical outcome for coaches, a decision tree model may be more applicable as the output does not consider all features. Rather, decision trees provide a practical, parsimonious rule set for coaches who may be focused on the most influential performance indicators.

Features or variables are representative aspects of data that should be relevant, in that they have an influence on the model result with a function that is not assumed by the rest (Ladha & Deepa, 2011). Performance indicators that were a direct function of scoring in AFLW, including shots at goal, goal assist, behind assist and goal accuracy were not included in this study. These variables would potentially trivialise the process of determining performance skills which influence match success. Modelling quarter points scored produced the lowest prediction errors on the larger feature sets GEE (n = 169) and decision tree (n = 143) for both algorithms. But this was only a slight improvement from using the smaller relative values feature sets (n = 13). A larger data set could facilitate improved feature extraction and selection engineering for better representation of the data characteristics. More efficient algorithm processing and prediction accuracy (Liu et al., 2010) may also be increased. Further extracting distribution features, from individual athletes, demonstrates the structure contributions for AFLW teams. Interestingly, results suggest that in contrast to the AFL game, increased match skill performance contributions from key high performing individual athletes is more beneficial for team success. This is suggested by the higher percentile feature distributions contributing most strongly to the decision model (Figure 5.4-2). For example, the inside 50 P75, short kick P95, long kick P90 and ineffective kick P75 values.

Successful outcomes in the AFL involve relatively even performances from athletes across a team (Robertson, Gupta, et al., 2016). The comparatively higher performance contributions by key individual athletes to team success in the AFLW may be explained by the fact it is a new competition format and across many facets is still developing. As such, the level of game plan seen in the AFL competition (Alexander, Spencer, Mara, et al., 2019; Robertson, Gupta, et al., 2016; Woods et al., 2017) may yet be reasonable in the AFLW due to the variety of AFLW athlete game experience and skill maturity levels being contracted. The skill development of AFLW athletes, who have either recently progressed from junior competitions or transitioned from another sport and hence not marquee or high performing athletes may be also affected by the lower resourced professional support structures and training opportunities currently experienced in the AFLW. As opposed to the well-established AFL, where newly contracted athletes are highly coached, skill acquired and AF experienced before competing in AFL level matches (Haycraft et al., 2017). This may be partly because of lower coaching and sports science resource support in the AFLW competition relative to elite male AF. These factors could be contributing to individual athlete dominance in the AFLW, potentially preventing collective team contributions towards successful match outcome.

Comparison of the current results to AFLW match skill analysis during the 2017 season only (Black et al., 2018a) is difficult, due to the differentiating features sets used. In Black et al. (2018), variables with direct functions of scoring were used. In order to build upon this previous analysis (Black et al., 2018a), further data feature extraction from a larger sample size and revised statistical modelling was run in the present study. Breaking down the performance indicators to types of the variable, for example, including long, short and ineffective kicks allows for expansion of the key performance measures.

Practically, as the strongest features in the regression decision tree models relate to kick performance indicators, clubs may look to emphasise kick skill development. inside 50's, hit outs and contested possessions, by key athletes, contribute most to quarter success during matches. AFLW clubs may also look to compile teams with capable skilled kickers and recruit future athletes with current or potential strong kick skills (Stewart et al., 2007). Game plan development around a kick dominant ball movement strategy, particularly in hit-out clearances and efficient inside 50 entries may also be of match tactical advantage. Coaches may work specifically with key forward and midfield athletes to develop efficient plays and decision making from centre bounce to inside 50 entry possession chains, in order to maximise scoring opportunities. Improving an athlete's kick execution skills may also benefit kick delivery and mark success from a team member in contested possessions during matches. Analysis of match performance statistical information can also be applied off field in the athlete recruitment department (McIntosh et al., 2018b). As the AFLW is in its infancy, a greater understanding of team and individual contributions to winning may highlight what performance characteristics are beneficial towards maximising team success. Recruiters could make strategic decisions on selecting athletes that currently exhibit or have the potential to develop the key performance characteristics identified.

Future research may look to investigate the contextual variables around match play on the outcome such as travel requirements, days between matches and player interchange rotations per quarter. Specifically, given the current short home-and-away season, increased importance is on the outcome of each match for ladder positioning. Across different team sports, contextual variables influence match outcomes and performance indicators (Gómez et al., 2013; Oliveira et al., 2012), particularly in team field sports who play multi-round home-and-away seasons (García-Rubio, Gómez, Lago-Peñas, & Ibáñez Godoy, 2015; Ruano et al., 2016; Taylor et al., 2008). Furthermore, spatiotemporal data characteristics of players could be analysed to explain team behaviours in match play styles and tactics (Alexander, Spencer, Mara, et al., 2019) in the AFLW.

5.6 Conclusion

Quarter success in the AFLW was characterised by greater inside 50s as a relative to the opposition and key athletes in the 0.75 percentile performing inside 50s. Results suggest within the current AFLW competition, key athletes' skilled performances are contributing more to match success rather than a collective team effort as opposed to the AFL competition. Using machine learning methods in sport analytics to uncover practical information from athlete match performance statistics allows for analysis on how these athletes are contributing towards team success. Post-hoc reporting of results, in a comprehensible format for coaching staff, may provide a basis for training and match strategic planning.

The current study highlights the key performance skills contributing to AFLW team match success and indicates that efficient kick skills are important. As this thesis is aimed at taking a multi-disciplinary approach into the use of IMUs and analytical methods in AFLW performances and sport-specific movements recognition, it is beneficial to further define the kicking skills of elite female AF athletes. Currently it is unknown how the most common kick in AF, the drop punt, biomechanically compares to elite male AF athletes which has been heavily researched. Evidence from similar soccer kicking research which found key differences between the biomechanics of male and female kick executions (Alcock, 2010; Sakamoto et al., 2014) creates scope for investigation in AF kicking.

Chapter Six: Biomechanical characteristics of elite female Australian Rules football preferred and non-preferred drop punt kicks

---

Conference proceeding copied from the original published article: **Cust, E. E.**, Ball, K., Sweeting, A. J., & Robertson, S. (2019). Biomechanical characteristics of elite female Australian rules football preferred and non-preferred drop punt kicks. Proceedings of the 7th International Conference on Sport Sciences Research and Technology Support - Volume 1: IcSPORTS, 32–37. https://doi.org/10.5220/0008066300320037 (Appendix D).

6.1 Abstract

While Australian Rules kick biomechanics has been researched considerably, there is yet to be focus specifically on women participants. Elite female Australian Rules football drop punt kick characteristics were collected from 15 elite female participants for both the preferred and non-preferred legs. All participants undertook a 20-kick protocol captured by a 3-dimensional motion analysis camera system. Preferred leg kicks produced faster foot velocities prior to foot-ball contact, 18.0 ± 1.2 m.s^{-1} preferred, 16.2 ± 1.3 m.s^{-1} non-preferred, and faster ball velocities post foot-ball contact, 24.7 ± 1.4 m.s^{-1} preferred, 21.6 ± 2.0 m.s^{-1} non-preferred. Differences in movement patterns of the hip and knee segments were shown between kick leg preferences; hip angular velocity 94.4 ± 75.9°/s preferred and 126.2 ± 66.3°/s non-preferred, knee angular velocity 1384.8 ± 415.2°/s preferred and 1013.6 ± 230.2°/s non-preferred. Research results identified the changes in elite women's drop punt kick mechanics in comparison to leg preference, which can be viewed against senior and junior men's Australian football kick analysis findings. The current research information

could be of benefit to practitioners in linking targeted field coaching cues and conditioning programs tailored to identified kick skill and movement deficiencies.

6.2 Introduction

The National Women's Australian Rules Football competition (AFLW) is in its fourth year of operation, yet there has been no reported biomechanical analysis of women's kicking. In Australian Rules football (AF), efficient kick performance has been identified as a strong contributor towards team match success (Black et al., 2018a; Robertson, Gupta, et al., 2016).

In AF the drop punt is the most commonly performed kick due to the flight accuracy and ease of catching for the receiver (Ball, 2008). Across the six phases of a drop punt (Ball, 2008), several kinematic factors have been found to influence the success, efficiency, and accuracy of performance. Prominently, higher kick leg foot velocities prior to ball contact have a major influence on the kick distance (Ball, 2008; Ball et al., 2013; Peacock et al., 2017) and ball velocities (Ball, 2008; Peacock & Ball, 2017, 2016) achieved. The flight path accuracy of the ball is determined primarily by the combination of the flight characteristics imparted on the ball during the foot-to-ball contact phase (Peacock and Ball, 2018; Peacock *et al.*, 2018). Differences in kick biomechanics have been found between the preferred and non-preferred leg in men's AF kicks (Ball, 2011; Smith et al., 2009) and soccer (Nunome et al., 2006). The ability to kick proficiently on both legs and over long distances in AF is a tactical advantage (Ball, 2008, 2011) in the dynamic unpredictable nature of match play. Biomechanical assessment may be an important information source for individual athlete skill profiling to identify areas of deficiencies for drop punts kicks.

The kick impact and technical components of men’s kicking across several athlete levels has already been established allowing for quantified information to further develop kick skills on a team and individual basis. To address the lack of quantitative information in women’s AF kick biomechanics, 3-dimensional optoelectronic motion analysis was undertaken. Conducting this research is important for broadening the sport science support invested in the new AFLW competition, with the intention of improving athlete kick skill and therefore team match performances. The aim of this research was to analyse the biomechanical characteristics of elite female AF drop punts for both the preferred and non-preferred kick legs. The outcomes of this research can inform the technical aspects of distance kicking in women’s AF to aid in athlete kick skill development, as well as links with strength and conditioning and injury.

6.3 Methods

6.3.1 Participants

Fifteen elite female AF athletes provided written informed consent to participate in this research. Of the participants, twelve were contracted to an AFLW team and three were competing at a high standard in their respective State based competition. The University’s Human Research Ethics Committee approved the study (application number 0000025654).

6.3.2 Research procedures

Athletes undertook a drop punt kick protocol as part of a broader test battery. Ten drop punts were undertaken for maximum distance and intensity on each leg. Maximal kicks were performed into a net situated 30 m from the kick launch area. Prior to undertaking the protocol, each athlete completed a dynamic warmup including jogging, dynamic stretches and five 20 m submaximal kicks on each leg. All athletes wore their regular football boots and used official AFLW match balls (Sherrin, Scoresby, Australia). The testing was conducted in purpose built indoor football training facility on artificial turf.

Drop punt kicks were captured by a 10-camera optoelectronic motion analysis system (MAS) capturing at 100 Hz (T-40 series, Vicon Nexus v2, Oxford, UK). Previous assessment of sampling rates had found low maximum error ranges for kick parameters from 500 Hz to 100 Hz (Coventry et al., 2015). Cameras were set up as an arc around the testing area and mounted at varying heights in order to allow full capture of the kick and ball flight movements. 35 reflective markers (diameter: 14 mm) were taped to each athlete at anatomical landmarks as per previous kick research (Blair, Duthie, et al., 2018), shown in Figure 6.3-1. Four reflective makers were attached to the football (Figure 6.3-2) to create a coordinate system and establish the ball centre.

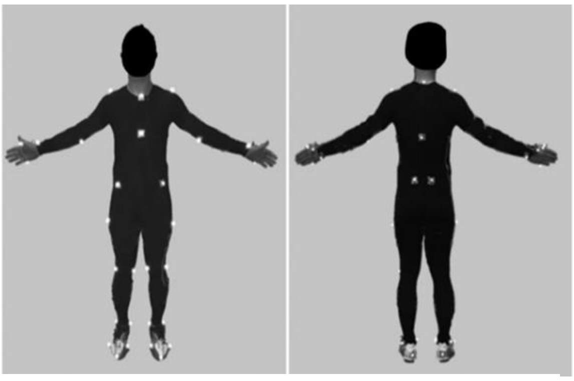

Figure 6.3-1. Athlete marker set-up.

Figure 6.3-2. Football reflective marker positions.

6.3.3 Data analysis

Raw motion analysis data were digitised in Nexus (v.2.0, Vicon, Oxford, UK) and processed in Visual 3D (C-motion, Inc. Germantown, USA). Data were pre-processed through a polynomial interpolation (order: 3) and smoothed using a low-pass fourth-order Butterworth filter (cut-off frequency: 10 Hz) (Ball, 2008, 2011); frequent in-lab evaluation of VICON data using spectral and residual analyses.

A total of 300 drop punts (150 preferred and 150 non-preferred kicks) were analysed for ball velocity values. The trials with the highest preferred and non-preferred ball velocities were selected for each athlete for full kinematic analysis in this study, as final ball velocities are the reflection of efficacy in impact characteristics applied to the ball (Peacock, Ball and Taylor, 2017; Peacock and Ball, 2018a). A total of nine drop punt kick parameters were analysed from the MAS data, see Table 6.1, based on previous technical parameters assessed in AF kick performance (Ball, 2008; Ball et al., 2010; Smith et al., 2009). Processed data for each parameter were exported to a custom Excel file and the group mean and standard deviation (SD) calculated for each preferred and non-preferred kick parameter. Paired t-tests were computed for each parameter with statistical significance set at $p < 0.05$. The effect size for each measure for between-group distances was calculated using Cohen's d statistic indicating a small or trivial (d = 0–0.2), moderate (d = 0.2–0.5), large (d = 0.5–0.8), and very large (d . 0.8) effect (Hopkins et al., 2009).

Table 6.1. Definitions of measured kick parameters.

Parameter	Definition
Foot velocity prior to ball contact ($m.s^{-1}$)	Linear velocity of the foot segment measured from the head of the 5th metatarsal
Ball velocity post foot contact ($m.s^{-1}$)	Linear velocity of the ball segment
Ball: foot velocity ratio	Ball velocity at release divided by foot velocity at initial impact
Support leg knee flexion (°)	Degree of flexion of the support leg at ball contact
Knee angle at ball contact (°)	Angle between the thigh and shank of kick leg
Knee angular velocity (°/s)	Angular velocity of the knee joint of kick leg
Hip angle at ball contact (°)	Angle between the thigh and the trunk on the anterior aspect of the participant
Hip angular velocity (°/s)	Angular velocity of the hip segment
Pelvis linear velocity ($m.s^{-1}$)	Linear velocity of the pelvis segment

6.4 Results

Table 6.2 reports the mean data kinematic parameters of the foot, knee, hip, and ball segments. The preferred leg produced significantly greater foot velocity, ball velocity, knee angular velocity, and pelvis linear velocity, and a significantly smaller hip angle and hip angular velocity in comparison to the non-preferred leg. The maximum foot velocities achieved were 20.9 $m.s^{-1}$ and 17.7 $m.s^{-1}$ on the preferred and non-preferred legs, respectively. The maximum ball velocities achieved were 27.0 $m.s^{-1}$ and 25.5 $m.s^{-1}$ on the preferred and non-preferred legs, respectively.

Table 6.2. Impact characteristics for preferred and non-preferred drop punt distance kicks for elite women's AF. Data reported as mean and standard deviation values and results of statistical tests comparing preferred and non-preferred leg kicks.

Parameter	Preferred leg		Non-preferred leg		p	Effect size (d)
	mean	SD	mean	SD		
Foot velocity ($m.s^{-1}$)	18.9	1.2	16.2	1.3	<0.001*	2.2 Very large
Ball velocity ($m.s^{-1}$)	24.7	1.4	21.6	2.0	<0.001*	1.8 Very large
Ball: foot velocity ratio	1.31	0.11	1.33	0.07	0.59	0.2 Small
Support leg knee flexion (°)	37.0	11.3	41.0	8.3	0.25	0.4 Moderate
Knee angle at ball contact (°)	50.7	12.2	57.7	13.5	0.13	0.5 Moderate
Knee angular velocity (°/s)	1384	415	1014	230	0.02*	1.1 Very large
Hip angle at ball contact (°)	34.3	13.5	48.8	15.8	0.01*	1.0 Very large
Hip angular velocity (°/s)	94	76	126	66	0.04*	0.7 Large
Pelvis linear velocity ($m.s^{-1}$)	1.7	0.4	1.4	0.5	0.03*	0.6 Large

* Significant difference ($p < 0.05$)

6.5 Discussion

The current research on women's elite Australian Rules football kick biomechanics reports the first analysis of its type to further the understandings of kick skill execution. Results showed that preferred leg kicks were characterised by faster foot velocities prior to ball contact, greater knee angular velocities, pelvis linear velocities, and smaller hip angular velocities. Linking information from biomechanical analysis with field coaching cues and conditioning programs may be beneficial for individualised athlete kick skill development.

Elite female AF athletes in this study produced higher foot and ball velocities on their preferred leg kicks. Foot and ball velocities for elite women were lower than the reported values for senior elite men (Ball, 2008, 2011; Smith et al., 2009) and junior elite men (Ball et al., 2010) AF athletes. Preferred leg drop punt kicks in elite senior men have shown foot velocities of 26.5 ± 2.5 m.s^{-1} and ball velocities of 32.6 ± 4.4 m.s^{-1} (Smith et al., 2009). Relation could also be drawn to kick distances achieved by women and men as foot velocity has shown strong correlation association with ball flight distance (Ball, 2008; Peacock et al., 2017). Elite female soccer athletes have reported foot velocities of 17.70 ± 1.92 m.s^{-1} (instep kicks) and 17.45 ± 1.59 m.s^{-1} (curve kicks), and ball velocities of 22.62 ± 1.71 m.s^{-1} (instep kicks) and 21.51 ± 1.33 m.s^{1} (curve kicks) (Alcock et al., 2012).

The ball-to-foot velocity ratio is a measure of kick impact efficiency and is widely reported on in AF (Smith, Ball and MacMahon, 2009; Ball, Smith and MacMahon, 2010; Ball *et al.*, 2013; Peacock and Ball, 2018b) and soccer research (Nunome et al., 2018; Sakamoto & Asai, 2013; Shinkai et al., 2009). The present study showed no difference for ball-to-foot ratio between the preferred (1.31) and non-preferred legs (1.33), which has previously been reported in male AF kick research (Smith et al., 2009). This may indicate that greater ball velocities on the preferred leg are the result of a faster leg swing as attributed by faster foot velocities and knee angular velocities in applying greater force onto the ball (Nunome et al., 2006; Smith et al., 2009) (Table 6.2). Differences in body mass have also been reported to affect the ball-to-foot ratio, which may confound comparisons between male, female, junior, and senior playing groups (Shinkai et al., 2013).

Differences in movement patterns were shown between kick leg preferences. Overall, the preferred leg achieved greater knee angular velocity and pelvis linear velocity, and smaller hip angle and hip angular velocity (Table 6.2). As the non-preferred leg produced larger hip angular velocities and hip angles, this may suggest that greater use of the thigh and hip segments were recruited. The change in movement pattern between the kick leg types may indicate the need for more stability via dominant hip control on non-preferred kicks. Another factor could also be related to the speed of run-up in approach towards the kick execution on each leg, although this was not measured in this study. Also, the result of less efficient use of sequential summation or transfer of momentum (Ball, 2011) as indicated by the lower knee angular velocity on non-preferred leg kicks. In comparison to senior AF male athletes (Ball, 2011), greater mean knee and hip angles, and knee angular velocities were achieved for both preferred and non-preferred kicks by elite women AF athletes. Although, lower hip angular velocities were produced in comparison to reported male AF athletes, 56 ± 65°/s preferred leg and 138 ± 81°/s non-preferred leg (Ball, 2011).

Technical differences in kick strategies have been demonstrated for thigh dominant or knee dominant kickers during maximal distance kicking (Ball, 2008) and further supported during goal kicking constraints tasks (Blair, Robertson, et al., 2018). Although kick performance indicators of foot velocity and kick distance were not significantly different between each approach suggesting similar kick performance outcomes can be achieved with either movement strategy (Ball, 2008). Looking into the thigh-knee angular velocity continuum, Ball (2008) sorted the participant data to provide indicative values for those athletes who use a thigh or knee dominant strategy for preferred leg distance kicking. In comparison, post-hoc evaluation of the current elite women's data was undertaken using the hip and knee angular velocities. Further analysis showed 14 out of the 15 athletes would be considered using a knee dominant strategy on their preferred leg. In contrast, on the non-preferred leg, the majority of the group would be classified hip dominant with data from 10 athletes of 15 indicating this. This trend is consistent with findings in the men's data where on preferred leg kicks there is increased contribution from the knee segment and lower hip or thigh involvement. The opposite shown on non-preferred leg kicks with greater hip segment contribution than the knee for force generation through the kick motion (Ball, 2011).

The support leg is important in maintaining stability through the kick motion and plays a role in the performance quality of a kick (Ball, 2013). The current results showed a moderate non-significant effect of less knee flexion in the supporting leg at ball contact occurred on preferred leg kicks, 37 ± 11.3° to non-preferred kicks, 41 ± 8.3°. These are in contrast to results found in elite males across maximal kicks which showed greater support leg flexion on preferred leg kicks, 43 ± 6°, than non-preferred leg kicks, 41 ± 1° (Ball, 2013). Although it has been suggested that greater support leg knee flexion leads to a lower centre of gravity and hence stability in the motion allowing for improved kick accuracy (Dichiera et al., 2006). Further results by Ball (2013) indicated that a more extended support leg knee on stance kick phase that was maintained to the ball contact phase related to higher foot velocities and an improved drop punt ball flight distance. During match play athletes are repeatedly required to perform kicks with constraints against both distance and accuracy, most commonly in goal kicking (Blair, Robertson, et al., 2018). Kicking kinematics measured across changing distance on goal kicks showed that increased distances resulted in greater knee extension on the support leg during the stance phase (large effect size), and moderately higher foot velocities, shank, and knee angular velocities (Blair, Duthie, et al., 2018). The authors noted potential technical difference for tasks in the literature when both kick skill accuracy and distance constraints were combined. Suggesting this related to the research protocols used with accuracy tasks performed over shorter distances at lower speeds compared to research on maximal distance kicking causing athlete to adopt differing techniques to suit each task (Blair, Robertson, et al., 2018). For example, during maximal distance high impact kicks the athlete adopts a more upright position through the torso and consequently a higher hip position to generate the faster foot velocities required (Ball, 2013). Further work to assess how these variables influence elite women's kick performance considering the altered match play styles and therefore kick constraints compared to the men's game (Cust, Sweeting, Ball, Anderson, et al., 2019) would be of skill technique coaching benefit.

As foot velocity prior to ball impact is strongly correlated with drop punt kick distance (Ball, 2008, 2011; Peacock & Ball, 2017) and used as a strategy to control the kick outcome (Peacock et al., 2017), a focus on improving an athlete's ability to generate high foot velocities on both legs would benefit kick skills for in-match tactics (Ball, 2008, 2011). Furthermore, if footballers dominantly kick on one leg, the increased repetition loading may create imbalances in hip and lower limb strength which could affect skill performance and increase asymmetry load related injury (Hart et al., 2013, 2014). As the current results show differences in the use of lower limb segments between the two legs, there is potential for muscle asymmetries to develop (Ball, 2011; Hart et al., 2014). Strategies such as training the non-preferred leg to recruit greater lower limb involvement through skilled coaching cues and targeted conditioning programs may again be of benefit to improving kick skill performance across both legs for tactical advantage in matches. Research has indicated that combined technical and strength-based interventions for AF athletes in training for the drop punk kick serves as a constructive approach to performance improvements (Ball, 2008; Hart et al., 2014).

Further research should progress assessment of the support leg mechanisms (Ball, 2013) and kinematic characteristics of the kick impact phase for elite women AF athletes in relation to kick accuracy (Peacock et al., 2017). Greater understandings into the underlying mechanisms for the differences between both preferred and non-preferred leg kicks for elite women, and between male and female kinematics during kick execution would be important to further quantify. As different movement approaches exist for kick execution, future research looking at the relationships between knee and thigh (or hip) strategies for distance kicks and kick accuracy would be of benefit to kick skill coaching and individual conditioning. Knowing individual athlete movement strategies would directly affect coaching and conditioning due to the different muscle recruitment processes for generating forces for each approach (Ball, 2008). In-depth information within this field could provide links to improve practices in women's AF kick skill coaching, individual athlete injury patterns related to repeated kick execution, and targeted strength and conditioning practices.

6.6 Conclusion

The biomechanical characteristics of elite female Australian Rules football drop punts kicks for both the preferred and non-preferred legs were quantified. Preferred leg kicks produced faster foot velocities prior to ball contact, greater knee angular velocities, pelvis linear velocities, and smaller hip angular velocities. Movement differences were found in hip and lower limb segments between both kick legs as greater knee angular velocity and pelvis linear velocity characterised preferred leg kicks, yet a higher hip angular velocity on non-preferred leg kicks. Improved understandings of women's AF kick skill via kinematic technical analysis could be of benefit in linking with targeted field coaching cues and conditioning programs tailored to identified kick skill and movement deficiencies.

The collated knowledge from Chapters Five and Six highlights the importance of kick skills in AF and the technical components of them. Linking to the information of Chapter Three which presented several applications of IMUs for sport movement recognition and performance analysis, IMUs may be beneficial in AF training for both performance analysis and skill coaching purposes. As IMUs have not been researched for AF kicks in-situation, initial investigations looking at the signal characteristics of a kick, sensor hardware and body mount location in relation the movement, and data analysis methods are required to create new knowledge for future research improvements. The following chapter starts to look at how a common commercial IMU operates during in-situation AF kicking, then further assess the signal characteristics of an AF drop punt kick. The methods used in Chapter Seven are extended on in Chapter Eight to assess AF kick type recognition in an on-field environment as further evidence for how feasible the technology could be to integrate into AF training performance analysis programs.

Chapter Seven: Pilot analysis of IMUs in Australian Rules football kick movement detection

7.1 Introduction

Inertial measurement units (IMUs) can provide an accessible method to collect sport-specific movement data, for example, see Chapters Two and Three of the thesis. However, the feasibility to implement commercially available IMUs for Australian football (AF) kick monitoring requires further investigation. Specifically, the determining the content validity and interpretability of IMU sensors to measure the required kick variables and provide data for monitoring athlete skill performances would be considered an informative step in the sport sciences (Robertson, Kremer, Aisbett, Tran, & Cerin, 2017). The output from an IMU and any subsequent derived metrics need to be assessed for specificity of its intended measure and the degree to which practical meaning can be applied to warrant ongoing confidence of measurement properties (Robertson et al., 2017). A series of developmental investigations were undertaken to assess the potential use of IMUs for on-field AF kick monitoring. The overall aim of these pilot studies was to inform future methods and explore concepts for research in aptness towards recording measurement properties for AF kicking skills during training sessions.

A semi-automated method, using IMUs to quantify kick volumes and potentially kick type, could be informative in designing AF skilled training. Further, this approach may also aid in kick skill development based on evidence from Abdullah et al. (2020) and Wang et al. (2018) or in rehabilitation from injury monitoring as indicated by Mehta (2019) and Reilly et al. (2017a). The opportunity for enhanced data capture and processing of skill performance analysis during football training sessions may be possible by combining IMUs with data analytical methods for reducing manual data collection, as evident from the evaluations by Camomilla et al. (2018). Although further investigation is required to evaluate the feasibility for implementation of semi-automated kick recognition methods in AF training environments compared to current manual notational based methods. Particularly, in semi-controlled environments such as known kicking drills in an indoor training space, which may provide the basis for bridging the gap between lab-based and on-field research in working towards uncontrolled AF training environments such as small-sided games. The aim of the Part One of this pilot study was to investigate the sensor implementation and data detection analysis of kicks within a semi-controlled protocol using commercially available IMUs in order to inform of future method practices.

IMUs present as a tool for sport skill analysis on-field; for example, providing feedback on performance metrics during snowboard half-pipe competitions to coaches and athletes (Harding et al., 2008). Or, monitoring tennis hitting loads using a single IMU set-up system which classifies eight tennis stroke types during hitting practice (Ebner & Findling, 2019). Quantification of AF kicking volumes and possible performance variables using IMUs during training could provide greater performance insights for athlete coaching practices. Although, investigation of AF kick accelerometer ranges is needed to provide clear information on the IMU specification requirements when applying practically. Knowing the IMU dynamic range required to capture an AF kick without the output signal clipping would also be important when looking at detecting individual AF kick types or possibly devising kick intensity bands. Part Two of this pilot study and its methods investigated within the first phase in developing a human activity recognition system: data acquisition (Bulling et al., 2014). The data acquisition phase is defined by the choice, set-up, and constraints of wearable sensors that need to be considered in relation to the nature of activity being targeted. The aim of this pilot was to investigate the output signal of a high range analogue accelerometer mounted on the lower limb for AF drop punt kicks.

7.2 Methods

The methods and data analysis for studies Part One and Part Two of the research development are presented separately. Both testing sessions were conducted in an indoor, purpose-built football training facility on an artificial turf field at different time points, Part one in early 2018 and Part Two early in 2019. All participants were injury free at the commencement of the studies and provided written informed consent. Furthermore, all wore their regular football boots and official AFLW match balls were used during the kick protocols (Sherrin, Scoresby, Australia). The University's Human Research Ethics Committee approved (application number 0000025654) the methods for this study.

Part One involved ten female AF athletes (n = 4 AFLW listed and n = 6 State competition level at the time of study) who participated in a semi-controlled kick protocol. All participants indicated their right leg was their preferred kick leg. The participants wore a 3-axis IMU (accelerometer ± 16 *g*, gyroscope ± 2000°/s, magnetometer ± 1200 µT; IMeasureU BlueThunder sensor, Auckland, New Zealand) sampling at 500 Hz on the lateral aspect of both lower limbs, above the malleolus; see Figure 7.2-1. Several placements of the IMU were considered taking into the factors of athlete comfort, ability to perform the kick with minimal interference and minimal soft tissue artefacts (Camomilla et al., 2018). The evidence collated in Chapter Two section 2.3 and Chapter Three of this thesis on the placement of IMUs in regards to the movements being recorded also informed mounting the IMU on the lateral lower limb. A more direct measure of a kick can be obtained as it is close to the foot-to-ball impact, this concept is discussed in Nedergaard et al. (2017) in regards to foot-to-ground acceleration measurement. The improved accuracy of the signal kick measure by placing it close to the impact area may also aid in differentiating signal characteristics for the AF kick types. The IMUs were tightly secured using the manufacturer's recommended Velcro band. Recording control of the IMUs was performed using the manufacture's research application (IMU Research App; IMeasureU Auckland, New Zealand) on an iPad (Apple Inc., Cupertino, USA). A video camera was fixed and elevated in the corner of the testing field, circa 4 m from the ground to capture the whole test protocol as the ground truth reference for post-analysis. The trial protocol is outlined in Table 7.1 a schematic of the set-up in Figure 7.2-2. The protocol was run in a continuous progression from each kick with short rest between kick types. Briefly, it involved five repetitions each of drop punts at varying distances, opposite leg drop punts, grubber kicks, surge kicks, and snap kicks. A grubber kick involves kicking the football fast along the ground so it moves forward in a rolling motion usually used to prevent the opposition from marking kicks or attempting to rush close shots on goal from an angle. A snap is kicked off the inside of the boot and curves in the opposite direction. A surge kick is a quick kick for maximum height and distance used in play to clear the ball from congestion giving distance from the opposition. A standardised warm-up was completed before participants commenced the trials. Prior to starting the protocol, all participants completed three consecutive vertical jumps in front of the video camera in order to create a time point to sync the IMU and vision files during analysis.

