Measuring cricket fast bowling workload using inertial measurement units.

Joseph William McGrath

A thesis submitted to Auckland University of Technology in fulfilment of the requirements for the degree of Doctor of Philosophy

2022

Sports Performance Research Institute of New Zealand (SPRINZ)

Supervisors:

Dr Jono Neville

Dr Tom Stewart

Professor John Cronin

Thesis abstract

Fast bowlers have the highest incidence of injury compared to any other position in cricket. A bowling volume that is too high or too low – measured by the number of deliveries bowled in a session – can increase the chance of injury. Despite the importance and the simplicity of measuring bowling volume, it is rarely done at an amateur level due to the monotonous task of manually counting bowls and the effort required to analyse the data. Bowling volume by itself is also not a true measure of bowling workload, as it does not consider the intensity of each delivery. Bowling workload is now considered a combination of bowling volume and bowling intensity. Due to tactical and motivational reasons, fast bowlers will often bowl at different intensities during a match and in training. There are also variations in bowling technique, run-up speed, and anthropometrical characteristics of players. Therefore, the forces exerted on the body are not constant across players or deliveries.

This thesis explores whether an inertial measurement unit (IMU) can predict bowling volume and different intensity metrics – ball release speed, perceived intensity, and ground reaction forces (GRF). These metrics allow a more comprehensive picture of intensity, as each captures slightly different constructs related to bowling workload. This can provide researchers with a mechanism to determine possible links between workload and injury, leading to a more personalised approach to injury management, and give coaches and players a tool to monitor fatigue and performance.

In Chapter 2, a systematic literature review was conducted to examine methods for activity classification in court and field-based sports using IMUs. A key finding was that machine learning techniques had shown promising results across a range of sports. However, only user-defined algorithms had been used in cricket, meaning the application of machine learning had yet to be tested.

Chapter 3 was the first of five studies to develop a system that could estimate bowling workload. A standard IMU was positioned on the upper back in a training setting, and five different machine learning models were used to estimate bowling volume. When tested against outfield throws, several models achieved an F-score of 1.0, meaning perfect differentiation of bowling versus throwing. The analysis was repeated with several down-sampled datasets (i.e., 250 Hz to 25 Hz) to simulate a low-cost IMU that samples data less frequently. A minimal drop in accuracy was observed (F-score = 0.97).

In chapter 4, bowling intensity was quantified by predicting two metrics, (1) ball release speed, and (2) perceived intensity, using the same IMU as Chapter 3 located on the upper back. The gradient boosting algorithm (XGB) was the most consistent machine learning model for measuring ball release speed (mean absolute error (MAE) = 3.61 km/h at 25 Hz) and the perceived intensity zone (F-score = 0.88 at 25 Hz). The results were again consistent across different sampling frequencies, meaning a

range of different IMUs might be able to quantify these bowling intensity parameters, including consumer-grade wearables.

The aim of Chapter 5 was to examine whether an IMU placed on the dominant (bowling arm) and non-dominant wrist (instead of the upper back) could improve the previously observed results in Chapters 3 and 4. For practical application, a research-grade IMU (capable of measuring 100 g) was compared against a consumer-grade Apple Watch (32 g). XGB models had the best results across all bowling volume and bowling intensity measures. A slight improvement was observed compared to the previous study (bowling volume: F-score = 1.0; ball release speed: MAE = 2.76 km/h; perceived intensity: F-score = 0.92). There was no significant difference between the research-grade IMU and Apple Watch; however, IMUs on the dominant wrist classified perceived intensity significantly better than on the non-dominant wrist.

In Chapter 6, another component of bowling intensity was introduced – the GRF experienced during the front foot contact of the delivery. Peak force and loading rate, measured by a force plate, were significantly different across three perceived intensity zones in the horizontal and vertical axes (Cohen's *d* range = 0.14–0.45, $p < 0.01$). When ball release speed increased, peak force and loading rate also increased in the horizontal and vertical axes ($\eta_p^2 = 0.04–0.18$, $p < 0.01$). Lastly, moving from high to medium intensity, or medium to low intensity, was associated with a larger relative decrease in GRF compared to the relative decrease in ball release speed. For example, reducing bowling effort from high to medium intensity resulted in a 7–17% decrease in the horizontal GRF compared to only a 5% decrease in ball release speed. This finding could influence bowlers' strategies during an unlimited overs match, as they could conserve energy and reduce workload with only a small reduction in bowling speed.

Similar machine learning techniques as the previous chapters were used in Chapter 7 to estimate GRF. As earlier research had only used accelerometer data to estimate GRF in other sports, this study also assessed whether the addition of gyroscope data could improve accuracy. Research-grade IMUs were attached to the upper back and bowling wrist. A mean absolute percentage error (MAPE) of 22.1% for vertical and horizontal peak force, 24.1% for vertical impulse, and 32.6% and 33.6% for vertical and horizontal loading rates were observed, respectively. The linear support vector machine model had the most consistent overall results. In general, there were no significant differences between using data only from the accelerometer compared to data from the accelerometer and gyroscope. Although the results were similar to previous studies that estimated GRF, the magnitude of error would likely prevent its use in individual monitoring. However, due to the large differences in raw GRFs between participants, researchers may be able to help identify links among GRFs, injury, and performance by categorising values into levels (i.e., low and high).

It is hoped that the methods explored in this thesis can be used as a foundation for future applications that automatically estimate bowling workload across weeks or seasons. As access to smart devices is increasing in developing nations, such a system has the potential to reach most of the cricketing population.

Table of contents

Measuring cricket fast bowling workload using inertial measurement units. 1

Thesis abstract 2

Table of contents 5

List of figures 8

List of tables 8

List of images 10

Attestation of authorship 11

Co-authored works 12

Peer-reviewed journal publications: 12

Papers under review: 12

Research chapter contributions 13

Acknowledgements 15

Chapter 1 – Introduction 16

Background 16

Thesis rationale 17

Thesis organisation 19

Chapter 2 - Upper body activity classification using an inertial measurement unit in court and field-based sports: a systematic review. 22

Preface 22

Abstract 23

Key points 23

Introduction 24

Methods 24

Results 25

Discussion 33

Chapter 3 - Cricket fast bowling detection in a training setting using an inertial measurement unit and machine learning. 42

Preface 42

Abstract 43

Introduction 44

Methods 45

Results 48

Discussion 51

Appendix 53

Chapter 4 - Can an inertial measurement unit in combination with machine learning measure fast bowling speed and perceived intensity in cricket? 54

Preface 54

Abstract 55

Introduction 56

Methods 57

Results 60

Discussion 64

Appendix 67

Chapter 5 - Quantifying cricket fast bowling volume, speed and perceived intensity zone using an Apple Watch and machine learning. 72

Preface 72

Abstract 73

Introduction 74

Methods 75

Results 78

Discussion 81

Appendix 84

Chapter 6 - The relationship between bowling intensity and ground reaction forces in cricket fast bowling........ 86

Preface 86

Abstract........ 87

Introduction 88

Methods........ 89

Results........ 91

Discussion........ 95

Appendix 98

Chapter 7 - Can an inertial measurement unit, combined with machine learning, accurately measure ground reaction forces in cricket fast bowling. 99

Preface 99

Abstract........ 100

Introduction 101

Methods........ 102

Results........ 105

Discussion........ 107

Chapter 8 - General discussion 111

Research summary........ 111

Significance of findings 114

Study limitations and future directions 119

Conclusion........ 121

References 122

List of figures

Figure 1-1. Thesis structure and flow. 20

Figure 2-1: PRISMA flow diagram of the systematic search. 26

Figure 3-1: Explanation of the fielding drill during testing session two. 46

Figure 3-2: Top 10 feature importance plots illustrating RF (250 Hz) and XGB (50 Hz) models for the delivery phase. 50

Figure 4-1: Model results for ball release speed (BRS) and the predicted perceived intensity zone (PIZ) across all sampling frequencies. 61

Figure 4-2: Top 10 relative sensor features for XGB 250 Hz and XGB 50 Hz models for predicted perceived intensity zone (PIZ) and ball release speed (BRS). 62

Figure 5-1: Model results for ball release speed (BRS) and the predicted perceived intensity zone (PIZ) for different IMU devices. 80

Figure 6-1: The relationship between bowling intensity (perceived bowling intensity and ball release speed) and GRF. 93

Figure 6-2: Percentage decrease in GRF and ball release speed relative to maximum perceived effort. 95

Figure 7-1: Inertial measurement unit and force plate traces from a single delivery. 104

Figure 8-1: Opening screen. 115

Figure 8-2: Real-time summary for a delivery. 116

Figure 8-3: a) Viewing history main page. b) History option screen. 116

Figure 8-4: a) Raw data from a session. b) Session summary. 116

Figure 8-5: Examples of viewing summaries across a session (a and b), week (c) and year (d). 117

Figure 8-6: Personal insights into workload management. 118

List of tables

Table 2-1: Database search terms used. 25

Table 2-2: Summary of past studies that used IMUs to quantify sport-specific movements. 28

Table 2-3: Description of the performance measures used. 33

Table 3-1: Classification results for models trained using all phases and just the delivery phase (250 Hz). 48

Table 3-2: Confusion matrix comparing data from three phases and the delivery phase (250 Hz). 49

Table 3-3: Classification results for models trained using delivery phase data. 49

Table 3-4: Confusion matrix comparing data down-sampled to 150 Hz, 50 Hz, and 25 Hz. 50

Table 3-5: Features used. 53

Table 4-1: Mean ball speed prediction error at different sampling frequencies. 60

Table 4-2: The predicted perceived bowling intensity zone classification results at different sampling frequencies. 62

Table 4-3: Confusion matrix for the perceived intensity zone classification at different sampling frequencies. 63

Table 4-4: Initial feature set. 67

Table 4-5: Pairwise contrast results for each model for ball release speed. 68

Table 4-6: Pairwise contrast results for each model for the predicted perceived intensity zone. 69

Table 4-7: Pairwise contrast results between frequencies for each model for ball release speed. 70

Table 4-8: Pairwise contrast results between frequencies for each model for the perceived intensity zone. 71

Table 5-1: Results for predicting bowling volume. 78

Table 5-2: Mean ball release speed prediction error. 79

Table 5-3: The predicted perceived bowling intensity zone. 79

Table 5-4: Initial feature set. 84

Table 5-5: Pairwise contrast results for each model for bowling volume (BV), ball release speed (BRS), and the predicted perceived intensity zone (PIZ). 84

Table 5-6: Pairwise contrast results for each inertial measurement unit for bowling volume (BV), ball release speed (BRS), and the predicted perceived intensity zone (PIZ). 85

Table 6-1: Participants' physical characteristics and ball release speeds. 91

Table 6-2: Differences in GRF across the three perceived intensity zones. 94

Table 6-3: The relationship between GRF and ball release speed. 94

Table 6-4: Estimated means and 95% confidence intervals of GRF across the three perceived intensity zones. 98

Table 7-1: The mean, standard deviation, and range for ball release speed and ground reaction forces across intensity zones for all participants. 105

Table 7-2: Predicted peak force values. 106

Table 7-3: Predicted impulse values. 106

Table 7-4: Predicted loading rate values. 107

List of images

Image 5-1: IMU setup and location. 76

Attestation of authorship

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

Joseph McGrath, June 2022

Co-authored works

Peer-reviewed journal publications:

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. doi:10.1080/02640414.2018.1553270

McGrath, J., Neville, J., Stewart, T., & Cronin, J. (2020). Upper body activity classification using an inertial measurement unit in court and field-based sports: a systematic review. Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology, 235(2), 83–95. doi:10.1177/1754337120959754

McGrath, J., Neville, J., Stewart, T., Clinning, H., & Cronin, J. (2021). Can an inertial measurement unit (IMU) in combination with machine learning measure fast bowling speed and perceived intensity in cricket? Journal of Sports Sciences, 39(12), 1402–1409. doi:10.1080/02640414.2021.1876312

McGrath, J., Neville, J., Stewart, T., Clinning, H., Thomas, B., & Cronin, J. (2021). Quantifying cricket fast bowling volume, speed and perceived intensity zone using an Apple Watch and machine learning, Journal of Sports Sciences, 40(3), 323–330, doi:10.1080/02640414.2021.1993640

McGrath, J., Neville, J., Stewart, T., Lamb, M., Alway, P., King, M., & Cronin, J. (2022). The relationship between bowling intensity and ground reaction forces in cricket fast bowling. Journal of Sports Sciences. Journal of Sports Sciences, 40(14), 1602-1608, doi:10.1080/02640414.2022.2094561

Papers under review:

McGrath, J., Neville, J., Stewart, T., Lamb, M., Alway, P., King, M., & Cronin, J. (2022). Can an inertial measurement unit, combined with machine learning, accurately measure ground reaction forces in cricket fast bowling. Sports Biomechanics.

Research chapter contributions

Chapters 2–7 of this thesis are either published in peer-reviewed journals or are under review. The percentage contribution of each author is presented below.

Chapter 2: Upper body activity classification using an inertial measurement unit in court and field-based sports: a systematic review.

Joseph McGrath	80%
Jono Neville	9%
Tom Stewart	9%
John Cronin	2%

Chapter 3: Cricket fast bowling detection in a training setting using an inertial measurement unit and machine learning.

Joseph McGrath	80%
Jono Neville	9%
Tom Stewart	9%
John Cronin	2%

Chapter 4: Can an inertial measurement unit in combination with machine learning measure fast bowling speed and perceived intensity in cricket?

Joseph McGrath	80%
Jono Neville	7%
Tom Stewart	7%
Hayley Clinning	4%
John Cronin	2%

Chapter 5: Quantifying cricket fast bowling volume, speed and perceived intensity zone using an Apple Watch and machine learning.

Joseph McGrath	80%
Jono Neville	7%
Tom Stewart	7%
Hayley Clinning	2%
Bernd Thomas	2%
John Cronin	2%

Chapter 6: The relationship between bowling intensity and ground reaction forces in cricket fast bowling.

Joseph McGrath	80%
Jono Neville	7%
Tom Stewart	7%
Matt Lamb	2%
Mark King	2%
John Cronin	2%

Chapter 7: Can an inertial measurement unit, combined with machine learning, accurately measure ground reaction forces in cricket fast bowling.

Joseph McGrath	80%
Jono Neville	7%
Tom Stewart	7%
Matt Lamb	2%
Mark King	2%
John Cronin	2%

Joseph McGrath

Jono Neville

Tom Stewart

John Cronin

Acknowledgements

I would firstly like to acknowledge the enormous contribution from my two supervisors, Dr Jono Neville and Dr Tom Stewart. I am lucky to not only have you as mentors but good friends. I have learned some valuable research and technical skills that I will use for the rest of my research career. Thank you to Professor John Cronin for his support, friendship, and organising some fantastic writing retreats. To my partner, Hayley Clinning, thank you for your love, support, and expertise in data science. I am very fortunate to have a partner that could contribute to the technical aspects of the project. To my parents, Mike and Louise, who gave me love and support and provided me with an excellent education. Lastly, I would like to thank all the fast bowlers who participated in the studies.

The PhD was funded by Auckland University of Technology and New Zealand Cricket.

Chapter 1 – Introduction

Background

Fast bowlers have the highest incidence of injury compared to any other position in cricket.[1] The incidence of injury during a season where one or more matches are missed is 20.6%, which is more than twice as much as spin bowlers (6.7%), batsmen (7.4%), and wicketkeepers (4.7%).[2] This is coupled with an average length of time on the sidelines of 4.5 weeks.[3] Common injuries include stress fractures to the lumbar spine, hamstring strains, and rotator cuff tendon injuries.[4] Modifiable and non-modifiable risk factors intertwine to increase the likelihood of injury. Non-modifiable risk factors include genetic susceptibility to injury, while modifiable risk factors currently include bowling technique and bowling volume, measured by the number of deliveries bowled in a session.[5-7] Retrospective studies have shown that fast bowlers who bowl too many and too few deliveries during a week, month and year, have an increased chance of injury.[1,4,8-16]

Caution needs to be used when interpreting results from bowling volume studies due to researchers relying on participants to fill in bowling logs accurately.[5] This reason has caused some researchers to exclude data from training sessions which can make up a large proportion of a bowler's weekly workload.[13,17,18] A potential reason for the poor adherence to measuring bowling volume is the monotonous task of counting bowls manually. This, along with the effort taken to analyse data, could also be a reason why it is also not recorded regularly at any playing level.[5,17]

Bowling volume by itself is also not a true measure of bowling workload as it does not consider the intensity of each delivery.[5,19,20] Fast bowlers will often bowl at different intensities in match and training settings due to tactical and motivational reasons.[5,21-25] There are also variations between technique, run-up speed, and anthropometrical characteristics (e.g., height and weight).[5] Therefore, the forces exerted on the body are not constant between each bowl or player. Although there is a lack of consensus on the best way to measure bowling intensity,[5,19] current methods use a perceived intensity rating scale,[17,26] speed radar gun,[26] or a force plate built into the cricket pitch.[24] However, these methods are manually intensive or require considerable outlay, time, and expertise to set up.[5,20,24] Consequently, a lack of bowling intensity data exists, meaning that researchers have rarely been able to study the effects of bowling intensity (and therefore bowling workload) on injury.

A cost-effective, automatic, and portable method of recording bowling workload (bowling volume and bowling intensity) is needed. Ideally, this method would capture different measures of bowling intensity (i.e., ball release speed, ground reaction forces (GRF), and perceived intensity) as they assess slightly different constructs.[26] For example, due to slight changes in technique, which can alter the timing of the kinetic chain, the perceived effort could still be at maximum; however, ball release speed may not. Having a more accurate measure of bowling volume and different intensity metrics could assist researchers in understanding the cause of injury and offer a more personalised approach to bowling workload management.[5,19] In addition, intensity measures could be used to improve and track performance. For example, higher ball release speeds are linked to improved performance,[27-29] and certain high GRFs are linked to an increase in bowling speed.[6,30-32]

Thesis rationale

A potential solution to this problem is to use an inertial measurement unit (IMU). An IMU typically consists of an accelerometer, gyroscope, and magnetometer. An accelerometer measures linear acceleration (measured in g-force), the gyroscope measures angular velocity (degrees per second), and the magnetometer measures the strength and direction of the local magnetic field. IMUs are small (< 12 cm^3), lightweight (< 10 grams), cheap (< 50 USD), and accessible to most of the world's population due to being installed in the majority of smart devices (e.g., smartwatches and smartphones).[33] This is important as over half the cricketing population comes from developing countries.[34]

When determining the type of IMU to use, considerations need to be made around what sensors (e.g., accelerometer and gyroscope) to use, sampling frequency and sensor threshold limits. Although the gyroscope and accelerometer will likely play a crucial role in measuring bowling volume, ball release speed and the perceived intensity zone,[20] studies using IMUs to estimate GRF in other sports have not included data from the gyroscope.[35-38] Regarding sampling rate, low-sampling rate IMUs will use less power, require smaller memory storage, and are often inexpensive. However, as less data are recorded, key events (such as the point of ball release) could be missed, which could affect overall accuracy. The opposite is true for high-sampling rate IMUs – these are more expensive and produce larger datasets. This may require more computational power, which has implications for practical application, particularly on any smart device.

The location of the IMU on the bowler's body can change the g-forces and rotational velocities an IMU is subjected to during the delivery. Possible options in cricket include the upper back (T1) and the wrist of the bowling and non-bowling arm. The upper back is a more discrete location for the bowler and offers protection for the unit. As it is closer to the centre of mass, it may be the ideal location for measuring GRFs.[39,40] It is also subjected to less g-force and rotational movement. IMUs with standard sensor thresholds limits (e.g., accelerometer 16–32 g) may benefit from this location as no data would be lost due to thresholding. However, the information recorded by the sensor is not a specific representation of upper body rotation during a cricket bowl. Therefore, accuracy might diminish when classifying the difference between a bowl and a throw or estimating ball release speed. An IMU on the wrist will provide more data specific to the bowling arm; however, it will be subjected to high g-force and rotational velocity (upwards of 70 g and 1500 °/s). As mentioned previously, thresholding would occur without a high specification IMU. In particular, data would be lost from the crucial point of ball release where the highest forces are observed.[41,42] Another potential limitation is that some bowlers do not like wearing anything on their bowling wrist. As non-bowling arm speed is also correlated with ball release speed,[43] the non-dominant wrist may offer an alternative location.

The second important factor to consider is the type of algorithm used to estimate bowling workload. Researchers have successfully implemented user-defined algorithms (based on a set of user-specified rules) and machine learning approaches to classify movement patterns or estimate variables in sports.[44] Manually-defined algorithms offer simplicity and interpretability. However, whether these algorithms can distinguish between more complex movements or estimate metrics like ball release speed remains unclear.[44] There are several supervised machine learning approaches (e.g., random forest, neural networks) that take a set of input variables and map them to a particular outcome. The outcome can be a category (e.g., bowl or throw) or numeric value (e.g., ball release speed). The drawbacks of this approach include requiring more computational power due to the increased amount of data needed and the complex implementation and interpretation of machine learning models.[45]

Statement of purpose

The overarching question of the thesis is to determine whether an IMU can predict bowling workload through bowling volume and different measures of bowling intensity (ball release speed, the perceived intensity zone, and GRF). To achieve this aim, the thesis will be organised into several chapters, where the primary objectives are to:

1. To systematically review existing research that have used IMUs in court and field-based sports, focusing on determining the best approach to classifying bowling workload (Chapter 2).

2. Assess the accuracy of an IMU combined with machine learning to classify bowling workload by:
    a. Determining the accuracy of an IMU located on the upper back to predict bowling volume (Chapter 3).

b. Examining the accuracy of an IMU located on the upper back to predict ball release speed and the perceived intensity zone (Chapter 4).
c. Investigating whether accuracy improves when an IMU is positioned on the bowling and non-bowling wrist (Chapter 5).
d. Analysing whether GRFs change with increased ball release speed and the perceived intensity zone (Chapter 6).
e. Determining whether an IMU can predict GRFs (Chapter 7).

3. Compare the accuracy of a consumer-grade IMU (Apple Watch) to a research-grade IMU for measuring bowling workload (Chapter 5).

Thesis organisation

Context

The thesis format fulfils the guidelines for a thesis by publication pathway. Chapters 2 to 8 are presented in the same format as the final copy sent to the target journal. Therefore, some repetition of information occurs. Each chapter begins with a preface and ends with a summary of how the work contributes to the body of knowledge. The themes of each chapter follow a sequential progression that culminates in a cohesive whole. The thesis is organised into eight chapters (see Figure 1.1).

Figure 1-1. Thesis structure and flow.

Introduction	Chapter 1. Introduction	
	Chapter 2. Upper body activity classification using an inertial measurement unit in court and field-based sports: a systematic review.	

⬇

Research chapters	Chapter 3. Cricket fast bowling detection in a training setting using an inertial measurement unit and machine learning.	Chapter 4. Can an inertial measurement unit in combination with machine learning measure fast bowling speed and perceived intensity in cricket?
	Chapter 5. Quantifying cricket fast bowling volume, speed and perceived intensity zone using an Apple Watch and machine learning.	
	Chapter 6. The relationship between bowling intensity and ground reaction forces in cricket fast bowling.	Chapter 7. Can an inertial measurement unit, combined with machine learning, accurately measure ground reaction forces in cricket fast bowling.

⬇

Conclusion	Chapter 8. General discussion

Candidate contribution

This thesis fulfils the requirements of an Auckland University of Technology Doctoral degree through a significant and unique contribution to this field; classification and measurement of bowling volume and intensity using IMUs and machine learning. In collaboration with co-supervisors, the candidate developed the research questions in this thesis. This process has led the candidate to build considerable research skills and attitudes necessary to operate in academia.

The candidate collected all data for Chapters 2 to 5. Chapters 6 and 7 required specialised equipment that was only available in England. Due to COVID-19 travel restrictions, data were collected by Loughborough University in England under the training and guidance of the candidate. Data processing from all chapters were carried out by the candidate with assistance from Dr Jono Neville. Under the supervision of Dr Tom Stewart, the candidate carried out all machine learning and statistical analysis. The candidate conducted the write-up for all chapters with input from corroborating authors.

Chapter 2 - Upper body activity classification using an inertial measurement unit in court and field-based sports: a systematic review.

Preface

A systematic literature review was conducted to explore past research that classified upper body movements in court and field-based sports using inertial measurement units (IMU). This aided future chapters of the thesis by identifying equipment and techniques that worked successfully and possible gaps in the literature. It also provided the sports science field with a review that explained in a non-technical way how these systems work, as well as the benefits and limitations that need to be considered when using IMUs.

This paper was published in 2020 in the Journal of Sports Engineering and Technology, 235(2), 83-95.

Abstract

Inertial measurement units (IMUs) are becoming increasingly popular in activity classification and workload measurement in sports. This systematic literature review focuses on upper body activity classification in court or field-based sports. The aim is to provide sports scientists and coaches with an overview of the past research in this area and the processes and challenges involved in activity classification. The SPORTDiscus, PubMed, and Scopus databases were searched, resulting in 20 articles. Both manually defined algorithms and machine learning approaches have been used to classify IMU data with varying degrees of success. Manually defined algorithms may offer simplicity and reduced computational demand, whereas machine learning may benefit complex classification problems. Inter-study results show that no one machine learning model is best for activity classification; differences in sensor placement, IMU specification, and pre-processing decisions can affect model performance. Accurate classification of sporting activities could benefit players, coaches, and team medical personnel by providing an objective workload estimate. This could help prevent injuries, enhance performance, and provide valuable data to the coaching staff.

Key points

- IMUs, combined with manually-defined algorithms or machine learning, can classify upper body movements in court or field-based sports with varying levels of success.
- The number of IMUs, their placement location, and the sensor specifications need to be considered when developing a classification system, as these can affect accuracy.
- Classification accuracy is also affected by data processing (e.g., filtering, windowing, feature engineering) and the classification model used, regardless of the sport or action performed.
- As movement patterns can vary between elite and sub-elite athletes, a broad range of skill levels should be considered when training an algorithm for improved generalisation.

Introduction

Activity classification is a fundamental part of movement analysis for most athletes and coaches. It is important for enhancing performance, athlete monitoring, and injury prevention. Movement data for activity classification is often captured using video analysis, which can be expensive, have low portability, and be labour intensive to analyse.[46,47] In addition, the placement or angle of video cameras can lead to missed body movements or orientation issues.[48-50] An alternative approach that may overcome some of these limitations is using an inertial measurement unit (IMU). IMUs are wearable sensors that contain an accelerometer, gyroscope, and sometimes a magnetometer. The accelerometer measures linear acceleration (measured in g-force), the gyroscope measures angular velocity (degrees per second), and the magnetometer measures the strength and direction of the local magnetic field. This information can be used for activity classification, analysis of movement parameters, and measurement of physical load.

The IMU (including battery, processing unit, wireless unit and local storage) can now be made smaller than a matchbox, allowing comfortable attachment with minimal disturbance to normal movement.[51] The price can vary; however, consumer-grade IMUs can cost less than 100 USD, making them affordable for most athletes in developed nations. IMUs embedded in smartphones have recently been used for activity classification in basketball,[52] cross country skiing,[53] and soccer and hockey.[51] As smartphone subscriptions are expected to rise to 8.3 billion in 2023,[54] IMUs could be accessible to most of the world's population.

The present review focuses on research that classifies upper body movements in court or field-based sports using an IMU. The reason for this scope is that upper body movement patterns can have high variability between sports (e.g., cricket batting strokes, volleyball spike), therefore, providing additional challenges for classification. Lastly, this review will provide sports scientists and strength and conditioning coaches with an overview of the processes involved in activity classification and the various study design and data processing decisions that can affect classification performance. This information may assist with judging the applicability of IMU-based activity classification in different sporting contexts.

Methods

Search strategy

This systematic literature review followed the PRISMA guidelines. The databases PubMed, SPORTDiscus, and Scopus, were searched for articles dated from January 1st, 2000, to December 20th,

2018. The terms used for each database search can be observed in Table 2-1. The inclusion criteria consisted of full-length journal articles or published conference papers written in English. Articles must have presented classification results for an upper body movement in court or field-based sports using an IMU.

Table 2-1: Database search terms used.

PubMed
(((((((wearable OR microsensor OR device OR IMU OR inertial measurement unit OR triaxial OR sensor OR accelerometer OR gyroscope)) AND ((Movement detection OR event classification OR classification OR workload OR recognition OR analysis)) AND (cricket OR baseball OR softball OR handball OR American football OR discus OR volleyball OR tennis OR table tennis OR badminton OR throwing)) NOT (clinical OR review))))
SPORTDiscus
(wearable OR microsensor OR device OR IMU OR inertial measurement unit OR triaxial OR sensor OR accelerometer OR gyroscope) AND (Movement detection OR event classification OR classification OR workload OR recognition OR analysis) AND (cricket OR baseball OR softball OR handball OR American football OR discus OR volleyball OR tennis OR table tennis OR badminton OR throwing) NOT (clinical OR review)
Scopus
wearable OR microsensor OR device OR IMU OR "inertial measurement unit" OR triaxial OR sensor OR accelerometer OR gyroscope AND "Movement detection" OR "event classification" OR classification OR workload OR recognition OR analysis AND cricket OR baseball OR softball OR handball OR "American football" OR discus OR volleyball OR tennis OR table AND tennis OR badminton OR throwing AND NOT clinical OR review

Results

Selection process

A flow diagram of the process used to identify and select articles is detailed in Figure 2-1. The initial search resulted in 1,454 articles. The titles of these articles were then screened for duplicates and relevance to the inclusion criteria. In total, 79 article abstracts were screened, of which 43 were reviewed in their entirety. Out of these articles, 13 met the inclusion criteria. The reference lists of these articles were then reviewed, which netted an additional seven articles. This brought the total number of included articles to 20.