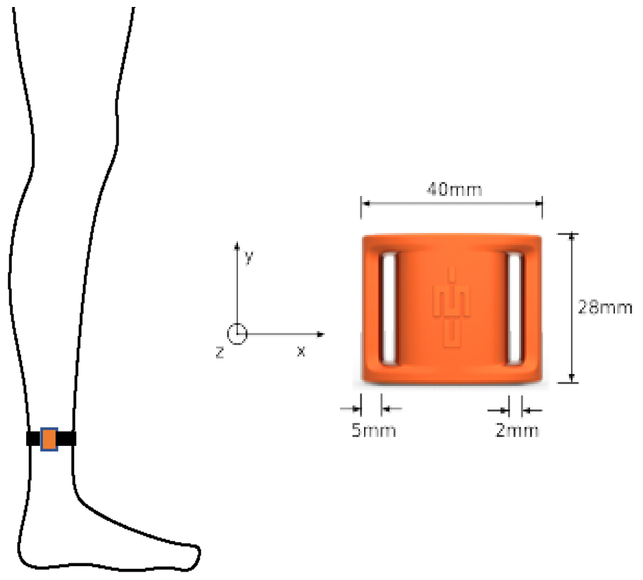

Figure 7.2-1. Part One IMU placement and axes orientation schematic.

Table 7.3. Protocol of study Part One.

Kick type	Repetitions	Kick distance target (m)	Kick approach
Drop punt	5	10	Step lead in, kick to athlete
Drop punt	5	20	Step lead in, kick to athlete
Drop punt	5	30	Run lead in, kick to net target
Grubber	5	10	Step lead in, kick to athlete
Opposite leg	5	10	Step lead in, kick to athlete
Surge	5	Maximum attempt	Run lead in, ground ball collect
Snap	5	20	Step lead in, kick to net target

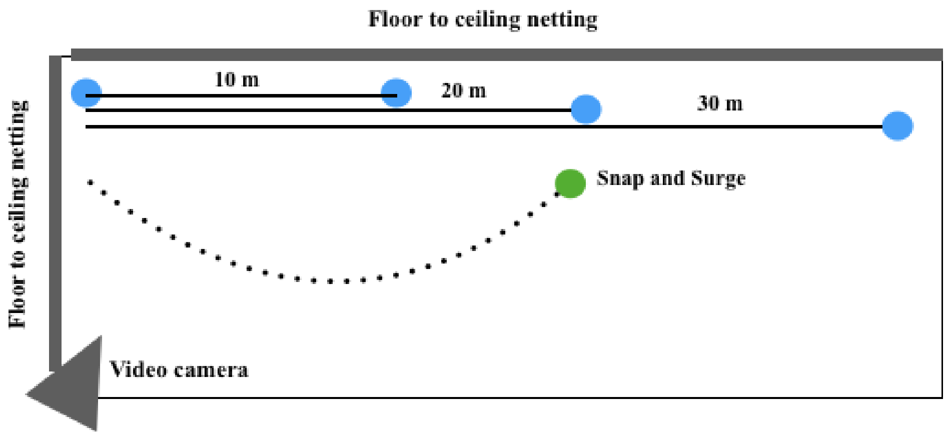

Figure 7.2-2. Schematic of study Part One protocol set-up.

Part Two involved 14 elite female AF athletes (n = 11 AFLW listed and n = 3 State competition level at the time of study) who participated in a controlled kick protocol. The test protocol was part of a larger biomechanical analysis which involved a 10-camera optoelectronic motion analysis system (T-40 series, Vicon Nexus v2, Oxford, UK); presented in Figure 7.2-3. The data required for this study was taken from when the participants performed ten maximal drop punt kicks on their preferred kick leg. Maximal kicks were performed into a net situated 30 m from the kick launch area. Prior to undertaking the protocol, each athlete completed a dynamic warmup including five 20 m submaximal kicks. One custom built 3-axis analogue accelerometer (Analog Devices ADXL377 chip, Norwood, Massachusetts) with a range of ± 200 *g* sampling at 500 Hz was mounted on the lateral aspect of the preferred kick leg lower limb above the malleoli. The orientation of the accelerometer axes in relation to the device's mounted leg position is shown in Figure 7.2-4. Both the accelerometer and the wire running to the reader box, which was positioned on the upper back within each participant's sports crop top, were securely taped to the skin using kinesiology tape. A video camera mounted on a tri-pod at circa 2 m from the ground to capture the relevant kicks as the ground truth reference for post-analysis. Again, as per part one methods, prior to starting the protocol all participants completed three consecutive vertical jumps in front of the video camera.

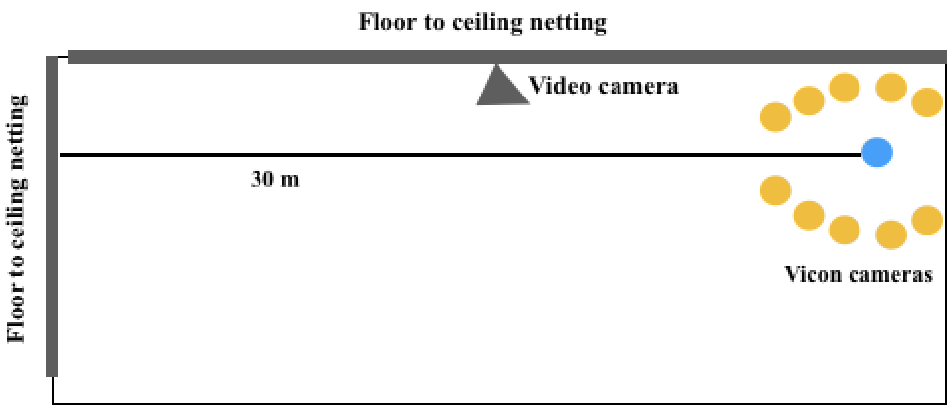

Figure 7.2-3. Schematic of study Part Two protocol set-up. The Vicon camera set-up was a part of a larger joint data capture session included in the schematic for context.

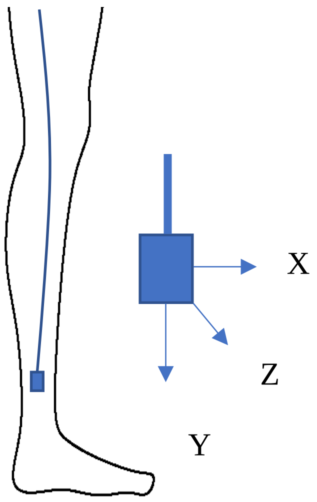

Figure 7.2-4. Study Part Two analogue accelerometer placement and axes orientation schematic.

7.2.1 Data analysis

In study Part One, raw accelerometer data were downloaded and exported into individual Excel spreadsheets via the manufacturer's research application and software programs (IMU Research App; IMeasureU Auckland, New Zealand). Each participant's trial was pre-processed using a 50-point moving average smoothing filter and run through an adapted peak feature detection script (Duarte, 2014) on the accelerometer y-axis with Python (Python Software Foundation, 2018). The main function of the detection script involved detecting peaks that are greater than a minimum peak height on the rising edge and not to record another kick for a minimum two seconds (1000 frames) after one had already been detected. A signal trace and identification of detected peaks was also produced as a visual check of all processed files. The video footage was manually coded by an experienced performance analyst using SportsCode Elite (Agile Sports Technologies, Inc., Lincoln, Nebraska) to create individual time-stamped activities for each participant. Timestamps from kick data and ground-truth video data files were aligned by the recorded vertical jumps performed first by each individual. A total of 329 right leg kicks and 47 left leg kicks were extracted for analysis. Kick detection results and numbers were recorded for each participant and collated to calculate the i) overall true positive, ii) true negative iii) false positive, iv) false negative, v) sensitivity, vi) precision, vii) F1 score, viii) accuracy percentage, ix) error rate results.

Analysis in study Part Two involved first determining the analogue accelerometer axes sensitivity offset; the ratio of change from device input to output signal. This was required to convert the raw analogue accelerometer data from device arbitrary bit values to *g*-force values. Once the data were converted to *g*-force, the accelerometer magnitude vector was also calculated for each file. Participant kick trial files were then run through the same adapted peak feature detection script (Duarte, 2014) as part one for each axis and the magnitude. The peak accelerations from each axis for the ten maximal kicks within the recorded protocol were extracted and exported to a Microsoft ExcelTM spreadsheet.

7.3 Results

The group mean results for Part One detection of right and left leg kicks are shown in Table 7.2. Detection of right leg kicks showed acceptable precision and robustness with an F1-score of 0.96. Results for left leg kicks were lower, showing a 51% accuracy rate, reasons discussed in the subsequent section.

Table 7.4. The mean group results of kick detection on the right and left legs.

Kick leg	Sensitivity	Precision	F1-score	Accuracy %
Right	0.96	0.95	0.96	92.0%
Left	0.91	0.53	0.63	50.9%

The accelerometer maximum impulse values from Part Two in *g*-forces for kick foot-to-ball point of impact are shown in Table 7.3. The maximum accelerometer magnitude calculated was 209.1 g. An example signal trace of a drop punt kick is presented in Figure 7.3-1 divided by each axis.

Table 7.5. Drop punt maximum impulse accelerometer values for point of impact.

Axis	Maximum impulse (± *g*-forces)
X	171.2
Y	99.3
Z	150.6

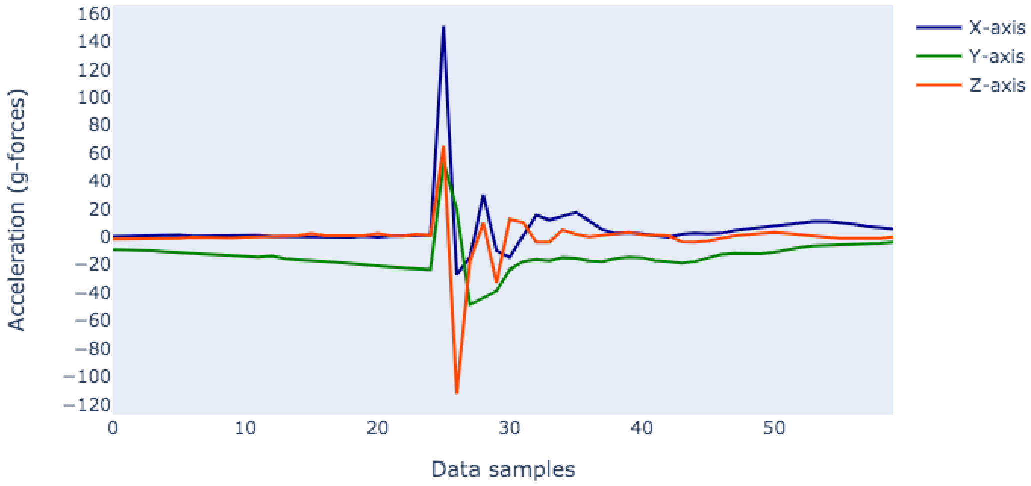

Figure 7.3-1. IMU signal plot traces a maximal drop punt kick. The plot y-axis represents the *g*-force values.

7.4 Discussion

The main aim of this study was to investigate the feasibility for implementation of a semi-automated AF kick recognition system under on-field controlled and semi-controlled training environments. Results flagged discussion around three main points: 1) data acquisition and data saturation, 2) sensor placement in relation to the kick motion, 3) practical applications in AF and future research investigations required in the field. These three points will be highlighted below and further discussed in Chapter Nine. The investigations in this study helped to inform the work in Chapter Eight by refining the location of the IMU on the kicking leg for on-field application, inform of the kick protocol appropriateness for a semi-controlled environment, and adapt the kick detection methods using the IMU data for the kick type recognition undertaken in the subsequent chapter.

The kick detection in Part One, showed reduced accuracy on non-dominant left leg kicks, although the higher sensitivity indicates the ability to identify all relevant instances. But the precision of the detection may suggest higher false positive occurrences. Reasons why the left leg kick detection is much lower were observed through further visual inspection of the identified instances. As all participants were right leg dominant, there were different movement patterns during the kick phases identified on the non-dominant left leg including, forefoot drags during the kick follow through, sharp landing on the left stance leg from a jump-up during right leg high velocity kicks, and forcefully landing the left leg on the final step into a right leg kick. Poor discrimination of kicks may also be related to the sensor accelerometer limitations of ± 16 *g*. Visual inspection confirmed data saturated at 156.9 m/s^2. As such, peak detection may not be able to differentiate between movements and kicks when data is saturating due to the complete signal range of each unique kick style being unknown. To quantify kick variables for athlete skill monitoring applications in AF, non-saturated complete signal data from IMUs may be of greater benefit for accurate analysis applications. Complete signal data without mathematical reconstruction of saturated signals could improve accurate kick detection, kick type classification models, and model feature engineering. The true accelerometer peak and subsequent windowed data around the detected kick could also be investigated in developing a kick intensity load measure for training-based applications. From the current investigations it is recommended that accelerometers being used for the purposes illustrated in this study should have a range of ± 200 *g*. This range is suggested based on the saturation seen in Part One and the maximum acceleration values shown in Part Two of the pilot studies. It may be possible that a 200 *g* accelerometer might not capture a complete drop punt kick signal from a highly skilled and powerful male AF athlete, but it is hypothesised that the majority of the signal would be captured to provide enough data for signal accurate feature extraction of kick types.

It is recommended that IMUs for AF kick analysis be placed on the lateral lower limb above the malleolus for data specificity regarding the movement. See Chapter Two section 2.3 for further discussion around inertial sensor placement and issues arising in the research literature for different movement and data applications. Mounting the sensor on the front of the lower shin may cause unwanted false positive data spikes if the ball hits the sensor in the event of a miss-placement during foot-to-ball contact in a kick; which occurred during these development studies. Also, having the sensor posteriorly on the back of the calf or calcaneal tendon may see issues created from the increased muscle and tendon movements. This could cause increased soft tissue artefacts or oscillations between the sensor and skin leading to measurement error (Camomilla et al., 2017).

The ideas presented in this pilot study and how further investigations addressing the issues identified could create potential for IMUs in the football industry for:

- Hardware specifications and data capture decisions: by determining the value ranges for kick types in a football code may allow for data capture decisions to be made to suit the purpose of the analysis application.
- Kick detection and classification: IMU data presents an opportunity as a potential practical and semi-automated method for on-field skill analysis of kick detection and classification via machine learning techniques.
- Improved training kick skill reporting: semi-automated methods for kick skill recognition could reduce the manual work and human error associated with notational

analysis. Manual notational analysis in AF involves considerate human resources for recording and detailed data labelling input; furthermore the process still presents issues in its validity and reliability (Barris & Button, 2008; Duthie et al., 2003). Developing more efficient methods of capturing skilled actions in AF would allow for the talent resources to be working in other performance analysis areas rather than the associated manual analysis.

- Kick skill analysis: implementing a semi-automated kick recognition system may access greater insights into the kick numbers and potentially types performed during training. This improvement could be beneficial for individual athlete kick development and training designs to replicate match demands.
- Kick load monitoring: by accessing greater insights, quantification of kick numbers could also provide benefits in athlete management for lower limb injuries and load control on return-to-play training programs.

7.5 Conclusion

The opportunity to extend on the potential of IMUs as a semi-automated method for on-field kick skill analysis in AF appears feasible. However, further research is required to adapt and validate this concept for practical operation in an AF program. Improved sensor hardware and analysis methods for IMU data AF kick detection and subsequent recognition applications will aid in developing this area. The subsequent Chapter Eight, uses the same commercial IMU device as Part One. Although, the IMU was applied in an on-field AF kick type protocol to assess the applicability of machine learning for sensor-based kick feature pattern recognition under three model condition types.

Chapter Eight: Classification of Australian football kick types in-situation via ankle-mounted inertial measurement units

This manuscript is currently under review in the Journal of Sports Sciences, and is copied from the original article submitted.

8.1 Abstract

The utility of inertial measurement units (IMUs) for sporting skill and performance analysis during training and competition is proving advantageous in enhancing the objectivity of athlete monitoring. This study aimed to classify Australian Rules football (AF) kick types in an applied environment using ankle-mounted IMUs. IMUs and video capture of a controlled protocol, including four kick types at varying distances, were recorded during a single testing session of female AF athletes (n = 20). Processed IMU data was modelled using support vector machine classifier, random forest, and k-nearest neighbour algorithms under a 2-Kick, 4-Kick, and kick distance (10, 20, 30 m) conditions. The random forest model showed highest results for overall classification accuracy (83% 2-Kick and 80% 4-Kick), test F1 score (0.76 2-Kick and 0.81 4-Kick), and AUC score (0.58 2-Kick and 0.60 4-Kick). Kick distance classification showed a model test and class weighted F1-score of 0.63 and overall accuracy of 64%, respectively. This study highlights the applied potential for semi-automated AF training kick detection and type classification monitoring using IMUs.

8.2 Introduction

Analysis of sport skilled performance via wearable technologies provide several advantages for athlete monitoring during training and competition. One such example is the use of inertial measurement units (IMUs) for sporting skill detection and movement performance assessment (Camomilla et al., 2018; Cust, Sweeting, Ball, & Robertson, 2019). The effectiveness of applied sport technology is reliant on feedback timing, feedback type, and the performance skill measured (Phillips et al., 2013). Furthermore, methodologies used in developing an activity recognition system using IMU data can impact overall recognition performance and differs, depending on the activity and sensor data type (Bulling et al., 2014; Cust, Sweeting, Ball, & Robertson, 2019).

Wearable IMUs and Global Positioning Sensor (GPS) devices are common in sport science as tools to collect biomechanical and spatiotemporal metrics for practical applications including athlete workload measures for return-to-play decision making and detection of head impacts indicating potential concussions (Seshadri et al., 2019). Further practices could extend to using IMUs as a specific skill coaching aid (Wang et al., 2018) or sport movement performance evaluations (Brock & Ohgi, 2017). The use of IMUs in baseball training demonstrates the potential as a monitoring tool for throwing workloads, in relation to arm injury (Mehta, 2019). In AF, kicking load is presently unable to be separated from other loads experienced by the athlete such as running (Boyd et al., 2013; Clarke et al., 2018). Therefore, IMUs may have the potential to be a tool for monitoring lower limb loads, including kicking and running.

Successful kick performances in AF matches by individual athletes are a strong contributor towards quarter and overall match wins (Black et al., 2018a; Cust, Sweeting, Ball, Anderson, et al., 2019; Robertson, Gupta, et al., 2016). Semi-automated AF kick detection in training situations may provide greater insight into athlete monitoring, compared to the current largely manual training performance analysis practices. To date, the use of IMUs in AF kicking has only focused on a mechanical kicking leg and the detection of ball contact, release, and intensity measures (Ellens et al., 2017). Identification of kick ball contact and release phases resulted in a 43% classification accuracy (Ellens et al., 2017). A linear regression identified footspeed intensity values from the maximum kick accelerations in the x and y axes , suggesting it may be possible to further progress a kick intensity measure through accelerometer data (Ellens et al., 2017). However, this testing was conducted in a controlled environment using only accelerometer data, and on a purpose designed mechanical leg that is not able to move in the frontal plane (Ellens et al., 2017). Accelerometers have also been utilised in an applied field soccer kick protocol. In this work, classification models using accelerometer and gyroscope data were trained and tested for a controlled kick protocol, before being tested on a small sided 11 vs. 11 uncontrolled match (Schuldhaus et al., 2015). A support vector machine algorithm demonstrated the highest results in classifying leg type, kick or other event during the controlled protocol (Schuldhaus et al., 2015). Mean classification performance results under the field match conditions were 89.5% for evaluation of other events such as tackling, fast running, and side steps; and 84.2% when classifying identified kick events as either a pass or a shot (Schuldhaus et al., 2015). The investigation of IMUs in an applied AF kicking protocol has yet to be tested with AF athletes.

The aim of the present study was to evaluate Australian football kick type classification models in an applied on-field environment using ankle-mounted IMUs. Specifically, the applicability of machine learning for sensor-based kick feature pattern recognition under three model condition types was assessed.

8.2 Methods

8.2.1 Participants

Twenty female Australian Rules football athletes (age range: 21 - 29 years) participated in the study, ten of whom were contracted to one AFLW team, and ten playing in the partnered State-based women's football team (Victorian Football League Women's, VFLW). All participants provided signed informed consent before proceeding with the research protocol. The University's Human Research Ethics Committee approved the study (application number 25654).

8.2.2 Design

The study design was a cross-sectional protocol collecting data during two testing session which were then processed for algorithm training and testing.

8.3 Methods

IMU and video capture of a controlled kick protocol were recorded during a single testing session. The IMU (IMeasureU BlueThunder sensor, Auckland, New Zealand), composed of a 3D accelerometer (± 16 *g*), 3D gyroscope (± 2000°/s), and 3D magnetometer (± 1200 μT) internally sampling at 500 Hz (Parrington et al., 2016; Whiteside et al., 2017). The IMU was mounted on the athletes' dominant kicking leg lower shank, superior to the lateral malleolus (Ellens et al., 2017). The sensor weighs 12 *g* and was strapped on using the manufacturer's recommended Velcro band. This commercial IMU has previously been used in sport-specific movement detection or recognition applications (Campbell et al., 2018; Ellens et al., 2017; Parrington et al., 2016; Walker et al., 2017; Whiteside et al., 2017). Recording control of the IMUs was undertaken using the manufacturer's research application on an iPad (Apple Inc., Cupertino, USA). The video camera (Sony, Tokyo, Japan) was elevated and positioned to capture the whole testing area. Two athletes undertook the protocol concomitantly in order to kick to each other. Prior to starting the first kick participants performed three vertical jumps then stood still in order create a defined spike in the data for synchronisation of the IMU and video data for analysis of true kicks. The protocol involved drop punts at 10, 20, and 30 m followed by snap kicks aiming for 20 m, grubber kicks for 10 m, and surge kicks for maximum height and distance (Figure 8.3-1). Further IMU data was taken as part of another separate test protocol where athletes individually performed drop punts kick repetitions at 10 m, 20 m, and 30 m at varying prescribed intensities. Camera vision was manually coded to create a timestamped log of instances for athlete, kick type, and kick distance by an experienced performance analyst using SportsCode Elite (Agile Sports Technologies, Inc., Lincoln, Nebraska). The log of instances for each athlete were exported to Microsoft Excel™ workbooks.

Pre-processing of the IMU data involved the accelerometer and gyroscope signals. All pre-processing was completed individually for each athlete's data using a single programmed script (Python Software Foundation, 2018). The data were filtered using a 50-point moving average technique (Whiteside et al., 2017). The magnitude vectors for both signal types were calculated, creating eight signal data types of x-, y-, z-axis and magnitude vector for accelerometer and gyroscope data. The accelerometer y-axis signal was used to parse for detecting kick instances in the continuous data signal. An external peak detection script was called to pass through the data and print the timestamps of identified peaks based on a defined threshold on the rising edge side of a peak, peak height, and not to detect a kick for two seconds after another. Identified peak timestamps were used to create windows of 250 data samples (0.5 seconds) each side of the peak across all eight signal types. Feature extraction on each 500 data sample of kick instances was undertaken yielding the mean, median, standard deviation, variance, minimum, maximum, skewness, kurtosis, and integral based on the Trapezoidal Rule (Kautz et al., 2017; Schuldhaus et al., 2015; Whiteside et al., 2017). The final labelled Pandas DataFrame (McKinney, 2010) of nine kick features for eight signal types, kick type, and kick distance was exported as a CSV file for each athlete's data. The IMU kick instance feature file and video instance log files were aligned and timestamps adjusted using the first vertical jump. True context kicks were extracted and collated in a single master CSV file. A total of 587 true kicks (drop punt: 461, grubber: 25, snap: 53, surge: 48) were extracted and collated with the associated signal features calculated.

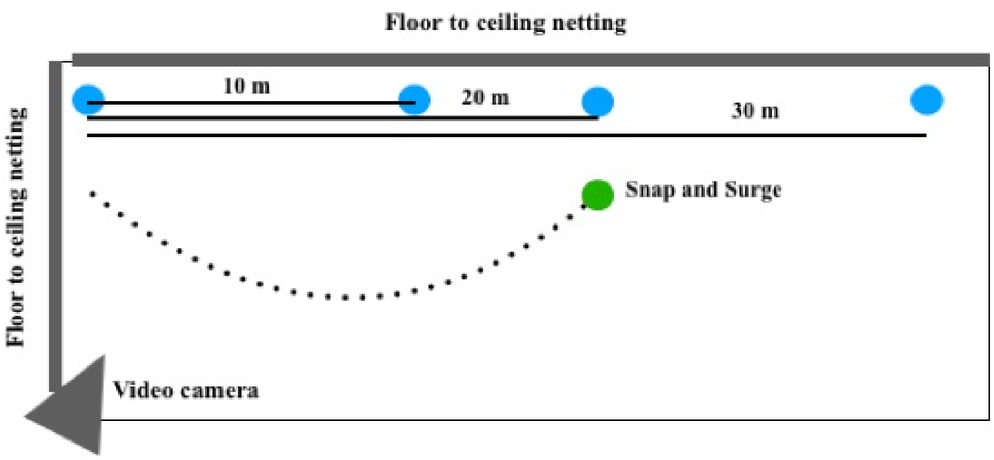

Figure 8.3-1. Kick protocol schematic.

8.3.1 Statistical analysis

Base model testing was run using seven machine learning classification algorithms. The default Scikit-learn hyperparameters (Pedregosa et al., 2011) were used to train support vector machine classifier (SVM), random forest (RF), k-nearest neighbour (KNN), decision tree (DT), multilayer perceptron (MLP), adaptive boosting classifier (AB), and gaussian naïve Bayes (NB) on the IMU features and kick classes. Data was segmented into a train-test split of 70% and 30%, previously used in lower samples sport movement recognition (Wang et al., 2018). Evaluation was by the weighted F1-score as a single evaluation metric for the balance between precision and recall on each class, which is preferred over classification accuracy that tends to underestimate the ability of the classifier on the smaller classes under class imbalance (Forman & Scholz, 2010). Results of F1-scores were: DT 0.73, KNN 0.72, RF 0.70, SVM 0.70, AB 0.69, NB 0.66, MLP 0.56. The SVC, KNN, and RF showed the highest results, after the DT which was not used due to its disadvantages in comparison to RF (Breiman, 2001a; Hastie et al., 2009). These algorithms have been used in similar sport-specific movement recognition achieving high results (Cust, Sweeting, Ball, & Robertson, 2019). Two conditions for training were defined: a 4-Kick multi-class classification of 1) drop punts, 2) surge, 3) grubber, 4) snap kicks, and a 2-Kick binary classification of 1) drop punts and 2) other kicks. The three classifiers were modelled, tuned, and tested separately for both kick class conditions. Data was standardised using a robust scaler, due to its use of IQR of data to scale and is robust to outliers. Then data was split into train-test sets of 70%/ 30%, and the training set further split into train/ validation sets using a stratified K-fold cross-validation (10-folds) for model training (Buckley et al., 2017; Chambers et al., 2019; Connaghan et al., 2011; Whiteside et al., 2017). The test set was withheld for final model testing. Classifier hyperparameters were individually tuned with a Bayesian optimiser (Bernardo & Smith, 2009; Snoek et al., 2012). Tuned model evaluations were trained and validation tested using F1-score. Final model testing on the withheld testing data set was evaluated using test and weighted F1-score, accuracy score percentage, confusion matrix metrics including precision and recall, and the area under curve (AUC) score. The weighted average calculates the metrics of each class and finds the average weighted by the support factor to account for any class imbalance in comparison to a macro average or treating all classes equally with a micro average. The AUC was included as a better alternative single number metric for evaluation of classifiers (Bulling et al., 2014; Ling et al., 2003). Feature importance ranking was calculated for the RF and feature extraction by taking the lowest contributing feature out of each iteration until a decrease in model performance occurred.

As the RF model showed the higher performance it was used to test the classification of kick distance classes 10 m, 20 m, 30 m for drop punt kicks only, following the same methods as described above. Data kick instance numbers in each class were: 976 for 10 m, 1696 for 20 m, and 1016 for 30 m.

The effect of class imbalance was investigated using over-sampling methods with the RF models. Three over-sampling methods were applied: naïve random over-sampling (ROS), synthetic minority over-sampling (SMOTE), and adaptive synthetic sampling (ADASYN) (Lemaître et al., 2017). Random over-sampling over samples existing minority classes with replacement, whereas SMOTE over-samples the minority class by creating synthetic examples (Chawla et al., 2002). The ADASYN adaptively creates minority data samples based on their distribution therefore more synthetic data is created in the harder to learn minority class samples than the easier to learn minority samples (He et al., 2008). There will be a small imbalance in the classes due to how ADASYN creates new data points from the minority classes according to the weighted distribution of their difficulty to classify (He et al., 2008). Over-sampling was applied to the training dataset features. Performance evaluations using the metrics previously listed were taken on the trained model using each over-sampling method. The highest performing method was further tuned using a Bayesian optimiser and re-tested using the withheld testing data set.

All experiments were run on an Apple MacBook Pro with an Intel Iris Plus Graphics 640 (2.3 GHz Intel Core i5), 8 GB RAM, and 64-bit operating.

8.4 Results

Figure 8.4-1 visually displays the evaluation metrics of each model for both kick conditions. The RF model showed higher results across both kick conditions for overall classification accuracy, Test F1 score, and AUC score. The KNN model had considerably higher positive recall and precision for under the 4-Kick condition, 0.67 precision and 0.78 recall. The train and test F1 scores for each model under the 2-Kick condition were: KNN train 0.70 and test 0.70; SVM train 0.74 and test 0.74; RF train 0.78 and test 0.81. The 4-Kick condition: KNN train 0.70 and test 0.70; SVM train 0.74 and test 0.72; RF train 0.72 and test 0.76.

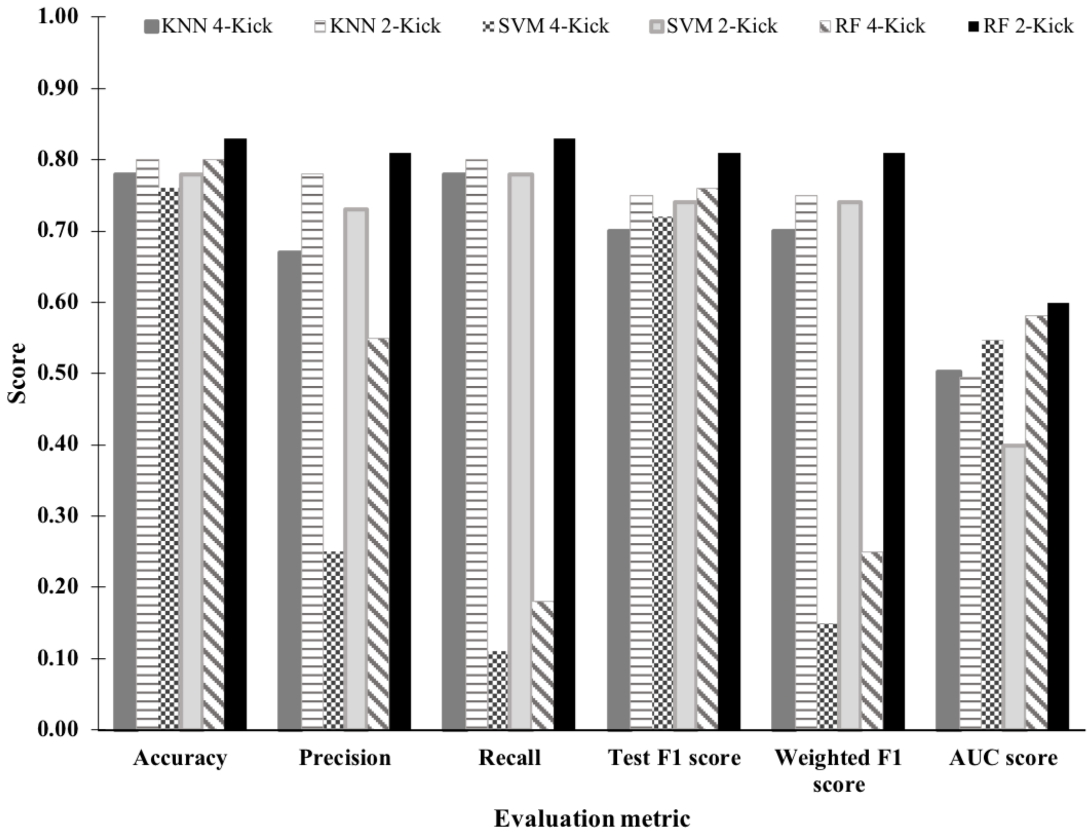

Figure 8.4-1. The comparison of model performances for kick classification in the 4-Kick and 2-Kick class conditions.

Reduced performance, based on the confusion matrix and evaluation metrics, were observed when distinguishing between the four kick types compared to the binary recognition as a direct function of more degrees of freedom in the larger class models; Tables 8.1 and 8.2. The RF under the 4-Kick condition was unable to detect each class well but could handle the drop punt class moderately well when detected, as shown by the higher precision scores and low recall scores; Table 8.1. The confusion matrix showed that out of the 1108 actual drop punt instances, the RF predicted correctly 1081 (97.6%) of them, and had the highest false positive rate testing positive for a drop punt but was a snap kick in 86 instances (7.8%). The RF under the 2-Kick condition handled the drop punt class well showing very high identification of drop punts from positive kick types with a recall of 0.95; Table 8.2. Of the

1108 actual drop punt instances, the RF predicted 1055 correctly (95.2%). The lower recall of the other kick class, 0.38, is evident as of the 301 other kick instances the RF predicted 115 correctly (38.2%) and tested negative for other kick when it actually was a class other kick class for 186 instances (Table 8.2).

Table 8.1. Confusion matrix and performance metrics achieved by the RF in the 4-Kick class condition.

Actual	Predicted				Precision	Recall	F1-Score	AUC score	Support
	Drop punt	Grubber	Surge	Snap					
Drop punt	1081	1	15	11	0.67	0.03	0.06	0.61	59
Grubber	56	2	1	0	0.46	0.23	0.31	0.52	116
Surge	83	0	27	6	0.59	0.19	0.29	0.60	126
Snap	86	0	16	24	0.00	0.00	0.00	0.59	0
				Weighted average	0.55	0.18	0.25	0.58	301

Table 8.2. Confusion matrix and performance metrics achieved by the RF in the 2-Kick class condition.

Actual	Predicted		Precision	Recall	F1-Score	Support
	Drop punt	Other kick				
Drop punt	1055	53	0.85	0.95	0.90	1108
Other kick	186	115	0.68	0.38	0.49	301
		Weighted average	0.81	0.83	0.81	1409
		AUC score	0.60			

Feature extraction on the RF models under both kick class conditions only showed marginal improvements in results. The only improvements made included in the 4-Kick condition by removing the two lowest features of standard deviation and mean, where precision increased 0.07 points, recall increased 0.01 points, and F1-score increased 0.02 points. In the 2-Kick condition by removing the first lowest feature being variance, the precision increased 0.01 points. Removal of further features caused decreases or no improvements in performances for both conditions in the RF model. Feature importance values for both RF models are shown in Figure 8.4-2.

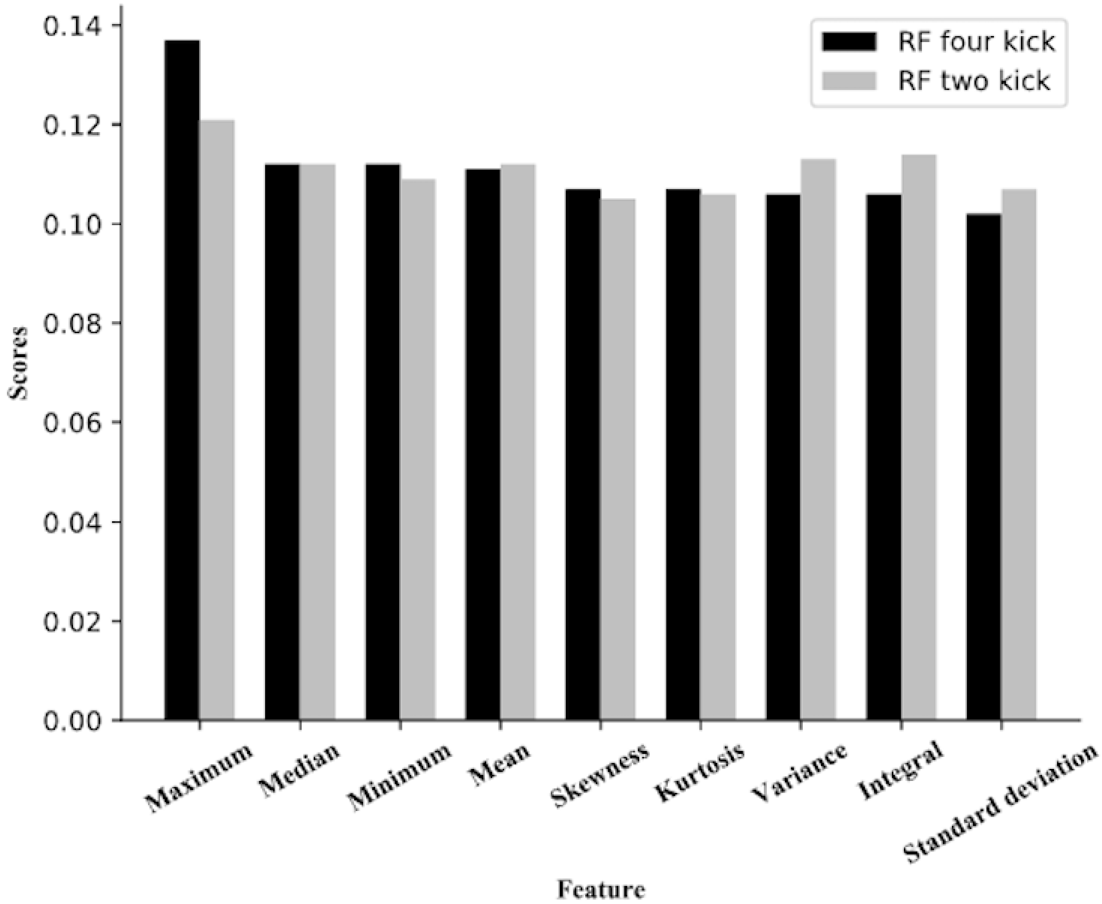

Figure 8.4-2. Standardised feature importance for random forest models in both kick conditions.

A class imbalance existed within the data with 78.5% instances reported as a drop punt. In the 4-Kick class condition, the ADASYN over-sampling method showed the best evaluation performance metrics; Table 8.3, and rebalanced and increased the y-training dataset originally from 3287 samples (drop punt: 2581; grubber: 140; surge: 269; snap: 297) to 10374 samples (drop punt: 2581; grubber: 2601; surge: 2561; snap: 2631). The model's ability to classify all kicks (accuracy score) decreased with oversampling. Although the recall ability greatly improved; weighted recall increased from 0.18 to 0.45, and the AUC increased from 0.58 to 0.70. In the 2-Kick class condition ROS was the best performing method; it rebalanced and increased the y-training dataset originally from 3287 samples (drop punt: 2581; other kick: 706) to 5162 samples (drop punt: 2581; other kick: 2581). Accuracy score also decreased slightly from 83.0% to 82.6%, and the AUC slightly rose from 0.60 to 0.62. The ability of the model to classify the minority class (recall) of ‘other kicks’ improved from 0.38 to 0.54; Table 8.4. Slight increases in precision and decreases in the recall and F1-score of the major class ‘drop punts’ showed a small improvement in the weighted F1-score from 0.81 to 0.81; Table 8.4.

Table 8.3. Confusion matrix and performance metrics achieved by the RF in the 4-Kick class condition rebalanced using ADASYN over-sampling method.

Actual	**Predicted**				**Precision**	**Recall**	**F1-Score**	**AUC curve**	**Support**
	Drop punt	**Grubber**	**Surge**	**Snap**					
Drop punt	868	72	73	94	0.28	0.48	0.35	0.75	60
Grubber	28	29	1	2	0.32	0.46	0.38	0.71	115
Surge	26	1	53	35	0.29	0.43	0.35	0.69	127
Snap	31	2	40	54	0.00	0.00	0.00	0.66	0
			Weighted average		0.30	0.45	0.36	0.70	302

Table 8.4. Confusion matrix and performance metrics achieved by the RF in the 2-Kick class condition rebalanced using Random over-sampling method.

Actual	**Predicted**		**Precision**	**Recall**	**F1-Score**	**Support**
	Drop punt	**Other kick**				
Drop punt	1001	106	0.88	0.90	0.89	1107
Other kick	139	163	0.61	0.54	0.57	302
	Weighted average		0.82	0.83	0.82	1409
	AUC Score		0.62			

Post-hoc kick distance recognition assessment showed a model test and class weighted F1-score of 0.63 and overall accuracy of 64%. The RF was able to correctly identify 20 m kicks the best, recall 0.73. The model found it harder to identify all positive samples completely in 10 m and 30 m kick classes, 0.61 and 0.51 recall score respectively. The ability to accurately predict kick distance type was low across all classes indicated by the poor precision scores; Table 8.5. Although the AUC scores are moderate suggesting model overall prediction ability potential to differentiate between positive and negative classes; Table 8.5. The model did present slight overfitting with a train F1-score of 0.644, higher than the test score.

Table 8.5. Confusion matrix and performance metrics achieved by the RF for kick distance classification.

Actual	**Predicted**			**Precision**	**Recall**	**F1-Score**	**AUC score**	**Support**
	10 m	**20 m**	**30 m**					
10 m	188	108	10	0.66	0.61	0.64	0.75	306
20 m	72	364	66	0.61	0.73	0.66	0.67	502
30 m	24	122	153	0.67	0.51	0.58	0.71	299
		Weighted average		0.64	0.64	0.63	0.71	1107

8.5 Discussion

This research developed multiple kick classification models using data collected from ankle-mounted IMUs with the intention of improving on-field kick monitoring. The performance of the RF model demonstrated higher evaluation metrics for the 2-Kick model and more favourable metrics in the 4-Kick model in comparison to the KNN and SVM algorithms. Overall, the findings indicate the potential for semi-automated AF training kick detection and type recognition monitoring using commercial IMUs.

The 4-Kick RF model had classification accuracy of 80%; but performed poorly on other metrics. This result compares relatively well to other sport skill recognition multi-classification problems, achieving accuracies between 80-90% (Brock et al., 2017; Groh et al., 2017; Holzemann & Van Laerhoven, 2018). For example, soccer on-field kick type classification using IMUs showed a mean class-dependent classification of 81.7% for passes and 86.7% for shots (Schuldhaus et al., 2015). Lower performances on other metrics may be due to the class imbalance as accuracy as an evaluation metric tends to undervalue how well an algorithm is doing on smaller classes (Forman & Scholz, 2010). The RF 4-Kick model showed very low recall for each class indicating trouble in correctly predicting the identified positive instance to the correct class. High miss-classification of surge and snap kicks; Tables 8.1 and 8.3, may be due to the similarities in how these technique types are performed by an athlete creating similar data signal features. The AUC score is better for comparing the overall performances of multiple classifiers than accuracy score (Ling et al., 2003), and tells how well the model is capable of distinguishing between classes (Hanley & McNeil, 1982). Where AUC is 0.5 for model prediction are almost random and an AUC of 1.0 represents a perfect classifier. When the AUC is approximately 0.5, the model has very little discrimination capacity to tell between different classes. The RF showed the best AUC, meaning there is a 60% (2-Kick) and 58% chance (4-Kick) that the model will distinguish between the classes. The KNN had no class separation capacity in both models, and the SVM had no capacity in the 2-Kick model and very slight (55%) in the 4-Kick. In terms of model class conditions and evaluation metrics relating to kick recognition, having a higher recall (sensitivity) output for the 2-Kick model would be more beneficial in minimising false negative instances where the model has predicated no kick when there actually was, therefore missing kick instances too often and not providing an accurate overall kick volume measure. The 4-Kick model would require higher precision in limiting false positives in order to prevent miss-classification of kick instances to other kick type classes.

Feature extraction on the RF models under both conditions did not result in marked improvements. The 4-Kick model saw slight improvements to precision, recall and AUC score after removing two of the lowest contributing features. This may suggest that all nine features are important to recognising kick type signal patterns for each class. More features extracted from the data such as the energy, correlations between axes, and percentiles (Kautz et al., 2017; Wundersitz, Josman, et al., 2015), or derived from other features may be required to improve modelling results. The degrees of freedom in a model are the independent variables on which the target depends on, those that are free to vary without impending the constraints of a model; $degrees\ of\ freedom = N - 1$. Knowing which variables the model depends upon is important in reducing model complexity, a common cause of overfitting, and can inform feature selection processes. Dimensionality reduction by feature selection methods to find the best representation of data in lower-dimensional spaces (Mannini & Sabatini, 2010) would be of benefit in larger datasets to minimise the bias-variance trade-off between setting the degrees of freedom. There was overfitting present, where performance is better on the training data than on the test data, in the 4-Kick KNN and SVM models, and the RF distance recognition model. Although only slight, the problem could be fixed by having more training data for each kick class. It is also considered that flexible machine learning models with more tuneable constraints are expected to produce improved re-substitution performance than less flexible models (Hawkins, 2004). Using oversampling did indicate improved ability to distinguish the minority classes, albeit slight, in the original dataset and may be considered a method for rebalancing data under repeated practical use in AF training sessions where it is likely the drop punt kicks are more frequently performed.

Initial kick distance recognition showed moderate precision for all classes and weighted averages for recall and F1-score suggesting that class detection is not well handled but has the potential to be trusted when the accurate prediction is made by the model. Classing kick distances by meters rather than m/s is more practical for coaches in adaption of short and long kicks to suit training drill designs for match play representation. Although, drop punt distance band recognition by signal feature pattern regularities may not be appropriate for Australian Rules due to the large variations in individual athletes foot-to-ball impact characteristics that define the kick (Peacock & Ball, 2018a, 2018b). Also the knee or thigh dominant kicking strategies adopted by athletes causing changes in pre-impact foot velocities (Ball, 2008; Cust, Ball, et al., 2019). Further investigation on a larger dataset including maximal kick efforts would be required to asses different feature engineering methods and variables such as the foot velocity through the kick phases, that could be adapted to identifying signal characteristics that define distances independent of the athlete kicking. Also, looking at the problem from a regression rather than a classification prediction approach may provide improved results.

The results from this study support a rationale for building several models types that will serve a different utility in an applied field setting. A basic kick detection model to identify actual kick instances from a continuous stream of IMU signal data could provide coaching information for example, changes in kick numbers for drills as the constraint of the drill are changed. A 2-Kick classification model (drop punt or other kick), could be applied in more sophisticated training monitoring to inform coaches on athlete skill performances and development across training session. The 4-Kick model may be more applicable for medical monitoring in quantifying kick rehabilitation progression for an athlete. This study presents a progression from previous work in using IMUs for AF kicks by Ellens et al. (2017) and an adaption from soccer kick type recognition work (Schuldhaus et al., 2015). In Ellens et al.

(2017), the same sensor was implemented on a mechanical fixed kick leg, although they did not include gyroscope data, and the model was not tested for in-situation AF kicking. A limitation of using the mechanical leg was that the z-axis of the accelerometer data was redundant due to the machine not being designed to move in the frontal plane (Ellens et al., 2017). The metrics derived from the z-axis of the foot during the foot-to-ball impact phase are important when determining the impact characteristics in AF drop punt kicks (Peacock & Ball, 2018b, 2018a). Following the work of Schuldhaus et al. (2015), progression of this current research could cover the recognition of the kicking and support leg if the athlete is wearing an IMU of each leg. Also testing the developed system during in-situation match simulation or small sided games.

The IMUs used were limited to ± 16 *g* and ± 2000°/s for the accelerometer and gyroscope respectively, which saw saturation in the kicking signal due to the high velocities and forceful impact of the foot in contact with the ball. This may have limited gaining representative data features for each kick type, reducing the ability of the algorithms to distinguish between kicks. Future work in this area would look towards developing a more tailored AF kick instance detection method that could create more automation in deducing kicks from a continuous stream of IMU data. A larger dataset would work in favour of assessing the use of deep learning algorithms which have the advantage of automated feature extraction from raw data in comparison to machine learning algorithms (Bengio, 2013; Lecun et al., 2015). Volleyball skill recognition using wrist worn IMUs have shown higher performances using deep learning methods in comparison to machine learning (Kautz et al., 2017), also in golf swing classification (Jiao, Bie, et al., 2018). Finally, the data was collected from female AF athletes; as kinematic differences have been found between elite male and female AF drop punt kicks (Ball, 2008, 2011; Cust, Ball, et al., 2019), further investigation incorporating male kick data should be undertaken to determine if the data source affects the recognition of kick types.

8.6 Conclusion

This study investigated the application of IMUs for on-field AF kick recognition providing a proof of concept for the advancement in this area towards an applied training volume monitoring tool. Specifically, it supports the notion of sport specific skill recognition in showing that kick types can be distinguished from one another and how this differs from a binary and multi-classifier model. The current research presents a progression from previous work and recommendations for advancement in this application for training implementation purposes in AF.

Chapter Nine: Discussion

This thesis firstly investigated applications of IMUs and various analytical methods to achieve sport-specific movement recognition, including classification models for AF kick type recognition using IMU data. Chapter Three presented a systematic review on the research methods of IMUs and computer vision for semi-automated sport movement recognition in on-field sporting contexts. The review demonstrated the practical application potentials for IMUs or computer vision technology as performance analysis tools. Furthermore, key areas of considerations were recommended for sport scientists around the adaption of data collection, data pre-processing, and model development methods in relation to the characteristics of the targeted sport and movement(s). Chapters Five and Six identified the unique characteristics of skilled performances and kick biomechanics of AFLW athletes. The findings of Chapter Five highlighted the strong influence of kick skill performance indicators towards team match success. As kicking was highlighted as a key AF skill for elite female athletes, further research was undertaken focusing on drop punt kick properties in the AFLW. Thus, Chapter Six quantified women's drop punt kick kinematics, recording differences in the movement patterns between dominant and non-dominant kick legs and also how AFLW kick movement patterns differ from their male counterparts. The results of this study have implications in skill training strategies for AFLW athletes as discussed below. Finally, Chapter Eight investigated the concept for how IMUs could be applied for AF kick recognition from the evidence collated in Chapters Three and Seven. The research in Chapter Eight was centred around an on-field application with the aim in progressing current AF training performance analysis data capture processes; a currently largely manual process.

Statistical and biomechanical analysis of athlete's skills have been pivotal for characterising physical, tactical, technical, and contextual factors across training and competition that: 1) contribute to achieving high performance, and 2) show differences between the performance aspects of men's and women's competitions. The new league format of the AFLW provides grounds for the skilled performances of elite female AF athletes to be understood separately from men's AF; as shown from the findings in Chapters Five and Six. This approach may help in progressing the AFLW towards a professional competition through women's specific skill coaching, potentially leading to improved athlete development pathways to cater for the expanding competition. The findings of Chapter Five are in-line with similar research on the skilled match performances in elite women's sport showing their unique characteristics in comparison to the an equivalent men's analysis, for example in AF (Clarke et al., 2018), basketball (Gómez et al., 2013), water polo (Gómez et al., 2014), and soccer (Pollard & Gómez, 2014). As an example in basketball, the stadium location of the match, the score lines at the end of each quarter, and the opposition quality had greater effects on women's basketball dynamics compared to the men's competition (Gómez et al., 2013). Similarly to the findings of Chapter Six, biomechanical differences have been found between men's and women's soccer kicks (Barfield et al., 2002; Navandarl et al., 2016; Sakamoto et al., 2014) and taekwondo kicks (Kazemi, Waalen, Morgan, & White, 2006; Li, Van, Zeng, & Wang, 2005; Pieter & Pieter, 1995). Several studies undertaken in women's soccer kicks (Alcock, 2010), highlighted important differences in defining coaching cues for women's kick technique based off their kinematic features. Therefore, findings from men's AF may not translate directly into women's AF; hence further independent research is required.

Female AF skill specific research findings are important for gaining greater understandings on individual and team performances in order to tailor coaching strategies. Key contributions from this thesis include that increased match skill performance contributions from key high performing individual athletes were more beneficial for team match success in the AFLW (Chapter Five). Also, changes in biomechanical movement patterns across both kick leg preferences were shown in women AF athletes compared to the findings of male AF kick research (Chapter Six). The drop punt is the main kick performed in AF and there are several technical factors that influence the quality and accuracy of the kick, such as, the foot-to-ball impact kinematics, the support leg dynamics, and the foot velocity prior to ball impact (Ball, 2008, 2013; Ball et al., 2013; Peacock et al., 2017; Peacock & Ball, 2018a). Understanding the drop punt kick movement patterns in the AFLW (see Chapter Six), may aid in tailoring kick development strategies for female athletes towards improvements in the kick skill performance factors contributing towards team match success. The female footballers involved in the research represented fair diversity across field playing positions, AF experience, age, and cross-sporting code competition experience. Higher foot velocities were present on the preferred leg which does have a strong correlation to ball flight distance. Considering that in matches being able to kick efficiently on both legs in a tactical advantage, addressing any specific strength deficiencies related to the kicking motion for the non-preferred kick leg would be a beneficial training strategy alongside normal skill practice. Furthermore, as it was shown that key high performing player's skills in AFLW contribute most to team overall match quarter success. Therefore, boosting the team's kicking performances would aid tactically in maintaining possession of the ball in matches and executing successful inside 50 passages which are strong contributing performance indicators in both AFLW and men's AF (Chapter Five). Separate analysis on dominant and non-dominant kick legs allows movement pattern differences to be identified for individual athlete profiles. For example, if an athlete presents with a lower foot velocity and larger hip and thigh segment recruitment patterns on the non-dominant kick leg compared to the dominant leg, it may suggest limitations in the transfer of momentum through the kick phases. As a result, muscle asymmetries may arise from the imbalance of kick numbers and differing movement muscle loading patterns potentially contributing to a higher injury risk and reduced kick skill ability in matches. Discovering the technical parameters of the movement pattern used by the athlete, either a knee strategy or thigh strategy, is important and may affect the skill coaching cues and conditioning training specifically for an athlete but at this stage is not associated with kick performance indicators as to which mechanism is better, if any (Ball, 2008). For example, as referred to in Ball (2008), an athlete who kicks with a hip or thigh dominant strategy will use the hip flexors more to generate foot speed and force, whilst a knee dominant kicker will rely more on the knee extensor muscles. Tailored kick coaching cues and strength programs addressing the identified kick kinematic deficiencies can therefore be implemented for an individual athlete's skill needs. Furthermore, considering the research undertaken in women's soccer kicks (Alcock, 2010; Barfield et al., 2002; Navandarl et al., 2016; Sakamoto et al., 2014), biomechanical analysis of women's rugby codes and Gaelic football kick types may also benefit from developments of skill performance characteristics for improved coaching practices. For example, Alcock et al. (2012) identified the key soccer coaching cues differentiating the ability to achieve accurate curved kicks compared to instep kicks for elite female athletes. The optoelectronic motion and data collection methodology for kinematic analysis used in these soccer studies and Chapter Six of this thesis could be adapted for women's specific kick investigation in other football codes.

Highlighted differences in athlete skilled contributions to team match success in the AFLW (Chapter Five) showed differences to the men's AF competition (Robertson, Gupta, et al., 2016). The analysis methods used in this thesis study could be extended to other field team sports, such as rugby and hockey, to understand any tactical differences between the men's and women's competitions for tailored training design and competition tactics. Also, finding imbalances in athlete skill levels within the AFLW places emphasis on the importance of skill development especially for cross-code and new draftee athletes with less AF match experience. Investigating athlete skilled performances within a sporting team could be important in determining levels of contribution towards team success, as this may impact the effectiveness of match tactical plays. Characterising skilled performances of female athletes during competition could also contribute towards defining attribute milestones for talent development pathways. For example, early kinematic analysis characterised the differences between high performing and less successful female freestyle cross-country skiers at the 1993 Winter Olympic Games (Gregory et al., 1994), and recently the physiological match demands between elite and sub-elite female AF athletes (Clarke et al., 2019). Quantifying skill characteristics of elite female athletes creates a testing benchmark for other athletes transitioning from junior and sub-elite divisions around understanding the performance profiles to work towards for on-field improvements.

Continued skill biomechanical and match performance research in the AFLW is a key research gap as the league continues to expand and the talent pathways broaden. A key area yet to be researched in the AFLW is spatiotemporal analysis of team behaviours which would be beneficial in determining collective team styles of play during different match phases and then again differentiating these from men's elite AF (Alexander, Spencer, Mara, et al., 2019; Alexander, Spencer, Sweeting, et al., 2019). With greater data sources available as the AFLW seasons continue, analysing how situational and contextual variables on skill performances in matches may be of benefit to how teams approach competitions considering the current shortened season format and relative importance on each game on ladder positioning (Oliveira et al., 2012; Ruano et al., 2016).

The capture of skilled actions in team-sport training and matches is commonly done so using notational analysis methods, which are quite laborious in nature. For example, manual video coding of athlete skills and interactions is not only arduous, but the subjective input can lead to human error and bias affecting consistency in reproducible results (Barris & Button, 2008; Duthie et al., 2003). The capture of discrete skilled actions in AF training such as kick numbers is currently manual, and also underutilised due to the increased resources required.

In order to undertake new research on IMUs in sport skill performance analysis, including movement recognition, it is important to assess the current knowledge and how this can be improved or translate to a new problem area. The application potential of IMUs for AF kick recognition aimed towards improved training applications in monitoring performances is currently unknown. Current evidence of the implementation of IMUs and computer vision for sport-specific movement recognition suggests the adaptation of experimental set-up, data pre-processing, and model development methods are best considered in relation to the characteristics of the targeted sport movement(s) (see Chapter Three). Focusing on AF, the use of commercial IMUs for kick recognition and the progression of practices in this application appear feasible as indicated by results from random forest modelling of AF kick features derived from IMU data which showed overall classification accuracies of ≥ 80% (Chapter Eight).

Investigating and progressing IMU skill recognition in team-sports could be integrated with routine GPS data for in-depth athlete training profiles. For example, in soccer and AF, knowing the changes in running and kicking profiles or loads based on the drills undertaken or when changing constraints in small-sided games could further specify individual athlete monitoring practices. Understanding the previous methods for semi-automated skill recognition in individual and team sports may be beneficial in adapting for improvements or implementing in a new sport application. The progression of this work via IMU or vision could also extend from performance analysis to athlete medical monitoring practices such as providing quantified loading information of a sporting skill on an injured area during a return-to-play protocol.

Chapters Two and Three indicated there are several challenges based on each sport that need to be considered when using IMUs or computer vision for sport movement recognition. For example, computer vision in AF is likely to suffer from continuous occlusion problems with the number of players on the field, and face challenges in image quality when covering the large field size (Faulkner & Dick, 2015; Thomas et al., 2017). In comparison to a tennis match between two players where views can be clear when using a fixed camera system (Shah et al., 2007; Zhu, Xu, Huang, Gao, et al., 2006). Overhead vision was used in boxing action recognition to alleviate the occlusion problems (Kasiri-Bidhendi et al., 2015; Kasiri et al., 2017). Moreover, the use of body worn IMUs or vision capture system markers in gymnastics or diving would likely interfere with the athlete's routine and balance; hence a context where vision-based methods are important (Díaz-Pereira et al., 2014; Nibali et al., 2017; Reily et al., 2017). Vision-based methods may also be advantageous in subjective performance evaluation sports (Nibali et al., 2017; Reily et al., 2017), or when differentiating between fine-grained action performances (Kasiri et al., 2017). Within AF, further research on the accuracy of kick recognition using IMUs within an uncontrolled training environment such as match-play would be beneficial. The combination of vision and IMUs; previously investigated in tennis (Conaire et al., 2010), may be a potential method in improving on the current lengthy manual vision coding for AF skill performance output data. With the increases in data availability, deep learning models could potentially be explored in future research with IMU data. The advantages in automated feature extraction may allow for higher-level data representations in identifying prominent and robust features (Goodfellow et al., 2016; Ordóñez & Roggen, 2016). However, deep learning models should be compared to machine learning models, to evaluate overall advantages and avoid arduous computational times. This is particularly important if near real-time deployment of an IMU based recognition system for capturing skilled activities is the aim. Lastly, at the time of data collection for this thesis, a ±16 g sensor was available as a commercial IMU and therefore utilised for kick type classification. Recent improvements in sensor hardware designs have made commercial ± 200 g accelerometer ranges available, therefore further investigation as to whether incorporating the full kick signal data improves the kick classification model ability is required.

The introduction of the AFLW competition and continual integration of technology and analytical methods in sport performance analysis presented knowledge deficits and areas for applied research. In summary, practically the work presented in this thesis firstly provides a reference for sport scientists on the methods and applications of IMUs and computer vision for sport movement recognition. Information on the current uses across several sports and the level of success can inform future practices. For example, highlighting a current proof-of-concept application of IMUs for AF kick recognition allows future developments to be undertaken based on the current feasibility assessment. Secondly, modelling AFLW match performance statistics by quarter outcomes provides specific information for in-game tactics and coaching; particularly demonstrating the importance of game plan development around a kick dominant ball movement strategy. The methods used in the modelling could also be applied across other elite women's team sports to identify the top performance indicators contributing towards match success. Also how individual athlete performances are contributing towards the team; therefore, impacting game tactics, athlete development, and recruitment decisions. Lastly, as biomechanical differences were found between elite male and female AF kicks, which was also identified in soccer and taekwondo research;

investigating if kinematic differences exist for key performance skills in other sports such as women's rugby and Gaelic football kicks may assist in tailoring coaching cues specific for women's characteristics.

9.1 Future directions for research

Several considerations have arisen from the research in this thesis for when enacting further development in the proof-of-concept use of IMUs for AF kick classification. These include improved hardware with a higher accelerometer range, data collection under different conditions, also training and testing models on a larger dataset. As hardware which was validated and appropriate for the applied aims of the research was unavailable at the time of this thesis data collection, further investigation and methodology development was beyond the scope of the thesis. An accelerometer reading greater than the current ±16 *g* would provide complete kick signal data for improved feature extraction towards defining the unique components of AF kick types. A more tailored signal processing kick detection algorithm could also be developed similar to the signal processing work on IMU semi-automated skill feedback in elite half-pipe snowboarding (Harding et al., 2008). By identifying that air-time was a key performance variable in the event judged score, Harding at el. (2008) used a two-pass signal processing method with power density in the frequency domain and a threshold-based search to detect and calculate air-time to minimise subjective bias in competition scoring. Improved kick signal processing may involve windowed kick technique stages to extract specific kinematic data based on the acceleration and angular velocity characterised. For example, the classification of legal and illegal cricket bowls used the signal profiled windows and tagged events of cricket bowling stages (Salman et al., 2017). Data should also be gained from both male and female AF athletes considering the established kinematic differences in drop punt kicks between AFL and AFLW athletes. Doing so may aid in signal feature extraction in capturing the fine-grained differences in kick signal features (Bulling et al., 2014; Ghojogh et al., 2019). Data collection and model testing should also be undertaken across controlled, semi-controlled, and uncontrolled environments such as small-sided games or full ground practice matches in order to assess the complete application capacity of an AF kick recognition system; as researched in soccer (Schuldhaus et al., 2015). Increased signal noise may likely result from the frequency and intensity of other activities in match play such as tackles, jumps, rapid changes of direction, and high-speed running. Hence model evaluations are also required in uncontrolled environment settings to account for changes in performance due to increased outside noise from the targeted movement(s). A complete data signal of a kick which hasn't been affected by saturation could also aid in the investigation of using IMUs as an on-field kick skill assessment and skill acquisition tool. Although laboratory validation against a 3D optoelectronic motion analysis system of a low-profile IMU set-up would be required. This may include trials of one IMU each leg that could capture the acceleration and angular velocity kinematic properties of the drop punt kick. Validation (Blair, Duthie, et al., 2018) and quantification (Blair, Robertson, et al., 2018) of kicking kinematics across football codes has previously been undertaken using a full body 17 IMU set-up. Both studies provide evidence for the use of IMUs to quantify biomechanical movements in a sporting setting; although the accelerometers used were also limited to ±16 *g* range. A low-profile IMU set-up that could be reliably used on-field for individual athlete drop punt kick skill assessments would be of interest to investigate towards an objective measure alongside vision and a coaches' subjective evaluation, where changes in skill execution could be made in-situation.