Figure 2-1: PRISMA flow diagram of the systematic search.

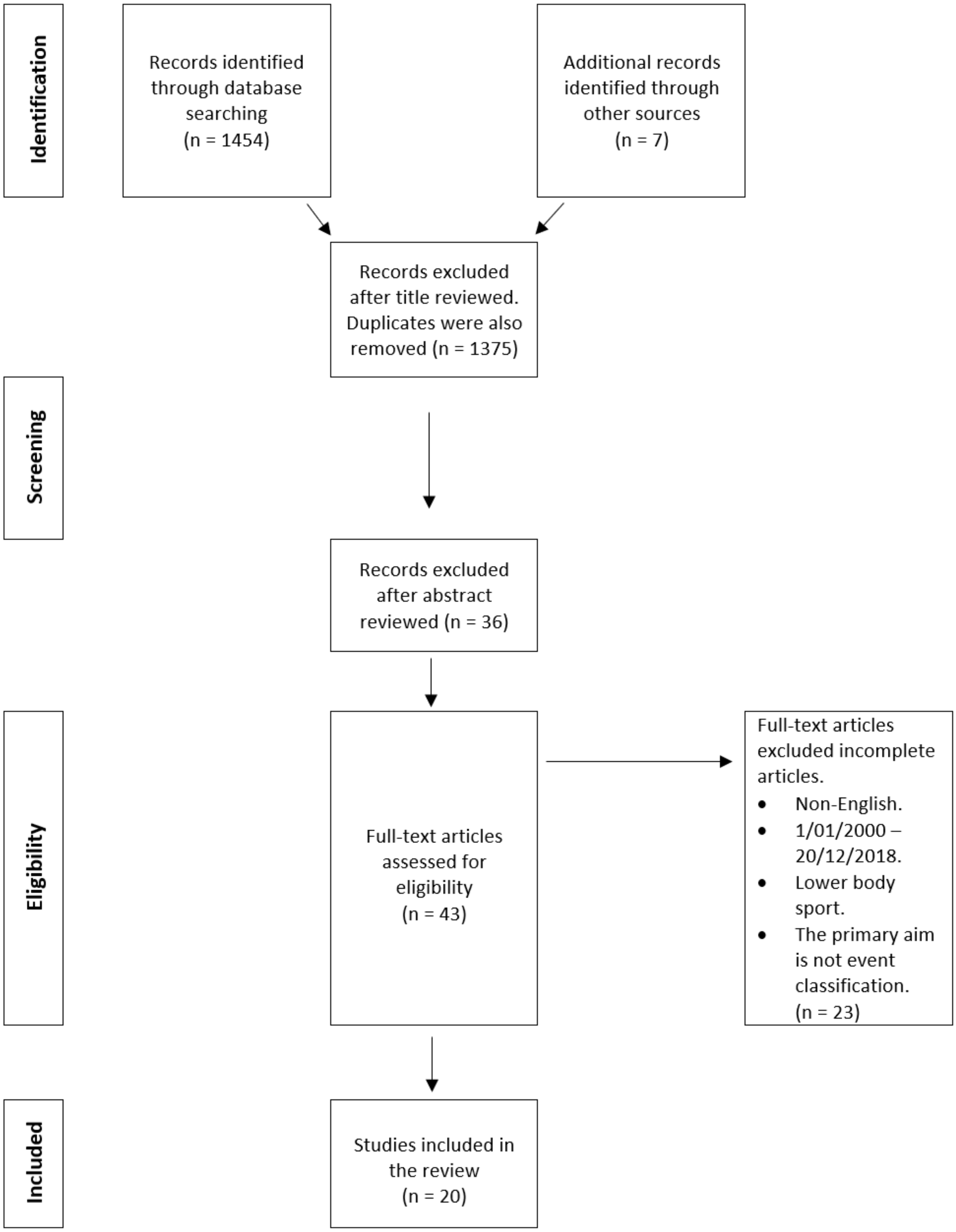


Table 2-2 summarises the study characteristics from the 20 research articles. Eight out of the 20 studies examined stroke classification in a racket sport (tennis,[48,55-60] badminton,[55,61] and squash[55]), while the remaining studies looked at classification in cricket fast bowling,[62-64] cricket batting,[65] baseball,[66,67] hockey,[51] volleyball,[67] beach volleyball,[68] and basketball.[49,52,69,70] Eight studies included both novice and experienced athletes, while the other studies used either novice participants (n= 1), experienced participants (n = 6), developmental athletes (n = 2) or did not state experience levels (n = 5). The study sample size ranged from 2 to 70, with the median being 12 participants.

Table 2-2: Summary of past studies that used IMUs to quantify sport-specific movements.

Reference	Study aim and Participants	IMU	Objectives (O)	Data signal processing (SP), Segmentation (S), Event detection (ED)	Feature computation (FC), Validation method (VM), Classification method (CM)	Results
Anand et al. (2017)[55]	Stoke detection and classification in tennis, badminton and squash. Tennis = 31 participants. Badminton = 34 participants. Squash = 5 participants.	Samsung smartwatch Gear S2. Triaxial ACC (±8 g) and GYR (±2000 deg·s^{-1}). SF = 100 Hz. IMU mounted on the dominant wrist.	O1: Stroke detection in all 3 sports. O2: Stoke classification (Tennis, 5 strokes; Badminton, 4 stokes; Squash, 2 strokes).	SP: Not specified. S: Event detection window (size not specified). ED: A combination of ACC and GYR thresholds, jerk based detection and shape-based detection of the time warp signal.	FC: ~2000 time domain and correlation-based. Reduced to an unspecified number by feature selection. VM: Two independent groups. CM: O1: Refer to ED; O2: Machine learning (LR, DCNN, BLSTM).	O1: F-score: Tennis = 92%, Badminton = 88%, Squash = 96%. O2: Average accuracy across sports: LR = 86.1%, DCNN = 88.1%, BLSTM = 88.9%.
Bai et al. (2016)[52]	Classification of shooting attempts in a one-to-one basketball game. 2 participants.	Microsoft Band & Android phone. Triaxial ACC, GYR. SF = 32 Hz & 100 Hz respectively. Microsoft Band device mounted on the dominant wrist and an Android phone placed in the trouser pocket.	O1: Detecting shot event. O2: Detecting shooter.	SP: Not Specified. S: Sliding window (length individualised to the player, 4-6 s, 1 s overlap). ED: N/A.	FC: Time and frequency domain features for Microsoft Band (66) and Android phone (15). VM: 10-fold cross-validation. CM: O1: Machine learning (RF, NB, CT, SVM, KNN); O2: Machine learning (RF, SVM, KNN).	O1: RF had the highest accuracy (88%). O2: SVM had the highest F-score (92.80%).
Connaghan et al. (2011)[48]	Stroke classification in tennis. 8 participants of different playing abilities.	TennisSense IMU system. Triaxial ACC ± 10 g, GYR ± 1200 deg·s^{-1} and MAG ± 6 Gauss. SF = not specified. IMU mounted distally on the dorsal surface of the forearm.	O1: Stroke and non-stroke detection during a competitive match. O2: Determining the best sensor modality for classifying data into 3 main strokes.	SP: Data was filtered for noise (filter not stated). S: Event detection window (size not specified). ED: ACC magnitude > 3 g.	FC: Unspecified. VM: O1: 10-fold cross-validation; O2: LOOCV. CM: Machine learning (NB).	O1: Accuracy = 85%, F-score = 77.61%. O2: Individually, the ACC had the highest accuracy (79%). A combination of all three sensors gave the best accuracy (90%).
Hölzemann and Van Laerhoven (2018)[69]	Classification of activities in basketball. 3 experienced participants (M).	Custom IMU (MPU-9250). Triaxial ACC. SF = 200 Hz (down-sampled to 25 Hz). IMU mounted on the wrist.	O1: Classify 4 movement activities related to basketball.	SP: Not specified. S: Sliding window (1 s). ED: N/A.	FC: 6 time-domain features. VM: LOOCV. CM: Machine learning (KNN, RF).	O1: RF had the highest accuracy (87.48%) and F-score (69.45%).

Kautz et al. (2017)[68]	Stroke classification in beach volleyball. 30 participants of different playing abilities. (M&F).	Bosch BMA280 IMU. Triaxial ACC (±16 g). SF = 39 Hz. GYR was not used. IMU mounted on the dorsal surface of the wrist of the dominant hand.	O1: Classifying data into 9 beach volleyball actions (1 null class). O2: Determine if accuracy improves with feature selection.	SP: High-pass Butterworth filter (8 Hz), then a low-pass Butterworth filter (3 Hz). S: Event detection window (4 s). ED: Max ACC signal > 5.1 $m \cdot s^{-2}$ & absolute ACC 200 ms before maximum.	FC: O1: 39 time and frequency domain features; O2: Reduced to 12 selected features by feature selection. VM: 10-fold cross-validation. CM: Machine learning (SVM, KNN, NB, CT, RF, VOTE, DCNN).	O1: DCNN had highest balanced accuracy (79.5%). SVM = 59.7%. O2: Accuracy with feature selection was consistently lower.
Khan et al. (2017)[65]	Classification of different cricket batting shots. 6 participants of different playing abilities (M&F).	4 IMUs. Triaxial ACC, GYR, MAG. SF = 100 Hz. IMUs mounted on the lower part of all 4 limbs.	O1: Five different levels of classification. 1 = Shot or no shot, 2 = hit of miss, 3 = shot direction, 4 = type of footwork used, and 5 = shot type.	SP: Not specified. S: Sliding window (1.6 s). ED: N/A.	FC: 464 time and frequency domain features. VM: LOOCV. CM: Machine learning (CT, KNN, SVM).	O1: No significant difference between type of classifier used and level of classification. Average F-score across all levels: SVM = 84.14%, KNN = 83.26%, DT = 79.24%.
Kos and Kramberger (2017)[56]	Stroke classification in tennis. 7 participants of different playing abilities.	IMU: Triaxial ACC (±16 g) and GYR (±2000 $deg \cdot s^{-1}$). SF = 1000 Hz. IMU was mounted on the posterior surface of the wrist.	O1: Classify into 3 main strokes.	SP: Not specified. S: Event detection window (size not specified). ED: Derivative average of all 3 ACC axes > 300.	FC: 4 time-domain features. VM: Not specified. CM: Manually defined algorithm (GYR max in y axis = backhand; GYR max in x & min in z = serve; GYR max in x & min in y + min in z < 1500 $deg \cdot s^{-1}$= serve, > 1500 $deg \cdot s^{-1}$= forehand).	O1: Accuracy: Serve = 98.8%, Forehand = 93.5% and backhand = 98.6%.
Kos et al. (2016)[57]	Stroke classification in tennis. 3 players of different playing abilities.	IMU: Triaxial ACC (±16 g) and GYR (±2000 $deg \cdot s^{-1}$). SF = 1000 Hz. IMU was mounted on the posterior surface of the wrist.	O1: Classify into 3 main strokes.	SP: Not specified. S: Event detection window (size not specified). ED: Derivative average of all three GYR axes > determined threshold.	FC: 5 time-domain features. VM: Not specified. CM: Manually defined algorithm (GYR max in y axis & min in x = backhand; GYR max in x & min in z = serve; GYR max in x & min in y = forehand).	O1: Accuracy: Serve = 100%, Forehand = 96.23% and backhand = 98.11%.
Mangiarotti et al. (2019)[70]	Classification of activities in basketball between two participants.	2 Custom IMU (MPU-6050). Triaxial ACC ± 16 g and GYR ±2000 $deg \cdot s^{-1}$. SF = 250 Hz.	O1: Classify 3 movement activities related to basketball.	SP: Butterworth low pass filter. S: Sliding window (1 s, 80% overlap).	FC: Unspecified number of time-domain features. VM: Non-independent training and testing group.	O1: F-score = 94.56%.

		IMUs were mounted on the posterior surface of the wrist.		ED: N/A.	CM: Machine learning (Combination of KNN and SVM).	
McGrath et al. (2019)[63]	Fast bowling detection in cricket. 17 experienced participants (M).	Sabel Sense IMU. Triaxial ACC ± 16 g and GYR ±2000 deg·s^{-1}. SF = 250 Hz. IMU was mounted to the upper posterior trunk between the shoulder blades.	O1: Bowling detection. O2: Determine the best window size. O3: Bowling detection with sampling frequency down-sampled to 125 Hz, 50 Hz and 25 Hz.	SP: 1Hz Butterworth low pass filter. S: Event detection window (10 s and 1 s). ED: Peaks > 500 deg·s^{-1} in the magnitude of GYR x, y and z axis.	FC: 282 time and frequency domain features. Reduced to 223 by feature selection. VM: Two groups (6 new participants in the testing group, 4 the same) CM: Machine learning (RF, SVM, PSVM, NN, XGB).	O1: F-score: SVM & PSVM = 100%. O2: No significant difference between 10 s and 1 s for any classifiers. O3: F-score: No significant difference at 25 Hz. XGB and RF = 97%.
McNamara et al. (2015)[62]	Fast bowling detection in cricket. Training phase: 12 experienced fast bowlers (M). Competition phase: 5 experienced fast bowlers (M).	Catapult MinimaxX S4 IMU. Triaxial ACC ± 10 g, GYR ± 1200 deg·s^{-}, and MAG. SF = 100 Hz. IMU was mounted to the upper posterior trunk between the shoulder blades.	O1: Classifying training data into bowling and non-bowling events. O2: Classifying competition data into bowling and non-bowling events.	SP & S not stated. ED: See CM.	FC: Not specified. VM: Not specified. CM: Manually defined algorithm. Back-foot contact (ACC) and peaks in the rotation speed (GYR) of the upper torso (> 500 deg·s^{-1}).	O1: Sensitivity = 99%, specificity = 98.1%. O2: Sensitivity = 99.5%, specificity = 74%.
Mitchell et al. (2013)[51]	Classification of activities in five-a-side soccer and field hockey using an IMU from a smartphone. Soccer = 15 amateur participants. Hockey = 17 elite participants.	1 Google Nexus One or HTC Desire smartphone. Triaxial ACC. SF = 16 Hz & 25 Hz respectively. IMU was mounted to the upper posterior trunk between the shoulder blades.	O1: Classifying data into 7 activities common to both sports. O2: Determining the best window length for both sports. O3: Determine the best mother wavelet function and DWT level.	SP: Filter designed by DWT. S: Event detection window (varying window length, 1 to 9 s). ED: Not specified.	FC: 42 descriptive features extracted utilising DWT. VM: Repeated (10x) cross-validation (66/33% train/test split). CM: Machine learning (SVM, KNN, NB, CT, NN).	O1: Hockey: NN had the highest F-score (82.3%), Soccer: NB (79.9%). Combination: Hockey = 88.8%, Soccer = 86.3%. O2: Hockey 7 s, Soccer 3 s. O3: Wavelet. Hockey = bior1.1, Soccer = rbio1.1 (Level 6 for both).
Mlakar and Luštrek (2017)[58]	Stroke classification in tennis. 5 experienced participants.	Catapult S5. GPS = 10 Hz, triaxial ACC ± 16 g, GYR ± 2000 deg·s^{-}, and MAG. SF = 100 Hz. IMU was mounted to the upper posterior trunk between the shoulder blades.	O1: Shot detection. O2: Classify into 3 main strokes.	SP: Butterworth bandpass filter (1.5 & 25 Hz) applied to windowed data. S: Sliding window (0.8 s and 1.2 s). ED: N/A.	FC: Unspecified number of time-based features. VM: LOOCV. CM: O1 & O2: Machine learning (RF).	O1: F-score = 94.1% O2: F-score: Forehand = 91%, Backhand = 92.1%, Serve = 99%.

Murray et al. (2016)[66]	Pitching and throwing detection in baseball. 17 development athletes.	Catapult MinimaxX S4 IMU. Triaxial ACC ± 10 g, GYR ± 1200 deg·s⁻, and MAG. SF = 100 Hz. IMU was mounted to the upper posterior trunk between the shoulder blades.	O1: Classifying training data into bowling and non-bowling events. O2: Classifying competition data into bowling and non-bowling events.	SP & S not stated. ED: See CM.	FC: Not specified. VM: Not specified. CM: Manually defined algorithm. Unspecified peaks in the rotation velocity (GYR), unspecified ACC measures.	O1: Sensitivity = 100%, specificity = 79.8%. O2: Sensitivity = 100%, specificity = 74.4%.
Nguyen et al. (2015)[49]	Classification of activities in basketball. 3 participants.	5 custom IMUs (LSM9DS0). Triaxial ACC. SF = 200 Hz (down-sampled to 40 Hz). IMUs mounted on the lower back, both feet and both knees.	O1: Classify 7 movement activities related to basketball.	SP: Low pass filter (15 Hz). S: Sliding window (3.2 s, 50% overlap). ED: ACC Z axis > 4 g.	FC: 80 time and frequency domain features. An unspecified number of cross-correlation coefficients. VM: LOOCV. CM: Machine learning (SVM).	O1: Overall accuracy = ~67.2%.
Rawashdeh et al. (2016)[67]	Classification of volleyball serves and baseball throws. 11 participants of different playing abilities.	Custom made IMU. Triaxial ACC ± 16 g, GYR ± 2000 deg·s^{-1}, MAG ± 8 Gauss. SF = 50 Hz. IMU was mounted to the medial surface of the upper arm of the dominant arm.	O1: Classifying data into volleyball serves and baseball throws. Non-events included shoulder rehab exercises.	SP: No specified. S: Event detection window (size not specified). ED: Event detection: Elevation of Arm > 45 deg and Angular Rate > 400 deg·s^{-1}.	FC: 81 time and frequency domain features. Reduced to 8 features by feature selection. VM: Cross-validation (40/60% train/test split). CM: Machine learning (DT).	O1: Accuracy = 86%.
Salman et al. (2017)[64]	Legality analysis of fast bowling in cricket. 14 medium to fast-paced bowlers (M).	3 Custom IMU. Triaxial ACC and GYR. SF = 150 Hz. IMU was mounted on the upper arm, forearm, and wrist.	O1: Classify the legality of bowling action. O2: Determine if accuracy improves with feature selection.	SP: Filter not stated. S: Event detection window (7 s). ED: GYR maxima in the x-axis.	FC: 120 time-domain features. Reduced to 21 by feature selection. VM: LOOCV. CM: Machine learning (SVM, K-NN, NB, RF, NN).	O1: F-score: NB = 80.7%, SVM = 89%, KNN = 90.3%, NN = 87%, RF = 86%. O2: Highest F-score was obtained after feature selection for all classifiers.
Srivastava et al. (2015)[59]	Stroke classification in tennis. 14 participants of different playing abilities.	Samsung smartwatch Gear S. Triaxial ACC (±8 g) and GYR (±2000 deg·s^{-1}). SF = 25 Hz. IMU mounted on the wrist.	O1: Classify into 3 main strokes. O2: Classify 3 main sub-strokes.	SP: Pan Tomkin's algorithm was used to isolate shot signal from noise. S: Sliding window (3*sampling rate). ED: ACC spike above an unspecified threshold.	FC: Not specified. VM: Not specified. CM: Machine learning (Phase 1: DTW. Phase 2: QDTW).	O1: Sensitivity: Professional = 99.6%, Novice = 99.3%. O2: Sensitivity: Professional = 90.7%, Novice = 86.2%.

Wang et al. (2016)[61]	Stroke classification in badminton. 12 club participants.	4 IMUs. Triaxial ACC ± 16 g, GYR ± 500 deg·s^{-1}, biaxial GYR ± 500 deg·s^{-1}. SF = 500 Hz. IMUs mounted to both wrists, waist, and right ankle.	O1: Classifying data into the 14 strokes and 4 non-stoke motions during a competitive match. O2: Determining the best segmentation approach.	SP: Low pass filter designed from a Wavelet function Coif5. S: Sliding window (1 s, 50% overlap) vs sliding window with high acceleration (> 1.5 g) vs event detection window (1 s). ED: ACC magnitude > 1.5 g & > 1 s removed from adjacent max.	FC: 120 time-domain and frequency-domain features. VM: 4-fold cross-validation. CM: Machine learning (HMM, NB, CT, SVM, LDF).	O1: HMM had the highest accuracy (97.96%). O2: The event detection window had the highest accuracy (97.96%).
Whiteside et al. (2017)[60]	Stroke classification in tennis. 19 development athletes (M&F).	IMeasureU IMU. Triaxial ACC ± 16 g, GYR ± 2000 deg·s^{-1}. SF = 500 Hz. IMU was mounted distally on the ventral surface of the forearm of the hitting limb.	O1: Classifying data into the 3 main stokes and false stroke. O2: Classifying data into 8 strokes and false stroke.	SP: Linear interpolation (smoothed using a 50-point moving average). S: Event detection window (1 s). ED: Max ACC signal > 25 m·s^{-2} & > 1.25 s removed from adjacent max. GYR > 400 deg·s^{-1} within 0.06 s either side of ACC max.	FC: 40 features (5 features x 8 waveforms). VM: 10-fold cross-validation. CM: Machine learning (SVM, DA, RF, KNN, CT, NN).	O1 & 2: SVM (cubic kernel) had the highest overall F-score (O1 = 94.33%, O2 = 90.01%).

Key: ACC = Accelerometer; bior1.1 = Biorthogonal 1.1; BLSTM = Bidirectional long short-term memory networks; DCNN = Deep convolutional neural networks; CT = Classification tree; DA = Discriminant analysis; DTW = Dynamic time warping; DWT = Discrete wavelet transform; F = Female; FBC = Feature based classification; GYR = Gyroscope; HMM = Hidden Markov model; KNN = K-nearest neighbour; LDF = Linear discriminant function; LOOCV = Leave one out cross-validation; LR = Logistic regression; MAG = Magnetometer; M = Male; N/A = Non-applicable; NB = Naive Bayesian classifiers; NN = Neural network; PSVM = Polynomial SVM; QDTW = Quaternion dynamic time warping; a_{rad} = Radial acceleration; RBF = Radial basis function; RF = Random forest; rbio 1.1 = Reverse biorthogonal 1.1 wavelet; SF = Sample frequency; SVM = Support vector machine; ω_{tan} = Tangential angular velocity; VOTE = Plurality voting; VM = Validation method; XGB = Gradient boosting.

Discussion

Evaluation metrics

Included studies used a range of metrics to evaluate the performance of their classification models. The classification of a binary event has four possible outcomes: (1) true positive (TP), where the event is labelled correctly as an event by the algorithm; (2) true negative (TN), where a non-event is labelled correctly as a non-event by the algorithm; (3) false positive (FP), where a non-event is incorrectly labelled as an event by the algorithm; and (4) false negative (FN), where an event is incorrectly labelled as a non-event by the algorithm. These outcomes can be presented in several ways, but commonly used methods include sensitivity (also known as recall), specificity (also known as the true negative rate), precision (also known as the positive predictive value), accuracy (the fraction of correctly identified events and non-events), F-score (the harmonic mean of precision and sensitivity), and balanced accuracy (the mean of sensitivity and specificity). Table 2-3 provides formulas and a description of each metric. Thirteen studies reported (or displayed classification matrices so the reader could calculate) more than one evaluation metric.[51,52,58,60,62-65,67-70] Depending on the nature of the classification problem, reporting only accuracy can misrepresent the results; data with imbalanced classes (e.g. a low number of observations in one activity class relative to another) can lead to inflated overall accuracy. Balanced accuracy and F-score are preferred in this regard. While the acceptable level of accuracy may depend on the application (e.g., a higher error rate might be acceptable for a system that predicts step counts for general consumers but unacceptable when quantifying bowling frequency in elite cricket players), the goal of fitting classification models is to achieve the best accuracy possible. An accurate classification model is particularly important when there is a relationship between activity parameters (e.g., load, volume, technique) and injury or performance.

Table 2-3: Description of the performance measures used.

Format	Formula	Description
Accuracy	$\frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}}$	The fraction of correctly identified events and non-events.
Sensitivity	$\frac{\mathrm{TP}}{\mathrm{TP+FN}}$	The proportion of correctly identified events, regardless of all non-events.
Specificity (Recall)	$\frac{\mathrm{TN}}{\mathrm{TN+FP}}$	The proportion of correctly identified non-events, regardless of all events.
Precision	$\frac{\mathrm{TP}}{\mathrm{TP+FP}}$	The proportion of correctly identified events relative to the total number of events detected.
F-score	$\frac{2 \text{ x precision x sensitivity}}{\text{precision} + \text{sensitivity}}$	The harmonic mean of precision and sensitivity.
Balanced accuracy	$\frac{\text{sensitivity} + \text{specificity}}{2}$	The mean of sensitivity and specificity.

Key: FN = False negative; FP = False positive; TN = True negative; TP = True positive.

IMU number

Using multiple IMUs means that sensors can be attached to both the trunk and extremities. This may provide complementary information to classification algorithms, therefore improving accuracy.[45] In this review, six studies[49,52,61,64,65,70] used multiple IMUs (between 2 and 5). Only one study examined individual versus combined accuracy and found that accuracy from a single IMU on the playing wrist was high (95%) when classifying four main badminton shots (forehand serve, backhand serve, forehand clear, and backhand clear).[61] However, when classifying ten additional sub-shots, the accuracy was significantly reduced (76%). When using a combination of four IMUs to classify the main and sub-shots, improved accuracy of 97% was observed. This may have been one of the reasons for the reduced accuracy when using a single IMU for classifying sub-shots in tennis,[59,60] compared to the results obtained by classifying main shots. Although multiple IMUs may improve classification performance, there are limitations regarding cost, simplicity of use, computational demand, movement restriction, and comfort, meaning a single IMU may still be preferred in a given context.

IMU specification

The three main specifications commonly reported for an IMU include: sampling frequency; the accelerometer, gyroscope, and magnetometer measurement range; and the number of axes each sensor can record. Triaxial IMUs can capture more data, improving classification accuracy, particularly for limb-based movements.[71] Sampling frequency ranged from 16 Hz to 1000 Hz in the reviewed studies, and sensor measurement range varied from 8 to 16 g for the accelerometer and 500 to 2000 degrees per second for the gyroscope. Although studies have shown that accuracy can be high when using a low sampling rate IMU,[51,59,63,68] only one study compared the accuracy of a high versus low sampling rate when detecting fast bowling events in cricket.[63] The accuracy did not significantly decrease when data were down-sampled from 250 Hz to 25 Hz (F-score 1.0 vs 0.97). It is important to note that this classification problem only had two possible outcomes (bowl or non-bowl), and the IMU was attached to the torso. Attaching an IMU with a low sampling rate or measurement range to an extremity may mean that a significant amount of information is lost during high-velocity movements. For example, a cricket bowl can generate upwards of 80 g at the wrist, but many IMUs are limited to 16 g. It is currently unclear how this impacts classification accuracy for different movements. Future studies should compare the minimum specifications required to inform the appropriate choice of IMU for each movement.

Lastly, it should also be mentioned that sensor resolution, which is seldom reported, can play an essential role in the accuracy of recorded data. IMU resolution refers to the smallest change that the sensor can detect in the activity that it is measuring. It is influenced by the full-scale range of the IMU.

For example, a higher resolution is desired when slight changes in acceleration are of interest without limiting the full range of the accelerometer. Most standard IMU units will have a minimum 8-bit resolution, with the vast majority being 16-bit or above.

IMU location

IMUs can be positioned at several locations on the body. This includes anywhere on the torso and all four extremities. The optimal placement location is likely specific to the sport under study.[72] Five studies mounted an IMU solely on the participants' torso,[51,58,62,63,66] while the remaining 15 studies had at least one IMU located on an extremity (e.g. wrist or ankle). A common location for the torso is between the shoulder blades on the upper thoracic spine.[51,58,62,63,66] The benefit of this location is that it does not impede movement, is protected from more common frontal impacts, and can be embedded in clothing.[60,73] These reasons make it a popular choice for sports that involve high impact forces on the body (e.g. rugby). It is also a convenient location for IMUs with lower accelerometer and gyroscope maximum ranges, as there is generally less acceleration and angular velocity than in a limb-based location.[51,63] Alternatively, IMUs with higher dynamic ranges have been successfully attached to extremities in sports with significant limb movement, such as tennis[60] and badminton.[61] These locations may provide an improved classification of limb movements, as the trunk is predominately exposed to linear acceleration and may not provide adequate spatiotemporal information to distinguish limb-based events.[74] This was evident in a study that examined badminton stroke classification,[61] which showed a wrist-mounted IMU of the dominant hand had an overall accuracy of 81.6% compared to 42.6% when placed on the waist. As IMUs with higher sampling frequency and measurement ranges are embedded in affordable wrist-worn consumer-grade products (e.g., the Apple Watch), developing classification algorithms for these systems may increase the accessibility to accurate activity classification.

Sensor modality

Activity classification algorithms can use data from just one sensor (i.e., the accelerometer, gyroscope, or magnetometer) or several sensors simultaneously. The magnetometer is used the least frequently of the three sensors present in most IMUs. This is because a controlled magnetic field without any distortion is required for optimal results,[75] especially with fast linear motions.[74] Two studies,[48,65] used data from all three sensors (accelerometer, gyroscope and magnetometer), while the remaining studies used either a combination of the accelerometer or gyroscope[52,56-64,66,67,70,76] or only the accelerometer[49,51,68,69]. Only one study investigated different combinations of sensors for classifying tennis strokes.[48] The accelerometer showed slightly higher classification accuracy (79%) compared to the gyroscope (76%) and the magnetometer data (76%), but 90% accuracy was achieved when combining all three sensors. The authors proposed that the lower accuracy observed using gyroscope data could be related to high variation in the temporal orientation of a given tennis stroke due to differing skill levels between participants. The importance of the gyroscope may depend on the movement being performed, the IMU placement position, and the gyroscope features that are quantified (see section 'Feature engineering and feature selection'). McGrath et al. (2019)[63] showed that when using a trunk-mounted sensor, gyroscope features consistently had higher importance for all classification models when detecting cricket bowls. This could have been because trunk rotation patterns are more distinct between activities (bowl and throw) compared to acceleration patterns. The low classification accuracy of volleyball strokes observed by Kautz et al. (2017)[68] may have been due to the omission of gyroscope data, as rotational patterns of different volleyball strokes are likely to be distinct. Although most studies do not report feature importance or compare sensor modalities, it is likely that accelerometer and gyroscope features are equally important for classifying movements in sports that have high rotational velocity.

Pre-processing

Pre-processing is an important step in the classification process, as it can improve accuracy and reduce the computational load for the classifier.[77,78] Pre-processing generally involves filtering raw data, data segmentation, feature engineering, and feature selection.

Filtering

Filtering makes raw data more representative of the desired activity by minimising the amount of data not associated with human movement tasks. This is often called signal artefact or noise and is usually caused by clothing, skin, muscle and other subcutaneous tissue movements.[79] A wide range of algorithmic approaches can convert raw data to time, frequency and discrete domains of representation so that it is easier to apply different filtering techniques.[77] Several filters have been specified, including moving average filters,[60] wavelet functions,[51,61] low pass filters,[49,58,63,68,70] and high pass filters.[58,68] Wavelet functions allow time-frequency characteristics to be examined, which help detect events involving a non-repetitive movement that varies over time (e.g. cricket batting).[78] Furthermore, wavelet functions can capture sudden changes in signals that are common in accelerometer and gyroscope readings.[77] There were no studies in this review that compared different filtering techniques. However, a low pass filter is the most common filter as the artefact is typically represented as high-frequency signals.[78] IMU filtering techniques must be not over-aggressive and preserve signal characteristics relevant to the activity of interest.[45]

Data segmentation

Data segmentation involves arranging data into windows (segments of data), so that features can be extracted. Despite the potentially large impact that different window types and sizes can have on classification accuracy,[45,71] the comparison of window techniques is rarely performed,[67,80] especially when classifying sporting events.

The sliding window technique is a common segmentation approach that requires no prior identification of potential events. Signals are divided into fixed-length windows that can be arranged sequentially or overlapping. A total of eight studies used this approach, in cricket,[65] basketball,[49,52,69,70] tennis,[58,59] and badminton.[61] Four of these studies used an overlapping sliding window (ranging from 16 to 50 % overlap).[49,52,61,70] This helps prevent one of the limitations of traditional sliding windows when activities occur at either end of the window segment.[61,81] Another potential limitation is a large amount of null data (data not of interest) retained. This can increase the error rate and the computation demand during the modelling process.[67] Sliding windows may be beneficial when classifying all activities within a given period, such as movements of interest and recovery periods (e.g., standing still).

An event-defined window is formed around a previously detected signal event. This windowing technique can decrease the processing time and improve model performance by only retaining data around potential events of interest. This type of window is often used in event classification involving sporadic movements,[80] and most of the studies (12 studies) used this approach. Event detection methods included using a predetermined threshold from either raw or derived accelerometer data,[48,49,56,59,61,68] raw or derived gyroscope data,[57,63,64,67] or a combination of the two.[55,60] However, a single threshold rule may not be enough to reliably detect events due to heightened noise around signal peaks and events with more than one peak. Five studies included additional rules as follows: two excluded peaks within a certain timeframe of an adjacent peak;[60,61] Kautz et al. (2017)[68] averaged absolute acceleration values of all three axes over a 200 ms interval before the maximum deflection occurred in a beach volleyball stroke; Rawashdeh et al. (2016)[67] only included volleyball and throwing events when the arm elevated above 45 degrees; and Anand et al. (2017)[55] used jerk and shaped based detection of the time warp signal for tennis, badminton and squash.

One study compared sliding and event-defined windows in badminton.[61] The event-defined window had the highest accuracy, lowest computation time and aided in describing the movement characteristics of badminton strokes due to the consistency of the event data in each window. Generally, improving event detection accuracy leads to more precise activity classification and reduces model complexity.[45,72]

Deciding on the appropriate window length to apply is another consideration. Only one study investigated the effects of different window durations (1 to 9 seconds) when using an event-defined window to classify activities common to both five-a-side football and field hockey.[51] Their results showed that the most optimal window length was 3 seconds for football and 7 seconds for hockey. When deciding on an appropriate window length, one must factor in the activity of interest and the type of window used. Choosing a too long window increases the chances of two or more activities occurring within one window and greatly increases classification difficulty.[51] A window that is too short may mean that the complete task is not included in the window. Viewing raw data may help determine appropriate window lengths.

Feature engineering and feature selection

Feature engineering or feature computation involves identifying and extracting quantitative measures from the data either manually or automatically. Most supervised machine learning models do not create their features, so their performance is based on the quality of the features computed. For signal data, it is common to compute features in the time and frequency domain, such as mean, standard deviation, and signal power.[72,78,82] Patterns among these features are used to distinguish between different movements or activities. Fourteen studies specified the number of features used, ranging from 4 to ~2000 features. Eight studies used only time-domain features,[55-58,60,64,69,70] while seven used a combination of features in the time and frequency domain.[49,52,61,63,65,67,68] No studies compared whether time or frequency domain features were more important for classifying certain movements; however, a broad range of features from both domains are likely to be beneficial. There is not enough evidence to propose an optimal feature set for the classification of specific movements, as it likely depends on the type of activity performed, in combination with IMU specifications and placement location.

Feature selection refers to reducing the number of features in the dataset while still representing the original feature set. A dataset with many features (i.e., high dimensionality) requires more training data to model parameter estimates,[45] increases computation time and can cause overfitting.[72,83] Five studies stated that they performed feature selection.[55,63,64,67,68] There are numerous ways to perform feature selection; a common method is to remove highly correlated features.[63] However, as shown by Kautz et al. (2017),[68] selecting too few features (39 features down to 12) severely reduced accuracy for all tested models. This may lead to insufficient information for the classification algorithm. It should be noted that some machine learning algorithms (e.g. certain neural networks (NN) and random forests (RF)) can learn the relevance of individual features and may not require feature selection.[78]

Activity classification

The reviewed studies observed two general activity classification approaches: manually defined algorithms and supervised machine learning techniques. Manually defined algorithms involve human-set rules that are usually based on previous observations. Data are classified by meeting or not meeting these sets of rules. Four studies used this approach to classify cricket fast bowling,[62] baseball,[66] and tennis strokes[56,57] with a moderate to high accuracy. The two tennis studies achieved high accuracy (> 93%) when classifying three main tennis strokes. Their algorithm was based on maximum gyroscope values among all three axes. McNamara et al. (2015)[62] used peaks in upper body rotational speed (gyroscope > 500 °/s deg·s^{-1}), front foot contact (accelerometer value unspecified), and GPS (value unspecified) to detect cricket bowls. Lastly, following a similar approach, Murray et al. (2016)[66] used unspecified gyroscope and accelerometer peaks to classify baseball pitches. Interestingly, specificity in both studies decreased when moving from a training environment to a game setting. This could be due to the unstructured nature of actions performed during competition – confusing the algorithm. Although manually defined algorithms may offer simplicity, interpretability, and reduced computational demand, future research needs to determine if this approach can distinguish between more complex movements (e.g., different tennis sub-shots).

Common supervised machine learning models used for activity classification include the support vector machine (SVM), linear discriminant analysis (LDA), logistic regression (LR), classification tree (CT), RF, k-nearest neighbour (KNN), naive Bayes (NB), and NNs. Supervised machine learning techniques were used in 16 of the studies to classify tennis,[48,55,58-60] badminton,[55,61] volleyball,[67] beach volleyball,[68] baseball,[67] cricket,[63-65] basketball,[49,52,69,70] hockey,[51] and squash.[55]

Two studies compared the difference between deep learning neural networks (DNNs) versus conventional machine learning approaches.[55,68] Like conventional NNs that mirror some of the basic principles of a biological brain, DDNs use more than one hidden layer (typically five to more than a hundred). Different DDNs have been used for activity classification, for example, long short-term memory networks (LSTM) and convolutional neural networks (CNNs). Recently, Kautz et al. (2017)[68] found significant improvement in balanced accuracy (+19.2%) when using CNNs compared to conventional machine learning approaches. In addition, Anand et al. (2017)[55] found improvements in CNNs and LSTM versus LR (88.1% & 88.9% vs 86.1%, respectively). DDNs can learn high-level features with more complexity and abstraction than NNs when using many samples to train.[84] Further work is needed to investigate the effectiveness of deep learning methods, especially with higher sampling rate IMUs, which produce more voluminous data.

By examining classification results among the reviewed studies, no single model is best for activity classification in sport. For example, the support vector machine (SVM) was effective at classifying shooting attempts in basketball,[52] cricket bowling,[63,64] and strokes in tennis,[60] badminton,[61] and cricket[65] (all > 84 % accuracy). However, it was poor at classifying movements in beach volleyball (59.7 % accuracy),[68] where movement patterns are similar to tennis and cricket (spike vs serve and bowl). This is consistent with what has been termed the 'no free lunch theorem', where machine learning models will not perform equally well on all problems.[85] Additional reasons include changes to the number of sensors used,[45] the sensor location,[45] the sampling frequency,[71] and pre-processing techniques.[45,51,68] The lack of a clear classification model is not unique to sport, as these results are consistent in activity classification involving tasks of everyday living.[45] Combining different models may also be a viable option for improving accuracy.[51,70] Mitchell et al. (2013)[51] fused the best model for each activity in hockey and five-a-side football and improved the F-score by 6% compared to the best individual model.

When fitting machine learning models, attention must be paid to avoid overfitting. Overfitting is the development of a model so specific to the training data that it can't be generalised to new data.[86] For most models, overfitting can be avoided by performing cross-validation. This is a process where a model is tested on an independent data set, ideally a separate sample from the training data.[68,87] In this review, only Anand et al. (2017)[55] stated using two independent groups, while McGrath et al. (2019)[63] had partially independent groups (six new participants, four from the training group). If it is not feasible to have two independent groups, other methods such as *k*-fold cross-validation can be used. In this case, the sample population is split into *k* groups (folds, e.g., 10-fold), and the model is trained on all groups except one, which is reserved for testing the model. This training and testing procedure is repeated so that each group acts as the test set. The results are then averaged. Leave-one-subject-out-cross-validation (LOOCV) is a popular form of *k*-fold where *k* is equal to the number of participants. The advantage of this method is that nearly all the data (all except one participant) is used for each training iteration, and thus produces relatively stable estimates. A total of five studies[48,52,60,61,68] performed a k-fold cross-validation with *k* ranging from 4–10, while six studies performed LOOCV.[48,49,58,64,65,69] Alternately, two studies[51,67] used a train/test split, where all data are randomly split into two groups: a training group and a testing group. The disadvantage of this method is that data from the same participant could be used to train and test the model, which may inflate accuracy estimates.

Additional considerations

It is important to note that many studies included in this review used small sample sizes (average = 12, range 3–34). It was common to have few participants performing a large number of repetitions of varying movements.[48,49,52,55-58,62,65,69,70] This increases the likelihood that the classification algorithm will not generalise well to other people or populations.[88] This is especially true when classifying sporting movements due to the variability in technique among athletes.[48,89] Training a classification algorithm on elite athletes may not generalise well to novice athletes and vice versa.[48] Therefore, caution must be used when interpreting results from studies that train their algorithm using small samples without varied skill levels. A detailed description of each movement may help to alleviate these issues, but this is not always presented. For example, Murray and Black[66] did not specify the different slot angles (i.e. over the top or side arm) that participants used when performing the throwing tasks in baseball. In addition, none of the tennis studies stipulated the type of serve used (i.e., a top spin or slice serve).

Avenues of future work

When developing a classification system, numerous factors need to be considered. These include: the number of IMUs; their specification; their placement location; and pre-processing decisions such as filtering, segmentation, feature engineering, and model selection. There are many combinations of these factors in the reviewed literature, but little formal comparison among them. This means that the optimal set of decisions to classify each movement type is unclear and will likely vary depending on the movement of interest. Given these findings, there are several recommendations for future research. Work that compares the minimum IMU specifications (e.g., sampling rate), sensor modality (e.g., accelerometer vs accelerometer + gyroscope), and IMU number and placement (e.g., single vs multiple IMUs; torso vs wrist placement) will assist researchers and practitioners with the selection of the most accurate and cost-effective device setup for their needs. The comparison of different data processing techniques is also warranted. No studies compared different filtering methods or whether time or frequency domain features were more important for classifying certain movements. A formal comparison of these decisions and presenting these results (as opposed to just presenting the most accurate model) will aid the progression of this field. Comparing deep learning frameworks to more traditional supervised learning methods is also warranted, given the promising results seen in other areas. Interestingly, as the success of a classification system in sport relies on the accuracy of classifying the individual, it is surprising that no study has looked at the effectiveness of individually-trained models (i.e., trained using one subject's data) relative to the generalisability of a single model trained on many subjects' data. This could be an effective approach in classifying upper body sporting movements, particularly those where technique can differ considerably between participants.

Conclusion

This systematic literature review examined upper body activity classification using an IMU in court or field-based sports. Both manually defined rule-based algorithms and machine learning approaches have been used to classify movements across several sports, with varying degrees of success. In addition to the choice of algorithm or model, a classification system includes the number of IMUs, their specifications and placement location, and numerous data processing decisions (e.g., data filtering, data segmentation, feature engineering and selection). Due to the limited number of studies that compare these decisions, no definitive approaches for classifying upper body movements can be recommended. However, this also presents an avenue for future work; formally comparing these decisions within the classification workflow will assist in identifying the optimal set of decisions to classify each movement of interest.

Chapter 3 - Cricket fast bowling detection in a training setting using an inertial measurement unit and machine learning.

Preface

The first part of delivering a system that can predict bowling workload is to classify a bowl versus other rotation arm movements seen in cricket (i.e., a throw). Without performing this first step, the intensity cannot be obtained as the algorithm would not know when a bowl occurred. The systematic literature review established that machine learning would likely provide the best results for measuring bowling workload. This is due to the complex biomechanical nature of a cricket bowl involving the whole body and how an outfield throw can mimic a bowl. However, the review found a lack of research detailing the best inertial measurement unit (IMU) specification, data processing technique, or machine learning algorithm for any given sport. The following chapter classified a cricket bowl using a standard IMU device (±16 G and ±2000 °/s) located on the upper back and compared the accuracy of several sampling frequencies and machine learning algorithms.

This paper was published in 2019 in the Journal of Sports Sciences, 37(11), 1220-1226.

Abstract

Purpose: Fast bowlers are at a high risk of overuse injuries. Specific bowling frequency ranges have negative or protective effects on fast bowlers. Inertial measurement units (IMUs) can quantify movements in sports; however, some consumer-grade sensors can be expensive for amateur athletes. As a large number of the world's population has access to an IMU (e.g., smartphones), a system that works on a range of different IMUs may increase the accessibility of automated workload monitoring in sports. Methods: Seventeen elite fast bowlers in a training setting were used to train and/or validate five machine learning models by bowling and performing fielding drills. The accuracy of machine learning models trained using data from all three bowling phases (pre-delivery, delivery and post-delivery) were compared to those trained using only the delivery phase at a sampling rate of 250 Hz. Next, models were trained using data down-sampled to 125 Hz, 50 Hz, and 25 Hz to mimic results from lower specification sensors. Results: Models trained using only the delivery phase showed similar accuracy (> 95%) to those trained using all three bowling phases. When delivery-phase data were down-sampled, the accuracy was maintained across all models and sampling frequencies (> 96%). The linear support vector machine had the shortest computational time among all models ($P < 0.001$). Conclusion: Machine learning can accurately classify fast bowling events from different sampling frequencies using data from just the delivery phase. No single machine learning model clearly outperformed the others.

Introduction

It has been reported that fast bowlers cover on average 22.6 km in a full day of cricket, which is between 20% to 80% more than other players.[90] During high-intensity bowling efforts, fast bowlers cover 1.8 to 7 times more distance and have at least 35% less recovery time than other positions.[10] It is no surprise that fast bowlers are the most injury-prone. Injury data from 2005 through 2014 at the state and national levels in Australia indicated that fast bowers had a 19.9% chance of injury each season, much higher than batsmen (7.2%), spin bowlers (7.2%) and wicketkeepers (5.1%).[2]

Researchers have linked acute and chronic bowling frequency to injury risk for fast bowlers across all ages,[4,8-15] with some bowling frequencies having adverse or protective effects. For example, Orchard et al. (2015a)[4] found that bowling greater than 50 overs in a match and a high previous season workload (> 400 overs) increases the chance of tendon injuries, whereas a high medium-term workload (> 150 overs in the last three months) was found to be protective. There can also be an increased risk of injury three to four weeks after acute overload, possibly due to damaging immature repair tissue.[12]

Despite the importance of monitoring acute and chronic bowling frequency, most age-group, club and elite bowlers do not undertake this consistently as it requires players and coaches to input and analyse data manually. A potential solution is to use a wearable inertial measurement unit (IMU) to automatically and accurately log bowling frequency.[62] Although IMUs are seen as a cost-effective monitoring tool for professional teams, consumer-grade sensors can be expensive (upwards of 500 USD), precluding their use by amateur athletes. As smartphone subscriptions are expected to rise to 8.3 billion in 2023,[54] IMUs are accessible to a large percentage of the world's population. Therefore, a solution that works on a range of different IMUs may increase the accessibility of automated workload monitoring. However, differences in IMU specifications may be problematic for identifying bowling motion.[71] For example, different sampling frequencies can lead to varying volumes of data passed to classification algorithms. Despite these limitations, recent advances in predictive modelling may help overcome these challenges. For example, a bowling detection algorithm based on back-foot contact and upper torso rotation movement greater than 500 degrees per second has been used.[62] Although specificity was high during training (98.1%), it decreased during a game situation (74%). An alternative approach is machine learning, where different algorithms allow a computer to learn patterns among data without being explicitly programmed. Inter-study results that have compared various machine learning models show that no single model is the best for event detection in sports.[72] Results may vary depending on the IMU sampling rate, how features are generated from the raw sensor data, and the type of activity being classified.[71]

A machine learning approach has not been used for event detection in cricket bowling. Therefore, this study aims to examine the accuracy of machine learning models for classifying bowling events by: (1)

comparing the accuracy of models trained using data from all three bowling phases (pre-delivery, delivery and post-delivery) versus just the delivery phase; and (2) determine the accuracy of each model when data are down-sampled to 125 Hz, 50 Hz, and 25 Hz to mimic results from sensors with different sampling frequencies.

Methods

Participants

Seventeen male fast bowlers (average age, 24 years) were recruited from the Auckland Cricket premier men's cricket competition. To be eligible, participants had to be over 16 years of age, classed as a fast or medium pace bowler (>115 km/h), and be fit and healthy at testing. Ethical approval was granted by the Auckland University of Technology Ethics Committee (16/383).

Design

This cross-sectional study consisted of testing session one and testing session two in a training setting.

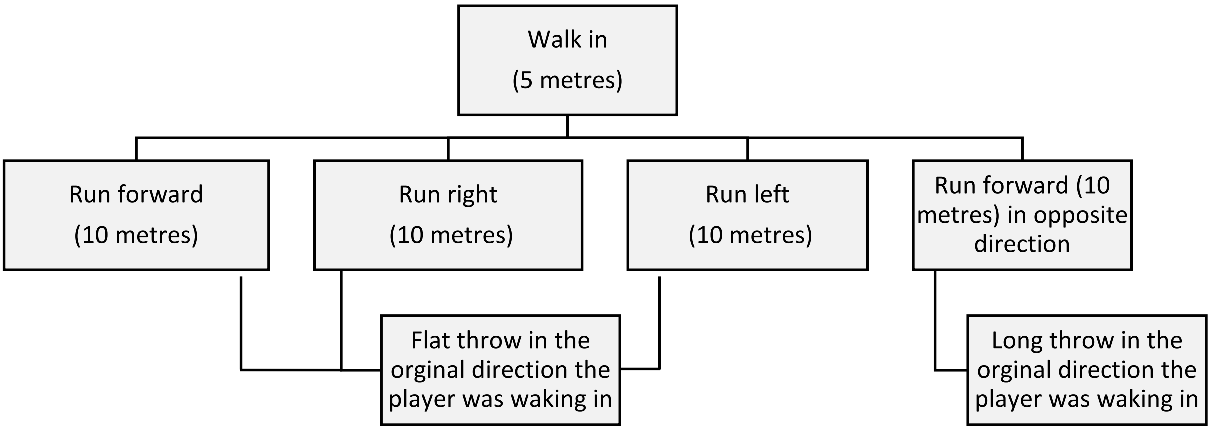
Testing session one

Data from 121 bowls were collected from 11 fast bowlers. In addition, non-bowling data consisted of 57 flat throws from a short run-up (5 metres) and 45 long throws from a long run-up (15 metres). Testing was conducted in artificial indoor or outdoor grass cricket nets. Participants were instructed to perform their regular warm-up and bowl and throw at their usual intensity.

Testing session two

Ten participants (six different bowlers from testing session one) were instructed to bowl 12 deliveries at a self-selected length at three different velocities (low = 75 to 84% of maximum, medium = 85% to 94% and high = 95% to 100%). Each participant's maximum bowling velocity was measured before the commencement of testing. Participants were asked to repeat deliveries that did not meet the velocity requirements, so all participants bowled a total of 401 deliveries. Ball velocity was measured using a Stalker radar gun (Radar Sales, Minneapolis, US), positioned behind the bowler at a sampling frequency of 250 Hz. The radar gun is a valid and reliable measure of velocity compared to a photocell system.[91] Non-bowling data included a fielding drill specific to a game situation, illustrated in Figure 3-1. This comprised of walking in, multidirectional running, followed by a flat or long throw. Participants repeated each throw three times and were asked to throw at their normal intensity. In total, there were 206 non-bowling events.

Figure 3-1: Explanation of the fielding drill during testing session two.



Equipment

One SABEL Sense IMU (SABEL Labs, Australia) consisting of a triaxial accelerometer (±16 G), gyroscope (±2000 °/s) and magnetometer (±1200 µT) (MPU-9150) was attached via a sports vest to each participant's upper posterior trunk between the shoulder blades. The sampling frequency of the IMU was 250 Hz. Pilot testing revealed that the ±16 G and ±2000 °/s were sufficient for capturing high-velocity rotational events when attached to the trunk. The IMU was orientated so that x was aligned with the medial-lateral axis, y aligned with the vertical axis and the z aligned with the anterior-posterior axis.

Data pre-processing and feature engineering

All IMU data were downloaded from the device using SABEL software (SABEL Labs, Australia). A fourth-order 1 Hz Butterworth low pass filter was used for baseline removal.[58,68] The magnitude of the gyroscope's x, y and z axis was calculated and peaks greater than 500 degrees per second were used to determine a possible bowling event. A 10-second event detection window was then used, which was further split into pre-delivery (5.5 seconds), delivery (1 second) and post-delivery (3.5 seconds) phases.

A total of 282 features from the time and frequency domains were computed using MATLAB (release 2017a, The MathWorks, Inc., MA, USA) from the individual axes and the magnitude of the accelerometer and gyroscope sensors. These features were similar to those used previously,[68] and included mean, standard deviation, maximum, minimum, skewness, kurtosis, amplitude, frequency, energy, the position of the minimum and maximum, and correlations between x, y and z axes. A complete list of features can be found in the Appendix (Table 3-5).

Model training and testing

Five machine learning models were evaluated, namely the random forest (RF), linear support vector machine (LSVM), polynomial SVM (PSVM), neural network (NN), and gradient boosting (XGB) algorithms. These models were chosen because they effectively classified other sporting activities with similar movement patterns to cricket bowling.[60,68] Data from testing session one were used to train and tune each model, while data from testing session two were treated as the validation set and used to evaluate model performance. All models were trained and tuned in R (R Core Team, Austria) using the 'train' function in the 'caret' package.

Due to the high dimensionality of the dataset (282 features), redundant features that were highly correlated with other features ($r > 0.95$) were removed, leaving the total number of features used for modelling at 223 (see Appendix 1). Data were then centred and scaled (normalised) for all models, apart from the RF, using the centre and scaling functions in the caret package. Optimal model tuning parameters for all models (e.g., polynomial degree for PSVM model, the number of trees in RF model) were determined using 10-fold cross-validation. The optimal values were chosen to maximise the receiver operating characteristic (ROC) metric. The final models were then used to predict bowling events on the test set (i.e., data from testing session two).

This process was initially undertaken using data sampled at 250 Hz for all three bowling phases (pre-delivery, delivery, and post-delivery) and then repeated using data from only the delivery phase. As models trained on just the delivery phase performed equally compared to all three phases, the delivery phase data were then down-sampled to 125 Hz, 50 Hz and 25 Hz to simulate sensors of different sampling rates.[71] The delivery phase data contained 75 features. The computer used to perform modelling consisted of an Intel® Xeon® CPU E3-1505M v6 @ 3.00 GHz, 32 GB of RAM and a 64-bit operating system.

Statistical analysis

The performance of each model was evaluated by computing the sensitivity, specificity, accuracy, and F-score of the predictions made on the test set. Sensitivity refers to the proportion of positive cases that are correctly identified (e.g., the proportion of bowling events identified as bowling), while specificity refers to the proportion of negative cases correctly identified (i.e., the proportion of non-bowling events identified as such). The accuracy is the proportion of correct predictions. The F-score[92] was computed as a supplementary measure because, unlike accuracy, it is not influenced by class distribution (our dataset contained more bowls than throws). To examine the contribution of each feature to model performance, the 'varImp' function in the 'caret' package was applied to each model. This function computes a ROC curve for each predictor and uses the area under the ROC curve as an indicator of feature importance. Models were compared using the accuracy metric and were deemed significantly different if the 95% confidence intervals did not overlap. The training times of each model were compared using an ANOVA and Tukey HSD post hoc test where applicable.

Results

Classification results for all the models when trained on data from all three bowling phases and data from only the delivery phase can be observed in Table 3-1. Overall, all models exhibited an accuracy > 95%. When all three phases were used for training, the SVM models had significantly higher accuracy than the other models (i.e., non-overlapping confidence intervals). There were no significant differences among models for the delivery phase or when comparing all phases with the delivery phase. The delivery phase models had shorter computational times than those that used all phases (mean difference = 49.9 seconds; P = 0.07).

Table 3-1: Classification results for models trained using all phases and just the delivery phase (250 Hz).

	Accuracy (95% CI)		Specificity		Sensitivity		F-score		Train time (sec)	
	All	Del	All	Del	All	Del	All	Del	All	Del
RF	0.95 (0.93–0.97)	0.95 (0.93–0.97)	1.00	1.00	0.93	0.93	0.93	0.93	84.6	43.8
LSVM	1.00 (0.99–1.00)	0.99 (0.98–1.00)	1.00	0.99	1.00	0.99	1.00	0.98	9.0	6.6
PSVM	1.00 (0.99–1.00)	0.99 (0.98–1.00)	1.00	0.99	1.00	0.99	1.00	0.99	201.0	88.8
XGB	0.96 (0.95–0.98)	0.98 (0.96–0.99)	1.00	0.99	0.95	0.97	0.95	0.97	49.8	34.8
NN	0.96 (0.94–0.97)	0.99 (0.97–0.99)	0.99	0.99	0.94	0.99	0.94	0.98	373.8	294.6

Key: All = data from all three phases; Del = data from just the delivery phase; LSVM = Linear support vector machine; NN = Neural network; PSVM = Polynomial SVM; RF = Random forest; XGB = Gradient boosting.