Chapter Ten: Conclusion

This thesis aimed to investigate across a multi-disciplinary approach on applications of wearable IMUs and computer analytical methods in sport performance analysis; and to assess how this knowledge translates practically to AFLW kick skill performance. Firstly, evidence on the growing emphasis and developments of IMUs and computer vision in the area of sport-specific movement recognition offer insights into the method applications across several sports. The data and results presented in this thesis contribute to the knowledge on the skilled performances of AFLW athletes in matches, and also profiles their drop punt kick biomechanics. Specifically, also highlighting key differences in match-play athlete skill contributions and kick kinematics in comparison to male AFL athletes. Findings indicate the importance of stand-alone sport science research in elite women's AF which could be generalised to other elite women's team sports. This thesis also demonstrates the potential of IMUs in AF kick recognition as an on-field training tool for several applications in kick skill development and volume monitoring. As improvements to device hardware are made available to suit the demands of AF kicks will prompt for further investigation in the area. Progression in the applications of IMUs for kick skill development and biomechanical evaluations in elite female AF athletes would be of interest for improving AFLW sport science support of skilled performance improvements.

References

Abdullah, M. A., Ibrahim, M. A. R., Shapiee, M. N. A. B., Razman, M. A. M., Musa, R. M., & Majeed, A. P. A. (2020). The classification of skateboarding trick manoeuvres through the integration of IMU and machine learning. In Z. Jamaludin & M. N. Ali Mokhtar (Eds.), *Intelligent Manufacturing and Mechatronics* (pp. 67–74). Springer Singapore. https://doi.org/10.1007/978-981-13-9539-0

Abe, S. (2005). Support vector machines for pattern classification. In *Advances in Pattern Recognition* (Second ed). Springer. https://doi.org/0.1007/978-1-84996-098-4

Adelsberger, R., & Tröster, G. (2013). Experts lift differently: classification of weight-lifting athletes. *2013 IEEE International Conference on Body Sensor Networks*. https://doi.org/10.1109/BSN.2013.6575458

Aggarwal, J. K., & Xia, L. (2014). Human activity recognition from 3D data: a review. *Pattern Recognition Letters*, *48*, 70–80. https://doi.org/10.1016/j.patrec.2014.04.011

Ahmadi, A., Mitchell, E., Richter, C., Destelle, F., Gowing, M., O'Connor, N. E., & Moran, K. (2015). Toward automatic activity classification and movement assessment during a sports training session. *IEEE Internet of Things Journal*, *2*(1), 23–32. https://doi.org/10.1109/JIOT.2014.2377238

Alcock, A. M. (2010). *Analysis of direct shots at goal from free kicks in elite women's football*. Southern Cross University.

Alcock, A. M., Gilleard, W., Hunter, A. B., Baker, J., & Brown, N. (2012). Curve and instep kick kinematics in elite female footballers. *Journal of Sports Sciences*, *30*(4), 387–394. https://doi.org/10.1080/02640414.2011.643238

Alexander, J. P., Spencer, B., Mara, J. K., & Robertson, S. (2019). Collective team behaviour of Australian Rules football during phases of match play. *Journal of Sports Sciences*, *37*(3), 237–243. https://doi.org/10.1080/02640414.2018.1491113

Alexander, J. P., Spencer, B., Sweeting, A. J., Mara, J. K., & Robertson, S. (2019). The influence of match phase and field position on collective team behaviour in Australian Rules football. *Journal of Sports Sciences*, *37*(15), 1699–1707. https://doi.org/10.1080/02640414.2019.1586077

Alsheikh, M. A., Selim, A., Niyato, D., Doyle, L., Lin, S., & Tan, H.-P. (2016). Deep activity recognition models with triaxial accelerometers. *The Workshops of the Thirtieth AAAI Conference on Artificial Intelligence Artificial Intelligence Applied to Assistive Technologies and Smart Environments: Technical Report WS-16-01 Deep*. https://doi.org/10.1016/j.appet.2015.08.014

Aminian, K., & Najafi, B. (2004). Capturing human motion using body-fixed sensors: outdoor measurement and clinical applications. *Computer Animation and Virtual*

Worlds, *15*(2), 79–94. https://doi.org/https://doi.org/10.1002/cav.2

Anand, A., Sharma, M., Srivastava, R., Kaligounder, L., & Prakash, D. (2017). Wearable motion sensor based analysis of swing sports. *Proceedings - 16th IEEE International Conference on Machine Learning and Applications, ICMLA 2017*, 261–267. https://doi.org/10.1109/ICMLA.2017.0-149

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, *26*(4), 217–238. https://doi.org/10.2165/00007256-199826040-00002

Attal, F., Mohammed, S., Dedabrishvili, M., Chamroukhi, F., Oukhellou, L., & Amirat, Y. (2015). Physical human activity recognition using wearable sensors. *Sensors (Switzerland)*, *15*(12), 31314–31338. https://doi.org/10.3390/s151229858

Baczkowski, K., Marks, P., Silberstein, M., & Schneider-Kolsky, M. E. (2006). A new look into kicking a football: an investigation of muscle activity using MRI. *Australasian Radiology*, *50*(4), 324–329. https://doi.org/10.1111/j.1440-1673.2006.01591.x

Bajric, R., Zuber, N., Skrimpas, G. A., & Mijatovic, N. (2016). Feature extraction using discrete wavelet transform for gear fault diagnosis of wind turbine gearbox. *Shock and Vibration*, 1–10. https://doi.org/10.1155/2016/6748469

Ball, K. (2008). Biomechanical considerations of distance kicking in Australian Rules football. *Sports Biomechanics*, *7*(1), 10–23. https://doi.org/10.1080/14763140701683015

Ball, K. (2011). Kinematic comparison of the preferred and non-preferred foot punt kick. *Journal of Sports Sciences*, *29*(14), 1545–1552. https://doi.org/10.1080/02640414.2011.605163

Ball, K. (2013). Loading and performance of the support leg in kicking. *Journal of Science and Medicine in Sport*, *16*(5), 455–459. https://doi.org/10.1016/j.jsams.2012.10.008

Ball, K., Ingleton, C., Peacock, J., & Nunome, H. (2013). Ball impact dynamics in the punt kick. *International Conference of Biomechanics in Sports*, 3–6.

Ball, K., Smith, J., & MacMahon, C. (2010). Kick impact characteristics of junior kickers. *International Symposium on Biomechanics in Sports: Conference Proceedings Archive*, *1*, 1–4.

Banos, O., Galvez, J. M., Damas, M., Pomares, H., & Rojas, I. (2014). Window size impact in human activity recognition. *Sensors (Switzerland)*, *14*(4), 6474–6499. https://doi.org/10.3390/s140406474

Barfield, W. R., Kirkendall, D. T., & Yu, B. (2002). Kinematic instep kicking differences between elite female and male soccer players. *Journal of Sports Science and Medicine*, *1*(3), 72–79.

Barris, S., & Button, C. (2008). A review of vision-based motion analysis in sport. *Sports Medicine*, *38*(12), 1025–1043. https://doi.org/10.2165/00007256-200838120-00006

Bayes, T. (1991). An essay towards solving a problem in the doctrine of chances. 1763. *MD Computing: Computers in Medical Practice*, *8*(3), 157.

Bengio, Y. (2013). Deep learning of representations: looking forward. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7978 LNAI*, 1–37. https://doi.org/10.1007/978-3-642-39593-2_1

Beniwal, S., & Arora, J. (2014). Classification and feature selection techniques in data mining. *International Journal of Engineering Research & Technology (IJERT)*, *1*(August 2012), 1–6.

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, *13*, 281–305.

Bernardo, J. M., & Smith, A. F. (2009). *Bayesian theory*. John Wiley & Sons.

Bertasius, G., Park, H. S., Yu, S. X., & Shi, J. (2017). Am I a baller? basketball performance assessment from first-person videos. *Proceedings of the IEEE International Conference on Computer Vision*, *October*, 2196–2204. https://doi.org/10.1109/ICCV.2017.239

Black, G. M., Gabbett, T. J., Johnston, R. D., Cole, M. H., Naughton, G., & Dawson, B. (2018a). A skill profile of the national women's Australian football league (AFLW). *Science and Medicine in Football*, *June 30*, 1–5. https://doi.org/10.1080/24733938.2018.1489140

Black, G. M., Gabbett, T. J., Johnston, R. D., Cole, M. H., Naughton, G., & Dawson, B. (2018b). The influence of physical qualities on activity profiles of female Australian football match play. *International Journal of Sports Physiology and Performance*, *13*, 524–529. https://doi.org/10.1123/ijspp.2016-0723

Blair, S., Duthie, G., Robertson, S., Hopkins, W., & Ball, K. (2018). Concurrent validation of an inertial measurement system to quantify kicking biomechanics in four football codes. *Journal of Biomechanics*, *73*, 24–32. https://doi.org/10.1016/j.jbiomech.2018.03.031

Blair, S., Robertson, S., Duthie, G., & Ball, K. (2018). The effect of altering distance on goal-kicking technique in Australian football. *ISBS Proceedings Archive*, *36*(1), 358–361. https://commons.nmu.edu/isbs/vol36/iss1/69

Bonney, N., Berry, J., Ball, K., & Larkin, P. (2019). The development of a field-based kicking assessment to evaluate Australian football kicking proficiency. *Research Quarterly for Exercise and Sport*, 1–10. https://doi.org/10.1080/02701367.2019.1647331

Boulesteix, A. L., & Schmid, M. (2014). Machine learning versus statistical modeling. *Biometrical Journal*, *56*(4), 588–593. https://doi.org/10.1002/bimj.201300226

Boyd, L. J., Ball, K., & Aughey, R. J. (2013). Quantifying external load in Australian football matches and training using accelerometers. *International Journal of Sports Physiology and Performance*, *8*(1), 44–51. https://doi.org/10.1123/ijspp.8.1.44

Breiman, L. (2001a). Random forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L. (2001b). Statistical modeling: the two cultures. *Statistical Science*, *16*(3), 199–215. https://doi.org/10.1214/ss/1009213726

Brice, S. M., Hurley, M., & Phillips, E. J. (2018). Use of inertial measurement units for measuring torso and pelvis orientation, and shoulder–pelvis separation angle in the discus throw. *International Journal of Sports Science and Coaching*, *13*(6), 985–992. https://doi.org/10.1177/1747954118778664

Brochu, E., Cora, V. M., & de Freitas, N. (2010). A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *ArXiv Preprint ArXiv:1012.2599*. http://arxiv.org/abs/1012.2599

Brock, H. (2018). Deep learning - accelerating next generation performance analysis systems? *Multidisciplinary Digital Publishing Institute Proceedings*, *2*(6), 303. https://doi.org/10.3390/proceedings2060303

Brock, H., & Ohgi, Y. (2017). Assessing motion style errors in ski jumping using inertial sensor devices. *IEEE Sensors Journal*, *17*(12), 3794–3804. https://doi.org/10.1109/JSEN.2017.2699162

Brock, H., Ohgi, Y., & Lee, J. (2017). Learning to judge like a human: convolutional networks for classification of ski jumping errors. *Proceedings of International Symposium on Wearable Computers (ISWC)*, 106–113. https://doi.org/10.1145/3123021.3123038

Brogden, C. M., Armstrong, R., Page, R., Milner, D., Norris, D., & Greig, M. (2018). Use of triaxial accelerometry during the dance aerobic fitness test. *Journal of Dance Medicine & Science*, *2*(3), 115–122. https://doi.org/https://doi.org/10.12678/1089-313X.22.3.115

Buckley, C., O'Reilly, M. A., Whelan, D., Vallely Farrell, A., Clark, L., Longo, V., Gilchrist, M. D., & Caulfield, B. (2017). Binary classification of running fatigue using a single inertial measurement unit. *EEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN 2017)*, 197–201. https://doi.org/10.1109/BSN.2017.7936040

Bulling, A., Blanke, U., & Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys*, *46*(3), 1–33. https://doi.org/10.1007/978-3-319-99007-1_55

Busa, M., & McGregor, S. J. (2008). The use of accelerometers to assess human locomotion. *Clinical Kinesiology*, *62*(4), 21–25.

Buthe, L., Blanke, U., Capkevics, H., & Troster, G. (2016). A wearable sensing system for timing analysis in tennis. *BSN 2016 - 13th Annual Body Sensor Networks Conference*, 43–48. https://doi.org/10.1109/BSN.2016.7516230

Bux, A., Angelov, P., & Habib, Z. (2017). Vision based human activity recognition: a review. *Advances in Intelligent Systems and Computing*, 341–371. https://doi.org/10.1007/978-3-319-46562-3_23

Bzdok, D., Altman, N., & Krzywinski, M. (2018). Statistics versus machine learning. *Nature Methods*, *15*(4), 233–234. https://doi.org/10.1038/nmeth.4642

Bzdok, D., Krzywinski, M., & Altman, N. (2017). Points of significance: machine learning a primer. *Nature Methods*, *14*, 1119–11120. https://doi.org/10.1016/j.physbeh.2017.03.040

Callaway, A. J. (2015). Measuring kinematic variables in front crawl swimming using accelerometers: a validation study. *Sensors (Switzerland)*, *15*(5), 11363–11386. https://doi.org/10.3390/s150511363

Camomilla, V., Bergamini, E., Fantozzi, S., & Vannozzi, G. (2018). Trends supporting the in-field use of wearable inertial sensors for sport performance evaluation: a systematic review. *Sensors (Switzerland)*, *18*(3), 873. https://doi.org/10.3390/s18030873

Camomilla, V., Bergamini, E., Fantozzi, S., & Vannozzi, G. (2015). In-field use of wearable magneto-inertial sensors for sports performance evaluation. *33rd International Conference on Biomechanics in Sports*, 1–4. https://doi.org/10.1557/opl.2014.

Camomilla, V., Dumas, R., & Cappozzo, A. (2017). Human movement analysis: the soft tissue artefact issue. *Journal of Biomechanics*, *62*, 1–4. https://doi.org/10.1016/j.jbiomech.2017.09.001

Campbell, R., Pease, D., & Cossens, P. (2018). Quantifying landing impacts during a leg strength circuit in male artistic gymnasts - a pilot study. *36th Conference of the International Society of Biomechanics in Sports*, 831–834. https://sprinz.aut.ac.nz/__data/assets/pdf_file/0014/203072/194_1331_Campbell.pdf

Cappozzo, A., Catani, F., Della Croce, U., & Leardini, A. (1995). Position and orientation in space of bones during movement: anatomical frame definition and determination. *Clinical Biomechanics*, *10*(4), 171–178.

Cardinale, M., & Varley, M. C. (2017). Wearable training-monitoring technology: applications, challenges, and opportunities. *International Journal of Sports Physiology and Performance*, *12*, 55–62. https://doi.org/https://doi.org/10.1123/ijspp.2016-0423

Cereatti, A., Bonci, T., Akbarshahi, M., Aminian, K., Barré, A., Begon, M., Benoit, D. L., Charbonnier, C., Dal Maso, F., Fantozzi, S., Lin, C. C., Lu, T. W., Pandy, M. G., Stagni, R., van den Bogert, A. J., & Camomilla, V. (2017). Standardization proposal of soft tissue artefact description for data sharing in human motion measurements. *Journal of Biomechanics*, *62*, 5–13. https://doi.org/10.1016/j.jbiomech.2017.02.004

Chambers, R. M., Gabbett, T. J., & Cole, M. H. (2018). Validity of a microsensor-based algorithm for detecting scrum events in rugby union. *International Journal of Sports Physiology and Performance*, *14*(2), 176–182. https://doi.org/https://doi.org/10.1123/ijspp.2018-0222

Chambers, R. M., Gabbett, T. J., Cole, M. H., & Beard, A. (2015). The use of wearable microsensors to quantify sport-specific movements. *Sports Medicine*, *45*(7), 1065–

1081. https://doi.org/10.1007/s40279-015-0332-9

Chambers, R. M., Gabbett, T. J., Gupta, R., Josman, C., Bown, R., Stridgeon, P., & Cole, M. H. (2019). Automatic detection of one-on-one tackles and ruck events using microtechnology in rugby union. *Journal of Science and Medicine in Sport*, *22*(7), 827–832. https://doi.org/10.1016/j.jsams.2019.01.001

Chapelle, O., Weston, J., & Schölkopf, B. (2003). Cluster kernels for semi-supervised learning. In S. Thrun, L. K. Saul, & B. Schölkopf (Eds.), *Advances in Neural Information Processing Systems* (pp. 601–608). MIT Press.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. . (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1613/jair.953

Chen, Y., & Xue, Y. (2015). A deep learning approach to human activity recognition based on single accelerometer. *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 1488–1492. https://doi.org/10.1109/SMC.2015.263

Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., & Van Calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of Clinical Epidemiology*, *110*, 12–22. https://doi.org/10.1016/j.jclinepi.2019.02.004

Cintia, P., Giannotti, F., Pappalardo, L., Pedreschi, D., & Malvaldi, M. (2015). The harsh rule of the goals: data-driven performance indicators for football teams. *Proceedings of the 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 1–10. https://doi.org/10.1109/DSAA.2015.7344823

Clarke, A. C., Couvalias, G., Kempton, T., & Dascombe, B. (2019). Comparison of the match running demands of elite and sub-elite women's Australian football. *Science and Medicine in Football*, *3*(1), 70–76. https://doi.org/10.1080/24733938.2018.1479067

Clarke, A. C., Ryan, S., Couvalias, G., Dascombe, B. J., Coutts, A. J., & Kempton, T. (2018). Physical demands and technical performance in Australian Football League Women's (AFLW) competition match-play. *Journal of Science and Medicine in Sport*, *21*(7), 748–752. https://doi.org/10.1016/j.jsams.2017.11.018

Conaire, C. Ó., Connaghan, D., Kelly, P., Connor, N. E. O., Gaffney, M., & Buckley, J. (2010). Combining inertial and visual sensing for human action recognition in tennis. *Proceedings of the First ACM International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams*, 51–56.

Connaghan, D., Kelly, P., O'Connor, N. E., Gaffney, M., Walsh, M., & O'Mathuna, C. (2011). Multi-sensor classification of tennis strokes. *Proceedings of IEEE Sensors*, 1437–1440. https://doi.org/10.1109/ICSENS.2011.6127084

Corazza, S., Mündermann, L., Gambaretto, E., Ferrigno, G., & Andriacchi, T. P. (2010). Markerless motion capture through visual hull, articulated ICP and subject specific model generation. *International Journal of Computer Vision*, *87*(1–2), 156–169. https://doi.org/10.1007/s11263-009-0284-3

Corbett, D. M., Sweeting, A. J., & Robertson, S. (2017). Weak relationships between stint duration, physical and skilled match performance in Australian football. *Frontiers in Physiology*, *8*(October), 1–8. https://doi.org/10.3389/fphys.2017.00820

Couceiro, M. S., Dias, G., Mendes, R., & Araújo, D. (2013). Accuracy of pattern detection methods in the performance of golf putting. *Journal of Motor Behavior, 45*(1), 37–53. https://doi.org/10.1080/00222895.2012.740100

Coutts, A. J., Quinn, J., Hocking, J., Castagna, C., & Rampinini, E. (2010). Match running performance in elite Australian Rules football. *Journal of Science and Medicine in Sport*, *13*(5), 543–548. https://doi.org/10.1016/j.jsams.2009.09.004

Coventry, E., Ball, K., Parrington, L., Aughey, R. J., & McKenna, M. (2015). Kinematic effects of a short-term fatigue protocol on punt-kicking performance. *Journal of Sports Sciences*, *33*(15), 1596–1605. https://doi.org/10.1080/02640414.2014.1003582

Cuesta-Vargas, A. I., Galán-Mercant, A., & Williams, J. M. (2010). The use of inertial sensors system for human motion analysis. *Physical Therapy Reviews*, *15*(6), 462–473. https://doi.org/10.1179/1743288X11Y.0000000006

Cust, E. E., Ball, K., Sweeting, A., & Robertson, S. (2019). Biomechanical Characteristics of Elite Female Australian Rules Football Preferred and Non-preferred Drop Punt Kicks. *Proceedings of the 7th International Conference on Sport Sciences Research and Technology Support*, 32–37. https://doi.org/10.5220/0008066300320037

Cust, E. E., Sweeting, A. J., Ball, K., Anderson, H., & Robertson, S. (2019). The relationship of team and individual athlete performances on match quarter outcome in elite women's Australian Rules football. *Journal of Science and Medicine in Sport*, *22*(10), 1157–1162. https://doi.org/10.1016/j.jsams.2019.05.004

Cust, E. E., Sweeting, A. J., Ball, K., & Robertson, S. (2019). Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance. *Journal of Sports Sciences*, *37*(5), 568–600. https://doi.org/10.1080/02640414.2018.1521769

Davids, K., Araújo, D., Vilar, L., Renshaw, I., & Pinder, R. (2013). An ecological dynamics approach to skill acquisition: implications for development of talent in sport. *Talent Development and Excellence*, *5*(1), 21–34.

Dawson, B., Hopkinson, R., Appleby, B., Stewart, G., & Roberts, C. (2004). Player movement patterns and game activities in the Australian Football League. *Journal of Science and Medicine in Sport*, *7*(3), 278–291. https://doi.org/10.1016/S1440-2440(04)80023-9

Deisenroth, M. P., Faisal, A. A., & Ong, C. S. (2018). *Mathematics for Machine Learning*. Cambridge University Press.

del Rosario, M. B., Redmond, S. J., & Lovell, N. H. (2015). Tracking the evolution of smartphone sensing for monitoring human movement. *Sensors (Switzerland)*, *15*(8), 18901–18933. https://doi.org/10.3390/s150818901

Derrick, T. (2004). Signal processing. In D. Robertson, C. Caldwell, J. Hamill, G. Kamen,

& S. Whittlesey (Eds.), *Research methods in biomechanics* (Second edi, pp. 282–284). Human Kinetics.

Derungs, A., Soller, S., Weishaupl, A., Bleuel, J., Berschin, G., & Amft, O. (2018). Regression-based, mistake-driven movement skill estimation in Nordic walking using wearable inertial sensors. *2018 IEEE International Conference on Pervasive Computing and Communications, PerCom 2018*. https://doi.org/10.1109/PERCOM.2018.8444576

Díaz-Pereira, M. P., Gómez-Conde, I., Escalona, M., & Olivieri, D. N. (2014). Automatic recognition and scoring of olympic rhythmic gymnastic movements. *Human Movement Science*, *34*(1), 63–80. https://doi.org/10.1016/j.humov.2014.01.001

Dichiera, A., Webster, K. E., Kuilboer, L., Morris, M. E., Bach, T. M., & Feller, J. A. (2006). Kinematic patterns associated with accuracy of the drop punt kick in Australian Football. *Journal of Science and Medicine in Sport*, *9*(4), 292–298. https://doi.org/10.1016/j.jsams.2006.06.007

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, *55*(10), 78–87. https://doi.org/10.1145/2347736.2347755

Dörge, H. C., Andersen, T. B., SØrensen, H., & Simonsen, E. B. (2002). Biomechanical differences in soccer kicking with the preferred and the non-preferred leg. *Journal of Sports Sciences*, *20*(4), 293–299. https://doi.org/10.1080/026404102753576062

Duarte, M. (2014). *Detect peaks* (1.0.6). http://nbviewer.ipython.org/github/demotu/BMC/blob/master/notebooks/DetectPeaks.ipynb%0A

Duthie, G., Pyne, D., & Hooper, S. (2003). The reliability of video based time motion analysis. *Journal of Human Movement Studies*, *44*(3), 259–271.

Ebner, C. J., & Findling, R. D. (2019). Tennis stroke classification: comparing wrist and racket as IMU sensor position. In ACM (Ed.), *17th International Conference on Advances in Mobile Computing & Multimedia (MoMM '19)* (pp. 1–10). https://doi.org/https://doi.org/10.1145/3365921.3365929

Edwards, S., White, S., Humphreys, S., Robergs, R., & O'Dwyer, N. (2019). Caution using data from triaxial accelerometers housed in player tracking units during running. *Journal of Sports Sciences*, *37*(7), 810–818. https://doi.org/10.1080/02640414.2018.1527675

Ellens, S., Blair, S., Peacock, J., Barnes, S., & Ball, K. (2017). Use of accelerometers in Australian Football to identify a kick. *35th International Conference on Biomechanics in Sport*, 218–221.

Ermes, M., Pärkkä, J., Mäntyjärvi, J., Korhonen, I., & Juha, P. (2008). Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *IEEE Transactions on Information Technology in Biomedicine*, *12*(1), 20–26. https://doi.org/10.1109/TITB.2007.899496

Eyobu, O. S., & Han, D. S. (2018). Feature representation and data augmentation for human activity classification based on wearable IMU sensor data using a deep LSTM neural network. *Sensors (Switzerland)*, *18*(9), 1–26. https://doi.org/10.3390/s18092892

Falbriard, M., Meyer, F., Mariani, B., Millet, G. P., & Aminian, K. (2018). Accurate estimation of running temporal parameters using foot-worn inertial sensors. *Frontiers in Physiology*, *9*, 610. https://doi.org/10.3389/fphys.2018.00610

Farrahi, V., Niemelä, M., Kangas, M., Korpelainen, R., & Jämsä, T. (2019). Calibration and validation of accelerometer-based activity monitors: a systematic review of machine-learning approaches. *Gait and Posture*, *68*, 285–299. https://doi.org/10.1016/j.gaitpost.2018.12.003

Farrow, D., & Robertson, S. (2017). Development of a skill acquisition periodisation framework for high-performance sport. *Sports Medicine*, *47*(6), 1043–1054. https://doi.org/10.1007/s40279-016-0646-2

Faulkner, H., & Dick, A. (2015). AFL Player Detection and Tracking. *2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, 1–8. https://doi.org/10.1109/DICTA.2015.7371226

Figo, D., Diniz, P. C., Ferreira, D. R., & Cardoso, J. M. P. (2010). Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing*, *14*(7), 645–662. https://doi.org/10.1007/s00779-010-0293-9

Fong, D. T. P., & Chan, Y. Y. (2010). The use of wearable inertial motionsensors in human lower limb biomechanics studies: a systematic review. *Sensors*, *10*(12), 11556–11565. https://doi.org/10.3390/s101211556

Forman, G., & Scholz, M. (2010). Apples-to-apples in cross-validation studies: pitfalls in classifier performance measurement. *ACM SIGKDD Explorations Newsletter*, *12*(1), 49–57. https://doi.org/10.1145/1882471.1882479

Forner-Cordero, A., Mateu-Arce, M., Forner-Cordero, I., Alcántara, E., Moreno, J. C., & Pons, J. L. (2008). Study of the motion artefacts of skin-mounted inertial sensors under different attachment conditions. *Physiological Measurement*, *29*(4). https://doi.org/10.1088/0967-3334/29/4/N01

Gabbett, T. J., Jenkins, D., & Abernethy, B. (2010). Physical collisions and injury during professional rugby league skills training. *Journal of Science and Medicine in Sport*, *13*(6), 578–583. https://doi.org/10.1016/j.jsams.2010.03.007

Gabbett, T. J., Jenkins, D. G., & Abernethy, B. (2011). Physical collisions and injury in professional rugby league match-play. *Journal of Science and Medicine in Sport*, *14*(3), 210–215. https://doi.org/10.1016/j.jsams.2011.01.002

Gabbett, T. J., Jenkins, D. G., & Abernethy, B. (2012). Physical demands of professional rugby league training and competition using microtechnology. *Journal of Science and Medicine in Sport*, *15*(1), 80–86. https://doi.org/10.1016/j.jsams.2011.07.004

García-Rubio, J., Gómez, M. Á., Lago-Peñas, C., & Ibáñez Godoy, S. J. (2015). Effect of match venue, scoring first and quality of opposition on match outcome in the UEFA champions league. *International Journal of Performance Analysis in Sport*, *15*(2), 527–

539. https://doi.org/10.1080/24748668.2015.11868811

Gastin, P. B., Mclean, O., Breed, R. V. P., & Spittle, M. (2014). Tackle and impact detection in elite Australian football using wearable microsensor technology. *Journal of Sports Sciences*, *32*(10), 947–953. https://doi.org/10.1080/02640414.2013.868920

Gastin, P. B., McLean, O., Spittle, M., & Breed, R. V. P. (2013). Quantification of tackling demands in professional Australian football using integrated wearable athlete tracking technology. *Journal of Science and Medicine in Sport*, *16*(6), 589–593. https://doi.org/10.1016/j.jsams.2013.01.007

Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd Editio). O'Reilly Media, Inc.

Ghojogh, B., Samad, M. N., Mashhadi, S. A., Kapoor, T., Ali, W., Karray, F., & Crowley, M. (2019). Feature selection and feature extraction in pattern analysis: a literature review. *ArXiv Preprint ArXiv:1905.02845*, *May 7*. http://arxiv.org/abs/1905.02845

Gløersen, Ø., Myklebust, H., Hallén, J., & Federolf, P. (2018). Technique analysis in elite athletes using principal component analysis. *Journal of Sports Sciences*, *36*(2), 229–237. https://doi.org/10.1080/02640414.2017.1298826

Godfrey, A., Conway, R., Meagher, D., & ÓLaighin, G. (2008). Direct measurement of human movement by accelerometry. *Medical Engineering and Physics*, *30*(10), 1364–1386. https://doi.org/10.1016/j.medengphy.2008.09.005

Gómez, M. Á., Lorenzo, A., Ibañez, S. J., & Sampaio, J. (2013). Ball possession effectiveness in men's and women's elite basketball according to situational variables in different game periods. *Journal of Sports Sciences*, *31*(14), 1578–1587. https://doi.org/10.1080/02640414.2013.792942

Gómez, M. Á., Serna, A. Dela, Lupo, C., & Sampaio, J. (2014). Effects of situational variables and starting quarter score in the outcome of elite women's water polo game quarters. *International Journal of Performance Analysis in Sport*, *14*(1), 73–83. https://doi.org/10.1080/24748668.2014.11868704

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. http://www.deeplearningbook.org

Gray, A. J., & Jenkins, D. G. (2010). Match analysis and the physiological demands of Australian football. *Sports Medicine*, *40*(4), 347–360. https://doi.org/10.2165/11531400-000000000-00000

Gregory, R. W., Humphreys, S. E., & Street, G. M. (1994). Kinematic analysis of skating technique of olympic skiers in the women's 30-km race. *Journal of Applied Biomechanics*, *10*(4), 382–392. https://doi.org/10.1123/jab.10.4.382

Groh, B. H., Fleckenstein, M., & Eskofier, B. M. (2016). Wearable trick classification in freestyle snowboarding. *BSN 2016 - 13th Annual Body Sensor Networks Conference*, 89–93. https://doi.org/10.1109/BSN.2016.7516238

Groh, B. H., Fleckenstein, M., Kautz, T., & Eskofier, B. M. (2017). Classification and visualization of skateboard tricks using wearable sensors. *Pervasive and Mobile Computing*, *40*, 42–55. https://doi.org/10.1016/j.pmcj.2017.05.007

Groh, B. H., Kautz, T., & Schuldhaus, D. (2015). IMU-based trick classification in skateboarding. *KDD Workshop on Large-Scale Sports Analytics*.