A confusion matrix detailing the number of correct and incorrect predictions for each model shown in Table 3-1 is presented in Table 3-2. The LSVM and PSVM correctly classified all bowling and non-bowling events when all three phases were used. More bowling events were misclassified compared to non-bowling events.

Table 3-2: Confusion matrix comparing data from three phases and the delivery phase (250 Hz).

			Predicted			
			All three phases		Delivery phase	
			Bowl	Non-Bowl	Bowl	Non-bowl
Observed	Bowl	RF	373	28	371	30
		LSVM	401	0	397	4
		PSVM	401	0	398	3
		XGB	379	22	390	11
		NN	378	23	395	6
	Non-bowl	RF	1	205	0	206
		LSVM	0	206	3	203
		PSVM	0	206	3	203
		XGB	0	206	2	204
		NN	2	204	2	204

Key: LSVM = Linear support vector machine; NN = Neural network; PSVM = Polynomial SVM; RF = Random forest; XGB = Gradient boosting.

The effects of down-sampling to 125 Hz, 50 Hz, and 25 Hz can be observed in Table 3-3. All models had an accuracy across all sampling rates of > 96%. Although these differences were not significant, a decline of approximately 1% was seen with each down-sampling step. There were no significant differences in accuracy among models within each of the three sampling rates. When comparing the average model training times across the four sampling frequencies, all models were significantly different from each other (all post hoc P < 0.001).

Table 3-3: Classification results for models trained using delivery phase data.

	Accuracy (95% CI)			Specificity			Sensitivity			F-score			Train time (sec)		
	125 Hz	50 Hz	25 Hz	125 Hz	50 Hz	25 Hz	125 Hz	50 Hz	25 Hz	125 Hz	50 Hz	25 Hz	125 Hz	50 Hz	25 Hz
RF	0.98 (0.96–0.99)	0.98 (0.97–0.99)	0.98 (0.97–0.99)	1.00	1.00	0.99	0.96	0.97	0.98	0.96	0.97	0.97	46.2	45.6	54.0
LSVM	0.99 (0.98–1.00)	0.98 (0.96–0.99)	0.96 (0.94–0.98)	0.99	0.96	0.92	0.99	0.98	0.98	0.99	0.96	0.94	5.4	4.8	4.8
PSVM	0.99 (0.97–0.99)	0.98 (0.96–0.99)	0.96 (0.94–0.98)	0.99	0.96	0.92	0.98	0.98	0.98	0.98	0.96	0.94	88.8	87.6	88.8
XGB	0.99 (0.98–1.0)	0.98 (0.97–0.99)	0.98 (0.97–0.99)	0.99	0.98	0.97	0.99	0.98	0.99	0.98	0.97	0.97	36.0	37.2	39.6
NN	0.99 (0.97–0.99)	0.98 (0.97–0.99)	0.96 (0.95–0.98)	0.99	0.95	0.92	0.99	1.00	0.99	0.98	0.97	0.95	292.2	297.6	295.8

Key: LSVM = Linear support vector machine; NN = Neural network; PSVM = Polynomial SVM; RF = Random forest; XGB = Gradient boosting.

The confusion matrix detailing the number of correct and incorrect predictions for each model shown in Table 3-3 is presented in Table 3-4. Bowling misclassification was similar across all sampling frequencies, while the number of misclassified non-bowling events increased as the sampling rate decreased.

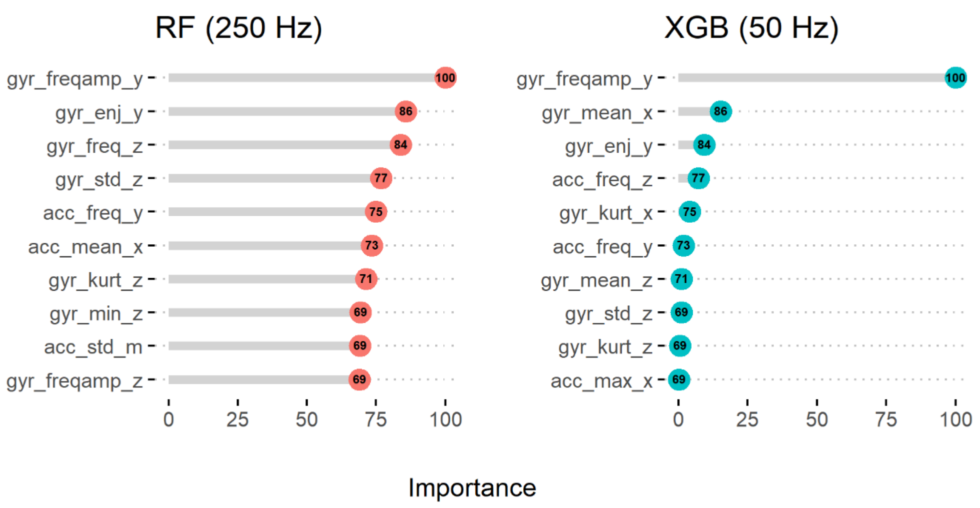
Table 3-4: Confusion matrix comparing data down-sampled to 150 Hz, 50 Hz, and 25 Hz.

			Predicted				
		125 Hz		50 Hz		25 Hz	
		Bowl	Non-Bowl	Bowl	Non-bowl	Bowl	Non-bowl
Observed: Bowl	RF	386	15	390	11	393	8
	LSVM	398	3	394	7	394	7
	PSVM	394	7	394	7	394	7
	XGB	397	4	393	8	395	6
	NN	396	5	399	2	395	6
Observed: Non-bowl	RF	0	206	0	206	3	203
	LSVM	3	203	8	198	16	190
	PSVM	2	204	8	198	16	190
	XGB	3	203	4	202	6	200
	NN	3	203	10	196	16	190

Key: LSVM = Linear support vector machine; NN = Neural network; PSVM = Polynomial SVM; RF = Random forest; XGB = Gradient boosting.

Figure 3-2 depicts important features for two models: the RF (250 Hz) and XGB (50 Hz) models trained using delivery phase data. These models were chosen because they performed well at their respective sampling frequencies, but all models had very similar feature importance rankings. The top 10 features are shown (out of 75) and are scaled to 100 to allow for comparison. The gyroscope's primary frequency amplitude in the y axis was the most important feature in both models. Gyroscope features, in general, were prominent among the top 10 in both models.

Figure 3-2: Top 10 feature importance plots illustrating RF (250 Hz) and XGB (50 Hz) models for the delivery phase.



Discussion

This study investigated whether machine learning models can accurately detect fast bowling events using a range of different IMU sampling frequencies. Although there were differences in accuracy among classifiers at certain sampling frequencies, these were small and, in most cases, non-significant. Every model showed high accuracy (> 95%), regardless of how many bowling phases were used to train the model or the sampling frequency used. These results suggest that when paired with machine learning algorithms, IMUs with different sampling frequencies have the potential to record bowling frequency, thereby forming the basis for monitoring bowling workload. This represents a significant progression from previous work and has practical applications in the sporting arena.

All tested machine learning models could classify bowling and non-bowling events with high accuracy, and no one model was consistently better than another. This corresponds with the 'no free lunch' theorem, which implies no one machine learning model will perform the best on all problems.[85] This is seen with the SVM models, which had the highest accuracy (and F-score) when trained using 250 Hz data but had the lowest accuracy when trained on down-sampled 25 Hz data. While SVMs performed well in the current study, two studies[51,68] found that SVMs were poor at classifying soccer and beach volleyball movements. The latter included movement patterns (shoulder extension) like a cricket bowl. It is known that small changes in data processing methodology, including data segmentation[80] and feature computation,[51,68] can positively or negatively affect model performance.

Our results illustrate that several features consistently showed higher importance (i.e., higher contribution to discerning bowling and non-bowling events) irrespective of the model used. In most cases, these features were derived from the gyroscope sensor. This may be because trunk acceleration patterns are similar during a bowl and a throw. However, this can differ significantly among bowlers due to the variability in body mass and bowling technique. The most important feature was the gyroscope's primary frequency amplitude on the y axis. This may indicate that a bowling action primarily comprises a single dominant frequency component from a long continuous rotation. In contrast, a throw includes a more complex rotation with a wider range of frequencies at shorter durations. Gyroscope energy in the y-axis was also an important feature, indicating that bowling causes a greater medial-lateral rotation than throwing (a greater duration and angular displacement in the torso is often associated with bowling due to a more extensive follow-through). The most prominent accelerometer feature was the primary frequency in the y-axis. During the delivery phase, this may equate to the back-foot landing followed by the front-foot landing, which might differ from a throw due to the longer duration between steps. Collectively, these features help to explain why data from only the delivery phase is needed to classify a bowl accurately. This has additional benefits, such as less computational expense and may reduce the chances of false positives. Our results show several models had higher specificity using delivery-phase data than data from all phases.

An important aspect of this study was that accuracy was maintained when data were down-sampled to simulate sensors with different sampling frequencies. Several other studies have used low sampling rate sensors[51,67,68] and have achieved modest results. Data sampled at a lower frequency has a higher chance of missing accelerometer and gyroscope peaks. Our results indicate that any loss of information did not significantly impact model performance. The accuracy observed in the current study is higher than the majority of previous studies that have examined event detection in sports.[48,49,51,65,67,68] However, a classification problem with only two possible values (i.e., a bowl or not) and testing under controlled conditions may have inflated the observed accuracy. Two recent studies[62,66] found a significant decrease in event detection accuracy during a game setting, so the random nature of actions performed during competition can likely confuse event detection models.

Practical applications

An IMU recording at different sampling frequencies combined with machine learning can accurately predict bowling frequency. This forms the basis for monitoring the bowling workload, which comprises bowling frequency and intensity. As research has found specific bowling frequency ranges can have a negative or protective effect on fast bowlers, this device could assist decision making to reduce injury rates and improve performance.

There are limitations to this study that the reader needs to be aware of when interpreting the findings. Firstly, of the 10 participants who took part in testing session two, which formed the validation dataset, four were also participants in testing session one. This means that the training and validation datasets were not completely independent, as training and testing on the same participants may have led to elevated accuracy. Another limitation of this work was assuming that down-sampling the data would be an accurate simulation of a low specification sensor, which may not be the case in reality.

Future work should also consider examining a broader range of participants, including children and females, as they have similar bowling injury rates to adult males.[93,94] As our study focused on senior cricketers, determining whether these models hold with less experienced participants will need to be established.[48] With bowling workload in mind, a future avenue of progression will be to deduce whether a measure of bowling intensity can be extracted from IMU data. When paired with bowling frequency, this would provide a comprehensive measure of workload. This is important as not all deliveries exert the same stress on the body.

Conclusions

This study investigated whether machine learning models could accurately predict fast bowling events using a range of different IMU sampling frequencies. All models exhibited high accuracy (>95%) regardless of the number of bowling phases used or the sampling frequency (25 Hz – 250 Hz). Features derived from the gyroscope sensor were significant contributors to the accuracy of each model: an important finding for future event-detection work. Future work will also need to evaluate the generalisation ability of these models in a game setting and test players of different ages, skill levels and bowling actions. The next step will be to examine whether a valid and reliable prediction of bowling intensity can be obtained using this IMU protocol, as this will better indicate the overall bowling workload and assist in preventing injury.

Appendix

Table 3-5: Features used.

Accelerometer				Gyroscope			
Feature	Phase one	Phase two	Phase three	Feature	Phase one	Phase two	Phase three
Mean	x, y, z, mag	x, y, z, ~~mag~~	x, y, z, mag	Mean	x, y, z, mag	x, y, z, mag	x, y, z, mag
SD	x, y, z, ~~mag~~	x, ~~y~~, ~~z~~, mag	x, y, z, ~~mag~~	SD	x, y, z, mag	~~x~~, y, z, mag	x, y, z, mag
Max	x, y, z, mag	x, y, z, mag	x, y, z, mag	Max	x, y, z, mag	x, y, z, mag	x, y, z, mag
Min	x, y, z, mag	x, y, z, mag	x, y, z, mag	Min	x, y, z, mag	x, y, z, mag	x, y, z, mag
Skewness	x, y, z, mag	x, y, z, mag	x, y, z, mag	Skewness	x, y, z, mag	x, y, z, ~~mag~~	x, y, z, mag
Kurtosis	x, y, z, mag	x, y, z, ~~mag~~	x, y, z, mag	Kurtosis	x, y, z, mag	x, y, z, mag	x, y, z, mag
Freq amp	~~x~~, ~~y~~, ~~z~~, ~~mag~~	~~x~~, ~~y~~, ~~z~~, ~~mag~~	~~x~~, ~~y~~, ~~z~~, ~~mag~~	Freq amp	x, y, z, ~~mag~~	x, y, z, mag	x, y, z, ~~mag~~
Freq	~~x~~, ~~y~~, ~~z~~, ~~mag~~	~~x~~, ~~y~~, ~~z~~, ~~mag~~	~~x~~, ~~y~~, ~~z~~, ~~mag~~	Freq	x, y, z, mag	x, y, z, mag	x, y, z, mag
Energy	~~x~~, ~~y~~, ~~z~~, ~~mag~~	~~x~~, ~~y~~, ~~z~~, ~~mag~~	~~x~~, ~~y~~, ~~z~~, ~~mag~~	Energy	x, y, z, ~~mag~~	x, y, z, mag	~~x~~, y, z, ~~mag~~
Pos of the max	x, y, z, mag	x, y, z, mag	x, y, z, mag	Pos of the max	x, y, z, mag	x, y, z, ~~mag~~	x, y, z, mag
Pos of the min	x, y, z, mag	x, y, z, mag	x, y, z, mag	Pos of the min	x, y, z, mag	x, y, z, mag	x, y, z, mag
Correlation	~~xy~~, ~~xz~~, ~~yz~~	~~xy~~, ~~xz~~, ~~yz~~	~~xy~~, ~~xz~~, ~~yz~~	Correlation	xy, xz, yz	xy, xz, yz	xy, xz, yz

Key: Features with a strikethrough (e.g., ~~mag~~) were removed due to being highly correlated with another feature ($r > 0.95$). Mag = magnitude of x, y and z.

Chapter 4 - Can an inertial measurement unit in combination with machine learning measure fast bowling speed and perceived intensity in cricket?

Preface

The previous chapter found that accurate delivery classification is possible using an inertial measurement unit (IMU) and machine learning. However, bowling volume was only half of the bowling workload equation as it does not give bowlers, coaches, and researchers insight into the intensity of each delivery. Therefore, the next chapter determined if the same IMU located on the upper back can accurately predict bowling intensity through ball release speed and the perceived intensity.

This paper was published in 2021 in the Journal of Sports Sciences, 39(12), 1402-1409.

Abstract

This study examined whether an inertial measurement unit (IMU), combined with machine learning, could accurately predict two indirect measures of bowling intensity – ball release speed and the perceived intensity zone. One IMU was attached to the thoracic back of 44 fast bowlers. Each participant bowled 36 deliveries at two different perceived intensity zones (low = 24 deliveries at 70 to 85% of maximum perceived bowling effort; high = 12 deliveries at 100% of maximum perceived bowling effort) in random order. IMU data (sampling rate = 250 Hz) was down-sampled to 125 Hz, 50 Hz, and 25 Hz to determine if sampling frequency affected model accuracy. Data were analysed using four machine learning models. A two-way repeated-measures ANOVA was used to compare the mean absolute error (MAE) and accuracy scores (separately) across the four models and four sampling frequencies. Gradient boosting models were the most consistent at measuring ball release speed (MAE = 3.61 km/h) and the perceived intensity zone (F-score = 88%) across all sampling frequencies. This method could be used to predict ball release speed and the perceived intensity zone, which may contribute to a better understanding of the overall bowling workload, which may help reduce injuries.

Introduction

Since the widespread implementation of T20 cricket (20 overs per innings) in 2005, match days in the elite cricketing calendar have increased by over 30% (2004 compared to 2015).[1,95] Although genetics, bowling technique and athlete conditioning impact injury rates, it is likely that increased bowling workloads explain why injuries have increased compared to the pre-T20 era.[1,96] In particular, fast bowlers have the highest incidence of injury during a season (20.6%) compared to wicketkeepers (4.7%), spin bowlers (6.7%) and batsmen (7.4%).[1] This is no surprise, considering fast bowlers cover between 20 to 80% more distance (22.6 km)[90], cover 1.8 to 7 times more distance at a higher intensity, and have 35% less recovery time[10] compared to other positions.

Research has linked acute and chronic bowling frequency (number of balls bowled) to injury risk for fast bowlers across all ages,[1,4,8-16] with some bowling frequencies having a negative or protective effect. However, bowling frequency alone may not provide an accurate picture of the overall bowling workload. A measure that could quantify bowling intensity may give a better understanding of the overall bowling workload as bowlers typically bowl at different intensities during training and competition, meaning that the stress exerted on bowlers will vary.[5,21-25] However, only a limited number of injury prevalence studies have included a measure of bowling intensity due to the cost of equipment, time constraints, and no consensus on the best methodology to assess it.[5,95]

Two studies that included a measure of intensity used a subjective recall of perceived exertion from an entire training session.[10,16] Although both studies showed a link with injury, obtaining these measures relies on consistent manual documentation after each training session and game: a limitation that is recognised in the literature.[5,17] To overcome this, inertial measurement units (IMUs) have been proposed as a user-friendly tool that can provide an objective indirect measurement of bowling intensity.[5,95]

Modern IMUs are wearable sensors that normally consist of a tri-axial accelerometer, gyroscope, and magnetometer. The accelerometer measures linear acceleration (measured in g-force), the gyroscope measures angular velocity (degrees per second), and the magnetometer measures the strength and direction of the local magnetic field. Of the three sensors used, the magnetometer is used the least frequently in activity classification and prediction.[97] The data obtained from these devices can be processed with either user-defined algorithms, machine learning techniques or a combination of the two.[97] Three studies have used an IMU to identify bowling deliveries and thus determine bowling frequency. McNamara et al. (2015)[62] applied a user-defined algorithm to data collected from a training setting (specificity = 98.1%) and a competition setting (specificity = 74%). While the other two studies used machine learning in a training and match setting with F-scores greater than 98%.[63,98]

Only two studies have used an IMU to indirectly quantify bowling intensity by estimating ball speed and analysing accelerometer data. The first study[99] used polynomial regression to determine the association between perceived intensity (at four different intensities), ball speed, and certain IMU outputs, including PlayerLoad. Although the PlayerLoad metric lacks standardisation in the published literature, [100] McNamara described it using the following equation:

$$\text{PlayerLoad} = \sqrt{\frac{(a_{y1}-a_{y-1})^2+(a_{x1}-a_{x-1})^2+(a_{z1}-a_{z-1})^2}{100}}$$

where x, y, and z are the three acceleration axes. Their results showed that perceived intensity (R = 0.83 ± 0.19) and relative ball speed (R = 0.82 ± 0.20) were associated with peak PlayerLoad. The last study[101] developed a method to predict ball speed using accelerometer data and a user-defined algorithm. Although this approach showed promise, they could not predict true speed as they did not have access to a radar gun.

IMUs have been used in other sports to estimate ball speed. Regression analysis was used to estimate baseball pitch speed with an accuracy of ± 10 km/h.[102] Four recent studies have used machine learning algorithms to predict peak ball speed from handball throws.[103-106] An error range of 23.4–30.4 km/h was reported using linear regression and a 250 Hz accelerometer.[106] The three other studies found promising results with a mean absolute error (MAE) of between 3.8 km/h and 6.3 km/h. All three studies used an IMU with a high sampling frequency (500 Hz–1125 Hz), which is higher than most consumer-grade IMUs. The extent to which lower sampling frequencies impact speed prediction is unknown; however, this may explain the large error observed by Skejø et al. (2018).[106] Interestingly, Van den Tillaar et al. (2020)[103] was the only study to use gyroscope data. Their analysis showed that one gyroscope feature (gyroscope amplitude in the z axis) was the most important feature for predicting speed. Therefore, the current study aims are: (1) to predict ball speed and the perceived intensity zone using IMU data and machine learning; (2) to determine if model accuracy is affected by IMU sampling frequency; and (3) to examine the most important IMU signal features for predicting these outcomes.

Methods

Participants

Forty-four male fast bowlers (mean age = 23 years) were recruited from the Auckland, Surrey and Sussex club and county (first-class) cricket competitions. Playing ability ranged from sub-elite (35 players) to elite (played first-class cricket; 9 players). To be eligible, participants had to be over 16 years of age, classed as a pace bowler (bowls regularly at 100% intensity from a long run-up), and be fit and healthy at the time of testing. Informed consent was obtained from all participants before testing, and ethical approval was granted by the Auckland University of Technology Ethics Committee (AUTEC Reference 19/47).

Design

This study used a cross-sectional design with data collected from a single testing session.

Testing session

Data from 1,584 bowls were collected from 44 fast bowlers (29 deliveries were later removed due to speed data not being captured). Participants performed their regular warm-up and were instructed to bowl a total of 36 deliveries at two different perceived intensity zones (low = 70 to 85% of maximum perceived bowling effort (24 deliveries), high = 100% of maximum perceived bowling effort (12 deliveries)) in random order at a self-determined bowling length. Ball release speed was measured using a Stalker II radar gun (Radar Sales, Minneapolis, US) positioned behind the bowler. The radar gun samples at 250 Hz and has been shown as a valid and reliable measure of speed compared to a photocell system.[91] Testing was conducted in either artificial indoor cricket nets or outdoor grass and artificial cricket nets.

Equipment

One SABELSense IMU (SABEL Labs, Australia) consisting of a triaxial accelerometer (±16 G), gyroscope (±2000 °/s) and magnetometer (±1200 μT) (MPU-9150) was attached via a sports vest to each participant's upper posterior trunk between the shoulder blades. The IMU was orientated so that x was aligned with the medial-lateral axis of the body, y was aligned with the vertical axis, and z was aligned with the anterior-posterior axis. Before each testing session, a factory calibration was performed for the gyroscope, and a standard six axes calibration was performed for the accelerometer. The sampling frequency was set at 250 Hz.

Data pre-processing and feature computation

All IMU data were downloaded from the device using the SABEL software (SABEL Labs, Australia). Bowling events were detected by identifying gyroscope peaks greater than 500 °/s. To help identify gyroscope peaks, a fourth-order Butterworth 25 Hz low pass filter was used. A 10-second event detection window was used, where each event was split into pre-delivery (starting 5.5 seconds before the gyroscope peak), delivery (0.5 seconds before and after the gyroscope peak) and post-delivery (finishing 3.5 seconds after the gyroscope peak) phases. Both bowling event detection rule and window type were the same as McGrath et al. (2019).[63]

For feature extraction, a fourth-order 1 Hz Butterworth high pass filter was used to remove the effects of gravity from each channel of the accelerometer. A total of 282 features from the time and frequency domains were computed using MATLAB (release 2019a, The MathWorks, Inc., MA, USA) from the individual axes and the magnitude of the accelerometer and gyroscope sensors. These features were similar to those used previously by Kautz et al. (2017)[68] and McGrath et al. (2019).[63] These included the mean, standard deviation, maximum, minimum, skewness, kurtosis, amplitude, frequency, energy, the position of the minimum and maximum, and correlations between the x, y and z axes (see Appendix, Table 4-4).

Raw IMU features were compared to relative features during preliminary analysis, calculated by dividing each feature by the corresponding feature from the participant's maximum speed delivery. The relative features were found to have superior results and were thus used for further analysis. Finally, the IMU data were down-sampled to 125 Hz, 50 Hz and 25 Hz to simulate sensors running at lower frequencies,[71] and the feature generation steps were repeated.

Model training and testing

For both ball release speed and the perceived intensity zone, four machine learning models were evaluated, namely random forest (RF), linear support vector machine (LSVM), polynomial SVM (PSVM), and gradient boosting (XGB). These were chosen because they have been previously effective at classifying cricket bowling frequency.[63,98] Model training was done separately on each sampling frequency (250 Hz, 125 Hz, 50 Hz, 25 Hz; total = 64 models).

Due to the high dimensionality of the dataset, redundant features that were highly correlated with other features (r > 0.95) were removed. Data were then centred and scaled for all models apart from RF. Optimal model hyperparameters were determined using 10-fold cross-validation. The values were chosen based on maximising the receiver operating characteristic metric for the perceived intensity zone models and optimising MAE for ball release speed models. The final models for ball release speed and the perceived intensity zone were evaluated using leave-one-participant-out cross-validation. All machine learning models were trained and tuned in R (R Core Team, Austria) using the train function in the 'caret' package.[107] To examine the contribution of each feature to model performance, the 'varImp' function in the 'caret' package was applied to each model.

Statistical analysis

The metrics MAE, mean absolute percentage error, and root mean square error were used to examine the performance of each model for predicting ball release speed. For analysing the perceived intensity zone, models were evaluated using accuracy, sensitivity, specificity, and F-score.[92] The results from each cross-validation iteration were used within a two-way repeated-measures ANOVA to compare the MAE and accuracy scores (separately) across the four models and four sampling frequencies. Model assumptions (i.e., no significant outliers, dependant variable normality, sphericity) were checked before fitting each model using the 'afex' R package. Both models violated the sphericity assumption and were thus adjusted using the Greenhouse-Geisser sphericity correction. Estimated means and pairwise contrasts (between models and sampling frequencies) were estimated using the 'emmeans' package, with multiple comparisons adjusted using the Holm method. A priori alpha of 0.05 was used for all analyses.

Studies that use machine learning for activity classification tasks generally do not calculate sample size because the number of participants is only one component of creating an accurate model. The volume of data collected from each participant and the heterogeneity of the sample is also important.[108] However, when comparing accuracy among the models and sampling frequencies, a post-hoc power analysis was performed in the G*Power software. Using a within-factors repeated measures ANOVA, a sample size of 44 subjects, a type I error rate of 0.05, a Greenhouse-Geisser sphericity correction of 0.67, and a correlation among repeated measurements of 0.7, there was an 83% power to detect a difference in accuracy among sampling frequencies (Cohen's $f = 0.17$, $\eta_p^2 = 0.027$) and >99% power to detect a difference between models ($f = 0.59$, $\eta_p^2 = 0.26$).

Results

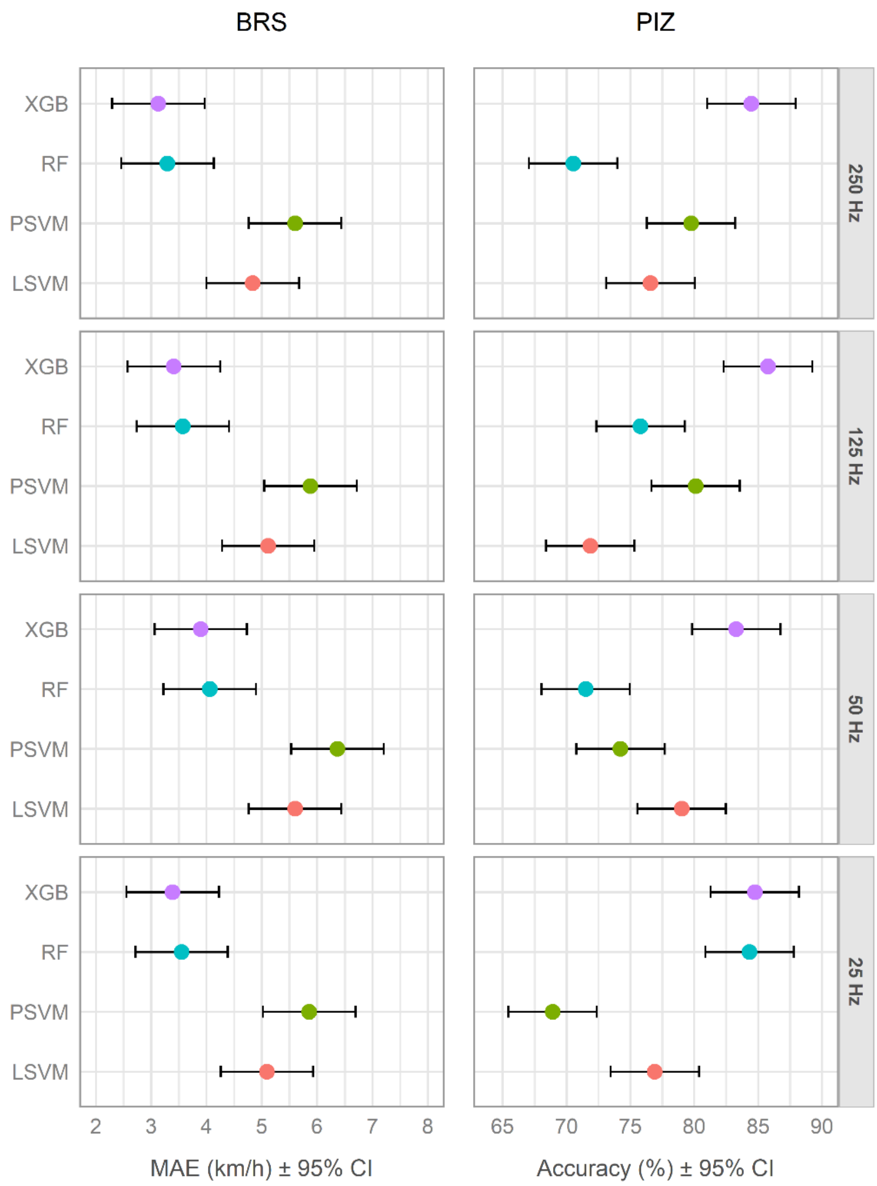
The MAE and mean accuracy (± 95% confidence intervals) for ball release speed and the perceived intensity zone, respectively, for each sampling frequency, are illustrated in Figure 4-1. Results for mean ball release speed error at different sampling frequencies are shown in Table 4-1. The pairwise contrast results for each model and the sampling frequency for ball release speed and the perceived intensity zone are presented in the Appendix (Tables 4-5 - 4-8). All models had an MAE of less than 6 km/h at 250 Hz. XGB and RF had the lowest MAE at 250 Hz with 3.52 and 3.46, respectively. XGB and RF had significantly lower MAE compared to PSVM at all sampling frequencies (all contrasts $p < 0.05$) and a lower MAE than LSVM at 25 Hz (both $p < 0.001$). When contrasting individual models across sampling frequencies, LSVM had significantly higher MAE at 25 Hz compared to 125 Hz (difference = 2.02 km/h, $p = 0.031$) and 250 Hz (2.22 km/h, $p = 0.016$). No other models were significantly different across the sampling frequencies.

Table 4-1: Mean ball speed prediction error at different sampling frequencies.

	MAE				MAPE				RMSE			
	250Hz	125Hz	50Hz	25Hz	250Hz	125Hz	50Hz	25Hz	250Hz	125Hz	50Hz	25Hz
RF	3.52	3.53	3.62	3.86	3.52	3.54	3.62	3.85	4.66	4.69	4.77	5.05
LSVM	4.37	4.58	5.17	6.57	4.36	4.51	6.03	4.97	5.97	6.46	7.37	6.90
PSVM	5.57	6.62	5.69	5.91	5.53	6.61	5.57	5.73	9.86	12.78	9.39	9.43
XGB	3.46	3.30	3.46	3.61	3.43	3.27	3.42	3.59	4.53	4.28	4.51	4.71

Key: LSVM = Linear support vector machine; MAE = Mean absolute error; MAPE = Mean absolute percentage error; PSVM = Polynomial SVM; RF = Random forest; RMSE = Root mean square error; XGB = Gradient boosting.

Figure 4-1: Model results for ball release speed (BRS) and the predicted perceived intensity zone (PIZ) across all sampling frequencies.



Key: LSVM = Linear support vector machine; MAE = Mean absolute error; PSVM = Polynomial SVM; RF = Random forest; XGB = Gradient boosting.

The predicted perceived intensity zone results at different sampling frequencies are compared in Table 4-2. At 125 Hz, XGB had significantly higher accuracy (86%) compared to every other model (all $p < 0.05$). At 250 Hz, 50 Hz and 25 Hz, XGB had a significantly higher accuracy than all but one model at each sampling frequency. Interestingly, some models performed better at lower sampling frequencies. RF had significantly higher (84%) accuracy at 25 Hz compared to all other frequencies (all $p < 0.001$), and LSVM had significantly better (79%) accuracy at 50 Hz compared to 125 Hz (72%, p =

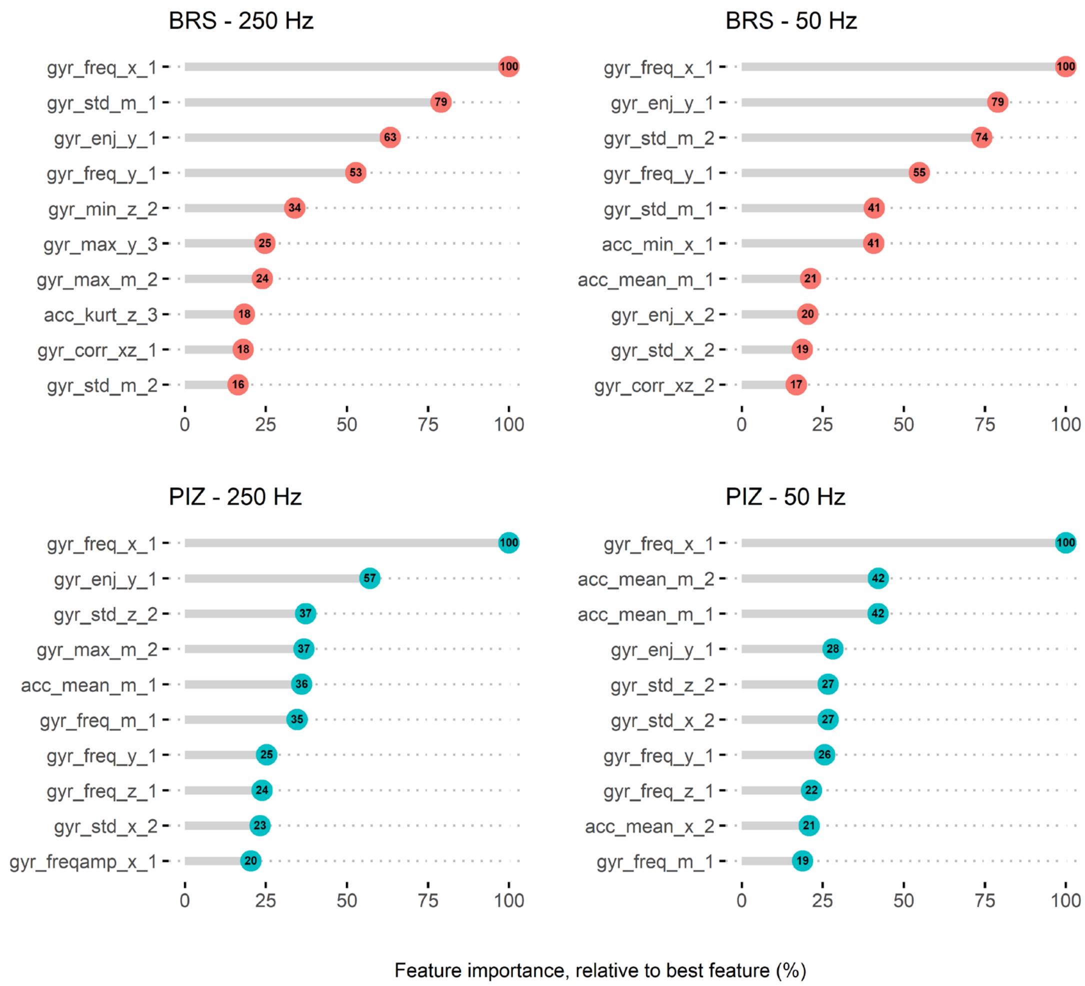
0.002). Finally, PSVM was the most variable, with significantly different accuracy (range = 69–80%) across all sampling frequencies (all $p < 0.05$), except 125–250 Hz.

Table 4-2: The predicted perceived bowling intensity zone classification results at different sampling frequencies.

	Accuracy				Sensitivity				Specificity				F-score			
	250Hz	125Hz	50Hz	25Hz	250Hz	125Hz	50Hz	25Hz	250Hz	125Hz	50Hz	25Hz	250Hz	125Hz	50Hz	25Hz
RF	70%	76%	71%	84%	77%	82%	79%	91%	58%	63%	57%	70%	78%	82%	79%	89%
LSVM	76%	72%	79%	77%	77%	72%	83%	78%	75%	72%	71%	73%	81%	77%	84%	82%
PSVM	80%	80%	74%	69%	89%	91%	93%	99%	61%	59%	37%	09%	85%	86%	83%	81%
XGB	84%	86%	83%	85%	86%	88%	86%	88%	81%	81%	77%	77%	88%	89%	87%	88%

Key: LSVM = Linear support vector machine; PSVM = Polynomial SVM; RF = Random forest; XGB = Gradient boosting.

Figure 4-2: Top 10 relative sensor features for XGB 250 Hz and XGB 50 Hz models for predicted perceived intensity zone (PIZ) and ball release speed (BRS).



Key: acc = Accelerometer; corr = Correlation; enj = Energy; gyr = Gyroscope; freq = Frequency; freqamp = Frequency amplitude; kurt = Kurtosis; mag = Magnitude of x, y and z axis; std = Standard deviation; x = x axis; y = y axis; z = z axis; 1 = Pre-delivery; 2 = Delivery; 3 = Post-delivery.

A confusion matrix, which details the correct and incorrect perceived intensity zone predictions, is shown in Table 4-3. Sensitivity was greater than 72% for all models. In particular, PSVM and XGB had all values above 85%. Only RF and LSVM had any considerable differences in sensitivity between sampling frequencies. RF showed the highest sensitivity at 25 Hz (91%) and the lowest at 250 Hz (77%), while LSVM at the highest at 50 Hz (83%) and the lowest at 125 Hz (72%). There were considerable changes in specificity between the sampling frequencies for RF and PSVM. RF had the largest specificity at 25 Hz (70%) and the lowest at 50 Hz and 250 Hz (57% and 58%, respectively). PSVM dramatically decreased specificity from a high of 61% at 250 Hz to a low of 9% at 25 Hz. Overall, XGB had the lowest MAE for the prediction of ball release speed and the best overall perceived intensity zone results compared to the other models for all sampling frequencies.

Table 4-3: Confusion matrix for the perceived intensity zone classification at different sampling frequencies.

			Predicted							
			250 Hz		125 Hz		50 Hz		25 Hz	
			low	high	low	high	low	high	low	high
Observed	low	RF	795	240	849	186	813	222	945	90
		LSVM	797	238	742	293	858	177	812	223
		PSVM	922	113	937	98	964	71	1026	9
		XGB	893	142	912	123	893	142	915	120
	high	RF	221	299	192	328	223	297	155	365
		LSVM	128	392	148	372	152	368	138	382
		PSVM	203	317	212	308	330	190	474	46
		XGB	101	419	100	420	120	400	119	401

Key: LSVM = Linear support vector machine; PSVM = Polynomial SVM; RF = Random forest; XGB = Gradient boosting.

The top 10 sensor features for all XGB models can be observed in Table 4-4. Features derived from the gyroscope during the pre-delivery and delivery phase dominate the list. The most important feature in seven of the eight models was the gyroscope's primary frequency in the x-axis during the pre-delivery (run-up) phase. The gyroscope’s primary frequency in the y-axis followed this during the pre-delivery phase, which was a top 10 feature in all XGB models apart from one. The relative importance of these features for the perceived intensity zone and ball release speed at 250 Hz are illustrated in Figure 4-2.

Discussion

This study investigated whether an IMU located on the thoracic spine can accurately predict ball release speed and the perceived intensity zone. With a MAE of 3.61 km/h for ball release speed, and an F-score of 88% for the perceived intensity zone, XGB was the best performing model for both intensity measures, even when down-sampled. Although XGB was the most accurate, it should be noted that other models performed similarly at specific frequencies. In regards to measuring ball release speed, these results could only be compared to the three handball studies, where an improvement was seen compared to the results obtained by Gençoğlu and Gümüş (2020)[104] (MAE = 6.73 km/h) and Skejø et al. (2020)[105] (MAE = 4.75 km/h) and were similar to Van den Tillaar et al. (2020)[103] (MAE = 3.78 km/h). It is important to note that these studies used IMUs with a high sampling frequency between 500–1200 Hz. It appears that an IMU with a low sampling frequency (e.g., 25 Hz) can maintain adequate performance, as the results from the down-sampled data do not show a decreasing trend in accuracy. This may have important implications, as this could mean that readily available consumer-grade IMUs (e.g., smartphones and watches) may be able to predict bowling workload accurately.

Unlike the study by Gençoğlu and Gümüş (2020),[104] where MAE was similar in all models, performance varied considerably for both ball release speed and the perceived intensity zone. For example, RF worked well in determining ball release speed but struggled to predict the perceived intensity zone at higher frequencies (250 Hz, 125 Hz, and 50 Hz). This is consistent with the no free lunch theorem, which states that no machine learning model will perform best on all problems.[85] Changes in the way data are pre-processed, the number of features extracted, the amount of data supplied to the model, and the model tuning parameters can affect results.[97]

It is challenging to determine why XGB worked best in the current study. The XGB algorithm has been widely used in winning solutions for several high-profile machine learning and data mining challenges.[109] It consists of multiple decision trees which are fit sequentially, each of them aiming to explain the error resulting from the previous tree until no further improvements can be made.[109] The most important reason for its success is its scalability in all scenarios like classification and prediction problems.[109] Gençoğlu and Gümüş[104] was the only similar study to compare gradient boosting against other models (support vector machine and generalised linear models) but did not find any improvement. It is important to note that they did not use any gyroscope data, which may explain these differences. Due to a lack of studies investigating similar measures, it is difficult to determine why some models do not work well. More studies need to display results comparing multiple machine learning classifiers to examine trends.[97]

Features derived from the gyroscope during the pre-delivery and delivery phase dominate the feature importance list in Table 4. The two most important features for all XGB models were the gyroscope's primary frequencies in the x and y-axis during the pre-delivery (run-up) phase. This might give an indication of run-up speed from the bowler. This is consistent with Salter et al. (2007),[43] who found that run-up speed, the angular velocity of the bowling arm, and the vertical velocity of the non-bowling arm accounted for 80% of the within-bowler variation in ball release speed. Feature importance differed from Van den Tillaar et al. (2020),[103] with no similar top 10 features for measuring handball throw speed when using a RF model. This might highlight the difference in run-ups between both studies, as participants threw the handball from a standing, running, or jumping start. Features from the post-delivery phase (follow-through) appear less important, with only two of these features seen in the top 10 for two separate models.

Practical applications

An IMU recording at different sampling frequencies combined with machine learning can estimate ball release speed with a MAE of 3.61 km/h and the perceived intensity zone with an F-score of 88%. When combined with models that predict bowling frequency,[63,98] this suggests that an IMU could be used to better estimate bowling workload in a training setting, as not all deliveries exert the same level of stress to the bowler. As results were similar across all sampling frequencies, this has the potential to be applied to a range of readily available IMUs (e.g., smart devices), which will increase accessibility. It also has the potential to be automated, avoiding the need to manually record the number of bowls and the session's rating of perceived exertion. If there is a widespread use of such a device, researchers will not have to rely on bowling frequency as the only measure of bowling workload. This may improve understanding of the relationship between workload and injury and help inform the development of individualised bowling workload and return to play protocols. It could also enable players and coaches to monitor loading to help prevent injuries, improve performance, and assist coaches with player selection.

Limitations

There are several study limitations that the reader should be aware of when interpreting the findings. Firstly, data processing and model implementation were not done in real-time in conditions that reflect on-field use. Although the authors could not find any studies comparing real-time versus delayed model implementation, it would be less beneficial to the player or coach if they were not receiving immediate feedback. This is because real-time feedback has the potential to improve motivation, learning, and performance.[110] In addition, models were not implemented in game situations, where there are more random actions (e.g. throwing, catching) that could confuse the models.[62] Furthermore, as the optimal model used relative features, each participant must perform a maximum effort delivery before the model can be implemented.

Secondly, ball release speed and perceived bowling intensity may not represent the stress exerted on the bowler's body. Although studies have examined ground reaction forces (GRF) in fast bowlers, the authors could not find any studies examining the difference in GRF between maximal and submaximal effort between the same bowlers. Future studies should determine whether submaximal efforts reduce forces that contribute to injury in fast bowlers.[24]

Lastly, as the IMU was attached to the thoracic back, it is not completely reflective of distal limb kinematics. An IMU with this mounting position may miss information pertinent to ball speed. If future studies use a wrist-mounted IMU, they will need to consider the higher acceleration and angular velocity values, which may mean that IMUs with low sampling frequencies or measurement ranges might not be as accurate compared to the current study.[97]

Conclusion

This study investigated whether an IMU, combined with machine learning, can accurately predict ball release speed and the perceived intensity zone. Although model performance varied considerably, XGB had the most consistent results for predicting ball release speed and the perceived intensity zone at all sensor sampling frequencies. Performance did not change for most models when the data were down-sampled, indicating that IMUs with a sampling frequency of at least 25 Hz can predict ball release speed and the perceived intensity zone with similar accuracy to an IMU sampling at 250 Hz. Features derived from the gyroscope were the most important for predicting ball release speed and the perceived intensity zone. Future projects will need to test whether this methodology can work in match settings and examine the efficacy of real-time implementation to give immediate feedback to players or coaching staff.

Appendix

Table 4-4: Initial feature set.

	Accelerometer				Gyroscope		
Feature	Pre-delivery	Delivery	Post-delivery	Feature	Pre-delivery	Delivery	Post-delivery
Mean	x, y, z, mag	x, y, z, mag	x, y, z, mag	Mean	x, y, z, mag	x, y, z, mag	x, y, z, mag
Standard deviation	x, y, z, mag	x, y, z, mag	x, y, z, mag	Standard deviation	x, y, z, mag	x, y, z, mag	x, y, z, mag
Maximum	x, y, z, mag	x, y, z, mag	x, y, z, mag	Maximum	x, y, z, mag	x, y, z, mag	x, y, z, mag
Minimum	x, y, z, mag	x, y, z, mag	x, y, z, mag	Minimum	x, y, z, mag	x, y, z, mag	x, y, z, mag
Skewness	x, y, z, mag	x, y, z, mag	x, y, z, mag	Skewness	x, y, z, mag	x, y, z, mag	x, y, z, mag
Kurtosis	x, y, z, mag	x, y, z, mag	x, y, z, mag	Kurtosis	x, y, z, mag	x, y, z, mag	x, y, z, mag
Frequency amplitude	x, y, z, mag	x, y, z, mag	x, y, z, mag	Frequency amplitude	x, y, z, mag	x, y, z, mag	x, y, z, mag
Frequency	x, y, z, mag	x, y, z, mag	x, y, z, mag	Frequency	x, y, z, mag	x, y, z, mag	x, y, z, mag
Energy	x, y, z, mag	x, y, z, mag	x, y, z, mag	Energy	x, y, z, mag	x, y, z, mag	x, y, z, mag
Position of the maximum	x, y, z, mag	x, y, z, mag	x, y, z, mag	Position of the maximum	x, y, z, mag	x, y, z, mag	x, y, z, mag
Position of the minimum	x, y, z, mag	x, y, z, mag	x, y, z, mag	Position of the minimum	x, y, z, mag	x, y, z, mag	x, y, z, mag
Correlation	xy, xz, yz	xy, xz, yz	xy, xz, yz	Correlation	xy, xz, yz	xy, xz, yz	xy, xz, yz

Table 4-5: Pairwise contrast results for each model for ball release speed.

SF (Hz)	Contrast	Estimate	Confidence intervals (95%)	Adj. p.value
250	RF - LSVM	-0.85	-2.79, 1.08	0.650
	RF - PSVM	**-2.04**	**-3.97, -0.10**	**0.027**
	RF - XGB	0.05	-1.88, 1.98	0.948
	LSVM - PSVM	-1.18	-3.11, 0.75	0.423
	LSVM - XGB	0.90	-1.03, 2.83	0.650
	PSVM - XGB	**2.08**	**0.15, 4.02**	**0.027**
125	RF - LSVM	-1.04	-2.98, 0.89	0.307
	RF - PSVM	**-3.08**	**-5.01, -1.15**	**0.000**
	RF - XGB	0.22	-1.71, 2.16	0.760
	LSVM - PSVM	**-2.04**	**-3.97, -0.11**	**0.022**
	LSVM - XGB	1.27	-0.67, 3.20	0.250
	PSVM - XGB	**3.30**	**1.37, 5.24**	**0.000**
50	RF - LSVM	-1.54	-3.47, 0.39	0.105
	RF - PSVM	**-2.07**	**-4.00, -0.14**	**0.024**
	RF - XGB	0.14	-1.80, 2.07	0.942
	LSVM - PSVM	-0.53	-2.46, 1.41	0.942
	LSVM - XGB	1.68	-0.25, 3.61	0.087
	PSVM - XGB	**2.20**	**0.27, 4.14**	**0.016**
25	**RF - LSVM**	**-2.73**	**-4.66, -0.80**	**0.001**
	RF - PSVM	**-2.04**	**-3.98, -0.11**	**0.016**
	RF - XGB	0.25	-1.69, 2.18	0.736
	LSVM - PSVM	0.68	-1.25, 2.62	0.696
	LSVM - XGB	**2.98**	**1.04, 4.91**	**0.000**
	PSVM - XGB	**2.29**	**0.36, 4.22**	**0.007**

Key: LSVM = Linear support vector machine; PSVM = Polynomial SVM; RF = Random forest; SF = Sample frequency; XGB = Gradient boosting. Bold text indicates a significant difference ($p < 0.05$).

Table 4-6: Pairwise contrast results for each model for the predicted perceived intensity zone.

SF (Hz)	Contrast	Estimate	Confidence intervals (95%)	Adj. p.value
250	**RF - LSVM**	**-6.05**	**-11.71, -0.38**	**0.015**
	RF - PSVM	**-9.23**	**-14.89, -3.57**	**0.000**
	RF - XGB	**-13.95**	**-19.61, -8.29**	**0.000**
	LSVM - PSVM	-3.18	-8.85, 2.48	0.137
	LSVM - XGB	**-7.90**	**-13.56, -2.24**	**0.001**
	PSVM - XGB	-4.72	-10.38, 0.94	0.055
125	RF - LSVM	3.93	-1.73, 9.60	0.088
	RF - PSVM	-4.31	-9.97, 1.35	0.088
	RF - XGB	**-9.98**	**-15.64, -4.32**	**0.000**
	LSVM - PSVM	**-8.25**	**-13.91, -2.59**	**0.001**
	LSVM - XGB	**-13.91**	**-19.58, -8.25**	**0.000**
	PSVM - XGB	**-5.67**	**-11.33, -0.01**	**0.025**
50	**RF - LSVM**	**-7.51**	**-13.17, -1.85**	**0.002**
	RF - PSVM	-2.72	-8.38, 2.94	0.203
	RF - XGB	**-11.78**	**-17.44, -6.12**	**0.000**
	LSVM - PSVM	4.79	-0.87, 10.45	0.077
	LSVM - XGB	-4.27	-9.93, 1.39	0.092
	PSVM - XGB	**-9.06**	**-14.72, -3.40**	**0.000**
25	**RF - LSVM**	**7.42**	**1.76, 13.09**	**0.001**
	RF - PSVM	**15.42**	**9.76, 21.08**	**0.000**
	RF - XGB	-0.41	-6.07, 5.26	0.849
	LSVM - PSVM	**8.00**	**2.34, 13.66**	**0.001**
	LSVM - XGB	**-7.83**	**-13.49, -2.17**	**0.001**
	PSVM - XGB	**-15.83**	**-21.49, -10.17**	**0.000**

Key: LSVM = Linear support vector machine; PSVM = Polynomial SVM; RF = Random forest; SF = Sample frequency; XGB = Gradient boosting. Bold text indicates a significant difference ($p < 0.05$).

Table 4-7: Pairwise contrast results between frequencies for each model for ball release speed.

SF (Hz)	Contrast (Hz)	Estimate	Confidence intervals (95%)	Adj. p.value
RF	250 - 125	-0.01	-1.96, 1.93	1.000
	250 - 50	-0.10	-2.05, 1.85	1.000
	250 - 25	-0.34	-2.29, 1.60	1.000
	125 - 50	-0.09	-2.03, 1.86	1.000
	125 - 25	-0.33	-2.28, 1.62	1.000
	50 - 25	-0.25	-2.19, 1.70	1.000
LSVM	250 - 125	-0.20	-2.15, 1.74	0.855
	250 - 50	-0.79	-2.73, 1.16	0.855
	250 - 25	**-2.22**	**-4.17, -0.27**	**0.016**
	125 - 50	-0.58	-2.53, 1.36	0.855
	125 - 25	**-2.02**	**-3.96, -0.07**	**0.031**
	50 - 25	-1.43	-3.38, 0.51	0.206
PSVM	250 - 125	-1.06	-3.00, 0.89	0.902
	250 - 50	-0.13	-2.08, 1.82	1.000
	250 - 25	-0.35	-2.30, 1.59	1.000
	125 - 50	0.93	-1.02, 2.87	1.000
	125 - 25	0.71	-1.24, 2.65	1.000
	50 - 25	-0.22	-2.17, 1.72	1.000
XGB	250 - 125	0.16	-1.79, 2.11	1.000
	250 - 50	-0.01	-1.96, 1.94	1.000
	250 - 25	-0.15	-2.09, 1.80	1.000
	125 - 50	-0.17	-2.12, 1.77	1.000
	125 - 25	-0.31	-2.25, 1.64	1.000
	50 - 25	-0.14	-2.08, 1.81	1.000

Key: LSVM = Linear support vector machine; PSVM = Polynomial SVM; RF = Random forest; SF = Sample frequency; XGB = Gradient boosting. Bold text indicates a significant difference ($p < 0.05$).

Table 4-8: Pairwise contrast results between frequencies for each model for the perceived intensity zone.

SF (Hz)	Contrast (Hz)	Estimate	Confidence intervals (95%)	Adj. p.value
RF	**250 - 125**	**-5.27**	**-10.48, -0.06**	**0.023**
	250 - 50	-0.98	-6.19, 4.23	0.619
	250 - 25	**-13.81**	**-19.02, -8.60**	**0.000**
	125 - 50	4.29	-0.92, 9.50	0.059
	125 - 25	**-8.54**	**-13.75, -3.33**	**0.000**
	50 - 25	**-12.83**	**-18.04, -7.62**	**0.000**
LSVM	250 - 125	4.71	-0.50, 9.92	0.068
	250 - 50	-2.44	-7.66, 2.77	0.644
	250 - 25	-0.34	-5.55, 4.87	0.863
	125 - 50	**-7.16**	**-12.37, -1.95**	**0.002**
	125 - 25	-5.05	-10.26, 0.16	0.053
	50 - 25	2.11	-3.11, 7.32	0.644
PSVM	250 - 125	-0.35	-5.56, 4.86	0.859
	250 - 50	**5.53**	**0.32, 10.74**	**0.015**
	250 - 25	**10.84**	**5.63, 16.05**	**0.000**
	125 - 50	**5.88**	**0.67, 11.09**	**0.012**
	125 - 25	**11.19**	**5.98, 16.40**	**0.000**
	50 - 25	**5.32**	**0.10, 10.53**	**0.015**
XGB	250 - 125	-1.30	-6.51, 3.91	1.000
	250 - 50	1.19	-4.02, 6.40	1.000
	250 - 25	-0.27	-5.48, 4.94	1.000
	125 - 50	2.49	-2.72, 7.70	1.000
	125 - 25	1.03	-4.18, 6.24	1.000
	50 - 25	-1.45	-6.67, 3.76	1.000

Key: LSVM = Linear support vector machine; PSVM = Polynomial SVM; RF = Random forest; SF = Sample frequency; XGB = Gradient boosting. Bold text indicates a significant difference ($p < 0.05$).

Chapter 5 - Quantifying cricket fast bowling volume, speed and perceived intensity zone using an Apple Watch and machine learning.

Preface

The previous two chapters found that an inertial measurement unit (IMU) located on the upper back can provide accurate bowling volume and bowling intensity estimates. Chapter 5 extends these results by using an IMU located on the wrist. In theory, gathering information from the wrist should give machine learning models more relevant data from the bowling hand, improving results – especially for measuring ball release speed. However, this method may come with complications of higher forces and rotational speeds, which means that most IMUs will reach their measurement thresholds. Therefore, a high-range, research grade IMU was also compared to a consumer-grade IMU (Apple Watch). The Apple Watch was chosen to determine if this method was feasible using consumer-grade wearables, potentially reaching a wider audience. As many bowlers do not feel comfortable with a device on their bowling wrist, the non-dominant wrist was also tested.

This paper was published in 2021 in the Journal of Sports Sciences, 40(3), 323-330.

Abstract

This study examined whether an inertial measurement unit (IMU) and machine learning models could accurately predict bowling volume, ball release speed, and the perceived intensity zone. Forty-four male pace bowlers wore a high measurement range, research-grade IMU (SABELSense) and a consumer-grade IMU (Apple Watch) on both wrists. Each participant bowled 36 deliveries, split into two different perceived intensity zones (low = 70 – 85% of maximum bowling effort, high = 100% of maximum bowling effort). Ball release speed was measured using a radar gun. Four machine learning models were compared. Gradient boosting models had the best results across all measures (bowling volume: F-score = 1.0; ball release speed: mean absolute error = 2.76 km/h; the perceived intensity zone: F-score = 0.92). There was no significant difference between the SABELSense and Apple Watch on the same hand when measuring bowling volume, ball release speed, and the perceived intensity zone. A significant improvement in classifying the perceived intensity zone was observed for IMUs located on the dominant wrist. For all measures, there was no added benefit of combining IMUs on the dominant and non-dominant wrists.