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, *3*(3), 1157–1182. https://doi.org/10.1016/j.aca.2011.07.027

Guyon, I., Gunn, S., Nikravesh, M., & Zadeh, L. A. (2008). Feature extraction: foundations and applications. In *Studies in Fuzziness and Soft Computing, Volume 207* (pp. 1–25). Springer. https://doi.org/10.1007/978-3-540-35488-8_1

Hachaj, T., Ogiela, M. R., & Koptyra, K. (2015). Application of assistive computer vision methods to Oyama karate techniques recognition. *Symmetry*, *7*(4), 1670–1698. https://doi.org/10.3390/sym7041670

Hadlow, S. M., Pinder, R. A., & Sayers, M. G. L. (2017). Influence of football size on kicking performance in youth Australian rules footballers. *Journal of Sports Sciences*, *35*(18), 1808–1816. https://doi.org/10.1080/02640414.2016.1239023

Hafer, J. F., & Boyer, K. A. (2017). Variability of segment coordination using a vector coding technique: reliability analysis for treadmill walking and running. *Gait and Posture*, *51*, 222–227. https://doi.org/10.1016/j.gaitpost.2016.11.004

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*, 29–36. https://doi.org/10.1080/02634938208400381

Harding, J. W., Small, J. W., & James, D. A. (2008). Feature extraction of performance variables in elite half-pipe snowboarding using body mounted inertial sensors. *BioMEMS and Nanotechnology III*, 1–12. https://doi.org/10.1117/12.759259

Harrop, K., & Nevill, A. (2014). Performance indicators that predict success in an English Professional League One soccer team. *International Journal of Performance Analysis in Sport*, *14*(3), 907–920. https://doi.org/10.1080/24748668.2014.11868767

Hart, N. H., Nimphius, S., Cochrane, J. L., & Newton, R. U. (2013). Leg mass characteristics of accurate and inaccurate kickers - an Australian football perspective. *Journal of Sports Sciences*, *31*(15), 1647–1655. https://doi.org/10.1080/02640414.2013.793377

Hart, N. H., Nimphius, S., Spiteri, T., & Newton, R. U. (2014). Leg strength and lean mass symmetry influences kicking performance in Australian football. *Journal of Sports Science and Medicine*, *13*(1), 157–165.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference and prediction* (Second). Springer. https://doi.org/10.1007/978-0-387-84858-7

Hawkins, D. M. (2004). The problem of overfitting. *Journal of Chemical Information and Computer Sciences*, *44*(1), 1–12. https://doi.org/10.1021/ci0342472

Haycraft, J. A. Z., Kovalchik, S., Pyne, D. B., & Robertson, S. (2017). Physical characteristics of players within the Australian Football League participation pathways: a systematic review. *Sports Medicine - Open*, *3*(1). https://doi.org/10.1186/s40798-017-0109-9

He, H., Bai, Y., Garcia, E., & Li, S. (2008). ADASYN: adaptive synthetic sampling approach for imbalanced learning. *IEEE World Congress on Computational Intelligence*, *3*, 1322– 1328. https://doi.org/10.1109/IJCNN.2008.4633969

Heasman, J., Dawson, B., Berry, J., & Stewart, G. (2008). Development and validation of a player impact ranking system in Australian football. *International Journal of Performance Analysis in Sport*, *8*(3), 156–171. https://doi.org/10.1080/24748668.2008.11868457

Hecht-Nielsen, R. (1989). the Backpropagation Neural Network. *Proceedings Of The International Joint Conference On Neural Networks*, *1*, 593–605. https://doi.org/10.1109/IJCNN.1989.118638

Herold, M., Goes, F., Nopp, S., Bauer, P., Thompson, C., & Meyer, T. (2019). Machine learning in men's professional football: current applications and future directions for improving attacking play. *International Journal of Sports Science & Coaching*. https://doi.org/10.1177/1747954119879350

Higham, D. G., Hopkins, W. G., Pyne, D. B., & Anson, J. M. (2014). Performance indicators related to points scoring and winning in international rugby sevens. *Journal of Sports Science and Medicine*, *13*(2), 358–364.

Hiscock, D., Dawson, B., Heasman, J., & Peeling, P. (2012). Game movements and player performance in the Australian Football League. *International Journal of Performance Analysis in Sport*, *12*(3), 531–545. https://doi.org/10.1080/24748668.2012.11868617

Hochreiter, S., & Schmidhuber, J. J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1–32. https://doi.org/10.1162/neco.1997.9.8.1735

Hodges, N., & Williams, A. M. (Eds.). (2012). *Skill acquisition in sport: research, theory and practice*. Routledge.

Holzemann, A., & Van Laerhoven, K. (2018). Using wrist-worn activity recognition for basketball game analysis. *Proceedings of the 5th International Workshop on Sensor-Based Activity Recognition and Interaction*. https://doi.org/10.1145/3266157.3266217

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine*, *30*(1), 1–15. https://doi.org/10.2165/00007256-200030050-00006

Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. In *Medicine and Science in Sports and Exercise* (Vol. 41, Issue 1, pp. 3–12). https://doi.org/10.1249/MSS.0b013e31818cb278

Horton, M., Gudmundsson, J., Chawla, S., & Estephan, J. (2014). Classification of passes in football matches using spatiotemporal data. *ArXiv Preprint ArXiv:1407.5093*. https://doi.org/10.1145/3105576

Howe, S. T., Aughey, R. J., Hopkins, W. G., Stewart, A. M., & Cavanagh, B. P. (2017). Quantifying important differences in athlete movement during collision-based team sports: accelerometers outperform global positioning systems. *4th IEEE International Symposium on Inertial Sensors and Systems, INERTIAL 2017 - Proceedings*, 1–4. https://doi.org/10.1109/ISISS.2017.7935655

Hughes, M. (2015). Sports analysis. In *Essentials of Performance Analysis in Sport* (pp. 141–151). Routledge.

Hughes, M. (2017). Notational analysis – a mathematical perspective. *International Journal of Performance Analysis in Sport*, *4*(2), 97–139. https://doi.org/10.1080/24748668.2004.11868308

Hughes, M., & Bartlett, R. M. M. (2002). The use of performance indicators in performance analysis. *Journal of Sports Sciences*, *20*(10), 739–754. https://doi.org/10.1080/026404102320675602

Hughes, M., & Franks, I. M. (2004). *Notational analysis of sport: systems for better coaching and performance in sport* (2nd ed.). Routledge Publishing.

Hughes, M., Franks, I. M., Franks, I. M., & Dancs, H. (2007). *Essentials of performance analysis : an introduction*. Taylor & Francis Group.

Hulin, B. T., Gabbett, T. J., Johnston, R. D., & Jenkins, D. G. (2017). Wearable microtechnology can accurately identify collision events during professional rugby league match-play. *Journal of Science and Medicine in Sport*, *20*(7), 638–642. https://doi.org/10.1016/j.jsams.2016.11.006

Ibrahim, M. M. S., Muralidharan, S., Deng, Z., Vahdat, A., Mori, G., & Science, C. (1971). A Hierarchical Deep Temporal Model for Group Activity Recognition. *Cvpr*, 1971–1980. https://doi.org/10.1109/CVPR.2016.217

Ignatov, A. (2018). Real-time human activity recognition from accelerometer data using convolutional neural networks. *Applied Soft Computing Journal*, *62*, 915–922. https://doi.org/10.1016/j.asoc.2017.09.027

Jackson, K. (2016). *Assessing player performance in Australian football using spatial data* [Swinburne University of Technology, Melbourne]. http://hdl.handle.net/1959.3/422155

Janecek, A., Gansterer, W. N., Demel, M., & Ecker, G. F. (2008). On the relationship between feature selection and classification accuracy. *New Challenges for Feature Selection in Data Mining and Knowledge Discovery*, *4*, 90–105. https://doi.org/10.1007/978-3-540-77226-2_19

Jang, J., Ankit, A., Kim, J., Jang, Y. J., Kim, H. Y., Kim, J. H., & Xiong, S. (2018). A unified deep-learning model for classifying the cross-country skiing techniques using wearable gyroscope sensors. *Sensors (Switzerland)*, *18*(11), 1–25. https://doi.org/10.3390/s18113819

Janidarmian, M., Fekr, A. R., Radecka, K., & Zilic, Z. (2017). A comprehensive analysis on wearable acceleration sensors in human activity recognition. *Sensors (Switzerland)*, *17*(3). https://doi.org/10.3390/s17030529

Jensen, U., Blank, P., Kugler, P., & Eskofier, B. M. (2016). Unobtrusive and energy-efficient swimming exercise tracking using on-node processing. *IEEE Sensors Journal*, *16*(10), 3972–3980. https://doi.org/10.1109/JSEN.2016.2530019

Jensen, U., Prade, F., & Eskofier, B. M. (2013). Classification of kinematic swimming data with emphasis on resource consumption. *2013 IEEE International Conference on Body Sensor Networks, BSN 2013*. https://doi.org/10.1109/BSN.2013.6575501

Jensen, U., Schmidt, M., Hennig, M., Dassler, F. A., Jaitner, T., & Eskofier, B. M. (2015). An IMU-based mobile system for golf putt analysis. *Sports Engineering*, *18*(2), 123–133. https://doi.org/10.1007/s12283-015-0171-9

Jerri, A. J. (1977). The Shannon sampling theorem - its various extensions and applications: a tutorial review. *Proceedings of the IEEE*, *65*(11), 1565–1596. https://doi.org/10.1109/PROC.1977.10771

Ji, S., Xu, W., Yang, M., Yu, K., & Xu, W. (2012). 3D convolutional neural networks for human action recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(1), 221–231. https://doi.org/10.1109/TPAMI.2012.59

Jiao, L., Bie, R., Wu, H., Wei, Y., Ma, J., Umek, A., & Kos, A. (2018). Golf swing classification with multiple deep convolutional neural networks. *International Journal of Distributed Sensor Networks*, *14*(10). https://doi.org/10.1177/1550147718802186

Jiao, L., Wu, H., Bie, R., Umek, A., & Kos, A. (2018). Multi-sensor Golf Swing Classification Using Deep CNN. *Procedia Computer Science*, *129*, 59–65. https://doi.org/10.1016/j.procs.2018.03.046

Johnston, R. J., Watsford, M. L., Pine, M. J., Spurrs, R. W., Murphy, A., & Pruyn, E. C. (2012). Movement demands and match performance in professional Australian football. *International Journal of Sports Medicine*, *33*(2), 89–93. https://doi.org/10.1055/s-0031-1287798

Jolliffe, I. T. (2002). Principal Component Analysis, Second Edition. In *Encyclopedia of Statistics in Behavioral Science* (2nd ed., Vol. 30, Issue 3). Springer. https://doi.org/10.2307/1270093

Jones, N. M. P., Mellalieu, S. D., & James, N. (2004). Team performance indicators as a function of winning and losing in rugby union. *International Journal of Performance Analysis in Sport*, *4*(1), 61–71. https://doi.org/10.1080/24748668.2004.11868292

Joshi, K. D., Sarker, S., & Sarker, S. (2007). Knowledge transfer within information systems development teams: Examining the role of knowledge source attributes. *Decision Support Systems*, *43*(2), 322–335. https://doi.org/10.1016/j.dss.2006.10.003

Kapela, R., McGuinness, K., Swietlicka, A., & O'connor, N. E. (2014). Real-time event detection in field sport videos. *Advances in Computer Vision and Pattern Recognition*, *71*, 293–316. https://doi.org/10.1007/978-3-319-09396-3_14

Kapela, R., Świetlicka, A., Rybarczyk, A., Kolanowski, K., & O'Connor, N. E. (2015). Real-time event classification in field sport videos. *Signal Processing: Image Communication*, *35*, 35–45. https://doi.org/10.1016/j.image.2015.04.005

Karantonis, D. M., Narayanan, M. R., Mathie, M., Lovell, N. H., & Celler, B. G. (2006). Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE Transactions on Information Technology in Biomedicine*, *10*(1), 156–167. https://doi.org/10.1109/TITB.2005.856864

Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., Fei-Fei, L. L., & Li, F. F. (2014). Large-scale video classification with convolutional neural networks. In IEEE (Ed.), *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 1725–1732). https://doi.org/10.1109/CVPR.2014.223

Kasiri-Bidhendi, S., Fookes, C., Morgan, S., Martin, D. T., & Sridharan, S. (2015). Combat sports analytics: boxing punch classification using overhead depth imagery. *2015 IEEE International Conference on Image Processing (ICIP)*, 4545–4549. https://doi.org/10.1109/ICIP.2015.7351667

Kasiri, S., Fookes, C., Sridharan, S., & Morgan, S. (2017). Fine-grained action recognition of boxing punches from depth imagery. *Computer Vision and Image Understanding*, *159*, 143–153. https://doi.org/10.1016/j.cviu.2017.04.007

Kautz, T. (2017). Acquisition, filtering and analysis of positional and inertial data in sports. In *FAU University Press*. Friedrich-Alexander-Universität Erlangen-Nürnberg.

Kautz, T., Groh, B. H., Hannink, J., Jensen, U., Strubberg, H., & Eskofier, B. M. (2017). Activity recognition in beach volleyball using a deep convolutional neural network. *Data Mining and Knowledge Discovery*, *31*(6), 1678–1705. https://doi.org/10.1007/s10618-017-0495-0

Kavanagh, J. J., & Menz, H. B. (2008). Accelerometry: a technique for quantifying movement patterns during walking. *Gait and Posture*, *28*(1), 1–15. https://doi.org/10.1016/j.gaitpost.2007.10.010

Kazemi, M., Waalen, J., Morgan, C., & White, A. R. (2006). A profile of Olympic taekwondo competitors. *Journal of Sports Science and Medicine*, *5*(CSSI-1), 114–121.

Ke, S. R., Thuc, H., Lee, Y. J., Hwang, J. N., Yoo, J. H., & Choi, K. H. (2013). A review on video-based human activity recognition. In *Computers* (Vol. 2). https://doi.org/10.3390/computers2020088

Keaney, E. M., & Reid, M. (2018). Quantifying hitting activity in tennis with racket sensors: new dawn or false dawn? *Sports Biomechanics*. https://doi.org/10.1080/14763141.2018.1535619

Kelly, D., Coughlan, G. F., Green, B. S., & Caulfield, B. (2012). Automatic detection of collisions in elite level rugby union using a wearable sensing device. *Sports Engineering*, *15*(2), 81–92. https://doi.org/10.1007/s12283-012-0088-5

Kempton, T., Sullivan, C., Bilsborough, J. C., Cordy, J., & Coutts, A. J. (2015). Match-to-match variation in physical activity and technical skill measures in professional

Australian football. *Journal of Science and Medicine in Sport*, *18*(1), 109–113. https://doi.org/10.1016/j.jsams.2013.12.006

Khalid, S., Khalil, T., & Nasreen, S. (2014). A survey of feature selection and feature extraction techniques in machine learning. *Proceedings of 2014 Science and Information Conference, SAI 2014*, 372–378. https://doi.org/10.1109/SAI.2014.6918213

Khan, A. M., Siddiqi, M. H., & Lee, S. W. (2013). Exploratory data analysis of acceleration signals to select light-weight and accurate features for real-time activity recognition on smartphones. *Sensors (Switzerland)*, *13*(10), 13099–13122. https://doi.org/10.3390/s131013099

Kobsar, D., Osis, S. T., Hettinga, B. A., & Ferber, R. (2014). Classification accuracy of a single tri-axial accelerometer for training background and experience level in runners. *Journal of Biomechanics*, *47*(10), 2508–2511. https://doi.org/10.1016/j.jbiomech.2014.04.017

Kohavi, R., John, G. H., John, H., & John, G. H. (2011). Wrappers for feature subset selection. *Artificial Intelligence*, *97*(97), 273–324. https://doi.org/10.1016/S0004-3702(97)00043-X

Kos, M., & Kramberger, I. (2017). A wearable device and system for movement and biometric data acquisition for sports applications. *IEEE Access*, *5*, 6411–6420. https://doi.org/10.1109/ACCESS.2017.2675538

Kos, M., Ženko, J., Vlaj, D., & Kramberger, I. (2016). Tennis stroke detection and classification using miniature wearable IMU device. *International Conference on Systems, Signals, and Image Processing*, *2016-June*(May). https://doi.org/10.1109/IWSSIP.2016.7502764

Kotsiantis, S. B. B., Zaharakis, I. D., & Pintelas, P. E. (2007). Supervised machine learning: a review of classification techniques. *Informatica*, *31*, 249–268. https://doi.org/10.1115/1.1559160

Krizhevsky, A., Hinton, G. E., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances In Neural Information Processing Systems*, 1–9. https://doi.org/http://dx.doi.org/10.1016/j.protcy.2014.09.007

Kunze, K., Bahle, G., Lukowicz, P., & Partridge, K. (2010). Can magnetic field sensors replace gyroscopes in wearable sensing applications? *International Symposium on Wearable Computers (ISWC)*, 3–6. https://doi.org/10.1109/ISWC.2010.5665859

Ladha, L., & Deepa, T. (2011). Feature selection methods and algorithms. *International Journal on Computer Science and Engineering (IJCSE)*, *3*(5), 1787–1797. https://doi.org/10.1145/2513228.2513313

Lago-Peñas, C., Lago-Ballesteros, J., & Rey, E. (2011). Differences in performance indicators between winning and losing teams in the UEFA Champions League. *Journal of Human Kinetics*, *27*(1), 135–146. https://doi.org/10.2478/v10078-011-0011-3

Lai, D. T. H., Hetchl, M., Wei, X., Ball, K., & Mclaughlin, P. (2011). On the difference in swing arm kinematics between low handicap golfers and non-golfers using wireless inertial sensors. *Procedia Engineering*, *13*, 219–225. https://doi.org/10.1016/j.proeng.2011.05.076

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

LeCun, Y, Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *November*, 1–46. https://doi.org/10.1109/5.726791

LeCun, Yann, Bottou, L., Orr, G. B., & Müller, K. R. (1998). Efficient backprop. In *Neural Networks: Tricks of the Trade* (Vol. 1524, pp. 9–50).

Lees, A., Asai, T., Andersen, T. B., Nunome, H., & Sterzing, T. (2010). The biomechanics of kicking in soccer: a review. *Journal of Sports Sciences*, *28*(8), 805–817. https://doi.org/10.1080/02640414.2010.481305

Lemaître, G., Nogueira, F., & Aridas, C. K. (2017). Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning. *Journal of Machine Learning Research*, *18*(17), 1–5. http://jmlr.org/papers/v18/16-365.html

Lever, J., Krzywinski, M., & Altman, N. (2016). Points of significance: model selection and overfitting. *Nature Methods*, *13*(9), 703–704. https://doi.org/10.1038/nmeth.3968

Li, F., Shirahama, K., Nisar, M. A., Köping, L., & Grzegorzek, M. (2018). Comparison of feature learning methods for human activity recognition using wearable sensors. *Sensors (Switzerland)*, *18*(2), 1–22. https://doi.org/10.3390/s18020679

Li, J., Tian, Q., Zhang, G., Zheng, F., Lv, C., & Wang, J. (2018). Research on hybrid information recognition algorithm and quality of golf swing. *Computers and Electrical Engineering*, 1–13. https://doi.org/10.1016/j.compeleceng.2018.02.013

Li, Y., Van, F., Zeng, Y., & Wang, G. (2005). Biomechanical analysis on roundhouse kick in taekwondo. *ISBS-Conference Proceedings Archive*, *1*, 391–394. https://ojs.ub.uni-konstanz.de/cpa/article/view/830

Liao, W. H., Liao, Z. X., & Liu, M. J. (2003). Swimming style classification from video sequences. *16th IPPR Conference on Computer Vision, Graphics and Image Processing*, *Cvgip*, 226–233.

Liaw, A., & Wiener, M. (2002). Classification and regression by Random Forest. *R News*, *2*(December), 18–22. https://doi.org/10.1177/154405910408300516

Ling, C. X., Huang, J., & Zhang, H. (2003). AUC: a statistically consistent and more discriminating measure than accuracy. *International Joint Conference on Artificial Intelligence (IJCAI)*, *3*, 519–524.

Liu, H., & Motoda, H. (2012). *Feature Selection for Knowledge Discovery and Data Mining*. Springer Science & Business Media. https://doi.org/10.1007/978-1-4615-5689-3

Liu, H., Motoda, H., Setiono, R., & Zhao, Z. (2010). Feature selection: an ever evolving frontier in data mining. *JMLR: Workshop and Conference Proceedings 10: Feature Selection in Data Mining*, 4–13. https://doi.org/10.1007/BF00807254

Liu, T., Inoue, Y., & Shibata, K. (2009). Measurement of soft tissue deformation to improve the accuracy of a body-mounted motion sensor. *Journal of Medical Devices, Transactions of the ASME*, *3*(3), 035001. https://doi.org/10.1115/1.3212558

Lu, W. L., Okuma, K., & Little, J. J. (2009). Tracking and recognizing actions of multiple hockey players using the boosted particle filter. *Image and Vision Computing*, *27*(1–2), 189–205. https://doi.org/10.1016/j.imavis.2008.02.008

Luinge, H. J. (2002). *Inertial Sensing of Human Movement* [Twente University Press]. https://doi.org/10.1007/s13398-014-0173-7.2

Ma, R., Yan, D., Peng, H., Yang, T., Sha, X., Zhao, Y., & Liu, L. (2018). Basketball movements recognition using a wrist wearable inertial measurement unit. In IEEE (Ed.), *1st IEEE International Conference on Micro/Nano Sensors for AI, Healthcare, and Robotics (NSENS)* (pp. 73–76).

Magalhaes, F. A. de, Vannozzi, G., Gatta, G., & Fantozzi, S. (2015). Wearable inertial sensors in swimming motion analysis: a systematic review. *Journal of Sports Sciences*, *33*(7), 732–745. https://doi.org/10.1080/02640414.2014.962574

Mannini, A., Intille, S. S., Rosenberger, M., Sabatini, A. M., & Haskell, W. (2013). Activity recognition using a single accelerometer placed at the wrist or ankle. *Medicine and Science in Sports and Exercise*, *45*(11), 2193–2203. https://doi.org/10.1249/MSS.0b013e31829736d6

Mannini, A., & Sabatini, A. M. (2010). Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors*, *10*(2), 1154–1175. https://doi.org/10.3390/s100201154

Mansbridge, N., Mitsch, J., Bollard, N., Ellis, K., Miguel-Pacheco, G. G., Dottorini, T., & Kaler, J. (2018). Feature selection and comparison of machine learning algorithms in classification of grazing and rumination behaviour in sheep. *Sensors (Switzerland)*, *18*(10), 1–16. https://doi.org/10.3390/s18103532

Mayagoitia, R. E., Nene, A. V, & Veltink, P. H. (2002). Accelerometer and rate gyroscope measurement of kinematics: an inexpensive alternative to optical motion analysis systems. *Journal of Biomechanics*, *35*(4), 537–542. https://doi.org/10.1016/S0021-9290(01)00231-7

McGrath, J. W., Neville, J., Stewart, T., & Cronin, J. (2019). Cricket fast bowling detection in a training setting using an inertial measurement unit and machine learning. *Journal of Sports Sciences*, *37*(11), 1220–1226. https://doi.org/10.1080/02640414.2018.1553270

McIntosh, A., Willmott, C., Patton, D. A., Mitra, B., Brennan, J. H., Dimech-Betancourt, B., Howard, T. S., & Rosenfeld, J. V. (2019). An assessment of the utility and functionality of wearable head impact sensors in Australian football. *Journal of Science and Medicine in Sport*, *22*(7), 784–789. https://doi.org/10.1016/j.jsams.2019.02.004

McIntosh, S., Kovalchik, S., & Robertson, S. (2018a). Examination of player role in the Australian Football League using match performance data. *International Journal of Performance Analysis in Sport*, *18*(3), 451–462. https://doi.org/10.1080/24748668.2018.1486116

McIntosh, S., Kovalchik, S., & Robertson, S. (2018b). Validation of the Australian Football League player ratings. *International Journal of Sports Science and Coaching*, *13*(6), 1064–1071. https://doi.org/10.1177/1747954118758000

McIntosh, S., Kovalchik, S., & Robertson, S. (2019). Comparing subjective and objective evaluations of player performance in Australian Rules football. *PLoS ONE*, *14*(8), 1–16. https://doi.org/https:// doi.org/10.1371/journal.pone.0220901

McKinney, W. (2010). Data structures for statistical computing in python. *Proceedings of the 9th Python in Science Conference*, *445*, 51–56.

McNamara, D., Gabbett, T. J., Blanch, P., & Kelly, L. (2018). The relationship between variables in wearable microtechnology devices and cricket fast-bowling intensity. *International Journal of Sports Physiology and Performance*, *13*(2), 135–139. https://doi.org/10.1123/ijspp.2016-0540

McNamara, D., Gabbett, T. J., Chapman, P., Naughton, G., & Farhart, P. (2015). The validity of microsensors to automatically detect bowling events and counts in cricket fast bowlers. *International Journal of Sports Physiology and Performance*, *10*(1), 71–75. https://doi.org/10.1123/ijspp.2014-0062

Mehta, S. (2019). Relationship between workload and throwing injury in varsity baseball players. *Physical Therapy in Sport*, *40*, 66–70. https://doi.org/10.1016/j.ptsp.2019.08.001

Minnen, D., Westeyn, T. L., Starner, T., Ward, J. A., & Lukowicz, P. (2006). Performance metrics and evaluation issues for continuous activity recognition. *Proc. Int. Workshop on Performance Metrics for Intelligent Systems*, 141–148. https://doi.org/10.1145/1889681.1889687

Mitchell, E., Monaghan, D., & O’Connor, N. E. (2013). Classification of sporting activities using smartphone accelerometers. *Sensors (Switzerland)*, *13*(4), 5317–5337. https://doi.org/10.3390/s130405317

Mohammed, M., Khan, M., & Bashier, E. (2016). *Machine Learning: Algorithms and Applications*. CRC Press.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*, *6*(7). https://doi.org/10.1371/journal.pmed.1000097

Montoliu, R., Martín-Félez, R., Torres-Sospedra, J., Martínez-Usó, A., Torres-Sospedra, O., & Martínez-Usó, A. (2015). Team activity recognition in association football using a bag-of-words-based method. *Human Movement Science*, *41*, 165–178. https://doi.org/10.1016/j.humov.2015.03.007

Mooney, R., Corley, G., Godfrey, A., Quinlan, L. R., & ÓLaighin, G. (2015). Inertial sensor technology for elite swimming performance analysis: a systematic review. *Sensors (Switzerland)*, *16*(1), 1–55. https://doi.org/10.3390/s16010018

Moreno, E., Gomez, M. A., Lago-Peñas, C., & Sampaio, J. E. (2013). Effects of starting quarter score, game location, and quality of opposition in quarter score in elite women's basketball. *Kinesiology*, *45*(1), 48–54.

Navandarl, A., Veigal, S., Garcia, C., Ruedal, J., Torres, G., Horrol, D. C., & Navarrol, E. (2016). Gender effects on limb dominance in kicking biomechanics in elite soccer players. *34th International Conference on Biomechanics in Sports*. https://ojs.ub.uni-konstanz.de/cpa/article/viewFile/6933/6228

Nedergaard, N. J., Robinson, M. A., Eusterwiemann, E., Drust, B., Lisboa, P. J., & Vanrenterghem, J. (2017). The relationship between whole-body external loading and body-worn accelerometry during team-sport movements. *International Journal of Sports Physiology and Performance*, *12*(1), 18–26. https://doi.org/10.1123/ijspp.2015-0712

Neville, J., Rowlands, D., Wixted, A., & Jamesa, D. (2011). Determining over ground running speed using inertial sensors. *Procedia Engineering*, *13*, 487–492. https://doi.org/10.1016/j.proeng.2011.05.119

Niazi, A. H., Yazdansepas, D., Gay, J. L., Maier, F. W., Ramaswamy, L., Rasheed, K., & Buman, M. P. (2017). Statistical analysis of window sizes and sampling rates in human activity recognition. *HEALTHINF 2017 - 10th International Conference on Health Informatics, Proceedings*, *5*(Biostec), 319–325. https://doi.org/10.5220/0006148503190325

Nibali, A., He, Z., Morgan, S., & Greenwood, D. (2017). Extraction and classification of diving clips from continuous video footage. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2017-July*, 94–104. https://doi.org/10.1109/CVPRW.2017.18

Nicoella, D. P., Torres-Ronda, L., Saylor, K. J., & Schelling, X. (2018). Validity and reliability of an accelerometer-based player tracking device. *PLoS ONE*, *13*(2), e0191823. https://doi.org/doi.org/10.1371/journal.pone.0191823 February

Norris, M., Anderson, R., & Kenny, I. C. (2014). Method analysis of accelerometers and gyroscopes in running gait: a systematic review. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology*, *228*(1), 3–15. https://doi.org/10.1177/1754337113502472

Nunome, H., Ikegami, Y., Kozakai, R., Apriantono, T., & Sano, S. (2006). Segmental dynamics of soccer instep kicking with the preferred and non-preferred leg. *Journal of Sports Sciences*, *24*(5), 529–541. https://doi.org/10.1080/02640410500298024

Nunome, H., Inoue, K., Watanabe, K., Iga, T., & Akima, H. (2018). Dynamics of submaximal effort soccer instep kicking. *Journal of Sports Sciences*, *36*(22), 2588–2595. https://doi.org/10.1080/02640414.2018.1470216

Nweke, H. F., Teh, Y. W., Al-garadi, M. A., & Alo, U. R. (2018). Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: state of the art and research challenges. *Expert Systems with Applications*, *105*, 233–261. https://doi.org/10.1016/j.eswa.2018.03.056

O'Donoghue, P. (2014). *An Introduction to Performance Analysis of Sport* (1st ed.). Routledge.

O'Reilly, M. A., Whelan, D., Chanialidis, C., Friel, N., Delahunt, E., Ward, T., & Caulfield, B. (2015). Evaluating squat performance with a single inertial measurement unit. *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks, BSN 2015*. https://doi.org/10.1109/BSN.2015.7299380

O'Reilly, M. A., Whelan, D. F., Ward, T. E., Delahunt, E., & Caulfield, B. (2017). Technology in strength and conditioning tracking lower-limb exercises with wearable sensors. *Journal of Strength and Conditioning Research*, *31*(6), 1726–1736.