Introduction

Fast bowlers are at a heightened risk of injury compared to other positions in cricket.[1] The main modifiable risk factors include bowling technique and bowling volume (the number of bowls performed in training and during a match).[5,6] The relationship between bowling volume and injury have been extensively researched. Retrospective studies have found that too many or too few deliveries during a week, month, or year increases the chance of injury.[1,4,8-16]

Although it is hard to dispute the link between bowling volume and injury, researchers have stated that a more precise understanding of injury risk might be possible by monitoring bowling intensity.[5] This is because the stress or forces experienced by the bowler can vary for each delivery, and across training and match settings. However, only two studies have included a measure of bowling intensity.[10,16] Although both studies found a link with injury, the authors did not directly measure the bowling spell intensity, as the whole training session's rate of perceived exertion was divided by bowling volume. Bowling intensity is often not recorded by researchers, players, and coaching staff because of the difficulty in capturing this information reliably and no accepted definition or measure of bowling intensity.[5,17]

If bowling volume and intensity can be recorded effortlessly with minimal equipment outlay, this may inform a player's training decisions, reduce injury rates, and improve performance – particularly when recorded over a day, month, year, or multiple seasons. Researchers could also use this information to examine more precise relationships between bowling volume, intensity, and injury. A possible practical solution to measure bowling volume and intensity is to use an inertial measurement unit (IMU). An IMU usually consists of an accelerometer, gyroscope, and magnetometer. An accelerometer measures linear acceleration (measured in g-force), the gyroscope measures angular velocity (degrees per second), while the magnetometer measures the strength and direction of the local magnetic field. IMU's have a low relative cost and are accessible to most of the world's population through smart devices (i.e., smartphones and smartwatches).[97] This is important as most of the cricketing population lives in developing nations.

Recently, IMUs located on the cervical or thoracic spine have been accurate when measuring bowling volume, ball release speed, and the perceived intensity zone (i.e., 70 to 85%, and 100% of perceived max intensity).[62,63,98,111] McNamara et al. (2015)[62] used a user-defined algorithm to predict bowling volume and found a sensitivity of 99% and a specificity of 98.1% during training. However, specificity significantly decreased to 74% during a match situation. This may be due to the unstructured nature of actions performed during competition. Jowitt et al. (2020)[98] and McGrath et al. (2019)[63] used a machine learning approach to quantify bowling volume. In a training setting, both studies found an F-score above 98%. Interestingly, Jowitt et al. (2020)[98] were also able to show a high F-score (99%) during competition.

Recently, McGrath et al. (2021)[111] used an IMU located on the lower cervical spine to quantify bowling intensity by predicting ball release speed and the perceived intensity zone. Results showed a mean absolute error (MAE) of 3.61 km/h for ball release speed and an F-score of 88% for the perceived intensity zone. Although these results are promising, an IMU located on the spine does not consider arm speed, which is closely related to ball release speed.[43,112] Therefore, a wrist-based IMU may provide machine learning models with more relevant information to improve accuracy. However, a challenge of placing an IMU on a wrist is that there will be higher acceleration and angular velocity measurements compared with an IMU located on the trunk.[97] This may mean that an IMU with a higher sampling frequency and sensor measurement range might be needed to accurately predict bowling volume and intensity through a wrist-based IMU. In addition, some bowlers do not like to wear anything on their bowling wrist and may only be comfortable with an IMU attached to their non-dominant wrist. Given these challenges, the aims of this study are: 1) to investigate whether a high measurement range IMU (±100 g) located on the dominant and non-dominant wrists can accurately predict bowling volume, ball release speed, and the perceived intensity zone; and 2) to determine if a consumer-grade device with a lower measurement range (an Apple Watch; ±16 g) can predict the same metrics.

Methods

Participants

Forty-four male pace bowlers from the Auckland, Surrey, and Sussex club and county cricket competitions were recruited. To be eligible, participants needed to be 16 years or older, free of injury at testing and classed as a pace bowler. Informed consent from a legal guardian was required for any players aged 16 years. Written informed consent was obtained from all participants before the commencement of the study. The mean age was 23 years (SD ± 6), and playing ability ranged from sub-elite (35 players) to elite (having played first-class cricket, nine players). The Auckland University of Technology Ethics Committee (AUTEC Reference 19/47).

Design

This study used a cross-sectional design with data collected from a single testing session.

Testing session

After completing their regular warm-up, each participant bowled six practice bowls at low intensity (70 to 85% of maximum perceived bowling effort) with verbal feedback on bowling speed for familiarisation. They were then instructed to perform two tasks. Task one included bowling 36 deliveries at two different perceived intensity zones (low = 24 deliveries, high = 12 deliveries at 100% of maximum perceived bowling effort) in random order, at a self-determined run-up length, bowling length and bowling line. Task two required participants to perform 12 throws. The first six throws involved participants walking in five metres, running 10 metres to a stationary ball, and throwing a flat throw at 100% intensity over a 30-metre distance. The last six throws involved the same steps, but instead of picking up a stationary ball, participants were given a 'pop' pass where they would catch the ball at waist height before throwing it. The throwing task was designed to be a "worst-case scenario" as it closely resembled a delivery. Testing was conducted in either artificial indoor cricket nets or outdoor grass and artificial cricket nets.

Equipment

Two SABELSense IMUs (SABEL Labs, Australia) were attached to the posterior part of the dominant and non-dominant wrist (slightly proximal to where a regular watch would be positioned). Each SABELSense IMU consisted of a low measurement range triaxial accelerometer (±16 g), gyroscope (±2000 °/s) and magnetometer (±1200 µT) (MPU-9150) and a high measurement range accelerometer (±100 g). The sampling frequency was 250 Hz for both low and high measurement range IMUs. In addition, two Apple Watches (Series 4) consisting of a triaxial accelerometer (±32 g) and gyroscope (±2000 °/s) were attached to each participant's dominant and non-dominant wrist (directly distal to the SABELSense, please see Image 1). The Apple Watch has a sampling frequency of 100 Hz.

Image 5-1: IMU setup and location.

Ball release speed was quantified using a speed radar gun (Stalker II radar gun; Radar Sales, Minneapolis, US) positioned behind the bowler. The radar gun is a valid and reliable measure of speed compared to a photocell system.[91]

Data pre-processing and feature computation

For the SABELSense IMUs, raw data were downloaded from the devices using the SABEL software (SABEL Labs, Australia). For the Apple Watch, raw data were first saved to the device by an application called SensorLog (SensorLog, Germany) and then transferred to a computer as a CSV file. Event detection of bowls and throws were performed by calculating the magnitude of the gyroscope's x, y, and z axes before identifying peaks greater than 500 °/s.[63,111] An event detection window of 10 seconds was used to isolate each event. The window was broken into three phases, which included pre-delivery (starting 5.5 seconds before the gyroscope peak), delivery (0.5 seconds before and after the gyroscope peak) and post-delivery (finishing 3.5 seconds after the gyroscope peak). Features were then extracted from the time and frequency domains using MATLAB (release 2020b, The MathWorks, Inc., MA, USA). In total, 282 features were computed from the individual axes and the magnitude of the accelerometer and gyroscope sensors. These features were similar to those used previously by Kautz et al. (2017)[68] and McGrath et al. (2019)[63] and included the mean, standard deviation, maximum, minimum, skewness, kurtosis, amplitude, frequency, energy, the position of the minimum and maximum, as well correlations between x, y and z axes (please see Appendix, Table 5:4). Relative features were then calculated by dividing each feature by the corresponding feature from the participant's maximum speed delivery. This was shown to have superior results compared to raw data.[111]

Model training and testing

Four machine learning models were evaluated for each IMU when predicting bowling volume, ball release speed, and the perceived intensity zone, namely random forest (RF), linear support vector machine (LSVM), polynomial SVM (PSVM), and gradient boosting (XGB). These were chosen because they have been previously effective at classifying bowling volume, ball release speed and the perceived intensity zone with IMUs located on the thoracic or lumbar back.[63,98,111] In addition, the SABELSense dominant and SABELSense non-dominant features were analysed together to determine if combined data improved results.

All machine learning models were trained and tuned in R (R Core Team, Austria) using the 'train' function in the 'caret' package (Kuhn, 2008). Redundant features that were highly correlated with other features ($r > 0.95$) were removed due to the high dimensionality of the dataset. Data were then centred and scaled for all models apart from RF. Optimal model hyperparameters were determined using 10-fold cross-validation. The optimal values were chosen based on maximising the receiver operating characteristic metric for bowling volume and the perceived intensity zone models and optimising MAE for ball release speed models. The final models for both bowling volume, ball release speed and the perceived intensity zone were evaluated using leave-one-participant-out cross-validation. To examine the contribution of each feature to model performance, the 'varImp' function in the 'caret' package was applied to each model.

Statistical analysis

For analysing bowling volume and the perceived intensity zone, models were evaluated using accuracy, sensitivity, specificity, and F-score.[92] For predicting ball release speed, the metrics MAE, mean absolute percentage error, and root mean square error were used to examine the performance of each model. The results from each cross-validation iteration were used within a two-way repeated-measures ANOVA to compare the MAE and accuracy scores (separately) across the four models and four IMU positions. Model assumptions (i.e., no significant outliers, dependant variable normality, sphericity) were checked before fitting each model using the 'afex' R package. Both models violated the sphericity assumption and were thus adjusted using the Greenhouse-Geisser sphericity correction. Estimated means and pairwise contrasts (between models and between corresponding IMUs (e.g., SABELSense dominant and Apple Watch dominant) were estimated using the 'emmeans' package, with multiple comparisons adjusted using the Holm method. An a priori alpha of 0.05 was used for all analyses.

For our analyses comparing accuracy among the models and IMUs, a post hoc power analysis was performed in the G*Power software. Using a within-factors repeated measures ANOVA, our sample size of 44 subjects, a type I error rate of 0.05, a correlation among repeated measurements of 0.7, and a Greenhouse–Geisser sphericity correction (epsilon, ε), we had >99% power to detect a difference in accuracy among the four models (Cohen's ƒ = 0.57, $\eta p^2 = 0.247$, ε = 0.45) and >99% power to detect a difference among the five IMUs (ƒ = 0.41, $\eta p^2 = 0.145$, ε = 0.7).

Results

The results for predicting bowling volume, ball release speed, and the perceived intensity zone are shown in Tables 5-1, 5-2, and 5-3, respectively. The MAE and mean accuracy (± 95% confidence intervals) for ball release speed and the perceived intensity zone, respectively, are illustrated in Figure 5-1. Finally, confusion matrices for bowling volume and the perceived intensity zone are shown in Tables 5-4 and 5-5, respectively.

Table 5-1: Results for predicting bowling volume.

	Accuracy					Sensitivity					Specificity					F-score				
	SSD	AWD	SSND	AWND	SSD & SSND	SSD	AWD	SSND	AWND	SSD & SSND	SSD	AWD	SSND	AWND	SSD & SSND	SSD	AWD	SSND	AWND	SSD & SSND
RF	0.94	1.00 Sμ	^~ 0.96 μ	0.88	~ 0.96	0.96	1.00	0.97	0.94	0.98	0.88	1.00	0.93	0.73	0.90	0.96	1.00	0.97	0.92	0.98
LSVM	0.98	0.98 μ	~ 0.95	0.90	~ 0.98	0.99	0.99	0.99	0.97	1.00	0.94	0.96	0.83	0.71	0.95	0.98	0.99	0.97	0.94	0.99
PSVM	* 0.98 ¨³	0.98 μ	0.87	0.91 ¨	0.87	0.99	0.99	0.96	0.98	0.97	0.96	0.96	0.61	0.67	0.56	0.99	0.99	0.92	0.94	0.92
XGB	* 0.99	0.99	^~ 0.97	*^~ 0.97	~ 1.00	1.00	1.00	0.98	0.99	1.00	0.98	0.98	0.96	0.92	1.00	1.00	0.99	0.98	0.98	1.00

Key: AWD = Apple Watch dominant; AWND = Apple Watch non-dominant; LSVM = Linear support vector machine; PSVM = Polynomial SVM; RF = Random forest; SSD = SABELSense dominant; SSND = SABELSense non-dominant; XGB = Gradient boosting; * = Significantly better to RF; ^ = Significantly better to LSVM; ~ = Significantly better to PSVM; w = Significantly better to XGB; S = Significantly better to SSD; ¢ = Significantly better to AWD; ¨ = Significantly better to SSND; = μ Significantly better to AWND; ³ = Significantly better to SSD & SSND

Table 5-2: Mean ball release speed prediction error.

	MAE					MAPE					RMSE				
	SSD	AWD	SSND	AWND	SSD & SSND	SSD	AWD	SSND	AWND	SSD & SSND	SSD	AWD	SSND	AWND	SSD & SSND
RF	3.06	3.45	^ ~ 4.39	4.16	^ ~ 3.32	3.04	3.42	4.38	4.16	3.30	4.12	4.52	5.69	5.52	4.36
LSVM	3.82 ¨³	4.79	6.69	5.51	6.43	3.76	4.67	6.58	5.48	6.33	6.63	8.49	11.45	9.68	11.86
PSVM	3.82 ¨	4.75	6.97	5.59	5.18	3.76	4.65	6.83	5.58	5.14	6.68	8.08	12.36	8.34	9.40
XGB	2.76	3.31	^ ~ 4.12	^ ~ 3.87	^ ~ 2.84	2.74	3.26	4.08	3.87	2.81	3.61	4.37	5.50	5.13	3.72

Key: AWD = Apple Watch dominant; AWND = Apple Watch non-dominant; LSVM = Linear support vector machine; MAE = Mean absolute error; MAPE = Mean absolute percentage error; PSVM = Polynomial SVM; RF = Random forest; RMSE = Root mean square error; SSD = SABELSense dominant; SSND = SABELSense non-dominant; XGB = Gradient boosting. * = Significantly better to RF; ^ = Significantly better to LSVM; ~ = Significantly better to PSVM; w = Significantly better to XGB; S = Significantly better to SSD; ¢ = Significantly better to AWD; ¨ = Significantly better to SSND; = μ Significantly better to AWND; ³ = Significantly better to SSD & SSND

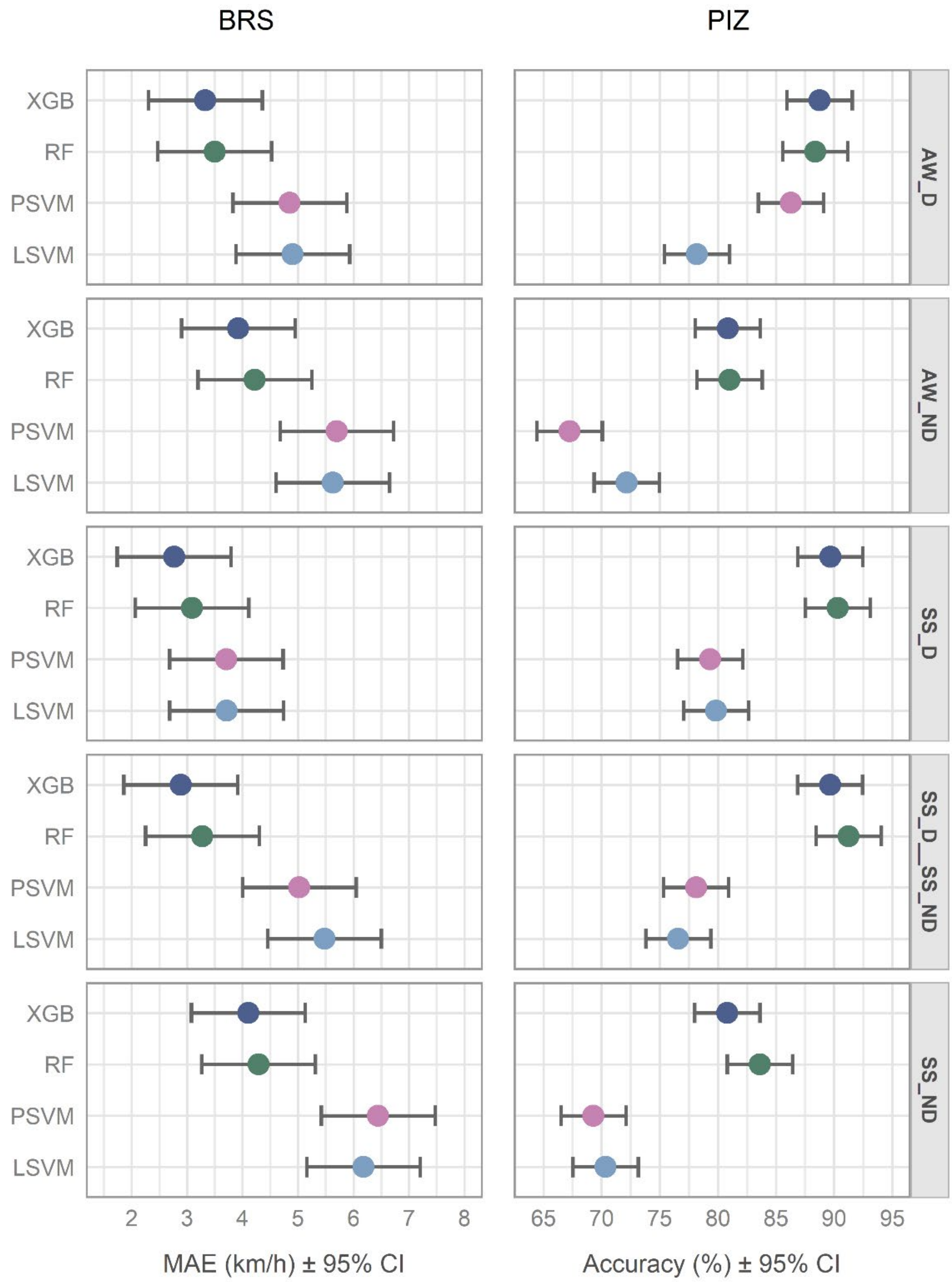
Table 5-3: The predicted perceived bowling intensity zone.

	Accuracy					Sensitivity					Specificity					F-score				
	SSD	AWD	SSND	AWND	SSD & SSND	SSD	AWD	SSND	AWND	SSD & SSND	SSD	AWD	SSND	AWND	SSD & SSND	SSD	AWD	SSND	AWND	SSD & SSND
RF	^ ~ 0.90 ¨	^ 0.88 μ	^ ~ 0.84	^ ~ 0.81	^ ~ 0.91	0.94	0.92	0.90	0.88	0.94	0.84	0.82	0.71	0.68	0.85	0.93	0.91	0.88	0.86	0.93
LSVM	0.80 ¨	0.79 μ	0.70	~ 0.73	0.77	0.82	0.78	0.71	0.74	0.76	0.75	0.81	0.68	0.86	0.77	0.84	0.83	0.76	0.78	0.81
PSVM	0.79 ¨	^ 0.87 Sμ	0.69	0.67	0.78	0.79	0.91	0.87	0.90	0.79	0.79	0.78	0.35	0.74	0.77	0.84	0.90	0.79	0.79	0.83
XGB	^ ~ 0.90 ¨	^ 0.89 μ	^ ~ 0.81	^ ~ 0.81	^ ~ 0.90	0.91	0.90	0.81	0.83	0.90	0.87	0.88	0.79	0.71	0.89	0.92	0.92	0.85	0.85	0.92

Key: AWD = Apple Watch dominant; AWND = Apple Watch non-dominant; LSVM = Linear support vector machine; PSVM = Polynomial SVM; RF = Random forest; SSD = SABELSense dominant; SSND = SABELSense non-dominant; XGB = Gradient boosting. * = Significantly better to RF; ^ = Significantly better to LSVM; ~ = Significantly better to PSVM; w = Significantly better to XGB; S = Significantly better to SSD; ¢ = Significantly better to AWD; ¨ = Significantly better to SSND; = μ Significantly better to AWND; ³ = Significantly better to SSD & SSND

For bowling volume, the XGB model had the best overall results (> 97% accuracy across all IMUs), and there were no significant differences between the IMUs when using XGB. For ball release speed, the XGB and RF models had the best overall results across all IMUs. In particular, XGB for the SABELSense dominant sensor had a MAE of ± 2.76 km/h. There were no significant differences between the IMUs for the RF and XGB models.

Figure 5-1: Model results for ball release speed (BRS) and the predicted perceived intensity zone (PIZ) for different IMU devices.



Key: LSVM = Linear support vector machine; MAE = Mean absolute error; PSVM = Polynomial SVM; RF = Random forest; XGB = Gradient boosting. The colour palette in this figure was produced using Manu R package (www.g-thomson.github.io/manu) and is inspired by the New Zealand Kererū.

Lastly, the perceived intensity zone had a similar trend, with RF and XGB having the best overall results (> 90% accuracy). These two models were significantly better than the LSVM and the PSVM (except for the Apple Watch dominant IMU). There was a significant improvement for the IMU placed on the dominant wrist (SABELSense dominant and Apple Watch dominant) compared to the corresponding

IMU placed on the non-dominant wrist. Apart from the PSVM, where Apple Watch dominant was significantly better than SABELSense dominant, there were no significant differences between IMUs. There was no added benefit for combining IMUs (SABELSense dominant & SABELSense non-dominant).

Discussion

XGB and the dominant wrist-based IMUs had the best overall results across all measures. If using an Apple Watch, which is a readily available consumer-grade device, and the XGB model, bowling volume can be predicted with 99% accuracy, ball release speed can be predicted with an error of ± 3.31 km/h, and the perceived intensity zone can be predicted with 89% accuracy. Little difference was observed between an Apple Watch and a high measurement range IMU (SABELSense), with only a slight non-significant improvement in ball release speed (± 2.76 km/h compared to ± 3.31 km/h). Although the acceleration likely exceeded the threshold limits on the Apple Watch, these results suggest that a high measurement range IMU is not necessary to measure the high acceleration and angular velocity seen at the wrist. This method of measuring bowling volume, ball release speed and the perceived intensity zone could be applied to most IMUs, making it more cost-effective and accessible to players. It was also evident that two IMUs on both the dominant and non-dominant wrist did not improve performance. This also has cost-saving benefits, with players only needing one IMU. Interestingly, the non-dominant wrist IMUs showed comparable results to the dominant wrist IMUs, with only significant differences in measuring the perceived intensity zone. Therefore, bowlers who do not like to wear anything on their dominant wrist while bowling can still have bowling volume, ball release speed and the perceived intensity zone measured with little change in accuracy.

The results obtained from this study are comparable to McGrath et al. (2021)[111] and McGrath et al. (2019),[63] which both used similar machine learning methods for an IMU attached to the lower cervical spine. There was only a slight improvement in the current study for ball release speed (MAE ± 2.76 km/h compared to ± 3.3 km/h) and the perceived intensity zone (F-score 93% compared to 89%) and no difference in predicting bowling volume. This was surprising, as a wrist-based IMU should be more reflective of distal limb kinematics.[97] Future studies could determine if a smartphone located on the lower cervical spine could produce similar results. As smart devices are more accessible each year in developing nations,[113] this could mean that an accurate measure of bowling volume and intensity could be available to most of the cricketing population. What was evident in all studies was that XGB was the best overall model for quantifying the variables of interest. Although it is unclear why this is the case, a reason for its success in many machine learning challenges is its scalability in classification and prediction problems.[109]

Practical applications

If the models created from this study can be implemented inside a smartwatch application, this could provide a cost-effective and easy way of measuring bowling volume, ball release speed and the perceived intensity zone. Practical applications exist for the player, coach, and researcher. For example, the bowling speed predictor could determine if a change of technique has been successful without setting up expensive equipment, therefore improving player performance. The application could track the player's bowling volume and intensity throughout the week, month, and season, guiding them when to bowl, to reduce the chances of injury. Players could also better understand how each perceived intensity zone affects ball release speed and fatigue levels, thus influencing strategies during bowling spells of different durations. Coaches could be aided in team selection by using all three measures to track a player's progress to determine if they are match fit or can bowl at a higher speed for a longer duration. Researchers could look at data from all three measures to establish better determinants of injury compared to the most commonly used workload measure, bowling volume.

Although higher model accuracy is always better, there is no defined accuracy threshold that a device must have to be classified as an effective tool for measuring bowling volume, ball release speed and the perceived intensity zone in cricket. However, the accuracy presented in this study may be enough for a bowler to gain meaningful feedback on performance. For example, a MAE of 3.3 km/h is likely to be less than what most batsmen could perceivably recognise between two deliveries.

Limitations

There are several limitations that the reader should be aware of. Firstly, the algorithms and IMUs were not tested in a game setting. McNamara et al. (2015)[62] found that random actions of bowlers during a match setting may have confused their user-defined algorithm leading to reduced accuracy. However, this may have been a limitation of their user-defined algorithm, as Jowitt et al. (2020)[98] found no reduction in performance during a game setting when using machine learning.

Secondly, the results obtained from the Apple Watch are not generalisable to all smartwatches. Smartwatches have a range of different IMUs built into them. In addition, depending on the software installed and applications opened, the proportion of power given to the IMU can vary, leading to variation in IMU performance.

Lastly, although bowling volume has been shown to correlate with injury, no studies have examined the relationship between ball release speed, perceived intensity zone or ground reaction forces on injury rates. This is likely from a lack of data because of the cost of specialised equipment and the expertise needed to measure these bowling workload parameters.

Conclusion

This study found that an IMU placed on either the dominant or non-dominant wrist, combined with machine learning, could accurately predict bowling volume, ball release speed and the perceived intensity zone. The XGB was the most accurate model across all measures. There was no significant difference between the high measurement range IMU (SABELSense) and the Apple Watch on both the dominant and non-dominant wrists. There was also no significant difference between the dominant wrist and non-dominant wrist for bowling volume and ball release speed; however, a significant difference was observed when measuring the perceived intensity zone. There was no benefit to combining IMUs (SABELSense dominant and SABELSense non-dominant) for all measures. If the model can be implemented in a smartwatch application, this has the potential to improve performance, aid in selection, and reduce injury.

Appendix

Table 5-4: Initial feature set.

Accelerometer				Gyroscope			
Feature	Pre-delivery	Delivery	Post-delivery	Feature	Pre-delivery	Delivery	Post-delivery
Mean	x, y, z, mag	x, y, z, mag	x, y, z, mag	Mean	x, y, z, mag	x, y, z, mag	x, y, z, mag
Standard deviation	x, y, z, mag	x, y, z, mag	x, y, z, mag	Standard deviation	x, y, z, mag	x, y, z, mag	x, y, z, mag
Maximum	x, y, z, mag	x, y, z, mag	x, y, z, mag	Maximum	x, y, z, mag	x, y, z, mag	x, y, z, mag
Minimum	x, y, z, mag	x, y, z, mag	x, y, z, mag	Minimum	x, y, z, mag	x, y, z, mag	x, y, z, mag
Skewness	x, y, z, mag	x, y, z, mag	x, y, z, mag	Skewness	x, y, z, mag	x, y, z, mag	x, y, z, mag
Kurtosis	x, y, z, mag	x, y, z, mag	x, y, z, mag	Kurtosis	x, y, z, mag	x, y, z, mag	x, y, z, mag
Frequency amplitude	x, y, z, mag	x, y, z, mag	x, y, z, mag	Frequency amplitude	x, y, z, mag	x, y, z, mag	x, y, z, mag
Frequency	x, y, z, mag	x, y, z, mag	x, y, z, mag	Frequency	x, y, z, mag	x, y, z, mag	x, y, z, mag
Energy	x, y, z, mag	x, y, z, mag	x, y, z, mag	Energy	x, y, z, mag	x, y, z, mag	x, y, z, mag
Position of the maximum	x, y, z, mag	x, y, z, mag	x, y, z, mag	Position of the maximum	x, y, z, mag	x, y, z, mag	x, y, z, mag
Position of the minimum	x, y, z, mag	x, y, z, mag	x, y, z, mag	Position of the minimum	x, y, z, mag	x, y, z, mag	x, y, z, mag
Correlation	xy, xz, yz	xy, xz, yz	xy, xz, yz	Correlation	xy, xz, yz	xy, xz, yz	xy, xz, yz

Table 5-5: Pairwise contrast results for each model for bowling volume (BV), ball release speed (BRS), and the predicted perceived intensity zone (PIZ).

		SSD			AWD			SSND			AWND			SSD & SSND		
	Contrast	Est	CI (95%)	Adj. p.value	Est	CI (95%)	Adj. p.value	Est	CI (95%)	Adj. p.value	Est	CI (95%)	Adj. p.value	Est	CI (95%)	Adj. p.value
BV	RF - LSVM	-2.55	-6.85, 1.75	0.341	1.83	-2.47, 6.13	1.000	**4.43**	**0.13, 8.74**	**0.013**	-3.04	-7.34, 1.26	0.143	-2.71	-7.01, 1.6	0.193
	RF - PSVM	**-5.13**	**-9.43, -0.82**	**0.008**	1.63	-2.68, 5.93	1.000	**9.70**	**5.39, 14**	**<0.001**	-3.23	-7.53, 1.08	0.143	**8.35**	**4.05, 12.65**	**<0.001**
	RF - XGB	**-6.33**	**-10.63, -2.03**	**0.001**	1.70	-2.6, 6.01	1.000	-1.39	-5.69, 2.92	0.394	**-9.61**	**-13.91, -5.31**	**<0.001**	-4.08	-8.38, 0.23	0.037
	LSVM - PSVM	-2.57	-6.88, 1.73	0.341	-0.20	-4.51, 4.1	1.000	**5.26**	**0.96, 9.57**	**0.004**	-0.19	-4.49, 4.12	0.909	**11.05**	**6.75, 15.36**	**<0.001**
	LSVM - XGB	-3.78	-8.08, 0.53	0.082	-0.12	-4.43, 4.18	1.000	**-5.82**	**-10.12, -1.52**	**0.001**	**-6.57**	**-10.87, -2.27**	**<0.001**	-1.37	-5.67, 2.93	0.399
	PSVM - XGB	-1.20	-5.51, 3.1	0.460	0.08	-4.22, 4.38	1.000	**-11.08**	**-15.38, -6.78**	**<0.001**	**-6.38**	**-10.69, -2.08**	**<0.001**	**-12.43**	**-16.73, -8.12**	**<0.001**
BRS	RF - LSVM	-0.62	-2.35, 1.1	1.000	-1.41	-3.13, 0.32	0.124	**-1.89**	**-3.62, -0.17**	**0.011**	-1.41	-3.13, 0.32	0.094	**-2.20**	**-3.93, -0.48**	**0.004**
	RF - PSVM	-0.62	-2.34, 1.1	1.000	-1.35	-3.08, 0.37	0.124	**-2.16**	**-3.88, -0.43**	**0.005**	-1.48	-3.2, 0.25	0.094	**-1.75**	**-3.47, -0.02**	**0.022**
	RF - XGB	0.32	-1.4, 2.04	1.000	0.17	-1.55, 1.89	1.000	0.18	-1.54, 1.91	1.000	0.30	-1.42, 2.02	1.000	0.39	-1.33, 2.11	0.971
	LSVM - PSVM	0.00	-1.72, 1.73	1.000	0.06	-1.67, 1.78	1.000	-0.26	-1.99, 1.46	1.000	-0.07	-1.8, 1.65	1.000	0.45	-1.27, 2.18	0.971
	LSVM - XGB	0.94	-0.78, 2.67	0.886	1.58	-0.15, 3.3	0.095	**2.08**	**0.35, 3.8**	**0.006**	**1.70**	**-0.02, 3.43**	**0.045**	**2.59**	**0.87, 4.31**	**<0.001**
	PSVM - XGB	0.94	-0.78, 2.66	0.886	1.52	-0.2, 3.24	0.099	**2.34**	**0.62, 4.06**	**0.002**	**1.78**	**0.05, 3.5**	**0.039**	**2.14**	**0.41, 3.86**	**0.004**
PIZ	RF - LSVM	**10.47**	**6.52, 14.41**	**<0.001**	**10.16**	**6.21, 14.11**	**<0.001**	**13.26**	**9.31, 17.21**	**<0.001**	**8.85**	**4.9, 12.79**	**<0.001**	**14.65**	**10.7, 18.6**	**<0.001**
	RF - PSVM	**10.98**	**7.04, 14.93**	**<0.001**	2.08	-1.86, 6.03	0.326	**14.30**	**10.36, 18.25**	**<0.001**	**13.74**	**9.8, 17.69**	**<0.001**	**13.11**	**9.16, 17.06**	**<0.001**
	RF - XGB	0.65	-3.3, 4.59	1.000	-0.38	-4.33, 3.56	0.797	2.81	-1.14, 6.76	0.120	0.15	-3.8, 4.09	0.922	1.61	-2.34, 5.55	0.564
	LSVM - PSVM	0.52	-3.43, 4.46	1.000	**-8.08**	**-12.03, -4.13**	**<0.001**	1.05	-2.9, 4.99	0.484	**4.90**	**0.95, 8.84**	**0.002**	-1.54	-5.49, 2.41	0.564
	LSVM - XGB	**-9.82**	**-13.77, -5.87**	**<0.001**	**-10.54**	**-14.49, -6.6**	**<0.001**	**-10.45**	**-14.4, -6.5**	**<0.001**	**-8.70**	**-12.65, -4.75**	**<0.001**	**-13.04**	**-16.99, -9.1**	**<0.001**
	PSVM - XGB	**-10.34**	**-14.28, -6.39**	**<0.001**	-2.47	-6.41, 1.48	0.296	**-11.50**	**-15.44, -7.55**	**<0.001**	**-13.60**	**-17.54, -9.65**	**<0.001**	**-11.51**	**-15.45, -7.56**	**<0.001**

Key: Adj = Adjusted; CI = Confidence intervals; Est = Estimate; LSVM = Linear support vector machine; PSVM = Polynomial SVM; RF = Random forest; SF = Sample frequency; XGB = Gradient boosting. Bold text indicates a significant difference ($p < 0.05$).

Table 5-6: Pairwise contrast results for each inertial measurement unit for bowling volume (BV), ball release speed (BRS), and the predicted perceived intensity zone (PIZ).

		SSD vs. AWD			SSD vs. SSND			SSD vs. SSD & SSD			AWD vs. AWND			SSND vs. AWND		
	Contrast	Est	CI (95%)	Adj. p.value	Est	CI (95%)	Adj. p.value	Est	CI (95%)	Adj. p.value	Est	CI (95%)	Adj. p.value	Est	CI (95%)	Adj. p.value
BV	RF	**-6.92**	**-11.69, -2.15**	**0.001**	-2.80	-7.57, 1.96	0.257	-2.81	-7.58, 1.96	0.257	**11.95**	**7.18, 16.72**	**<0.001**	**7.83**	**3.06, 12.6**	**<0.001**
	LSVM	-2.54	-7.31, 2.23	0.338	4.18	-0.59, 8.95	0.095	-2.96	-7.73, 1.81	0.326	**7.08**	**2.31, 11.85**	**0.001**	0.36	-4.41, 5.13	0.846
	PSVM	-0.17	-4.94, 4.6	0.926	**12.02**	**7.25, 16.79**	**<0.001**	**10.66**	**5.9, 15.43**	**<0.001**	**7.10**	**2.33, 11.87**	**<0.001**	**-5.09**	**-9.86, -0.32**	**0.012**
	XGB	1.11	-3.66, 5.88	1.000	2.14	-2.63, 6.91	1.000	-0.56	-5.33, 4.21	1.000	0.63	-4.13, 5.4	1.000	-0.39	-5.16, 4.38	1.000
BRS	RF	-0.41	-2.11, 1.29	1	-1.20	-2.9, 0.5	0.347	-0.19	-1.89, 1.52	1.000	-0.72	-2.43, 0.98	1.000	0.07	-1.64, 1.77	1.000
	LSVM	-1.19	-2.9, 0.51	0.211	**-2.47**	**-4.17, -0.77**	**0.001**	**-1.77**	**-3.47, -0.06**	**0.030**	-0.72	-2.43, 0.98	0.547	0.55	-1.15, 2.26	0.547
	PSVM	-1.14	-2.85, 0.56	0.251	**-2.74**	**-4.44, -1.03**	**<0.001**	-1.31	-3.02, 0.39	0.186	-0.85	-2.55, 0.85	0.395	0.75	-0.96, 2.45	0.395
	XGB	-0.56	-2.26, 1.14	1	-1.34	-3.04, 0.37	0.214	-0.12	-1.82, 1.58	1.000	-0.59	-2.3, 1.11	1.000	0.18	-1.52, 1.89	1.000
PIZ	RF	1.95	-3.21, 7.11	0.657	**6.70**	**1.54, 11.87**	**0.003**	-0.93	-6.09, 4.23	0.657	**7.37**	**2.2, 12.53**	**0.001**	2.61	-2.55, 7.77	0.573
	LSVM	1.65	-3.51, 6.81	0.733	**9.50**	**4.34, 14.66**	**<0.001**	3.26	-1.9, 8.42	0.310	**6.05**	**0.89, 11.21**	**0.010**	-1.80	-6.96, 3.36	0.733
	PSVM	**-6.95**	**-12.11, -1.79**	**0.002**	**10.03**	**4.87, 15.19**	**<0.001**	1.20	-3.96, 6.36	0.609	**19.03**	**13.86, 24.19**	**<0.001**	2.05	-3.11, 7.21	0.609
	XGB	0.92	-4.24, 6.08	1.000	**8.87**	**3.71, 14.03**	**<0.001**	0.03	-5.13, 5.19	1.000	**7.89**	**2.73, 13.06**	**<0.001**	-0.05	-5.21, 5.11	1.000

Key: Adj. = Adjusted using the Holm correction; CI = Confidence intervals; Est = Estimate; LSVM = Linear support vector machine; PSVM = Polynomial SVM; RF = Random forest; SF = Sample frequency; XGB = Gradient boosting. Bold text indicates a significant difference (p < 0.05).

Chapter 6 - The relationship between bowling intensity and ground reaction forces in cricket fast bowling.

Preface

Chapters 4 and 5 demonstrated that it was possible to predict several measures of bowling intensity, namely ball release speed and perceived intensity. However, these measures do not measure any forces the body is subjected to. Exposure to repeated high magnitude ground reaction forces (GRF) may be a significant cause of injury to fast bowlers. However, the expensive laboratory equipment required to measure this precluded research in this area. Although past studies in other sports have found that it is possible to predict GRF through an IMU and machine learning, it first needed to be established if GRF was a measure of bowling intensity. Therefore, this study examined whether GRF, measured by a force plate, was associated with bowling intensity measured through ball release speed and perceived bowling intensity. If GRF varied with different bowling efforts, it could influence bowling strategy. Chapters 6 and 7 were collaborative studies in conjunction with Loughborough University. The facilities at Loughborough University allowed the measurement of GRF during the delivery phase through a force plate embedded into the bowling crease.

This paper has been accepted in the Journal of Sports Sciences and is currently in press (as of May 2022).

Abstract

This study examined the relationship between perceived bowling intensity, ball release speed, and ground reaction force (measured by peak force, impulse, and loading rate) in male pace bowlers. Twenty participants each bowled 36 deliveries, split evenly across three perceived intensity zones: low = 70% of maximum perceived bowling effort, medium = 85%, and high = 100%. Peak force and loading rate were significantly different across the three perceived intensity zones in the horizontal and vertical directions (Cohen's *d* range = 0.14–0.45, $p < 0.01$). When ball release speed increased, peak force and loading rate also increased in the horizontal and vertical directions (η_p^2 = 0.04–0.18, $p < 0.01$). Lastly, bowling at submaximal intensities (i.e., low – medium) was associated with larger decreases in peak horizontal force (7.9–12.3% decrease), impulse (15.8–21.4%) and loading rate (7.4–12.7%) compared to decreases in ball release speed (5.4–8.3%). This may have implications for bowling strategies implemented during training and matches, particularly for preserving energy and reducing injury risk.

Introduction

A pace bowler has the highest injury prevalence compared to any other playing position in cricket.[1] Genetic susceptibility, technique, and bowling volume are the three main components linked to increased injury risk.[5] Bowling volume – the number of deliveries bowled in a session – has been used to track a player's bowling workload over time (e.g., match, week, season) to understand injury risk.[5,20] A bowling volume that is too high or too low over a week, month, or year has been linked to an increase in injury.[1,4,8,9,11-16,114] However, to gain a proper understanding of bowling workload, a metric needs to consider the intensity of each delivery, something that bowling volume does not. This is because not all deliveries are bowled at the same intensity; therefore, the stress exerted on the body is never constant.[5,99,111]

A potential method of measuring the intensity of a delivery is by recording the ground reaction forces (GRF) generated by the bowler during the delivery phase – usually from the front foot.[115] Typical GRF measurements include peak force, impulse, and loading rates in the vertical and horizontal (braking) axis. In contrast to recording bowling volume, measuring GRF in cricket involves a laboratory setting with a force plate embedded within a bowling crease. This specialised setup has caused a lack of longitudinal data to be collected, meaning that researchers have yet to confirm a link between GRF and injury. However, pace bowlers can generate large GRFs of up to nine times body weight during the delivery phase.[116] Therefore, it has been proposed that exposure to these repeated high magnitude ground impacts, combined with the spinal rotation observed during a delivery, may be a significant cause of injury, especially in the lower body.[116-118]

Only one study has investigated intra-athlete changes in GRF during a bowling spell.[30] Their results showed that different bowling lengths (i.e., short ball, length ball, or full ball) did not change GRF. However, as bowling intensity remained constant, no studies have investigated how a change in delivery intensity (particularly submaximal deliveries) affects GRF. This information could be valuable as previous research has shown that decreasing perceived intensity from 100% to 85% only slightly reduced ball release speed.[20,63,111,119,120] A potential reason ball release speed does not change to the same extent as perceived effort could be that slightly lower perceived effort improves timing and technique, therefore dampening the effects. This could explain why some bowlers, on occasion, would bowl a faster delivery at lower perceived intensities as opposed to 100%.[120] In addition, bowling at maximum speeds may not be a typical match bowl for some bowlers, which could compromise technique and timing and affect ball release speed. If GRF decreases similarly to perceived effort, it might cause pace bowlers to rethink their level of bowling effort during training and a match to potentially reduce the risk of injury without too much change in performance.[28,121]

The aims of this study were (1) to examine how vertical and horizontal GRF changes across different zones of perceived effort, (2) to determine the relationship between GRF and ball release speed, and

(3) to explore the relative decreases in GRF and ball release speed across different zones of perceived effort.

Methods

Participants

Twenty male pace bowlers from the Loughborough University cricket academy were recruited. Participants were 18 years or older and healthy at the time of testing. The mean age was 19.4 years (SD = 1.23), and playing ability ranged from sub-elite (19 players) to elite (1 player, having played first-class cricket). Eighteen players were right-handed bowlers. Ethics was granted by Loughborough University's Ethics Committee (reference 2020-2274-1855).

Using the PASS 15 software and a repeated measures analysis, we estimated that 15 subjects would allow us to detect a difference in GRF of 1.5 bodyweights (BW) among three perceived intensity zones, with 90% power and a type I error rate of 0.05. This was based on an F test with a single three-level within-subject factor (with estimated means for the three perceived intensity zones: low = 4.5 BW, medium = 5.25 BW, high = 6 BW), a between-subject standard deviation of 1.5 BW, a conservative autocorrelation among the repeated measurements of 0.5, with a compound symmetry structure. We obtained our estimated means from a previous systematic review, showing the mean peak vertical GRF in cricket was 5.8 BW, with a between-subject SD of 1.3 BW.[122] The sample size was increased to 20 participants to mitigate potential data loss.

Design

This study used a cross-sectional design with data collected from a single testing session. All data were collected on an indoor artificial pitch at the National Centre for Sport and Exercise Medicine (NCSEM) biomechanics laboratory at Loughborough University.

Testing session:

Each participant firstly had their height and body mass measured. After performing their regular warm-up, participants were instructed to bowl four practice deliveries at a low (70% of maximum) and medium (85%) perceived intensity zone. The instructions given to each player for the 70% perceived intensity zone was to bowl a "nice and easy delivery at 7 out of 10 intensity." For the 85% perceived intensity zone, instructions were given to keep the intensity in the middle of the low zone and high (100%) zone. A total of 36 deliveries split evenly between three perceived intensity zones were bowled by each participant. The recovery time between overs was three minutes. The three perceived intensity zones were bowled in random order, and participants were allowed to choose the line and length for each delivery.[30] Force plates were positioned at the popping crease to record the GRF of the front foot during the delivery phase. In total, 639 deliveries were recorded for analysis, with 81 deliveries omitted due to equipment failure (22 deliveries) or participants not landing with their front foot on the force place (59 deliveries).

Equipment

Ball release speed was calculated using an 18-camera retro-reflective motion analysis system (Vicon, MX13, OMG Plc, Oxford, UK) sampling at 300 Hz. Two reflective markers were placed on the ball. The velocity was calculated using the change in displacement from the first two frames after ball release, divided by the change in the time between frames.[123] GRF were measured using two Kistler force platforms (located next to each other to increase the area for a valid front foot contact) sampling at 1000 Hz (Type 9287B, Kistler AG, Switzerland).

Data processing

To determine the start and end of a delivery, the magnitude was first calculated by taking the square root of the sum of the x, y, and z-axis squares. The peak force of the magnitude was identified as the maximum value recorded during the front foot contact. A dynamic window for each delivery was created to determine the start and end of front foot contact. The window started at the first sample >= 35 N retrospectively from the peak magnitude and ended when it returned to <=35 N post peak magnitude. 35 N was chosen by visual inspection of the force plate data. Specifically, to avoid triggering false positives from the natural fluctuation in the force plate readings. The data from each delivery were then visualised to identify and remove errored trials (i.e., partial foot contacts). Raw data from the y-axis (horizontal) and z-axis (vertical) within each window were used to calculate peak force, impulse, and loading rates using custom code created in MATLAB R2021a. The impulse was calculated for the horizontal and vertical axis using trapezoidal numerical integration to determine the area under the curve. The loading rate was calculated by dividing the peak force by the time from initial foot contact to the time of the peak force.[116] Each measure of GRF was then normalised to each participant's body weight and expressed in bodyweights.[122]

Statistical analysis

A series of linear mixed-effects models were used to examine the relationship between measures of GRF and the perceived intensity zone (Aim 1) and ball release speed (Aim 2). Each measure of GRF (peak force, impulse, and loading rate) for each axis (horizontal and vertical) were treated as the dependent variables (separately). The perceived intensity zone (low – high) and ball release speed ($km \cdot h^{-1}$) were specified as fixed effects (separately), while the subject was added as a random intercept to account for the repeated measurements (~36 deliveries) for each participant.

To explore the relative decreases in GRF and ball release speed across different zones of perceived effort (Aim 3), the same model specification was used, except ball release speed and GRF measures were treated as the dependant variables (separately), while the perceived intensity zone was added as a fixed effect. For these models, the GRF and ball release speed measures were normalised relative to each participant's best delivery as a percentage of their maximum. This was done so the relative decrease of each variable from high to medium and medium to low could be compared.

All models were fitted using the *lme4* R package. Using the *performance* R package, each model's approximate residual normality and heteroscedasticity were confirmed by visualising the Q-Q and other residual plots. Model-estimated means and 95% confidence intervals were calculated using the *emmeans* R package. Pairwise contrasts among the three perceived intensity zones were adjusted using the Holm correction. The Cohen's *d* effect size for these contrasts was approximated using the *t* statistic ($d = 2t/\sqrt{df_{error}}$). The partial eta squared (η_p^2) effect size was also approximated from the *t* statistic in models where ball release speed (a continuous measure) was specified as a fixed effect. The level of significance for all analyses was set at $p < 0.05$.

Results

The characteristics of the participants are summarised in Table 6-1. As expected, there was an increase in mean ball release speed with a corresponding increase in the perceived intensity zone. Standard deviations for mean ball release speed were also similar across all the perceived intensity zones.

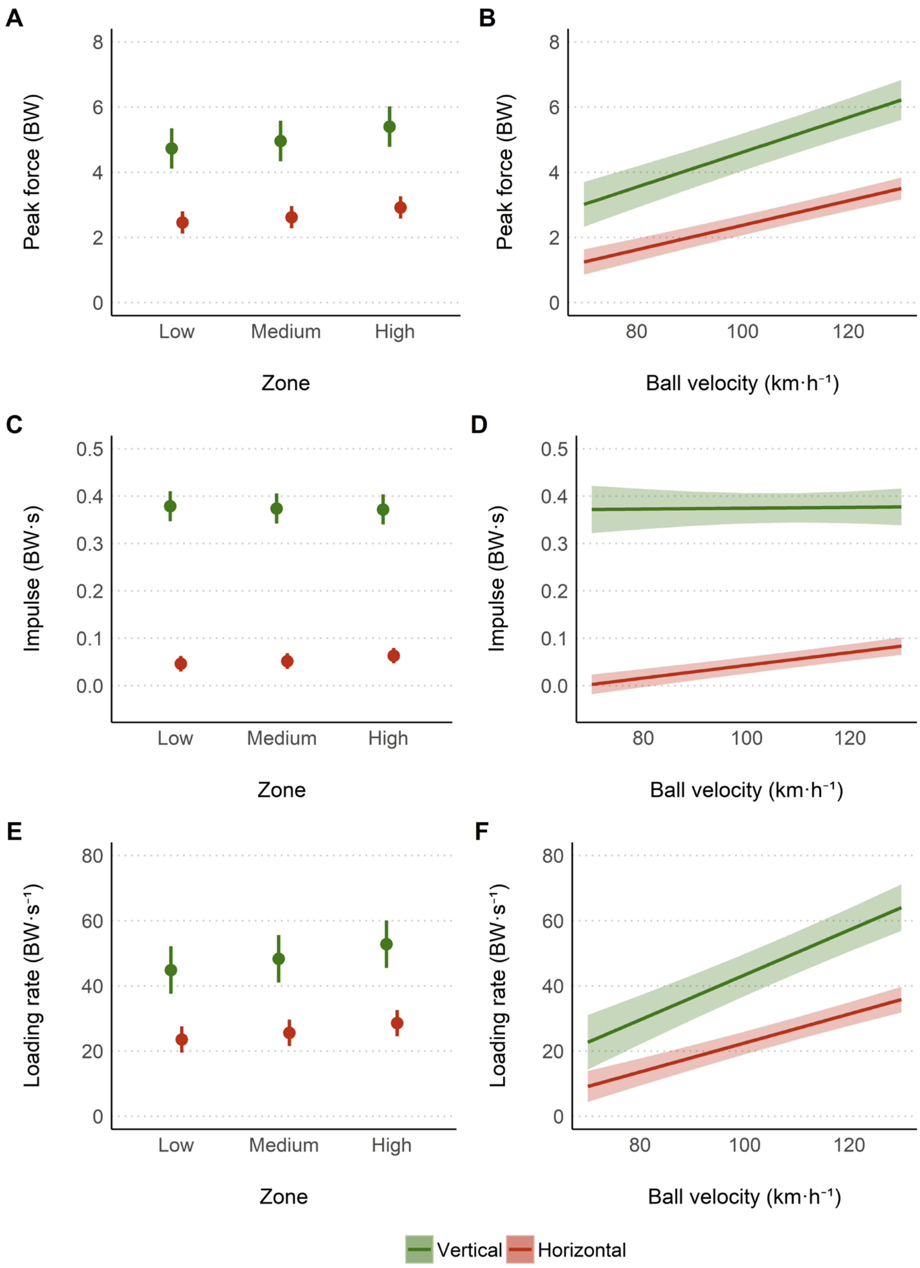
Table 6-1: Participants' physical characteristics and ball release speeds.

Variable	Zone	Mean	SD
Height (cm)		183.5	7.3
Body mass (kg)		78.5	10.5
Ball release speed (km·h^{-1})	Low	103.7	7.4
	Medium	107	6.7
	High	112.8	6.4
Ball release speed (% of maximum speed)	Low	89.2	5.0
	Medium	91.9	4.2
	High	97.3	1.8

The estimated means for peak force, impulse, and loading rate across the three perceived intensity zones and their relationship with ball release speed can be seen in Figure 6-1. The GRF on the vertical axis was greater than the horizontal axis by 60–63% for peak force, 144–153% for impulse, and 60–62% for loading rate. For all three GRF measures, there was a clear positive association with ball release speed, apart from impulse in the vertical direction. The numerical estimated means and confidence intervals for these figures can be found in the Appendix (Table 6-4).

The pairwise contrasts between each zone are presented in Table 6-2. Peak force and loading rate (for the vertical and horizontal axis) increased across all three zones (all $p < 0.01$, *d* range = 0.14–0.34). No significant difference was found for impulse in the vertical axis. In general, larger differences and effect sizes were observed between medium and high than low and medium (*d* range = 0.11–0.16 vs 0.19–0.29).

Figure 6-1: The relationship between bowling intensity (perceived bowling intensity and ball release speed) and GRF.



The relationship between (**A**) peak force and the perceived intensity zone, (**B**) peak force and ball release speed, (**C**) impulse and the perceived intensity zone, (**D**) impulse and ball release speed, (**E**) loading rate and the perceived intensity zone, and (**F**) loading rate and ball release speed. Values and error bars represent model-estimated means and 95% confidence intervals. The colour palette in this figure was produced using the *Manu* R package (www.g-thomson.github.io/manu) and is inspired by the New Zealand Kākāriki.

Table 6-2: Differences in GRF across the three perceived intensity zones.

		Horizontal				Vertical			
	Contrast	**Diff**	**95% CI**	**p**	***d***	**Diff**	**95% CI**	**p**	***d***
Peak force (BW)	Low – Med	-0.16	-0.24, -0.08	< .001	0.16	-0.23	-0.36, -0.10	< .001	0.14
	Low – High	-0.46	-0.54, -0.38	< .001	0.45	-0.67	-0.80, -0.54	< .001	0.41
	Med – High	-0.30	-0.38, -0.22	< .001	0.29	-0.44	-0.57, -0.31	< .001	0.27
Impulse (BW·s)	Low – Med	-0.01	-0.01, 0.00	0.006	0.11	0.00	-0.01, 0.02	0.955	0.03
	Low – High	-0.02	-0.02, -0.01	< .001	0.34	0.01	-0.01, 0.02	0.925	0.04
	Med – High	-0.01	-0.02, -0.01	< .001	0.23	0.00	-0.01, 0.02	0.955	0.01
Loading rate ($BW·s^{-1}$)	Low – Med	-2.04	-3.15, -0.93	< .001	0.15	-3.44	-5.26, -1.61	< .001	0.15
	Low – High	-5.02	-6.13, -3.90	< .001	0.36	-7.92	-9.76, -6.08	< .001	0.35
	Med – High	-2.98	-4.10, -1.85	< .001	0.21	-4.48	-6.34, -2.63	< .001	0.19

Key: CI = Confidence interval; d = Cohen's d; Diff = Difference; p = p-value. The degrees of freedom for all contrasts were all 600.

The relationship between GRF and ball release speed can be observed in Table 6-3. Apart from impulse in the vertical axis, every GRF measure significantly increased as ball release speed increased (all $p < 0.01$, η_p^2 range = 0.10–0.18). On the horizontal axis, for every 1 $km·h^{-1}$ increase in ball release speed, peak force increased by 3.76 BW_{100}, impulse by 0.13 $BW_{100}·s$, and loading rate by 44.45 $BW_{100}·s^{-1}$. On the vertical axis, for every 1 $km·h^{-1}$ increase in ball release speed, there was a significant increase in peak force by 5.34 BW_{100} and loading rate by 68.9 $BW_{100}·s^{-1}$. These relationships can be visualised in Figure 6-1.

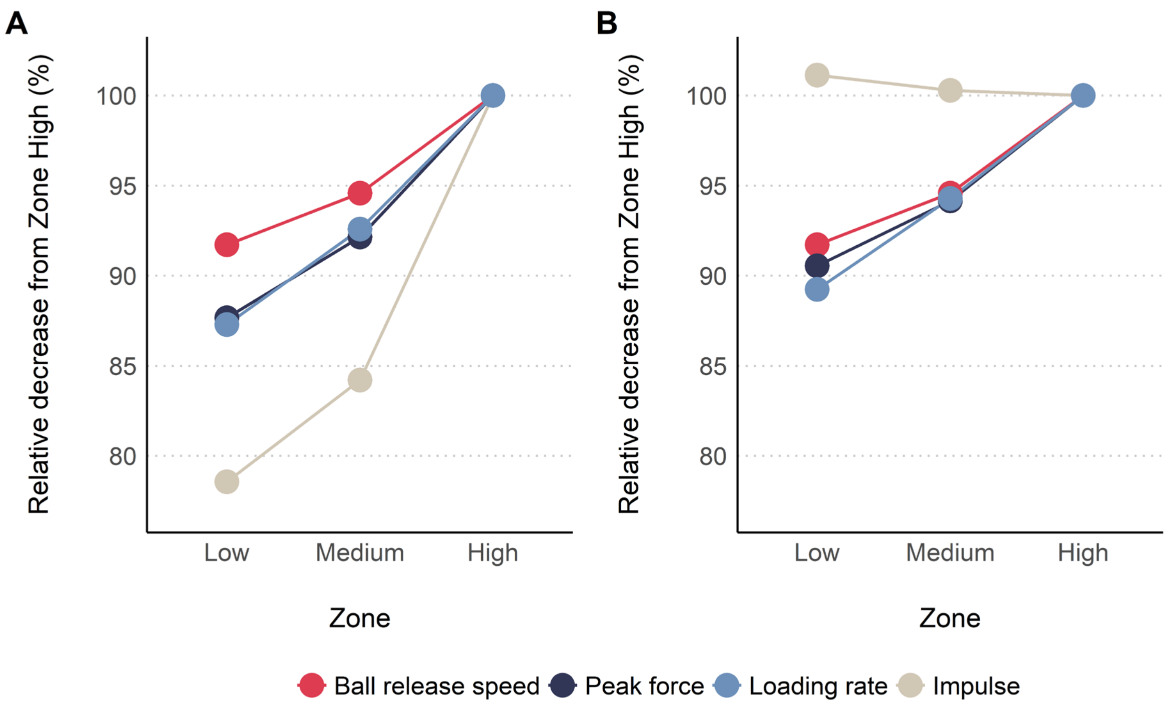
Table 6-3: The relationship between GRF and ball release speed.