O'Reilly, M., Caulfield, B., Ward, T., Johnston, W., & Doherty, C. (2018). Wearable inertial sensor systems for lower limb exercise detection and evaluation: a systematic review. *Sports Medicine*, *48*(5), 1221–1246. https://doi.org/10.1007/s40279-018-0878-4

Oliveira, T., Gómez, M. Á., & Sampaio, J. (2012). Effects of game location, period, and quality of opposition in elite handball performances. *Perceptual and Motor Skills*, *114*(3), 783–794. https://doi.org/10.2466/30.06.pms.114.3.783-794

Op De Beéck, T., Meert, W., Schütte, K., Vanwanseele, B., & Davis, J. (2018). Fatigue prediction in outdoor runners via machine learning and sensor fusion. In ACM (Ed.), *24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining: Vol. July* (pp. 606–615). https://doi.org/10.1016/0003-4975(94)91354-4

Orchard, J., Walt, S., McIntosh, A., & Garlick, D. (2002). Muscle activity during the drop punt kick. *Science and Football IV*, 32–43.

Ordóñez, F. J., & Roggen, D. (2016). Deep convolutional and LSTM recurrent neural networks for mutlimodal wearable activity recognition. *Sensors*, *16*(1), 115. https://doi.org/10.3390/s16010115

Orloff, H., Sumida, B., Chow, J., Habibi, L., Fujino, A., & Kramer, B. (2008). Ground reaction forces and kinematics of plant leg position during instep kicking in male and female collegiate soccer players. *Sports Biomechanics*, *7*(2), 238–247. https://doi.org/10.1080/14763140701841704

Pal, S. K., & Mitra, S. (1992). Multilayer perceptron, fuzzy sets, and classification. *IEEE Transactions on Neural Networks*, *3*(5), 683–697.

Parmar, N., James, N., Hughes, M., Jones, H., & Hearne, G. (2017). Team performance indicators that predict match outcome and points difference in professional rugby league. *International Journal of Performance Analysis in Sport*, *17*(6), 1044–1056. https://doi.org/10.1080/24748668.2017.1419409

Parrington, L., Phillips, E., Wong, A., Finch, M., Wain, E., & MacMahon, C. (2016). Validation of inertial measurement units for tracking 100m sprint data. *34th International Conference of Biomechanics in Sports*, 442–445.

Peacock, J. C. A., & Ball, K. (2017). The relationship between foot-ball impact and flight characteristics in punt kicking. *Sports Engineering*, *20*(3), 221–230. https://doi.org/10.1007/s12283-017-0237-y

Peacock, J. C. A., & Ball, K. (2018a). Is there a sweet spot on the foot in Australian football drop punt kicking? *Journal of Sports Sciences*, *37*(4), 467–476. https://doi.org/10.1080/02640414.2018.1505408

Peacock, J. C. A., & Ball, K. (2018b). Kick impact characteristics of accurate Australian football drop punt kicking. *Human Movement Science*, *61*(July), 99–108. https://doi.org/10.1016/j.humov.2018.07.009

Peacock, J. C. A., & Ball, K. (2018c). Strategies to improve impact efficiency in football kicking. *Sports Biomechanics*, *3141*(May), 1–14. https://doi.org/10.1080/14763141.2018.1452970

Peacock, J. C. A., & Ball, K. (2018d). The influence of joint rigidity on impact efficiency and ball velocity in football kicking. *Journal of Biomechanics*, *71*, 245–250. https://doi.org/10.1016/j.jbiomech.2018.02.015

Peacock, J. C. A., & Ball, K. (2016). The impact phase of drop punt kicking: validation and experimental data of a mechanical kicking limb. *34th International Conference of Biomechanics in Sports*, 913–916.

Peacock, J. C. A., Ball, K., & Taylor, S. (2017). The impact phase of drop punt kicking for maximal distance and accuracy. *Journal of Sports Sciences*, *35*(23), 2289–2296. https://doi.org/10.1080/02640414.2016.1266015

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: machine learning in python. *Journal of Machine Learning Research*, *12*, 2825–2830. https://doi.org/10.1007/s13398-014-0173-7.2

Pelham, T. W., Robinson, M. G., & Holt, L. E. (2006). Assessing human movement with accelerometry. *Work*, *27*(1), 21–28.

Peng, K., Zhao, Y., Sha, X., Ma, W., Wang, Y., & Li, W. J. (2018). Accurate recognition of volleyball motion based on fusion of MEMS inertial measurement unit and video analytic. *IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, 440–444.

Pernek, I., Kurillo, G., Stiglic, G., & Bajcsy, R. (2015). Recognizing the intensity of strength training exercises with wearable sensors. *Journal of Biomedical Informatics*, *58*, 145–155. https://doi.org/10.1016/j.jbi.2015.09.020

Phillips, E., Farrow, D., Ball, K., & Helmer, R. (2013). Harnessing and understanding feedback technology in applied settings. *Sports Medicine*, *43*(10), 919–925. https://doi.org/10.1007/s40279-013-0072-7

Pieter, F., & Pieter, W. (1995). Speed and force in selected taekwondo techniques. *Biology of Sport*, *12*, 257–266.

Pinder, R. A., Davids, K., Renshaw, I., & Araújo, D. (2011). Representative learning design and functionality of research and practice in sport. *Journal of Sport and Exercise Psychology, 33*(1), 146–155. https://doi.org/10.1123/jsep.33.1.146

Plötz, T., Hammerla, N. Y., & Olivier, P. (2011). Feature learning for activity recognition in ubiquitous computing. *IJCAI International Joint Conference on Artificial Intelligence*, 1729–1734. https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-290

Pollard, R., & Gómez, M. Á. (2014). Comparison of home advantage in men's and women's football leagues in Europe. *European Journal of Sport Science*, *14*(SUPPL.1), S77–S83. https://doi.org/10.1080/17461391.2011.651490

Poppe, R. (2010). A survey on vision-based human action recognition. *Image and Vision Computing*, *28*(6), 976–990. https://doi.org/10.1016/j.imavis.2009.11.014

Prandoni, P., & Vetterli, M. (2005). *Signal Processing for Communications* (1st ed.). EPFL press. https://doi.org/10.1201/9781439808009

Preece, S. J., Goulermas, J. Y., Kenney, L. P. J., & Howard, D. (2009). A comparison of different feature extraction methods for classification of dynamic activities from accelerometer data. *IEEE Transactions on Biomedical Engineering*, *56*(3), 871–879. https://doi.org/10.1109/TBME.2008.2006190

Preece, S. J., Goulermas, J. Y., Kenney, L. P. J., Howard, D., Meijer, K., & Crompton, R. (2009). Activity identification using body-mounted sensors: review of classification techniques. In *Physiological Measurement* (Vol. 30, Issue 4, pp. R1–R33). https://doi.org/10.1088/0967-3334/30/4/R01

Provost, F., & Fawcett, T. (2001). Robust classification for imprecise environments. *Machine Learning*, *42*(3), 203–231. https://doi.org/10.1023/A:1007601015854

Pulling, C., Eldridge, D., & Lomax, J. (2017). Centre passes in the UK netball super league. *International Journal of Performance Analysis in Sport*, *16*(1), 389–400. https://doi.org/10.1080/24748668.2016.11868894

Python Software Foundation. (2018). *Python programming language* (3.6.6). https://www.python.org/

Qaisar, S., Imtiaz, S., Glazier, P., Farooq, F., Jamal, A., Iqbal, W., & Lee, S. (2013). A method for cricket bowling action classification and analysis using a system of inertial sensors. *International Conference on Computational Science and Its Applications*, 396–412. https://doi.org/10.1007/978-3-642-39649-6

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Ramanathan, V., Huang, J., Abu-El-Haija, S., Gorban, A., Murphy, K., & Fei-Fei, L. (2016). Detecting Events and Key Actors in Multi-person Videos. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 3043–3053. https://doi.org/10.1109/CVPR.2016.332

Raschka, S. (2018a). *Machine learning lecture notes: overfitting and underfitting*.

University of Wisconsin–Madison. http://stat.wisc.edu/~sraschka/teaching/stat479-fs2018/

Raschka, S. (2018b). Model evaluation, model selection, and algorithm selection in machine learning. *ArXiv Preprint ArXiv:1811.12808*. http://arxiv.org/abs/1811.12808

Rassem, A., El-Beltagy, M., & Saleh, M. (2017). Cross-country skiing gears classification using deep learning. *ArXiv Preprint ArXiv:1706.08924*, 1–14. http://arxiv.org/abs/1706.08924

Ravi, D., Wong, C., Lo, B., & Yang, G. (2017). A deep learning approach to on-node sensor data analytics for mobile or wearable devices. *IEEE Journal of Biomedical and Health Informatics*, *21*(1), 56–64. https://doi.org/10.1109/JBHI.2016.2633287

Reilly, M. A. O., Whelan, D. F., Ward, T. E., Delahunt, E., & Caulfield, B. (2017a). Classification of lunge biomechanics with multiple and individual inertial measurement units. *Sports Biomechanics*, *16*(3), 342–360. https://doi.org/10.1080/14763141.2017.1314544

Reilly, M. A. O., Whelan, D. F., Ward, T. E., Delahunt, E., & Caulfield, B. M. (2017b). Classification of deadlift biomechanics with wearable inertial measurement units. *Journal of Biomechanics*. https://doi.org/10.1016/j.jbiomech.2017.04.028

Reily, B., Zhang, H., & Hoff, W. (2017). Real-time gymnast detection and performance analysis with a portable 3D camera. *Computer Vision and Image Understanding*, *159*, 154–163. https://doi.org/10.1016/j.cviu.2016.11.006

Richards, J. G. (1999). The measurement of human motion: a comparison of commercially available systems. *Human Movement Science*, *18*(5), 589–602. https://doi.org/10.1016/S0167-9457(99)00023-8

Rindal, O. M. H., Seeberg, T. M., Tjønnås, J., Haugnes, P., & Sandbakk, Ø. (2018). Automatic classification of sub-techniques in classical cross-country skiing using a machine learning algorithm on micro-sensor data. *Sensors (Switzerland)*, *18*(1). https://doi.org/10.3390/s18010075

Ritchie, D., Hopkins, W. G., Buchheit, M., Cordy, J., & Bartlett, J. D. (2016). Quantification of training and competition load across a season in an elite Australian football club. *International Journal of Sports Physiology and Performance*, *11*(4), 474–479. https://doi.org/https://doi.org/10.1123/ijspp.2015-0294

Robertson, G., Caldwell, G. E., Hamill, J., Kamen, G., & Whittlesey, S. N. (2013). *Research Methods in Biomechanics* (2nd ed.). Human Kinetics.

Robertson, S., Back, N., & Bartlett, J. D. D. (2016). Explaining match outcome in elite Australian Rules football using team performance indicators. *Journal of Sports Sciences*, *34*(7), 637–644. https://doi.org/10.1080/02640414.2015.1066026

Robertson, S., Gupta, R., & McIntosh, S. (2016). A method to assess the influence of individual player performance distribution on match outcome in team sports. *Journal of Sports Sciences*, *34*(19), 1893–1900. https://doi.org/10.1080/02640414.2016.1142106

Robertson, S., Kremer, P., Aisbett, B., Tran, J., & Cerin, E. (2017). Consensus on measurement properties and feasibility of performance tests for the exercise and sport sciences: a Delphi study. *Sports Medicine - Open*, *3*(2), 1–10. https://doi.org/10.1186/s40798-016-0071-y

Robertson, S., Spencer, B., Back, N., & Farrow, D. (2018). A rule induction framework for the determination of representative learning design in skilled performance. *Journal of Sports Sciences*, 1–6. https://doi.org/10.1080/02640414.2018.1555905

Ronao, C. A., & Cho, S. (2016). Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Systems With Applications*, *59*, 235–244. https://doi.org/10.1016/j.eswa.2016.04.032

Ruano, M. A., Serna, A. D., Lupo, C., & Sampaio, J. E. (2016). Effects of game location, quality of opposition, and starting quarter score in the outcome of elite water polo quarters. *Journal of Strength and Conditioning Research*, *30*(4), 1014–1020. https://doi.org/10.1097/JSC.0000000000000234

Saba, T., & Altameem, A. (2013). Analysis of vision based systems to detect real time goal events in soccer videos. *Applied Artificial Intelligence*, *27*(7), 656–667. https://doi.org/10.1080/08839514.2013.787779

Sabatini, A. M. (2011). Estimating three-dimensional orientation of human body parts by inertial/magnetic sensing. *Sensors*, *11*(2), 1489–1525. https://doi.org/10.3390/s110201489

Sakamoto, K., & Asai, T. (2013). Comparison of kicking motion characteristics at ball impact between female and male soccer players. *International Journal of Sports Science & Coaching*, *8*(1), 63–76. https://doi.org/10.1260/1747-9541.8.1.63

Sakamoto, K., Sasaki, R., Hong, S., Matsukura, K., & Asai, T. (2014). Comparison of kicking speed between female and male soccer players. *Procedia Engineering*, *72*, 50–55. https://doi.org/10.1016/j.proeng.2014.06.011

Salman, M., Qaisar, S., & Qamar, A. M. (2017). Classification and legality analysis of bowling action in the game of cricket. *Data Mining and Knowledge Discovery*, *31*(6), 1706–1734. https://doi.org/10.1007/s10618-017-0511-4

Sampaio, J., Lago, C., Casais, L., & Leite, N. (2010). Effects of starting score-line, game location, and quality of opposition in basketball quarter score. *European Journal of Sport Science*, *10*(6), 391–396. https://doi.org/10.1080/17461391003699104

Santos-Gago, J. M., Ramos-Merino, M., Vallarades-Rodriguez, S., Álvarez-Sabucedo, L. M., Fernández-Iglesias, M. J., & García-Soidán, J. L. (2019). Innovative use of wrist-worn wearable devices in the sports domain: a systematic review. *Electronics (Switzerland)*, *8*(11), 1–29. https://doi.org/10.3390/electronics8111257

Schuldhaus, D., Zwick, C., Körger, H., Dorschky, E., Kirk, R., & Eskofier, B. M. (2015). Inertial sensor-based approach for shot/pass classification during a soccer match. *21st ACM KDD Workshop on Large-Scale Sports Analytics*, 1–4. https://www5.informatik.uni-erlangen.de/Forschung/Publikationen/2015/Schuldhaus15-ISA.pdf

Seel, T., Raisch, J., & Schauer, T. (2014). IMU-based joint angle measurement for gait analysis. *Sensors*, *14*(4), 6891–6909. https://doi.org/10.3390/s140406891

Seiffert, C., Khoshgoftaar, T. M., Van Hulse, J., Napolitano, A., & Raton, B. (2008). RUSBoost: improving classification performance when training data is skewed. *19th International Conference on Pattern Recognition*, 1–4. https://doi.org/10.1109/ICPR.2008.4761297

Seshadri, D. R., Li, R. T., Voos, J. E., Rowbottom, J. R., Alfes, C. M., Zorman, C. A., & Drummond, C. K. (2019). Wearable sensors for monitoring the physiological and biochemical profile of the athlete. *Npj Digital Medicine*, *2*(1), 72. https://doi.org/10.1038/s41746-019-0150-9

Sha, L., Lucey, P., Morgan, S., Pease, D., & Sridharan, S. (2013). Swimmer localization from a moving camera. *2013 International Conference on Digital Image Computing: Techniques and Applications, DICTA 2013*. https://doi.org/10.1109/DICTA.2013.6691533

Shah, H., Chokalingam, P., Paluri, B., & Pradeep, N. (2007). Automated stroke classification in tennis. *International Conference Image Analysis and Recognition*, 1128–1137.

Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: from theory to algorithms*. Cambridge University Press.

Sharma, M., Srivastava, R., Anand, A., Prakash, D., & Kaligounder, L. (2017). Wearable motion sensor based phasic analysis of tennis serve for performance feedback. *2017 IEEE International Conference on Acoustics, Speech and Siginal Processing*, 5945–5949.

Shepherd, J., James, D., Espinosa, H., Thiel, D., & Rowlands, D. (2018). A literature review informing an operational guideline for inertial sensor propulsion measurement in wheelchair court sports. *Sports*, *6*(2), 34. https://doi.org/10.3390/sports6020034

Shinkai, H., Nunome, H., Isokawa, M., & Ikegami, Y. (2009). Ball impact dynamics of instep soccer kicking. *Medicine and Science in Sports and Exercise*, *41*(4), 889–897. https://doi.org/10.1249/MSS.0b013e31818e8044

Shinkai, H., Nunome, H., Suito, H., Inoue, K., & Ikegami, Y. (2013). Cross-sectional change of ball impact in instep kicks from junior to professional footballers. *Science and Football VII*, 43–48.

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *ICLR 2015*, 1–14.

Smith, J., Ball, K., & MacMahon, C. (2009). Foot-to-ball interaction in preferred and non-preferred leg Australian Rules kicking. In R. Anderson, D. Harrison, & I. Kenny (Eds.), *Proceedings of the 27th International conference on biomechanics in sports* (pp. 650–653).

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. *Advances in Neural Information Processing Systems*, 2951–2959.

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, *45*(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Soro, A., Brunner, G., Tanner, S., & Wattenhofer, R. (2019). Recognition and repetition counting for complex physical exercises with deep learning. *Sensors (Switzerland)*, *19*(3). https://doi.org/10.3390/s19030714

Spangler, R., Rantalainen, T., Gastin, P. B., & Wundersitz, D. W. T. (2018). Inertial sensors are a valid tool to detect and consistently quantify jumping. *International Journal of Sports Medicine*, *39*(10), 802–808. https://doi.org/10.1055/s-0044-100793

Sprager, S., & Juric, M. B. (2015). Inertial sensor-based gait recognition: a review. *Sensors (Switzerland)*, *15*(9), 22089–22127. https://doi.org/10.3390/s150922089

Srivastava, R., Patwari, A., Kumar, S., Mishra, G., Kaligounder, L., & Sinha, P. (2015). Efficient characterization of tennis shots and game analysis using wearable sensors data. *2015 IEEE SENSORS - Proceedings*. https://doi.org/10.1109/ICSENS.2015.7370311

Stein, M., Janetzko, H., Lamprecht, A., Breitkreutz, T., Zimmermann, P., Goldlücke, B., Schreck, T., Andrienko, G., Grossniklaus, M., & Keim, D. A. (2018). Bring it to the pitch: combining video and movement data to enhance team sport analysis. *IEEE Transactions on Visualization and Computer Graphics*, *24*(1), 13–22. https://doi.org/10.1109/TVCG.2017.2745181

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

Sugomori, Y., Kaluža, B., Soares, F. M., & Souza, A. M. F. (2017). *Deep Learning - Practical Neural Networks with Java* (1st ed.). Packt Publishing.

Sullivan, C., Bilsborough, J. C., Cianciosi, M., Hocking, J., Cordy, J., & Coutts, A. J. (2014a). Match score affects activity profile and skill performance in professional Australian football players. *Journal of Science and Medicine in Sport*, *17*(3), 326–331. https://doi.org/10.1016/j.jsams.2013.05.001

Sullivan, C., Bilsborough, J. C., Cianciosi, M., Hocking, J., Cordy, J. T., & Coutts, A. J. (2014b). Factors affecting match performance in professional Australian football. *International Journal of Sports Physiology and Performance*, *9*(3), 561–566. https://doi.org/10.1123/IJSPP.2013-0183

Sutton, R. S. (1992). Introduction: the challenge of reinforcement learning. *Machine Learning*, *8*(3–4), 225–227. https://doi.org/10.1007/BF00992695

Sze, V., Chen, Y.-H. H., Yang, T.-J. J., & Emer, J. S. (2017). Efficient processing of deep neural networks: a tutorial and survey. *ArXiv Preprint*, *August*(2), 2295–2329.

https://doi.org/10.1109/JPROC.2017.2761740

Szegedy, C., & Ibarz, J. (2015). Scene classification with inception-7. In CVPR (Ed.), *Large-scale Scene Understanding Challenge Workshop (LSUN)* (p. 5). CVPR.

Tamura, T. (2014). Wearable inertial sensors and their applications. In E. Sazonov & M. R. Neuman (Eds.), *Wearable Sensors: Fundamentals, Implementation and Applications* (Issue 1, pp. 85–104). Academic Press. https://doi.org/10.1016/B978-0-12-418662-0.00024-6.

Taylor, J., Mellalieu, S., James, N., & Shearer, D. (2008). The influence of match location, quality of opposition, and match status on technical performance in professional association football. *Journal of Sports Sciences*, *26*(9), 885–895. https://doi.org/10.1080/02640410701836887

Taylor, L., Miller, E., & Kaufman, K. R. (2017). Static and dynamic validation of inertial measurement units. *Gait and Posture*, *57*(April), 80–84. https://doi.org/10.1016/j.gaitpost.2017.05.026

Teufl, W., Miezal, M., Taetz, B., Frohlichi, M., & Bleser, G. (2019). Validity of inertial sensor based 3D joint kinematics of static and dynamic sport and physiotherapy specific movements. *PLoS ONE*, *14*(2), 1–18. https://doi.org/10.1371/journal.pone.0213064

Thomas, G., Gade, R., Moeslund, T. B., Carr, P., & Hilton, A. (2017). Computer vision for sports: current applications and research topics. *Computer Vision and Image Understanding*, *159*, 3–18. https://doi.org/10.1016/j.cviu.2017.04.011

Titterton, D. H., & Weston, J. L. (2009). *Strapdown inertial navigation technology* (2nd ed.). AIAA.

Tora, M. R., Little, J. J., Chen, J., & Little, J. J. (2017). Classification of puck possession events in ice hockey. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 147–154. https://doi.org/10.1109/CVPRW.2017.24

Torres-Ronda, L., & Schelling, X. (2017). Critical process for the implementation of technology in sport organizations. *Strength and Conditioning Journal*, *39*(6), 54–59. https://doi.org/10.1519/SSC.0000000000000339

Uddin Ahamed, N., Kobsar, D., Benson, L., Clermont, C., Osis, S. T., & Ferber, R. (2019). Subject-specific and group-based running pattern classification using a single wearable sensor. *Journal of Biomechanics*. https://doi.org/10.1016/j.jbiomech.2019.01.001

van der Kruk, E., & Reijne, M. M. (2018). Accuracy of human motion capture systems for sport applications: state-of-the-art review. *European Journal of Sport Science*, *18*(6), 806–819. https://doi.org/10.1080/17461391.2018.1463397

Verheul, J., Gregson, W., Lisboa, P., Vanrenterghem, J., & Robinson, M. A. (2019). Whole-body biomechanical load in running-based sports: the validity of estimating ground reaction forces from segmental accelerations. *Journal of Science and Medicine in Sport*, *22*(6), 716–722. https://doi.org/10.1016/j.jsams.2018.12.007

Verheul, J., Warmenhoven, J., Lisboa, P., Gregson, W., Vanrenterghem, J., & Robinson, M. A. (2019). Identifying generalised segmental acceleration patterns that contribute to ground reaction force features across different running tasks. *Journal of Science and Medicine in Sport*, *22*(12), 1355–1360. https://doi.org/10.1016/j.jsams.2019.07.006

Victor, B., He, Z., Morgan, S., & Miniutti, D. (2017). Continuous video to simple signals for swimming stroke detection with convolutional neural networks. *ArXiv Preprint ArXiv:1705.09894*. https://doi.org/10.1111/j.1467-8330.1974.tb00606.x

Wagner, D., Kalischewski, K., Velten, J., & Kummert, A. (2017). Activity recognition using inertial sensors and a 2-D convolutional neural network. In IEEE (Ed.), *2017 10th International Workshop on Multidimensional (nD) Systems (nDS)* (pp. 1–6). https://doi.org/10.1109/NDS.2017.8070615

Wagner, J. F. (2018). About motion measurement in sports based on gyroscopes and accelerometers - an engineering point of view. *Gyroscopy and Navigation*, *9*(1), 1–18. https://doi.org/10.1134/S2075108718010091

Walker, C., Sinclair, P., Graham, K., & Cobley, S. (2017). The validation and application of inertial measurement units to springboard diving. *Sports Biomechanics*, *16*(4), 485–500. https://doi.org/10.1080/14763141.2016.1246596

Wang, Y., Zhao, Y., Chan, R. H. M., & Li, W. J. (2018). Volleyball skill assessment using a single wearable micro inertial measurement unit at wrist. *IEEE Access*, *6*, 13758–13765. https://doi.org/10.1109/ACCESS.2018.2792220

Ward, J. A., Lukowicz, P., & Gellersen, H.-W. W. (2010). Performance metrics for activity recognition. *ACM Transactions on Intelligent Systems and Technology*, *1*(1), 111–132. https://doi.org/10.1145/1889681.1889687

Whiteside, D., Cant, O., Connolly, M., & Reid, M. (2017). Monitoring hitting load in tennis using inertial sensors and machine learning. *International Journal of Sports Physiology and Performance*, *12*(9), 1212–1217. https://doi.org/https://doi.org/10.1123/ijspp.2016-0683

Windolf, M., Götzen, N., & Morlock, M. (2008). Systematic accuracy and precision analysis of video motion capturing systems-exemplified on the Vicon-460 system. *Journal of Biomechanics*, *41*(12), 2776–2780. https://doi.org/10.1016/j.jbiomech.2008.06.024

Wixted, A., Billing, D. C., & James, D. A. (2010). Validation of trunk mounted inertial sensors for analysing running biomechanics under field conditions, using synchronously collected foot contact data. *Sports Engineering*, *12*(4), 207–212. https://doi.org/10.1007/s12283-010-0043-2

Wixted, A., Portus, M., Spratford, W., & James, D. A. (2011). Detection of throwing in cricket using wearable sensors. *Sports Technology*, *4*(3–4), 134–140. https://doi.org/10.1080/19346182.2012.725409

Wolpert, D. H. (1996). The lack of a priori distinctions between learning algorithms. *Neural Computation*, *8*(7), 1341–1390. https://doi.org/10.1162/neco.1996.8.7.1341

Wolpert, D. H. (2012). What the no free lunch theorems really mean; how to improve search algorithms. In *SFI WORKING PAPER: 2012-10-17*.

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, *1*(1), 67–82. https://doi.org/10.1109/4235.585893

Woods, C. T., Jarvis, J., & Mckeown, I. (2019). Differences between elite and semi-elite Australian football conceptualised through the lens of ecological dynamics. *Sports*, *7*(7), 159. https://doi.org/doi.org/10.3390/sports7070159

Woods, C. T., Robertson, S., & Collier, N. F. (2017). Evolution of game-play in the Australian Football League from 2001 to 2015. *Journal of Sports Sciences*, *35*(19), 1879–1887. https://doi.org/10.1080/02640414.2016.1240879

Worsey, M., Espinosa, H., Shepherd, J., & Thiel, D. (2019a). A systematic review of performance analysis in rowing using inertial sensors. *Electronics*, *8*(11), 1304. https://doi.org/https://doi.org/10.3390/electronics8111304

Worsey, M., Espinosa, H., Shepherd, J., & Thiel, D. (2019b). Inertial sensors for performance analysis in combat sports: a systematic review. *Sports*, *7*(1), 28. https://doi.org/10.3390/sports7010028

Wundersitz, D. W. T., Gastin, P. B., Richter, C., Robertson, S., & Netto, K. J. (2015). Validity of a trunk-mounted accelerometer to assess peak accelerations during walking, jogging and running. *European Journal of Sport Science*, *15*(5), 382–390. https://doi.org/10.1080/17461391.2014.955131

Wundersitz, D. W. T., Gastin, P. B., Robertson, S., Davey, P. C., & Netto, K. J. (2015). Validation of a trunk-mounted accelerometer to measure peak impacts during team sport movements. *International Journal of Sports Medicine*, *36*(9), 742–746. https://doi.org/10.1055/s-0035-1547265

Wundersitz, D. W. T., Josman, C., Gupta, R., Netto, K. J., Gastin, P. B., & Robertson, S. (2015). Classification of team sport activities using a single wearable tracking device. *Journal of Biomechanics*, *48*(15), 3975–3981. https://doi.org/10.1016/j.jbiomech.2015.09.015

Yang, C. C., & Hsu, Y. L. (2010). A review of accelerometry-based wearable motion detectors for physical activity monitoring. *Sensors*, *10*(8), 7772–7788. https://doi.org/10.3390/s100807772

Yang, J. B., Nguyen, M. N., San, P. P., Li, X. L., & Krishnaswamy, S. (2015). Deep convolutional neural networks on multichannel time series for human activity recognition. *IJCAI International Joint Conference on Artificial Intelligence*, *2015-Janua*, 3995–4001.

Yao, B., & Fei-Fei, L. (2010). Modeling mutual context of object and human pose in human-object interaction activities. *Computer Vision and Pattern Recognition*, 17–24.

Young, C. M., Luo, W., Gastin, P. B., Tran, J., & Dwyer, D. B. (2019). The relationship between match performance indicators and outcome in Australian football. *Journal of Science and Medicine in Sport*, *22*(4), 467–471.

https://doi.org/10.1016/j.jsams.2018.09.235

Young, C., & Reinkensmeyer, D. J. (2014). Judging complex movement performances for excellence: a principal components analysis-based technique applied to competitive diving. *Human Movement Science*, *36*, 107–122. https://doi.org/10.1016/j.humov.2014.05.009

Yu, G., Jang, Y. J., Kim, J., Kim, J. H., Kim, H. Y., & Panday, S. B. (2016). Potential of IMU sensors in performance analysis of professional alpine skiers. *Sensors (Switzerland)*, *16*(4), 1–21. https://doi.org/10.3390/s16040463

Zaki, M. J., & Meira Jr, W. (2014). *Data mining and analysis: fundamental concepts and algorithms*. University Press.

Zambom-Ferraresi, F., Rios, V., & Lera-López, F. (2018). Determinants of sport performance in European football: what can we learn from the data? *Decision Support Systems*, *114*(March), 18–28. https://doi.org/10.1016/j.dss.2018.08.006

Zebin, T., Scully, P. J., & Ozanyan, K. B. (2016). Human activity recognition with inertial sensors using a deep learning approach. *Proc. of IEEE Sensors 2016*, *1*, 1–3. https://doi.org/10.1109/ICSENS.2016.7808590

Zeng, H., & Zhao, Y. (2011). Sensing movement: microsensors for body motion measurement. *Sensors*, *11*(1), 638–660. https://doi.org/10.3390/s110100638

Zeng, M., Nguyen, L. T., Yu, B., Mengshoel, O. J., Zhu, J., Wu, P., & Zhang, J. (2014). Convolutional neural networks for human activity recognition using mobile sensors convolutional neural networks for human activity. *Proceedings of the 6th International Conference on Mobile Computing, Applications and Services*, *July 2015*, 197–205. https://doi.org/10.4108/icst.mobicase.2014.257786

Zhang, C., Yang, F., Li, G., Zhai, Q., Jiang, Y., & Xuan, D. (2018). MV-Sports: a motion and vision sensor integration-based sports analysis system. *Proceedings - IEEE INFOCOM*, 1070–1078. https://doi.org/10.1109/INFOCOM.2018.8485910

Zhang, S., Wei, Z., Nie, J., Huang, L., Wang, S., & Li, Z. (2017). A review on human activity recognition using vision-based method. *Journal of Healthcare Engineering*, *2017*(1–31). https://doi.org/10.1155/2017/3090343

Zheng, Y., Liu, Q., Chen, E., Ge, Y., & Zhao, J. L. (2014). Time series classification using multi-channels deep convolutional neural networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *8485 LNCS*, 298–310. https://doi.org/10.1007/978-3-319-08010-9_33

Zhu, G., Xu, C., Huang, Q., & Gao, W. (2006). Action recognition in broadcast tennis video. *Proceedings - International Conference on Pattern Recognition*, *1*, 251–254. https://doi.org/10.1109/ICPR.2006.205

Zhu, G., Xu, C., Huang, Q., Gao, W., & Xing, L. (2006). Player action recognition in broadcast tennis video with applications to semantic analysis of sports game. *Proceedings of the 14th Annual ACM International Conference on Multimedia, MM*

2006, 431–440. https://doi.org/10.1145/1180639.1180728

Ziaeefard, M., & Bergevin, R. (2015). Semantic human activity recognition: a literature review. *Pattern Recognition*, *48*(8), 2329–2345. https://doi.org/10.1016/j.patcog.2015.03.006

Zimmermann, T., Taetz, B., & Bleser, G. (2018). IMU-to-segment assignment and orientation alignment for the lower body using deep learning. *Sensors (Switzerland)*, *18*(1), 1–35. https://doi.org/10.3390/s18010302

Appendix

Appendix A. Definitions of the 13 performance indicators used in the analyses undertaken in Chapter Five; *The relationship of team and individual athlete performances on match quarter outcome in elite women's Australian Rules football*.

Performance indicator	Definition
Clanger kick	A disposal which goes directly to an opposition athlete; a conceded free kick; dropped mark or fumble under no pressure.
Contested mark	A mark achieved while engaging in a contest.
Contested possession	A possession achieved as a result of winning a contest.
Effective long kick	A kick of more than 40 meters to a teammate that hits the intended target.
Effective short kick	A kick of led than 40 meters that results in the intended target retaining possession.
Handball	Disposing of the football by hitting it with the clenched fist of one hand, while holding it with the other.
Hit-out	A tap by a Ruckman after a ball up or bounce by the umpire.
Ineffective kick	Kicks that are not advantageous to the team, but do not directly turn the ball over to the opposition.
Inside 50	The act of running or passing the ball into the 50 m arc at the opposition's defensive end of the field.
Kick: Handball ratio	The number of kicks compared to handballs expressed as a ratio.
Tackles	Taking hold of an opposition athlete in possession of the ball, in order to impede his progress or to force him to dispose of the ball quickly.
Uncontested marks	Marks taken under no physical pressure from an opponent, including marks taken on a lead and from opposition kicks.
Uncontested possession	A possession achieved without having to engage in a contest.

Appendix B. Original published article in Chapter Three: ***Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance***

VICTORIA UNIVERSITY

THE NEW WAY TO DO UNI

OFFICE FOR RESEARCH TRAINING, QUALITY AND INTEGRITY

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION: PAPERS INCORPORATED IN THESIS

This declaration is to be completed for each conjointly authored publication and placed at the beginning of the thesis chapter in which the publication appears.

1. PUBLICATION DETAILS (to be completed by the candidate)

Title of Paper/Journal/Book: Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance

Surname: Cust First name: Emily

Institute: Institute for Health and Sport Candidate's Contribution (%): 70

Status:

Accepted and in press: ☐ Date:

Published: ☑ Date: 11-10-18

2. CANDIDATE DECLARATION

I declare that the publication above meets the requirements to be included in the thesis as outlined in the HDR Policy and related Procedures – policy.vu.edu.au.

Emily Cust Digitally signed by Emily Cust Date: 2020.02.17 14:33:02 +11'00'	17-2-2020
Signature	**Date**

3. CO-AUTHOR(S) DECLARATION

In the case of the above publication, the following authors contributed to the work as follows:

The undersigned certify that:

1. They meet criteria for authorship in that they have participated in the conception, execution or interpretation of at least that part of the publication in their field of expertise;
2. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

3. There are no other authors of the publication according to these criteria;

4. Potential conflicts of interest have been disclosed to a) granting bodies, b) the editor or publisher of journals or other publications, and c) the head of the responsible academic unit; and

5. The original data will be held for at least five years from the date indicated below and is stored at the following **location(s)**:

6 Wingan Avenue Camberwell, VIC

Name(s) of Co-Author(s)	Contribution (%)	Nature of Contribution	Signature	Date
Professor Sam Robertson	15	Study design, assistance with literature search strategy and paper structure, reviewing.	[redacted]	17-2-202 0
Dr Alice Sweeting	10	Reviewing and editing	[redacted]	17-2-2020
Dr Kevin Ball	5	Study design	[redacted]	17-2-2020

Updated: September 2019

Emily E Cust, Alice J Sweeting, Kevin Ball & Sam Robertson (2019) Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance, Journal of Sports Sciences, 37:5, 568-600, DOI: 10.1080/02640414.2018.1521769

Appendix C. Original published article in Chapter Five: ***The relationship of team and individual athlete performances on match quarter outcome in elite women's Australian Rules football.***

OFFICE FOR RESEARCH TRAINING, QUALITY AND INTEGRITY

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION: PAPERS INCORPORATED IN THESIS

This declaration is to be completed for each conjointly authored publication and placed at the beginning of the thesis chapter in which the publication appears.

1. PUBLICATION DETAILS (to be completed by the candidate)

Title of Paper/Journal/Book: The relationship of team and individual athlete performances on match quarter outcome in elite women's Australian Rules football

Surname: Cust First name: Emily

Institute: Institute for Health and Sport Candidate's Contribution (%): 60

Status:

Accepted and in press: ☐ Date:

Published: ☑ Date: October 2019

2. CANDIDATE DECLARATION

I declare that the publication above meets the requirements to be included in the thesis as outlined in the HDR Policy and related Procedures – policy.vu.edu.au.

Emily Cust Digitally signed by Emily Cust Date: 2020.03.03 14:23:44 +11'00'	03/03/2020
Signature	**Date**

3. CO-AUTHOR(S) DECLARATION

In the case of the above publication, the following authors contributed to the work as follows:

The undersigned certify that:

1. They meet criteria for authorship in that they have participated in the conception, execution or interpretation of at least that part of the publication in their field of expertise;

2. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

Victoria University ABN 83776954731
CRICOS Provider No. 00124K (Melbourne),
02475D (Sydney), RTO 3113

3. There are no other authors of the publication according to these criteria;

4. Potential conflicts of interest have been disclosed to a) granting bodies, b) the editor or publisher of journals or other publications, and c) the head of the responsible academic unit; and

5. The original data will be held for at least five years from the date indicated below and is stored at the following **location(s)**:

6 Wingan Avenue Camberwell, VIC

Name(s) of Co-Author(s)	Contribution (%)	Nature of Contribution	Signature	Date
Professor Sam Robertson	20	Study design, data analysis, reviewing and editing	[redacted]	03/03/2020
Dr Alice Sweeting	5	Data visualisations, reviewing and editing	[redacted]	03/03/2020
Dr Kevin Ball	5	Reviewing and editing	[redacted]	03/03/2020
Hamish Anderson	10	Data collection	[redacted]	03/03/2020

Updated: September 2019

Journal of Science and Medicine in Sport 22 (2019) 1157–1162

Contents lists available at ScienceDirect

Journal of Science and Medicine in Sport

journal homepage: www.elsevier.com/locate/jsams

Original research

The relationship of team and individual athlete performances on match quarter outcome in elite women's Australian Rules football

Emily E. Cust[a,b], Alice J. Sweeting[a,b], Kevin Ball[a], Hamish Anderson[b], Sam Robertson[a,b,*]

[a] Institute for Health and Sport (IHES), Victoria University, Australia
[b] Western Bulldogs Football Club, Australia

ARTICLE INFO

Article history:
Received 20 January 2019
Received in revised form 17 April 2019
Accepted 7 May 2019
Available online 10 May 2019

Keywords:
Sports analytics
AFLW
Machine learning
Performance analysis

ABSTRACT

Objectives: To evaluate the relationships between the athlete distribution of team performance indicators and quarter outcome in elite women's Australian Rules football matches.
Design: Retrospective longitudinal cohort analysis.
Methods: Thirteen performance indicators were obtained from 56 matches across the 2017 and 2018 Australian Football League Women's (AFLW) seasons. Absolute and relative values of 13 performance indicators were obtained for each athlete, in each quarter of all matches. Eleven features were further extracted for each performance indicator, resulting in a total of 169 features. Generalised estimating equations (GEE) and regression decision trees were run across the different feature sets and dependent variables, resulting in 22 separate models.
Results: The GEE algorithm produced slightly lower mean absolute errors across all dependent variables and feature sets comparative to the regression decision tree models. Quarter outcome was more accurately explained when considered as total points scored comparative to quarter score margin. Team differential and the 75th percentile of individual athlete Inside 50s were the strongest features included in the models.
Conclusions: Modelling performance statistics by quarter outcomes provides specific practical information for in-game tactics and coaching in relation to athlete performances each quarter. Within the current elite women's Australian Rules football competition, key high performing individual athletes' skilled performances within matches contribute more to success rather than a collective team effort.

Practical implications

- Identifying key contributing athlete match skills in the AFLW can practically inform coaches on training drills, athlete development, and tactical match approaches per quarter relative to their opposition.
- Current AFLW match success is influenced more by marquee athletes within a team as opposed to an even team contribution, which indicates the need for athlete development support and resources to be improved in women's football.
- Results position kick variables as strong performance indicators hence training plans should place emphasis on individual kick skill development.

1. Introduction

Match performance analysis in team sports can provide a greater understanding of the physical, technical and tactical characteristics athletes require to produce a successful competition outcome.[1] Analysis may help guide coaching staff on training practices that replicate and prepare athletes for the demands of competition.[2] Determining the form and function of events within the specifics of a sport for teams and individual athletes should inform the variables for quantification of performance and therefore the sport analytics approaches used to facilitate future coaching practise.[3] The relationship between match athlete performance indicators[1] in Australian Rules football (AF) have been investigated heavily in the literature across elite male teams[4–6], individual athlete contributions[7], and recently, elite women's teams.[8]

In 2017, AF established a national elite women's competition, the Australian Football League Women's (AFLW) in addition to the long running elite men's Australian Football League (AFL). The opening two seasons consisted of a seven-round home-and-away

* Corresponding author at: Institute for Health and Sport (IHES), Victoria University, P.O. Box 14428, Melbourne, VIC 8001, Australia.
E-mail address: sam.robertson@vu.edu.au (S. Robertson).

https://doi.org/10.1016/j.jsams.2019.05.004
1440-2440/

competition, incorporating eight teams. As the depth of talent and resources develop, the league has set plans for expansion to the competition. This in turn will provide further opportunities to investigate elite women's football training and match physical, technical and tactical areas. For example, information on athlete match demands may improve club training practices, assess the effectiveness of the rule changes implemented differently to the AFL competition, and inform league directors on the quality of development in the competition.

Research in women's AF is currently limited.[8–10] Recent research on the physical demands, technical performances and activity profiling across field playing positions of match-play in AFLW[9] has provided initial insights into match activity. There were no absolute differences between physical variables, based on match playing position, in the AFLW.[9] Furthermore, no positional group differences were noted for skill measures such as total kicks, handballs, contested possessions, uncontested possessions, and marks.[9] This is in contrast to several physical demands characteristic differences that have been observed across athlete match positions in the AFL.[11,12] The specificity of AFLW positional roles may not yet be established and consequently, athletes may be more homogenous in playing tactics and physical abilities comparative to AFL players.[9] Although there are inherent differences between the AFL and AFLW games such as amount of time and players on ground creating independent constraints between each competition. Currently focussing on the AFLW as an independent competition and quantifying match variables as the league matures may be more beneficial over a direct sport analytics comparison of the AFL and AFLW given the current game constraint differences. Match performance indicator analysis assessed the relationship between team skill involvements and match outcome in the first season of AFLW.[8] Match outcome, defined as win/ loss and score margin, indicated that higher uncontested possessions and Inside 50: goal score ratio were the strongest predictors for winning. Increased kick numbers and contested marks resulted in a higher team ladder position.[8]

Match success in the AFL has been linked to individual athlete skill efficiency rather than their physical activity profile.[13] Specifically, physical activity profiles may increase, yet skill involvements efficiency may decrease when teams lose a quarter.[14] An analysis inclusive of athlete skilled match performances, by individual match quarter and across feature derived performance distributions, is yet to be investigated in AF. A quarter by quarter approach could provide differentiated information about specific technical and tactical foci for coaches. Situational variables such as starting quarter score, quality of opposition, and whether the team is playing at a home or away ground have shown influence on elite women's team sport quarter performances.[15,16] Analysing by quarter could improve relevancy of results, given output may fluctuate across quarters for several reasons.[17] During quarter time breaks, coaches can address athletes directly. Knowledge or information transfer from the coach to the playing group should be of purpose, work in context of the current events and tie in with previously delivered knowledge the coach has provided prior to the match to maximise group understandings of the information.[18] Factors may affect the extent of knowledge transfer to the playing group between the restricted quarter time frame such as the coach's communication style, clarity of information, and a player's prior involvement in the match strategy system development.[18]

Quantifiable information about skill performances, in context of the match, could further justify changes to team playing strategies based on the current situation. With respect to influence on the team match outcome, quantification of individual athlete distributions have been linked to successful match outcome.[7] Specifically, lower 75th, 90th and 95th percentile values for team goals and higher 25th and 50th percentile values for disposals.[7] Measured athlete performance distribution information calculated by individuals rather than a team data as a whole could determine the influential basis for match success in the AFLW. Information may also convey whether success in the current AFLW game constitutes a more collective team-based effort or skewed to a few stronger individual athletes. Findings may inform match team selection to suit the current game style influence or opposition at play. This may be important as several new teams are introduced to the competition over the next few years making key athlete retention or attainment a challenge.

The primary aim of this study was to evaluate the relationship of AFLW athlete skill performance indicator distributions, to explain match quarter outcomes during the 2017 and 2018 seasons. Secondly, this study aimed to compare quarter outcome model error rates from separate machine learning approaches, based on the varied input feature set variables.

2. Methods

All match performance indicators were obtained from the AFL match statistics provider, Champion Data Pty Ltd. (Melbourne, Australia) online portal, Coaches Information Analysis (CIA). Data collection by Champion Data involves human recordings of the statistics by working at each match, as such the inter- and intra-reliability of the data is currently unknown.[7,19] Reliability and validity of the data has been assessed independently to determine the agreement between the Champion Data and author-coded values.[7] Reliability assessment showed very high agreement levels, intra-class correlation coefficient range 0.947–1.000. The validity of author's coding showed low absolute error in regards to the Champion Data, RMSE range 0.0–4.5,[7] indicating the expected absolute error points between each performance indicator for each game. A total of 56 matches across the 2017 and 2018 AFLW season were obtained and 13 discrete performance indicators were selected.[6–8] The definitions for each indicator are provided in Appendix A. Absolute values from every quarter ($n=224$), match ($n=56$), athlete ($n=154$), and all teams ($n=7$), across performance indicators, were extracted into a custom Excel™ spreadsheet. Quarter outcome (as win = 1 or loss = 0 or draw = 2), quarter score margin (points), match outcome (win/ loss) and match score margin (points) were recorded. Score points were recorded as both their absolute values and relative values to the opposition at play. The University's Human Research Ethics Committee approved the study (application number 0000025654).

Each athlete's contribution to their team's total were converted to a relative form, as a percentage of their team total for each match.[7] Features extracted for each performance indicator were the minimum, maximum, mean, standard deviation and percentiles, at 0.05, 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95, resulting in a total of 143 features (11 features $\times$ 13 performance indicators).[7] Features were collated with team name, round number (1–7), season (2017 or 2018), quarter number (1–4), quarter outcomes (loss, win or draw), and match outcome (loss, win or draw). The stability of the data performance profiles[20] was plotted and assessed by visual inspection, and deemed acceptable to model for comparison of analysis methods and reporting of results for practical feedback.

A total of 22 models were developed. Modelling of statistics by machine learning was performed for quarter points scored (absolute), and quarter point margin relative to the opposition (relative). Four features sets were used in separate models: total performance indicator values ($n=13$), performance indicator values relative ($n=13$) to the opposition, derived feature distribution values for each performance indicator ($n=143$), combined performance indicator total, relative and feature distribution values ($n=169$).

Regression decision trees were computed with Python version 3.6.6,[21] using the package Scikit-learn.[22] Data was split into a 70%

Table 1
Model results across data variables evaluated by mean absolute error (MAE).

Algorithm	Parameters	Model	Variables (number of features)	MAE (points)
Regressor decision tree	Total training data	1	Quarter margin and PI totals (13)	6.65
		2	Quarter margin and PI relative (13)	5.93
		3	Quarter margin and PI features distributions (143)	6.81
		4	Quarter margin and combined PI totals, relatives and feature distributions (169)	5.98
		5	Quarter score and PI totals (13)	4.47
		6	Quarter score and PI relative (13)	4.31
		7	Quarter score and PI feature distributions (143)	4.32
		8	Quarter score and combined PI totals, relatives and feature distributions (169)	4.10
Regressor decision tree	70% train set and 30% test set	9	Quarter margin and PI totals (13)	8.56
		10	Quarter margin and PI relative (13)	7.63
		11	Quarter margin and PI features distributions (143)	9.57
		12	Quarter margin and combined PI totals, relatives and feature distributions (169)	8.38
		13	Quarter score and PI totals (13)	5.60
		14	Quarter score and PI relative (13)	5.59
		15	Quarter score and PI feature distributions (143)	5.45
		16	Quarter score and combined PI totals, relatives and feature distributions (169)	5.74
Generalised estimating equations	Total training data	17	Quarter margin and PI totals (13)	7.13
		18	Quarter margin and PI relative (13)	6.18
		19	Quarter margin and PI features distributions (143)	6.03
		20	Quarter margin and combined PI totals, relatives and feature distributions (169)	5.12
		21	Quarter score and PI totals (13)	4.48
		22	Quarter score and PI relative (13)	4.64
		23	Quarter score and PI feature distributions (143)	3.83
		24	Quarter score and combined PI totals, relatives and feature distributions (169)	3.60

PI, performance indicator.

training set and 30% testing set. Each regressor tree was computed with a minimum sample split of 30 (>13.4% of total sample) and a maximum depth of five. Several model parameter combinations were tested to reduce the risk of overfitting whilst minimising error.[23] Regression trees were also computed using the whole training set for the four feature sets as a comparison. Generalised Estimating Equations (GEE) were also constructed separately in R[24] for each dependent variable and feature sets. Team (n = 7) was considered a fixed repeated measure and a greedy feature selection was implemented for feature selection in model construction. Model evaluation was based on the mean absolute error (MAE) computed from the withheld testing set, unless otherwise indicated.

3. Results

The MAE results for each model are presented in Table 1. The GEE produced lower MAE's than the decision trees (Table 1). Across both analysis approaches, the influence of performance indicators was more accurately explained by quarter score points, as opposed to quarter score margin, for all input feature set variables. The mean average difference between score margin and score points MAE results was 2.32 points (Table 1). Modelling performance statistics by quarter score points using the relative values feature set (n = 13) resulted in one of the lower MAE scores for both the GEE (3.83) and the decision tree (5.59). The lowest prediction errors for both models were on larger feature sets. The GEE MAE was 3.60 on the 169-feature set comprised of the combined total, relative and feature distribution values. The decision tree MAE was 5.45 on the 143-feature set comprised of the derived feature distribution values.

Rule outputs from the two regressor decision tree models, with the lowest MAE, are shown in Figs. 1 and 2. The relative performance indicator of team differential of Inside 50 values (Fig. 1) and feature distribution Inside 50s in the 75th percentile (Fig. 2) contributed most strongly to the models. Interpretation involves following the branches down, from the root node representing the outcomes for each test, to the final terminal node to define the regression decision rules for the model. For example, in Fig. 1 following down the right side, teams with relative Inside 50s greater than −3.5, relative kicks long greater than 5.5 scored more points per quarter, model prediction of 18.5 points based on 22 samples. Teams with higher contributions from more athletes to their Inside 50 count, short and long kicks, and lower contributions from more athletes to their ineffective kick counts are more successful per quarter (Fig. 2). See Figs. 1 and 2 for further examples for rule sets. The defined rules represent performance skill fulfilment requirements for teams to achieve a successful quarter score or score margin outcome.

4. Discussion

This study assessed the extent to which AFLW athlete skill performance distributions explain match quarter outcome across the first two seasons of the inaugural AFLW national competition. Key results indicate that modelling data by quarter score points total was more accurate compared to quarter score margin. Teams with more successful Inside 50 entries than their opposition likely scored more points in the quarter.

Modelling performance indicator data by quarter and not an entire match may allow for specific information and clearer relationships between the variables and success within different periods of a match.[16] During matches, coaches have the chance to address the playing group and reset tactics at quarter time breaks. Specific quarter-based skill influence information may aid in modifying individual athlete and team tactics, in comparison to the opposition as shown in elite women's basketball.[25] Therefore, breaking performance indicator data into the influence by quarter may provide targeted information for coaches during matches. As the league expands and more data becomes available, longitudinal comparisons would be of interest. In comparison of the two approaches, the GEE produced lower prediction errors across all data input variables. This may indicate that a simpler model approach is more appropriate for the current smaller dataset with relatively low feature dimensionality. However, to provide a practical outcome for coaches, a decision tree model may be more applicable as the output does not consider all features. Rather, decision trees provide a practical, parsimonious rule set for coaches who may be focused on the most influential performance indicators.

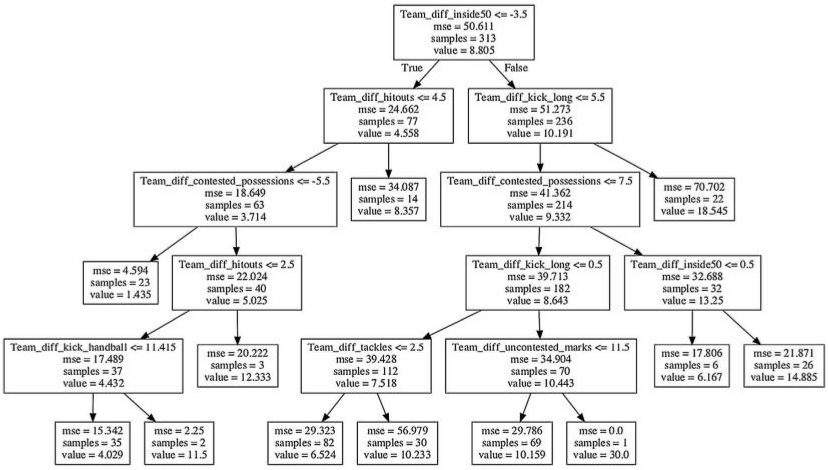

Fig. 1. Regressor decision tree output of model 14, quarter score points and performance indicator relative values. *diff*, differential, *mse*, mean sample error.

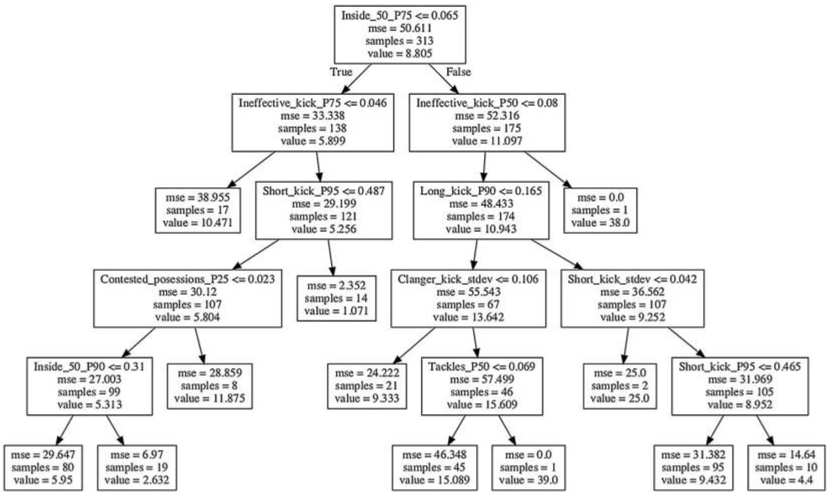

Fig. 2. Regressor decision tree output of model 15, quarter score points and performance indicator feature distributions. *mse*, mean sample error; *P25, P50, P75, P90, P95*, percentile level; *stdev*, standard deviation.

Features or variables are representative aspects of data that should be relevant, in that they have an influence on the model result with a function that is not assumed by the rest.[26] Performance indicators that were a direct function of scoring in AFLW, including shots at goal, goal assist, behind assist and goal accuracy were not included in this study. These variables would potentially trivialise the process of determining performance skills which influence match success. Modelling quarter points scored produced the lowest prediction errors on the larger feature sets GEE (n = 169) and decision tree (n = 143) for both algorithms. But this was only a slight improvement from using the smaller relative values feature sets (n = 13). A larger data set could facilitate improved feature extraction and selection engineering for better representation of the data characteristics. More efficient algorithm processing and prediction accuracy[27] may also be increased. Further extracting distribution features, from individual athletes, demonstrates the structure contributions for AFLW teams. Interestingly, results suggest that in contrast to the AFL game, increased match skill performance contributions from key high performing individual athletes is more beneficial for team success. This is suggested by the higher percentile feature distributions contributing most strongly to the decision model (Fig. 2). For example, the Inside 50 P75, short kick P95, long kick P90 and ineffective kick P75 values.

Successful outcomes in the AFL involve relatively even performances from athletes across a team.[7] The comparatively higher performance contributions by key individual athletes to team success in the AFLW may be explained by the fact it is a new competition format and across many facets is still developing. As such, the level of game plan seen in the AFL competition[7,28,29] may yet be reasonable in the AFLW due to the variety of AFLW athlete game experience and skill maturity levels being contracted. The skill development of AFLW athletes, who have either recently progressed from junior competitions or transitioned from another sport and hence not marquee or high performing athletes may be also affected by the lower resourced professional support structures and training opportunities currently experienced in the AFLW. As opposed to the well-established AFL, where newly contracted athletes are highly coached, skill acquired and AF experienced before competing in AFL level matches.[30] This may be partly because of lower coaching and sports science resource support in the AFLW competition relative to elite male AF. These factors could be contributing to individual athlete dominance in the AFLW, potentially preventing collective team contributions towards successful match outcome.

Comparison of the current results to AFLW match skill analysis during the 2017 season only[8] is difficult, due to the differentiating features sets used. In Black et al., variables with direct functions of scoring were used. In order to build upon this previous analysis,[8] further data feature extraction from a larger sample size and revised statistical modelling was run in the present study. Breaking down the performance indicators to types of the variable, for example, including long, short and ineffective kicks allows for expansion of the key performance measures.

Practically, as the strongest features in the regression decision tree models relate to kick performance indicators, clubs may look to emphasise kick skill development. Inside 50s, hit outs and contested possessions, by key athletes, contribute most to quarter success during matches. AFLW clubs may also look to compile teams with capable skilled kickers and recruit future athletes with current or potential strong kick skills.[5] Game plan development around a kick dominant ball movement strategy, particularly in hit-out clearances and efficient Inside 50 entries may also be of match tactical advantage. Coaches may work specifically with key forward and midfield athletes to develop efficient plays and decision making from centre bounce to Inside 50 entry possession chains, in order to maximise scoring opportunities. Improving an athlete's kick execution skills may also benefit kick delivery and mark success from a team member in contested possessions during matches. Analysis of match performance statistical information can also be applied off field in the athlete recruitment department.[19] As the AFLW is in its infancy, a greater understanding of team and individual contributions to winning may highlight what performance characteristics are beneficial towards maximising team success. Recruiters could make strategic decisions on selecting athletes that currently exhibit or have the potential to develop the key performance characteristics identified.

Future research may look to investigate the contextual variables around match play on the outcome such as travel requirements, days between matches and player interchange rotations per quarter. Specifically, given the current short home-and-away season, increased importance is on the outcome of each match for ladder positioning. Across different team sports, contextual variables influence match outcomes and performance indicators,[25,31] particularly in team field sports who play multi-round home-and-away seasons.[32–34] Furthermore, spatiotemporal data characteristics of players could be analysed to explain team behaviours in match play styles and tactics[29] in the AFLW.

5. Conclusion

Quarter success in the AFLW was characterised by greater Inside 50s as a relative to the opposition and key athletes in the 0.75 percentile performing Inside 50s. Results suggest within the current AFLW competition, key athletes' skilled performances are contributing more to match success rather than a collective team effort as opposed to the AFL competition. Using machine learning methods in sport analytics to uncover practical information from athlete match performance statistics allows for analysis on how these athletes are contributing towards team success. Post-hoc reporting of results, in a comprehensible format for coaching staff, may provide a basis for training and match strategic planning.

Acknowledgements

There was no financial assistance associated with this study.

References

1. Hughes MD, Bartlett RM. The use of performance indicators in performance analysis. *J Sport Sci* 2010; 20(10):739–754.
2. Pinder RA, Davids K, Renshaw I et al. Representative learning design and functionality of research and practice in sport. *J Sport Exerc Psychol* 2011; 33(1):146–155.
3. Hughes M. Sports analysis, In: *Essentials of performance analysis in sport Routledge.*, 2015. p. 141–151.
4. McIntosh S, Kovalchik S, Robertson S. Examination of player role in the Australian Football League using match performance data. *Int J Perform Anal Sport* 2018; 18(3):451–462.
5. Stewart MF, Mitchell H, Stavros C. Moneyball applied: econometrics and the identification and recruitment of elite Australian footballers. *Int J Sport Finance* 2007; 2(4):231–248.
6. Robertson S, Back N, Bartlett JD et al. Explaining match outcome in elite Australian Rules football using team performance indicators. *J Sports Sci* 2016; 34(7):637–644.
7. Robertson S, Gupta R, McIntosh S. A method to assess the influence of individual player performance distribution on match outcome in team sports. *J Sports Sci* 2016; 34(19):1893–1900.
8. Black GM, Gabbett TJ, Johnston RD et al. A skill profile of the national women's Australian football league (AFLW). *Sci Med Football* 2018; 30(June):1–5.
9. Clarke AC, Ryan S, Couvalias G et al. Physical demands and technical performance in Australian Football League Women's (AFLW) competition match-play. *J Sci Med Sport* 2018; 21(7):748–752.
10. Black GM, Gabbett TJ, Johnston RD et al. The influence of physical qualities on activity profiles of female Australian football match play. *Int J Sports Physiol Perform* 2018; 13:524–529.
11. Gray AJ, Jenkins DG. Match analysis and the physiological demands of Australian football. *Sport Med* 2010; 40(4):347–360.
12. Boyd LJ, Ball K, Aughey RJ. Quantifying external load in Australian football matches and training using accelerometers. *Int J Sports Physiol Perform* 2013; 8(1):44–51.
13. Sullivan C, Bilsborough JC, Cianciosi M et al. Factors affecting match running performance in professional Australian football. *Int J Sports Physiol Perform* 2014; 9(3):561–566.
14. Sullivan C, Bilsborough JC, Cianciosi M et al. Match score affects activity profile and skill performance in professional Australian Football players. *J Sci Med Sport* 2014; 17(3):326–331.
15. Moreno E, Gomez MA, Lago-Peñas C, Sampaio J et al. Effetcs of starting quarter score, game location, and quality of opposition in quarter score in elite women's basketball. *Kinesiology* 2013; 45(1):48–54.
16. Gómez MÁ, DelaSerna A, Lupo C, Sampaio J. Effects of situational variables and starting quarter score in the outcome of elite women's water polo game quarters. *Int J Perform Anal Sport* 2014; 14(1):73–83.
17. Sampaio J, Lago C, Casais L et al. Effects of starting score-line, game location, and quality of opposition in basketball quarter score. *Eur J Sport Sci* 2010; 10(6):391–396.
18. Joshi KD, Sarker S, Sarker S. Knowledge transfer within information systems development teams: examining the role of knowledge source attributes. *Decis Support Syst* 2007; 43(2):322–335.
19. McIntosh S, Kovalchik S, Robertson S. Validation of the Australian Football League player ratings. *Int J Sport Sci Coach* 2018; 13(6):1064–1071.
20. Hughes M. Notational analysis – a mathematical perspective. *Int J Perform Anal Sport* 2017; 4(2):97–139.
21. Foundation PS. *Python programming language [Internet]*, 2018.
22. Pedregosa F, Varoquaux G, Gramfort A et al. Scikit-learn: machine learning in python. *J Mach Learn Res* 2011; 12:2825–2830.
23. Hawkins DM. The problem of overfitting. *J Chem Inf Comput Sci* 2004; 44(1):1–12.
24. R Core Team. *R: a language and environment for statistical computing*, Vienna, Austria, R Foundation for Statistical Computing, 2018.

25. Gómez MA, Lorenzo A, Ibañez SJ et al. Ball possession effectiveness in men's and women's elite basketball according to situational variables in different game periods. *J Sports Sci* 2013; 31(14):1578–1587.
26. Ladha L, Deepa T. Feature selection methods and algorithms. *Int J Comput Sci Eng* 2011; 3(5):1787–1797.
27. Liu H, Motoda H, Setiono R et al. Feature selection: an ever evolving frontier in data mining, In: *JMLR: workshop and conference proceedings 10: feature selection in data mining Hyderabad, India.*, 2010. p. 4–13.
28. Woods CT, Robertson S, Collier NF. Evolution of game-play in the Australian Football League from 2001 to 2015. *J Sports Sci* 2017; 35(19):1879–1887.
29. Alexander JP, Spencer B, Mara JK et al. Collective team behaviour of Australian Rules football during phases of match play. *J Sports Sci* 2019; 37(3):237–243.
30. Haycraft JAZ, Kovalchik S, Pyne DB et al. Physical characteristics of players within the Australian Football League participation pathways: a systematic review. *Sport Med Open* 2017; 3(1).
31. Oliveira T, Gómez M, Sampaio J. Effects of game location, period, and quality of opposition in elite handball performances. *Percept Mot Skills* 2012; 114(3):783–794.
32. Taylor J, Mellalieu S, James N et al. The influence of match location, quality of opposition, and match status on technical performance in professional association football. *J Sports Sci* 2008; 26(9):885–895.
33. García-Rubio J, Gómez MÁ, Lago-Peñas C et al. Effect of match venue, scoring first and quality of opposition on match outcome in the UEFA champions league. *Int J Perform Anal Sport* 2015; 15(2):527–539.
34. Ruano MA, Serna AD, Lupo C et al. Effects of game location, quality of opposition, and starting quarter score in the outcome of elite water polo quarters. *J Strength Cond Res* 2016; 30(4):1014–1020.

Appendix D. Original published article in Chapter Six: ***Biomechanical characteristics of elite female Australian rules football preferred and non-preferred drop punt kicks.***

VICTORIA UNIVERSITY

THE NEW WAY TO DO UNI

OFFICE FOR RESEARCH TRAINING, QUALITY AND INTEGRITY

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION: PAPERS INCORPORATED IN THESIS

This declaration is to be completed for each conjointly authored publication and placed at the beginning of the thesis chapter in which the publication appears.

1. PUBLICATION DETAILS (to be completed by the candidate)

Title of Paper/Journal/Book: Biomechanical characteristics of elite female Australian rules football preferred and non-preferred drop punt kicks

Surname: Cust First name: Emily

Institute: Institute for Health and Sport Candidate's Contribution (%): 70

Status:

Accepted and in press: ☐ Date:

Published: ☑ Date: September 2

2. CANDIDATE DECLARATION

I declare that the publication above meets the requirements to be included in the thesis as outlined in the HDR Policy and related Procedures – policy.vu.edu.au.

Emily Cust Digitally signed by Emily Cust Date: 2020.02.17 14:34:46 +11'00'	17-2-2020
Signature	**Date**

3. CO-AUTHOR(S) DECLARATION

In the case of the above publication, the following authors contributed to the work as follows:

The undersigned certify that:

1. They meet criteria for authorship in that they have participated in the conception, execution or interpretation of at least that part of the publication in their field of expertise;
2. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

Victoria University ABN 83776954731
CRICOS Provider No. 00124K (Melbourne),
02475D (Sydney), RTO 3113

3. There are no other authors of the publication according to these criteria;

4. Potential conflicts of interest have been disclosed to a) granting bodies, b) the editor or publisher of journals or other publications, and c) the head of the responsible academic unit; and

5. The original data will be held for at least five years from the date indicated below and is stored at the following **location**(s):

6 Wingan Avenue Camberwell, VIC

Name(s) of Co-Author(s)	**Contribution (%)**	**Nature of Contribution**	**Signature**	**Date**
Dr Kevin Ball	20	Study design, data collection and analysis, reviewing and editing	[redacted]	17-2-202 0
Professor Sam Robertson	5	Reviewing		17-2-2020
Dr Alice Sweeting	5	Reviewing		17-2-2020

Updated: September 2019

Biomechanical Characteristics of Elite Female Australian Rules Football Preferred and Non-Preferred Drop Punt Kicks

Emily E Cust[1,2][a], Kevin Ball[1], Alice J Sweeting[1,2][b] and Sam Robertson[1,2][c]
[1]*Institute for Health and Sport (IHES), Victoria University, Footscray, 3011, Melbourne, Victoria, Australia*
[2]*Western Bulldogs Football Club, Whitten Oval, Footscray, 3011, Melbourne, Victoria, Australia*

Keywords: Australian Football, Kicking, Kinematics

Abstract: While Australian Rules kick biomechanics has been researched considerably, there is yet to be focus specifically on women participants. Elite female Australian Rules football drop punt kick characteristics were collected from 15 elite female participants for both the preferred and non-preferred legs. All participants undertook a 20-kick protocol captured by a 3-dimensional motion analysis camera system. Preferred leg kicks produced faster foot velocities prior to foot-ball contact, 18.0 ± 1.2 $m.s^{-1}$ preferred, 16.2 ± 1.3 $m.s^{-1}$ non-preferred, and faster ball velocities post foot-ball contact, 24.7 ± 1.4 $m.s^{-1}$ preferred, 21.6 ± 2.0 $m.s^{-1}$ non-preferred. Differences in movement patterns of the hip and knee joint segments were shown between kick leg preferences; hip angular velocity 94.4 ± 75.9 °/s preferred and 126.2 ± 66.3 °/s non-preferred, knee angular velocity 1384.8 ± 415.2 °/s preferred and 1013.6 ± 230.2 °/s non-preferred. Research results identified the changes in elite women's drop punt kick mechanics in comparison to leg preference, which can be viewed against senior and junior men's Australian football kick analysis findings. The current research information could be of benefit to practitioners in linking targeted field coaching cues and conditioning programs tailored to identified kick skill and movement deficiencies.

1 INTRODUCTION

The National Women's Australian Rules Football competition (AFLW) is in its fourth year of operation, yet there has been no reported biomechanical analysis of women's kicking. In Australian Rules football (AF), efficient kick performance has been identified as a strong contributor towards team match success (Robertson, Gupta and McIntosh, 2016; Black *et al.*, 2018).

In AF the drop punt is the most commonly performed kick due to the flight accuracy and ease of catching for the receiver (Ball, 2008). Across the six phases of a drop punt (Ball, 2008), several kinematic factors have been found to influence the success, efficiency, and accuracy of performance. Prominently, higher kick leg foot velocities prior to ball contact have a major influence on the kick distance (Ball, 2008; Ball *et al.*, 2013; Peacock, Ball and Taylor, 2017) and ball velocities (Ball, 2008; Peacock and Balll, 2016; Peacock and Ball, 2017) achieved. The flight path accuracy of the ball is determined primarily by the combination of the flight characteristics imparted on the ball during the foot-to-ball contact phase (Peacock and Ball, 2018; Peacock *et al.*, 2018). Differences in kick biomechanics have been found between the preferred and non-preferred leg in men's AF kicks (Smith, Ball and MacMahon, 2009; Ball, 2011) and soccer (Nunome *et al.*, 2006). The ability to kick proficiently on both legs and over long distances in AF is a tactical advantage (Ball, 2008, 2011) in the dynamic unpredictable nature of match play. Biomechanical assessment may be an important information source for individual athlete skill profiling to identify areas of deficiencies for drop punts kicks.

The kick impact and technical components of men's kicking across several athlete levels has already been established allowing for quantified information to further develop kick skills on a team and individual basis. To address the lack of quantitative information in women's AF kick

[a] http://orcid.org/0000-0001-6927-6329
[b] http://orcid.org/0000-0002-9185-6773
[c] http://orcid.org/0000-0002-8330-0011

biomechanics, 3-dimensional optoelectronic motion analysis was undertaken. Conducting this research is important for broadening the sport science support invested in the new AFLW competition, with the intention of improving athlete kick skill and therefore team match performances. The aim of this research was to analyse the biomechanical characteristics of elite female AF drop punts for both the preferred and non-preferred kick legs. The outcomes of this research can inform the technical aspects of distance kicking in women's AF to aid in athlete kick skill development, as well as links with strength and conditioning and injury.

2 METHODS

2.1 Participants

Fifteen elite female AF athletes provided written informed consent to participate in this research. Of the participants, twelve were contracted to an AFLW team and three were competing at a high standard in their respective State based competition. The University's Human Research Ethics Committee approved the study (application number 0000025654).

2.2 Research Procedures

Athletes undertook a drop punt kick protocol as part of a broader test battery. Ten drop punts were undertaken for maximum distance and intensity on each leg. Maximal kicks were performed into a net situated 30 m from the kick launch area. Prior to undertaking the protocol, each athlete completed a dynamic warmup including jogging, dynamic stretches and five 20 m submaximal kicks on each leg. All athletes wore their regular football boots and used official AFLW match balls (Sherrin, Scoresby, Australia). The testing was conducted in purpose built indoor football training facility on artificial turf.

Drop punt kicks were captured by a 10-camera optoelectronic motion analysis system (MAS) capturing at 100 Hz (T-40 series, Vicon Nexus v2, Oxford, UK). Previous assessment of sampling rates had found low maximum error ranges for kick parameters from 500 Hz to 100 Hz (Coventry *et al.*, 2015). Cameras were set up as an arc around the testing area and mounted at varying heights in order to allow full capture of the kick and ball flight movements. 35 reflective markers (diameter: 14 mm) were taped to each athlete at anatomical landmarks as per previous kick research (Blair, Duthie, *et al.*, 2018), shown in Figure 1. Four reflective makers were attached to the football (Figure 2) to create a coordinate system and establish the ball centre.

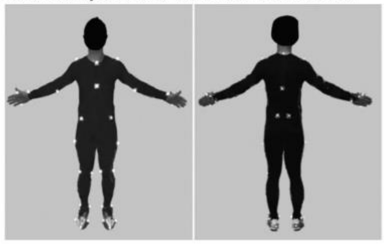

Figure 1: Athlete marker set-up

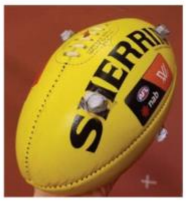

Figure 2: Football reflective marker positions.

2.3 Data Analysis

Raw motion analysis data were digitised in Nexus (v.2.0, Vicon, Oxford, UK) and processed in Visual 3D (C-motion, Inc. Germantown, USA). Data were pre-processed through a polynomial interpolation (order: 3) and smoothed using a low-pass fourth-order Butterworth filter (cut-off frequency: 10 Hz) (Ball, 2008, 2011; frequent in-lab evauation of VICON data using spectral and residual analyses).

A total of 300 drop punts (150 preferred and 150 non-preferred kicks) were analysed for ball velocity values. The trials with the highest preferred and non-preferred ball velocities were selected for each athlete for full kinematic analysis in this study, as final ball velocities are the reflection of efficacy in impact characteristics applied to the ball (Peacock, Ball and Taylor, 2017; Peacock and Ball, 2018a). A total of nine drop punt kick parameters were analysed from the MAS data, see Table 1, based on previous technical parameters assessed in AF kick performance (Ball, 2008; Smith, Ball and MacMahon, 2009; Ball, Smith and MacMahon, 2010).

Table 1: Definitions of measured kick parameters.

Parameter	Definition
Foot velocity prior to ball contact ($m.s^{-1}$)	Linear velocity of the foot segment measured from the head of the 5th metatarsal
Ball velocity post foot contact ($m.s^{-1}$)	Linear velocity of the ball segment
Ball: foot velocity ratio	Ball velocity at release divided by foot velocity at initial impact
Support leg knee flexion (°)	Degree of flexion of the support leg at ball contact
Knee angle at ball contact (°)	Angle between the thigh and shank of kick leg
Knee angular velocity (°/s)	Angular velocity of the knee joint of kick leg
Hip angle at ball contact (°)	Angle between the thigh and the trunk on the anterior aspect of the participant
Hip angular velocity (°/s)	Angular velocity of the hip segment
Pelvis linear velocity ($m.s^{-1}$)	Linear velocity of the pelvis segment

Processed data for each parameter were exported to a custom Excel file and the group mean and standard deviation (SD) calculated for each preferred and non-preferred kick parameter. Paired t-tests were computed for each parameter with statistical significance set at $p < 0.05$. The effect size for each measure for between-group distances was calculated using Cohen's d statistic indicating a small or trivial ($d = 0–0.2$), moderate ($d = 0.2–0.5$), large ($d = 0.5–0.8$), and very large (d . 0.8) effect (Hopkins *et al.*, 2009).

3 RESULTS

Table 2 reports the mean data kinematic parameters of the foot, knee, hip, and ball segments. The preferred leg produced significantly greater foot velocity, ball velocity, knee angular velocity, and pelvis linear velocity, and a significantly smaller hip angle and hip angular velocity in comparison to the non-preferred leg. The maximum foot velocities achieved were 20.9 $m.s^{-1}$ and 17.7 $m.s^{-1}$ on the preferred and non-preferred legs, respectively. The maximum ball velocities achieved were 27.0 $m.s^{-1}$ and 25.5 $m.s^{-1}$ on the preferred and non-preferred legs, respectively.

4 DISCUSSION

The current research on women's elite Australian Rules football kick biomechanics reports the first analysis of its type to further the understandings of kick skill execution. Results showed that preferred leg kicks were characterised by faster foot velocities prior to ball contact, greater knee angular velocities, pelvis linear velocities, and smaller hip angular velocities. Linking information from biomechanical analysis with field coaching cues and conditioning programs may be beneficial for individualised athlete kick skill development.

Elite female AF athletes in this study produced higher foot and ball velocities on their preferred leg kicks. Foot and ball velocities for elite women were lower than the reported values for senior elite men (Ball, 2008, 2011; Smith, Ball and MacMahon, 2009) and junior elite men (Ball, Smith and MacMahon, 2010) AF athletes. Preferred leg drop punt kicks in elite senior men have shown foot velocities of 26.5 ± 2.5 $m.s^{-1}$ and ball velocities of 32.6 ± 4.4 $m.s^{-1}$ (Smith, Ball and MacMahon, 2009). Relation could also be drawn to kick distances achieved by women and men as foot velocity has shown strong correlation association with ball flight distance (Ball, 2008; Peacock, Ball and Taylor, 2017). Elite female soccer athletes have reported foot velocities of 17.70 ± 1.92 $m.s^{-1}$ (instep kicks) and 17.45 ± 1.59 $m.s^{-1}$ (curve kicks), and ball velocities of 22.62 ± 1.71 $m.s^{-1}$ (instep kicks) and 21.51 ± 1.33 $m.s^{1}$ (curve kicks) (Alcock *et al.*, 2012).

The ball-to-foot velocity ratio is a measure of kick impact efficiency and is widely reported on in AF (Smith, Ball and MacMahon, 2009; Ball, Smith and MacMahon, 2010; Ball *et al.*, 2013; Peacock and Ball, 2018b) and soccer research (Shinkai *et al.*, 2009; Sakamoto and Asai, 2013; Nunome *et al.*, 2018). The present study showed no difference for ball-to-foot ratio between the preferred (1.31) and non-preferred legs (1.33), which has previously been reported in male AF kick research (Smith, Ball and MacMahon, 2009). This may indicate that greater ball velocities on the preferred leg are the result of a faster leg swing as attributed by faster foot velocities and knee angular velocities in applying greater force onto the ball (Nunome *et al.*, 2006; Smith, Ball and MacMahon, 2009) (Table 2). Differences in body mass have also been reported to affect the ball-to-foot ratio, which may confound comparisons between male, female, junior, and senior playing groups (Shinkai *et al.*, 2013).

Differences in movement patterns were shown between kick leg preferences. Overall, the preferred leg achieved greater knee angular velocity and pelvis linear velocity, and smaller hip angle and hip angular velocity (Table 2). As the non-preferred leg produced larger hip angular velocities and hip angles, this may suggest that greater use of the thigh and hip segments were recruited. The change in movement pattern between the kick leg types may indicate the need for more stability via dominate hip control on non-preferred kicks. Another factor could also be related to the speed of run-up in approach towards the kick execution on each leg, although this was not measured in this study. Also, the result of less efficient use of sequential summation or transfer of momentum (Ball, 2011) as indicated by the lower knee angular velocity on non-preferred leg kicks. In comparison to senior AF male athletes (Ball, 2011), greater mean knee and hip angles, and knee angular velocities were achieved for both preferred and non-preferred kicks by elite women AF athletes. Although, lower hip angular velocities were produced in comparison to reported male AF athletes, 56 ± 65 °/s preferred leg and 138 ± 81 °/s non-preferred leg (Ball, 2011).

Technical differences in kick strategies have been demonstrated for thigh dominant or knee dominate kickers during maximal distance kicking (Ball, 2008) and further supported during goal kicking constraints tasks (Blair, Robertson, *et al.*, 2018). Although kick performance indicators of foot velocity and kick distance were not significantly different between each approach suggesting similar kick performance outcomes can be achieved with either movement strategy (Ball, 2008). Looking into the thigh-knee angular velocity continuum, Ball (2008) sorted the participant data to provide indicative values for those athletes who use a thigh or knee dominant strategy for preferred leg distance kicking. In comparison, post-hoc evaluation of the current elite women's data was undertaken using the hip and knee angular velocities. Further analysis showed 14 out of the 15 athletes would be considered using a knee dominant strategy on their preferred leg. In contrast, on the non-preferred leg, the majority of the group would be classified hip dominant with data from 10 athletes of 15 indicating this. This trend is consistent with findings in the men's data where on preferred leg kicks there is increased contribution from the knee segment and lower hip or thigh involvement. The opposite shown on non-preferred leg kicks with greater hip segment contribution than the knee for force generation through the kick motion (Ball, 2011).

Table 2: Impact characteristics for preferred and non-preferred drop punt distance kicks for elite women's AF. Data reported as mean and standard deviation (SD) values and results of statistical tests comparing preferred and non-preferred leg kicks.

Parameter	**Preferred leg**		**Non-preferred leg**		***p***	**Effect Size (*d*)**
	mean	SD	mean	SD		
Foot velocity ($m.s^{-1}$)	18.9	1.2	16.2	1.3	<0.001*	2.2 Very large
Ball velocity ($m.s^{-1}$)	24.7	1.4	21.6	2.0	<0.001*	1.8 Very large
Ball: foot velocity ratio	1.31	0.11	1.33	0.07	0.59	0.2 Small
Support leg knee flexion (°)	37.0	11.3	41.0	8.3	0.25	0.4 Moderate
Knee angle at ball contact (°)	50.7	12.2	57.7	13.5	0.13	0.5 Moderate
Knee angular velocity (°/s)	13845	415	1014	230	0.02*	1.1 Very large
Hip angle at ball contact (°)	34.3	13.5	48.8	15.8	0.01*	1.0 Very large
Hip angular velocity (°/s)	94	76	126	66	0.04*	0.7 Large
Pelvis linear velocity ($m.s^{-1}$)	1.7	0.4	1.4	0.5	0.03*	0.6 Large

* Significant difference ($p < 0.05$)

The support leg is important in maintaining stability through the kick motion and plays a role in the performance quality of a kick (Ball, 2013). A moderate non-significant effect of less knee flexion in the supporting leg at ball contact occurred on preferred leg kicks, 37 ± 11.3 ° to non-preferred kicks, 41 ± 8.3 °. This is in contrast to results found in elite males across maximal kicks which showed greater support leg flexion on preferred leg kicks, 43 ± 6 °, than non-preferred leg kicks, 41 ± 11° (Ball, 2013). It has been suggested that greater support leg knee flexion leads to a lower centre of gravity and hence stability in the motion allowing for improved kick accuracy (Dichiera *et al.*, 2006). Although the findings of Dichiera et al. (2006) are in contrast to Ball (2013), where results indicated that a more extended support leg knee on stance kick phase which was maintained to ball contact phase related to higher foot velocities and an improved drop punt distance achieved. During match play, athletes are repeatedly required to perform kicks with constraints against both distance and accuracy, most commonly in goal kicking (Blair, Robertson, *et al.*, 2018). Kicking kinematics measured across changing distance on goal kicks showed that increased distances resulted in greater knee extension on the support leg during the stance phase (large effect size), and moderately higher foot velocities, shank, and knee angular velocities (Blair, Duthie, *et al.*, 2018). The authors noted potential technical difference for tasks in the literature when both kick skill accuracy and distance constraints were combined. Suggesting this related to the research protocols used with accuracy tasks performed over shorter distances at lower speeds compared to research on maximal distance kicking causing athlete to adopt differing techniques to suit each task (Blair, Robertson, *et al.*, 2018). For example, during maximal distance high impact kicks the athlete adopts a more upright position through the torso and consequently a higher hip position to generate the faster foot velocities required (Ball, 2013). Further work to assess how these variables influence elite women's kick performance considering the altered match play styles and therefore kick constraints compared to the men's game (Cust *et al.*, 2019) would be of skill technique coaching benefit.

Practically, as foot velocity is strongly correlated to drop punt kick distance (Ball, 2008, 2011; Peacock and Ball, 2017) and used as a strategy to control the kick outcome (Peacock, Ball and Taylor, 2017). A focus on improving an athlete's ability to generate high foot velocities on both legs would benefit overall kick skill and in-match tactical plays (Ball, 2008, 2011). Furthermore, if footballers dominantly kick on one leg, the increased repetition loading may create imbalances in hip and lower limb strength which could affect skill performance and increase asymmetry load related injury (Hart *et al.*, 2013, 2014). As the current results show differences in the use of joint segments between the two legs, there is potential for muscle asymmetries to develop (Ball, 2011; Hart *et al.*, 2014). Strategies such as training the non-preferred leg to recruit greater lower limb involvement through skilled coaching cues and targeted conditioning programs may again be of benefit to improving kick skill performance across both legs for tactical advantage in matches. Research has indicated that combined technical and strength-based interventions for AF athletes in training for the drop punk kick serves as a constructive approach to performance improvements (Ball, 2008; Hart *et al.*, 2014).

Further research should progress assessment of the support leg mechanisms (Ball, 2013) and kinematic characteristics of the kick impact phase for elite women AF athletes in relation to kick accuracy (Peacock, Ball and Taylor, 2017). Greater understandings into the underlying mechanisms for the differences between both preferred and non-preferred leg kicks for elite women, and between male and female kinematics during kick execution would be important to further quantify. As different movement approaches exist for kick execution, future research looking at the relationships between knee and thigh (or hip) strategies for distance kicks and kick accuracy would be of benefit to kick skill coaching and individual conditioning. Knowing individual athlete movement strategies would directly affect coaching and conditioning due to the different muscle recruitment processes for generating forces for each approach (Ball, 2008). In-depth information within this field could provide links to improve practices in women's AF kick skill coaching, individual athlete injury patterns related to repeated kick execution, and targeted strength and conditioning practices.

5 CONCLUSIONS

The biomechanical characteristics of elite female Australian Rules football drop punts kicks for both the preferred and non-preferred legs were quantified. Preferred leg kicks produced faster foot velocities prior to ball contact, greater knee angular velocities, pelvis linear velocities, and smaller hip angular velocities. Movement differences were found in hip and lower limb joint segments between both kick legs as greater knee angular velocity and pelvis linear velocity characterised preferred leg kicks, yet a higher hip angular velocity on non-preferred leg kicks. Improved understandings of women's AF kick skill via kinematic technical analysis could be of benefit in linking with targeted field coaching cues and conditioning programs tailored to identified kick skill and movement deficiencies.

ACKNOWLEDGEMENTS

The research authors would like to thank the research assistants and technician who assisted in the data collection, and the athletes who partook in the study.

REFERENCES

Alcock, A. M. *et al.* (2012) 'Curve and instep kick kinematics in elite female footballers', *Journal of Sports Sciences*, 30(4), 387–394.

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

Ball, K. (2011) 'Kinematic comparison of the preferred and non-preferred foot punt kick', *Journal of Sports Sciences*, 29(14), 1545–1552.

Ball, K. *et al.* (2013) 'Ball impact dynamics in the punt kick', in *International Conference of Biomechanics in Sports*. Taipei, Taiwan, 3–6.

Ball, K. (2013) 'Loading and performance of the support leg in kicking', *Journal of Science and Medicine in Sport*. Sports Medicine Australia, 16(5), 455–459.

Ball, K., Smith, J. and MacMahon, C. (2010) 'Kick impact characteristics of junior kickers', in *International Symposium on Biomechanics in Sports: Conference Proceedings Archive*. Michigan, USA, 1–4.

Black, G. M. *et al.* (2018) 'A skill profile of the national women's Australian football league (AFLW)', *Science and Medicine in Football*. Routledge, June 30, 1–5.

Blair, S., Duthie, G., *et al.* (2018) 'Concurrent validation of an inertial measurement system to quantify kicking biomechanics in four football codes', *Journal of Biomechanics*. Elsevier Ltd, 73, 24–32.

Blair, S., Robertson, S., *et al.* (2018) 'The effect of altering distance on goal-kicking technique in Australian football', *ISBS Proceedings Archive*. Auckland, New Zealand, 36(1), 358–361.

Coventry, E. *et al.* (2015) 'Kinematic effects of a short-term fatigue protocol on punt-kicking performance', *Journal of Sports Sciences*. Routledge, 33(15), 1596–1605.

Cust, E. E. *et al.* (2019) 'The relationship of team and individual athlete performances on match quarter outcome in elite women's Australian Rules football', *Journal of Science and Medicine in Sport*. Sports Medicine Australia.

Dichiera, A. *et al.* (2006) 'Kinematic patterns associated with accuracy of the drop punt kick in Australian Football', *Journal of Science and Medicine in Sport*, 9(4), 292–298.

Hart, N. H. *et al.* (2013) 'Leg mass characteristics of accurate and inaccurate kickers - an Australian football perspective', *Journal of Sports Sciences*, 31(15), 1647–1655.

Hart, N. H. *et al.* (2014) 'Leg strength and lean mass symmetry influences kicking performance in Australian football', *Journal of Sports Science and Medicine*, 13(1), 157–165.

Hopkins, W. G. *et al.* (2009) 'Progressive statistics for studies in sports medicine and exercise science', *Medicine and Science in Sports and Exercise*, 3–12.

Nunome, H. *et al.* (2006) 'Segmental dynamics of soccer instep kicking with the preferred and non-preferred leg', *Journal of Sports Sciences*, 24(5), 529–541.

Nunome, H. *et al.* (2018) 'Dynamics of submaximal effort soccer instep kicking', *Journal of Sports Sciences*. Routledge, 36(22), 2588–2595.

Peacock, J. *et al.* (2018) 'Is there a sweet spot on the foot in Australian football drop punt kicking?', *Journal of Sports Sciences*. Routledge, 37(4), 467–476.

Peacock, J. and Ball, K. (2018) 'Kick impact characteristics of accurate Australian football drop punt kicking', *Human Movement Science*. Elsevier, 61(July), 99–108.

Peacock, J., Ball, K. and Taylor, S. (2017) 'The impact phase of drop punt kicking for maximal distance and accuracy', *Journal of Sports Sciences*. Routledge, 35(23), 2289–2296.

Peacock, J. and Balll, K. (2016) 'The impact phase of drop punt kicking: validation and experimental data of a mechanical kicking limb', in *34th International Conference of Biomechanics in Sports*. Tsukuba, Japan: ISBS-Conference Proceedings, 913–916.

Peacock, J. C. A. and Ball, K. (2017) 'The relationship between foot-ball impact and flight characteristics in punt kicking', *Sports Engineering*. Springer London, 20(3), 221–230.

Peacock, J. C. A. and Ball, K. (2018a) 'Strategies to improve impact efficiency in football kicking', *Sports Biomechanics*, 1–14.

Peacock, J. C. A. and Ball, K. (2018b) 'The influence of joint rigidity on impact efficiency and ball velocity in football kicking', *Journal of Biomechanics*. Elsevier Ltd, 71, 245–250.

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

Sakamoto, K. and Asai, T. (2013) 'Comparison of kicking motion characteristics at ball impact between female and male soccer players', *International Journal of Sports Science & Coaching*, 8(1), 63–76.

Shinkai, H. *et al.* (2009) 'Ball impact dynamics of instep soccer kicking', *Medicine and Science in Sports and Exercise*, 41(4), 889–897.

Shinkai, H. *et al.* (2013) 'Cross-sectional change of ball impact in instep kicks from junior to professional footballers', in *Science and football VII*. Nagoya, Japan: Routledge, 43–48.

Smith, J., Ball, K. and MacMahon, C. (2009) 'Foot-to-ball interaction in preferred and non-preferred leg Australian Rules kicking', in Anderson, R., Harrison, D., and Kenny, I. (eds) *Proceedings of the 27th International conference on biomechanics in sports*. Ireland, 650–653.

Appendix E. Methods and model additional specifications from the original published article in Chapter Six: *Biomechanical characteristics of elite female Australian rules football preferred and non-preferred drop punt kicks.*

The participants undertook a static anatomical position pose as the static trial required for subject calibration prior to starting the kick trials. A ball only static trial was also conducted before kicking trials as well.

Ball markers, the coordinate system and subsequent model were determined based on previous research in the field (Peacock et al., 2017). Participant markers were determined from similar kicking research (Blair, Duthie, et al., 2018).

Raw motion analysis data were digitised in Nexus (v.2.0, Vicon, Oxford, UK) and processed in Visual 3D (C-motion, Inc. Germantown, USA). Data were pre-processed through a polynomial interpolation (order: 3) and smoothed using a low-pass fourth-order Butterworth filter (cut-off frequency: 10 Hz) (Ball, 2008, 2011; Blair, Duthie, et al., 2018; Peacock et al., 2017). The pre-processing methods were determined from previous research in the field which undertook spectral and residual analysis, and visual inspections of the data (Ball, 2008, 2011; Blair, Duthie, et al., 2018; Dichiera et al., 2006; Peacock et al., 2017). A six-degree-of-freedom model (Cappozzo et al., 1995) was created based on the position and orientation of the anatomical markers relative to the static trials, then model-based calculations were taken from the X-Y-Z Carden sequence (Lees et al., 2010).

All kicks (n = 300) were analysed from the kick foot toe-off (TO) until the frame before ball contact (BC) or at BC (Ball, 2008; Blair, Duthie, et al., 2018). Where BC corresponded to the peak linear velocity of the 5th metatarsal foot marker (Ball, 2011). The data was normalised through the pre-processing methods for the frames from TO to BC.

The nine parameters analysed (Chapter Six, Table 6.1) were taken from two different data phases as follows: the velocity parameters were measured at BC minus one frame, and all angles and displacement parameters were measured at BC. Furthermore, knee and hip ROM minimum and maximum values were taken from the x-axis, and foot speed was measured from the lateral foot marker.