	Horizontal				Vertical			
GRF	**Coefficient**	**95% CI**	**p**	**η_p^2**	**Coefficient**	**95% CI**	**p**	**η_p^2**
Peak force (BW_{100})	3.76	3.12, 4.40	< .001	0.18	5.34	4.31, 6.37	< .001	0.15
Impulse ($BW_{100}·s$)	0.13	0.10, 0.17	< .001	0.10	0.01	-0.10, 0.11	0.866	0.00
Loading rate ($BW_{100}·s^{-1}$)	44.45	35.83, 53.07	< .001	0.15	68.9	54.5, 83.29	< .001	0.13

Key: CI = Confidence interval; GRF = Ground reaction force; ηp2 = partial eta squared effect size
Note: The GRF measures have been multiplied by 100 (denoted BW100).

The relative decreases in GRF (in both the horizontal and vertical axis) and ball release speed across different zones of perceived effort are depicted in Figure 6-2. Bowling at submaximal intensities (i.e., low – medium) was associated with larger decreases in peak horizontal force (7.9–12.3% decrease), impulse (15.8–21.4%) and loading rate (7.4–12.7%) compared to decreases in ball release speed (5.4–8.3%). This trend was less apparent in the vertical direction.

Figure 6-2: Percentage decrease in GRF and ball release speed relative to maximum perceived effort.



Note: The relative percentage decrease in ball release speed and (**A**) horizontal GRF measures, and (**B**) vertical GRF measures. Values represent model-estimated means, expressed as a percentage difference from zone high estimated mean. The colour palette in this figure was produced using *Manu* R package (www.g-thomson.github.io/manu) and is inspired by the New Zealand Takahē.

Discussion

This study examined the relationship between bowling intensity and GRF in cricket pace bowlers. Both peak force and loading rate were significantly different among all three perceived intensity zones in the horizontal and vertical directions (Aim 1). All GRF measures in the horizontal axis increased significantly across low, medium and high. However, a larger difference was observed between medium and high compared to low and medium. An increase in ball release speed is associated with increases in peak GRF and loading rate (Aim 2) on both the horizontal and vertical axis. Lastly, moving from high to medium, or medium to low, was associated with a larger relative decrease in GRF on the horizontal axis compared to the relative decrease in ball release speed (Aim 3).

As the authors have not come across a similar study looking at GRFs across submaximal intensities, the results can only be compared to studies that have analysed GRFs between bowlers at maximum speeds. These studies have shown that horizontal peak force[6,31,124] or horizontal impulse[22] is most strongly associated with bowling speed. Our results also show that horizontal GRF explains a greater proportion of the variance in ball release speed compared to vertical GRF. This could be due to increased run-up speed at higher zones, leading to higher linear momentum about the front foot during the planting action prior to ball release.[123] More horizontal braking force is needed to slow the centre of mass and to stop the front foot. An increase in angular momentum results as the body rotates around the front leg faster, increasing the hand's linear velocity.[22,123,124] Interestingly, King et al. (2016)[22] also found a negative correlation between ball released speed and average vertical and horizontal loading rates (r = -0.452 and -0.484, respectively), which contradicts the finding of this study. A potential reason for the conflicting results between studies could be the difference in ability and technique between bowlers. For example, if a higher percentage of bowlers land on their heel, this could reduce loading rates as forces are spread over an extended period before reaching a peak – usually just before the forefoot contacts the ground.[22,123] There was also a difference in playing ability between the two studies, with King et al. (2016)[22] participants being all elite versus just one elite bowler in the current study.

This is the first study to show that GRF decreases more than ball release speed when going from a high to low perceived bowling intensity. It is hard to determine the exact causes; however, the relationships between GRFs and ball release speed might not be equal at submaximal intensities. A possible reason for this is to do with run-up speed. At low and medium perceived intensity, run-up speed might decrease – therefore decreasing horizontal GRF.[22] However, this might not affect the angular velocity of the bowling arm, and therefore ball release speed,[43] to the same proportion.

Practical applications and future study

The larger relative decrease in GRF compared to ball release speed when moving from maximal (high) to submaximal intensities (medium and low) could interest players and coaches. If a player's high delivery speed is 113 km/h (the average in this study), a decrease to the average medium speed (107 km/h) only results in a 5% decrease in speed, but a 7–17 % decrease in GRF on the horizontal axis. Previous studies have also shown this small decrease in speed from a reduced perceived intensity.[63,99,111,119] In particular, Perrett et al. (2021)[119] found a 0.25% drop in normalised release speed for every 1% decrease in prescribed intensity. More research is needed to explain the possible biomechanical reasons for this observation. If stronger links are made between higher GRFs and injury, this may influence bowlers' strategies during periods of high workloads, such as in a training session, to minimise fatigue and potentially reduce the chances of injury without the corresponding decrease in performance. It could also lead coaches to develop an individualised approach to bowling technique for players, emphasising reducing GRFs without decreasing ball release speed.[124]

As this study found a significant relationship between GRF and intensity, future studies could look at more user-friendly, cost-effective ways of measuring GRF so bowlers can obtain a more precise estimate of intensity. A potential solution could be to use an inertial measurement unit (IMU) in combination with machine learning. IMUs are installed in most smart devices (phones and watches); therefore, they are accessible to most cricketing populations.[97] IMUs and machine learning have been successfully used in cricket to predict ball release speed and the perceived intensity zone.[62,98,111,125]

Limitations

There are several limitations readers need to be aware of. Firstly, all testing was conducted on an artificial pitch, and therefore results may not be an accurate representation of all cricket pitches.[122] Currently, no studies have investigated how surface properties affect front-foot impact during pace bowling. This is due to the complexities involved with installing a force plate under a grass cricket pitch. Secondly, although this study generalised perceived effort as a combination of approach intensity (running speed) and arm rotation intensity, further insight could be gained by measuring run-up speed. A faster run-up speed has been associated with higher vertical and horizontal GRF[22] and may have been a key reason why GRF forces were lower at submaximal intensities. Future studies could look at submaximal run-up speeds with maximum bowling effort during the delivery phase and the relationship between GRF and bowling speed. Thirdly, this study did not test a wide variety of skill levels. Therefore, these results may not generalise to elite players. Lastly, the lack of familiarity with bowling at submaximal intensities may explain the smaller differences in GRFs between low and medium compared to medium and high perceived intensity zones. Although there was only a minor difference in standard deviation for each zone, participants may have found it hard to differentiate between 70% of maximum effort (low) and 85% of maximum effort (medium) compared to 100% maximum effort (high), which is a familiar intensity for bowlers.

Conclusion

A decrease in ball release speed and the perceived intensity zone was correlated with a reduction in peak force, impulse, and loading rate on the horizontal axis, and peak force and loading rate on the vertical axis. Lowering the perceived bowling intensity was associated with larger relative decreases in GRF on the horizontal axis compared to the relative decrease in ball release speed. This may have implications for players and coaches as a bowler can achieve a considerable decrease in GRF with only small changes in ball speed. If researchers can find consensus on how GRF affects injury, this finding could influence how players approach a bowling spell to decrease the chance of injury and maximise performance by reducing fatigue. When interpreting these results, caution should be made as all testing was done on an artificial pitch and may not be generalisable to a grass pitch.

Appendix

Table 6-4: Estimated means and 95% confidence intervals of GRF across the three perceived intensity zones.

		Horizontal			Vertical		
GRF	Zone	Estimated Mean	LCL	UCL	Estimated Mean	LCI	UCL
Peak force (BW)	Low	2.46	2.10	2.82	4.73	4.07	5.39
	Med	2.62	2.25	2.98	4.95	4.29	5.62
	High	2.92	2.56	3.28	5.40	4.74	6.06
Impulse (BW · s)	Low	0.05	0.03	0.06	0.38	0.35	0.41
	Med	0.05	0.03	0.07	0.37	0.34	0.41
	High	0.06	0.05	0.08	0.37	0.34	0.41
Loading rate (BW · s^{-1})	Low	23.54	19.26	27.82	44.86	37.12	52.60
	Med	25.58	21.30	29.86	48.30	40.55	56.04
	High	28.56	24.28	32.84	52.78	45.04	60.53

Key: GRF = Ground reaction force; LCI = Lower confidence limits; UCL = Upper confidence limits.

Chapter 7 - Can an inertial measurement unit, combined with machine learning, accurately measure ground reaction forces in cricket fast bowling.

Preface

Chapter 6 found that ground reaction forces (GRF) were positively associated with ball release speed and the perceived intensity zone. Therefore, GRFs can be used as a measure of bowling intensity. This study used machine learning to predict GRF from inertial measurement units located on the upper back and bowling wrist. If successful, this could initiate widespread monitoring of GRFs, which is currently only available in specialised laboratories. It could also allow researchers to understand better the relationships between GRF, injury and performance.

This paper has been submitted to the journal Sports Biomechanics and is currently under review (as of December 2022).

Abstract

This study examined whether an inertial measurement unit (IMU) could measure ground reaction force (GRF) during a cricket fast bowling delivery. Eighteen male fast bowlers had IMUs attached to their upper back and bowling wrist. Each participant bowled 36 deliveries, split into three different intensity zones: low = 70% of maximum perceived bowling effort, medium = 85%, and high = 100%. A force plate was embedded into the bowling crease to measure the ground truth GRF. Three machine learning models were used to estimate GRF from the IMU data. The best results from all models showed a mean absolute percentage error of 22.1% body weights (BW) for vertical and horizontal peak force, 24.1% for vertical impulse, 32.6% and 33.6% for vertical and horizontal loading rates, respectively. The linear support vector machine model had the most consistent results. Although results were similar to other papers that have estimated GRF, the error would likely prevent its use in individual monitoring. However, due to the large differences in raw GRFs between participants, researchers may be able to help identify links among GRF, injury, and performance by categorising values into levels (i.e., low and high).

Introduction

Cricket fast bowlers have more injuries than any other playing position.[1] It is commonly accepted that genetic susceptibility, technique, and bowling workload determine whether fast bowlers get injured.[5] However, the definition of bowling workload has come under debate recently.[5,19,20,111] Previous injury prevalence studies have used this term to describe bowling volume – the number of deliveries performed in a session. Although research has linked a high and low weekly, monthly and yearly bowling volume to injury,[126] it does not acknowledge that deliveries are bowled at different intensities, which may exert different levels of stress on the body.

Reasons for the omission of intensity data in research stem from the fact that there are no accepted measures of bowling intensity.[5] Potential measures include a rating of perceived intensity, ball release speed, and ground reaction forces (GRF), with the latter two requiring considerable outlay due to the specialist equipment required. A possible solution to overcome this barrier could be to use a wearable inertial measurement unit (IMU). An IMU typically consists of an accelerometer and gyroscope, which measure linear acceleration (measured in g-force) and angular velocity (degrees per second), respectively. Researchers have used IMU data to accurately predict bowling volume[62,63,98] and measures of bowling intensity – bowling velocity and perceived intensity.[111]

Although bowling speed and perceived intensity provide a more rounded measure of bowling workload, they do not provide information regarding the GRFs endured during the delivery phase. In fast bowlers, front foot GRFs are, on average, 6.7 times body weight (BW) for vertical GRF and 4.5 BW for horizontal GRF.[22,123] In cricket, GRF is measured in a specialised lab using a force plate embedded in the bowling crease. Typical measurements include peak force, impulse and loading rate during front foot contact, as these are related to performance in cricket.[22] The need for specialised equipment has meant that insufficient GRF data has been collected to investigate the long-term link with injury. It has been hypothesised that exposure to repeated high magnitude ground impacts combined with high spinal rotation may be a significant cause of acute and chronic injuries.[21,116-118] Therefore, an IMU-based estimate of GRF might provide a practical solution for monitoring GRF over time.

There has been a recent emergence of studies trying to predict GRF from IMU data in sports and physical activities. Callaghan et al. (2020)[24] compared GRFs derived from a force plate against IMU force signatures in cricket fast bowlers. Although there was mixed reliability (CV = 4.23–18.17%) between the two measures, it might be possible to model force plate-derived GRF from IMU data using more advanced modelling techniques, such as machine learning. Recently, machine learning has been used with IMU data to estimate GRF in sidestepping[36] (mean absolute percentage error (MAPE) = 19.7% BW), running[36] (MAPE = 29.7%), and ballet jumps[35] (unilateral landings, root-mean-square error (RMSE) = 0.42 BW; bilateral landings, RMSE = 0.39). To the authors' knowledge, no studies have used the IMU's gyroscope and the accelerometer to estimate GRF, which may improve results.[35]

Given the potential benefits of monitoring GRF in cricket fast bowlers and recent evidence of using IMU data to estimate GRF in other sports, this study aims to: (1) determine whether an IMU combined with machine learning can predict GRFs in fast bowlers; and (2) to determine if results improve with the addition of gyroscope data. Based on the existing literature, it was hypothesised that an IMU and machine learning would be able to measure GRF with similar accuracy to other studies. Furthermore, it was hypothesised that the addition of gyroscope data would improve all GRF measures.

Methods

Participants

Eighteen male pace bowlers from the (removed for anonymity) cricket academy were recruited. All participants were 18 years or older (mean age 19.4 ± 1.2 years), and 16 bowlers were right-handed bowlers. The mean height was 183 cm (SD = 7.3 cm), and the mean mass was 78.5 kg (SD = 10.5 kg). Participants had no reported injuries at testing, and written informed consent was obtained from each participant. Seventeen bowlers were sub-elite, playing at a premier club level, and one bowler was elite, having played first-class cricket. Ethics was granted by Loughborough University's Ethics Committee (reference 2020-2274-1855).

Design

This study used a cross-sectional design with data collected from a single testing session. All data were collected on an indoor artificial pitch at the National Centre for Sport and Exercise Medicine (NCSEM) biomechanics laboratory at Loughborough University, with sufficient space for a full run-up.

Testing session

Participants had their height and body mass measured and performed their regular warm-up. They were then instructed to bowl 36 deliveries at a chosen line and length, split evenly between three perceived intensity zones – Low = 70% of maximum perceived bowling effort, Medium = 85%, and High= 100% – in random order. These intensities were chosen as it they cover the most likely range that a bowler would perform at during training and a competitive match. The force plates were positioned at the popping crease to record the GRF of the front foot during the delivery phase. A total of 554 deliveries were recorded for analysis, with 94 deliveries omitted from nine participants due to the failure of an IMU to record (14 deliveries) or participants not landing with their front foot on the force plate (80 deliveries).

Equipment

Two Blue Trident IMUs (IMeasureU, Auckland, New Zealand) were attached to the upper back (around T1) and the bowling wrist (where a regular watch would be positioned). These locations were chosen due to their practicality, i.e., ease of application and should not get in the way when fielding. Each IMU consisted of a low measurement accelerometer (1122 Hz, ±16 g), a high measurement accelerometer (1600 Hz, ±250 g), and a gyroscope (1125 Hz, ±2000°/s). GRFs were measured using two Kistler force platforms sampling at 1000 Hz (Type 9287B, Kistler AG, Switzerland). Ball release speed was calculated using an 18-camera retro-reflective motion analysis system (Vicon, MX13, OMG Plc, Oxford, UK) sampling at 300 Hz. Two reflective markers were placed on the ball. The velocity was calculated using the change in displacement from the first two frames after ball release divided by the change in the time between the frames.

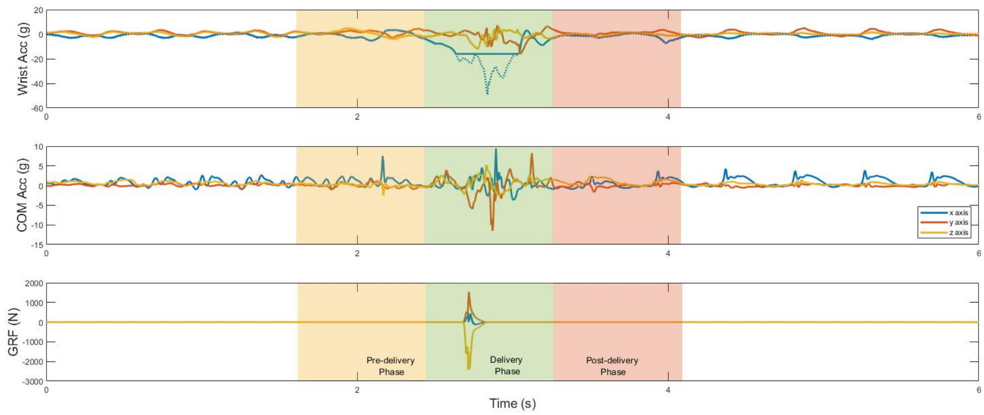
Data pre-processing and feature computation

Force plate: To determine the start and end of front foot contact on the force plate, the magnitude was first calculated by taking the square root of the sum of the x, y, and z-axis squares. A dynamic window for each delivery was created to determine the start and end of front foot contact. The window started at the first sample >= 35 N retrospectively from the peak magnitude and ended when it returned to <=35 N post peak magnitude. The 35 N threshold was chosen after observing the force plate data. If the force plate threshold was set to a more conventional 20 N, artefact caused by the steps prior to front foot contact would trigger the force plate in some bowlers. The data from each delivery were then visualised to identify and remove errored trials (i.e., partial foot contacts). Raw data from the y-axis (horizontal) and z-axis (vertical) were used to calculate peak force, impulse, and loading rates using a custom algorithm created in MATLAB R2021a. Specifically, for the horizontal and vertical axis, the impulse was calculated by determining the area under the curve using trapezoidal numerical integration, and the loading rate was calculated by dividing the peak force by the time from initial foot contact to the time of the peak force [116] GRFs were then normalised to each participant's body weight and expressed in bodyweights.[122]

IMUs: For both IMUs, a fourth-order 1 Hz Butterworth low pass filter was used for baseline removal.[58,68] Event detection of bowls was performed by calculating the magnitude of the gyroscope's x, y and z-axis and identifying peaks > 500°/s.[63,111] An event detection window of 3 seconds was used to isolate each event. The window was broken into pre-delivery (starting and ending 1.5 and 0.5 seconds before the gyroscope peak, respectively), delivery (0.5 seconds before and after the gyroscope peak), and post-delivery (starting and ending 0.5 and 1.5 seconds after the gyroscope peak, respectively). An example of the IMU and force plate traces from a single delivery can be seen in Figure 7-1.

Features were then extracted within each of the three phases from the time and frequency domains using MATLAB (release 2021b, The MathWorks, Inc., MA, USA). A total of 282 features were computed from the individual axes and the magnitude of the accelerometer and gyroscope channels. The features were similar to Kautz et al. (2017)[68] and included the mean, standard deviation, maximum, minimum, skewness, kurtosis, amplitude, frequency, energy, the position of the maximum, the position of the minimum, as well correlations between x, y and z axes.

Figure 7-1: Inertial measurement unit and force plate traces from a single delivery.



Model training and testing

Three machine learning models – random forest (RF), linear support vector machine (LSVM), and gradient boosting (XGB) – were used to predict GRF for each IMU separately. These were chosen because they have been previously effective at classifying bowling volume, ball release speed, and perceived intensity zone with IMUs located on the thoracic back, lumbar back, and wrist.[63,98,111]

All machine learning models were trained and tuned in R (R Core Team, Austria) using the caret package. As a pre-processing step, all features with zero variance and those that were highly correlated with other features ($r > 0.90$) were removed. Optimal model hyperparameters were determined using 10-fold cross-validation. The optimal values were chosen based on optimising root mean square error. Leave-one-participant-out cross-validation was used to evaluate the final models.

Statistical analysis

The accuracy of each model was expressed as the mean absolute error (MAE) and mean absolute percentage error (MAPE). The results from each cross-validation iteration were used within a two-way repeated-measures ANOVA to compare the MAE across the three models and two IMU positions. Model assumptions (i.e., no significant outliers, dependant variable normality, sphericity) were checked before fitting each model using the 'afex' R package. Both models violated the sphericity assumption and were adjusted using the Greenhouse-Geisser sphericity correction. Estimated means and pairwise contrasts (between models and between the two IMUs) were estimated using the 'emmeans' package, with multiple comparisons adjusted using the Holm method. An a priori alpha of 0.05 was used for all analyses.

Results

Table 7-1 shows the mean, standard deviation, and range for ball release speed and GRF across intensity zones for all participants. There was an increase in mean ball release speed, peak force in the horizontal and vertical axes, and loading rate in the horizontal and vertical axes with a corresponding increase in the perceived intensity zone. There was also a large range between participants for the average minimum and average maximum values across all GRFs.

Table 7-1: The mean, standard deviation, and range for ball release speed and ground reaction forces across intensity zones for all participants.

	Low			**Medium**			**High**		
Variable	Mean	SD	Range	Mean	SD	Range	Mean	SD	Range
Ball release speed (km·h-1)	104.46	5.96	31.79	107.53	5.56	31.79	112.8	6.32	30.04
Peak_y (BW)	2.6	0.71	4.31	2.73	0.73	4.31	3.08	0.87	4.93
Peak_ z (BW)	5.05	1.19	6.75	5.24	1.28	6.75	5.76	1.63	7.89
Impulse_y (BW·s)	0.05	0.04	0.22	0.05	0.04	0.22	0.06	0.04	0.23
Impulse_z (BW·s)	0.4	0.08	0.71	0.39	0.08	0.71	0.39	0.08	0.45
Loading rate_y (BW·s^{-1})	24.66	9.12	51.09	26.63	9.26	51.09	29.91	10.95	57.36
Loading rate_z (BW·s^{-1})	47.45	16.2	99.92	50.95	16.11	99.92	55.83	19.89	97.18
Average contact time (s)	0.37	0.09	0.6	0.36	0.07	0.47	0.35	0.08	0.39

Table 7-2 shows the model accuracy for predicting peak force from the upper back IMU and the wrist IMU. The results for accelerometer data and combined accelerometer and gyroscope data are shown. Although there were no significant differences between the models, LSVM generally performed better for predicting vertical (MAPE = 22.1%) and horizontal GRF (MAPE = 24.1%) using the upper back IMU, and vertical GRF (MAPE = 22.1%) using the wrist IMU. The upper back IMU tended to produce slightly less error (although non-significant) than the wrist. Similarly, the combined accelerometer and the gyroscope data tended to have slightly better results than the accelerometer data alone, although these differences were not significant.

Table 7-2: Predicted peak force values.

	MAPE				RMSE (BW)			
	Upper back		Wrist		Upper back		Wrist	
Model	Vertical (Z)	Horizontal (Y)	Vertical (Z)	Horizontal (Y)	Vertical (Z)	Horizontal (Y)	Vertical (Z)	Horizontal (Y)
Accelerometer only								
RF	24.9	27.5*	26.8	25.7	1.36	0.78	1.47	0.77
XGB	25.2	28.9	28.9	25.2	1.49	0.84	1.68	0.79
LSVM	22.1	25.2	22.1	26.7	1.30	0.75	1.27	0.79
Accelerometer and gyroscope								
RF	23.6	24.8	25.5	26.1	1.29	0.71	1.41	0.78
XGB	23.7	25.8	27.0	27.5	1.38	0.76	1.52	0.83
LSVM	22.4	24.1	23.1	27.4	1.31	0.73	1.34	0.82

Key: LSVM = Linear support vector machine; MAPE = Mean absolute percentage error; RF = Random forest; RMSE = Root mean square error; XGB = Gradient boosting. * = Significantly different from the corresponding model and axes in the accelerometer and gyroscope dataset.

Table 7-3 shows the model accuracy for predicted impulse values. For the vertical axis, the best results were seen with the upper back IMU and the LSVM model when using accelerometer and gyroscope data (vertical axis, MAPE = 16.2%). This model was also significantly better than the XGB model ($p = 0.003$) and the corresponding model for the wrist IMU ($p = 0.006$). Both the RF and XGB models produced better results using the wrist IMU, but the LSVM model tended to provide better results when using data from the upper back IMU. Regarding the horizontal axis, the variability among participants led to large MAPE scores when estimated from leave-one-subject-out cross-validation.

Table 7-3: Predicted impulse values.

	MAPE				RMSE (BW·s)			
	Upper back		Wrist		Upper back		Wrist	
Model	Vertical (Z)	Horizontal (Y)	Vertical (Z)	Horizontal (Y)	Vertical (Z)	Horizontal (Y)	Vertical (Z)	Horizontal (Y)
Accelerometer only								
RF	18.5	280	16.6*	288	0.07	0.04	0.06	0.04
XGB	19.6	280	18.1	233	0.08	0.04	0.07	0.04
LSVM	17.1^	296	20.2	221	0.07	0.05	0.08	0.04
Accelerometer and gyroscope								
RF	17.1~	255	16.4	249	0.07	0.04	0.06	0.04
XGB	21.6^	220	18.0	286	0.09	0.04	0.08	0.04
LSVM	16.2^~	228	19.3	215	0.06	0.04	0.07	0.04

Key: LSVM = Linear support vector machine; MAPE = Mean absolute percentage error; RF = Random forest; RMSE = Root mean square error; XGB = Gradient boosting. ^ = Significantly different from the corresponding model and axes when using the wrist IMU. ~ = Significantly different from XGB using the same IMU and dataset. * = Significantly different from LSVM using the same IMU and dataset.

Table 7-4 shows the model accuracy for predicting the loading rate. The LSVM model had the best vertical GRF estimate (MAPE = 32.6%), and the RF model had the best horizontal GRF estimate (MAPE

= 33.6%). In general, less error was observed when predicting vertical axis GRF and using the combined accelerometer and gyroscope datasets.

Table 7-4: Predicted loading rate values.

	MAPE				RMSE ($BW \cdot s^{-1}$)			
	Upper back		Wrist		Upper back		Wrist	
Model	Vertical (Z)	Horizontal (Y)	Vertical (Z)	Horizontal (Y)	Vertical (Z)	Horizontal (Y)	Vertical (Z)	Horizontal (Y)
Accelerometer only								
RF	35.1^	37.3	39.9*	38.5	17.1	9.51	19.7	9.92
XGB	37.5	40.9	38.7	41.6	19.2	10.95	20.9	11.32
LSVM	36.3	42.0*	33.3	39.2	18.9	10.86	17.3	10.35
Accelerometer and gyroscope								
RF	33.6^	33.6^~	39.2	39.5	16.7	8.88	19.7	10.14
XGB	35.2	38.2^	37.9	43.9	18.8	10.18	19.2	11.42
LSVM	32.6	36.9	36	40.8	16.8	9.7	18.1	10.57

Key: LSVM = Linear support vector machine; MAPE = Mean absolute percentage error; RF = Random forest; RMSE = Root mean square error; XGB = Gradient boosting. * = Significantly different from the corresponding model and axes in the accelerometer and gyroscope dataset. ^ = Significantly different from the corresponding model and axes when using the wrist IMU. ~ = Significantly different from XGB using the same IMU and dataset. * = Significantly different from LSVM using the same IMU and dataset.

Discussion

This study examined whether an IMU located either on the upper back or bowling arm could estimate GRFs with the assistance of machine learning. In all but two cases, the IMU located on the upper back had the best results for measuring peak force (MAPE = 22.1%, 24.1%), impulse (MAPE = 16.2%, RMSE = 0.04 BW·s) and loading rate (MAPE = 32.6%, 33.6%) in both the vertical and horizontal axis, respectively. This is not surprising as attenuation would occur as the force travels up through the lower limbs and into the trunk before progressing to the upper limbs. However, only a few cases showed a significant difference between IMU locations, which means the wrist can also be an effective IMU placement site. This may be partly explained by GRF being correlated with ball release speed[127] and the wrist being the most accurate location to measure ball release speed.[111,125]

The LSVM was the most consistent model, with the best result in 13 out of the 24 outcomes examined. RF, however, was the most consistent at measuring horizontal loading rate using both the upper back and the wrist IMUs. This was different from two previous studies that found XGB provided the best results when estimating ball release speed and perceived intensity zone.[111,125] It is difficult to unravel why XGB performed poorly compared to the other models on this dataset. However, this is consistent with what has been termed the “no-free lunch theorem,” where machine learning models will not perform equally well on all problems.[85]

To the authors' knowledge, this is the only study to use accelerometer and gyroscope data to predict GRFs. In seven of the 12 comparisons, the addition of gyroscope data improved the overall accuracy of results, although these differences were statistically non-significant. The gyroscope provides the models with more orientation and angular velocity information, which may have helped estimate overall force.[35] Furthermore, the accelerometer only data had the best overall results for peak force in the vertical axis. This goes against our hypothesis that gyroscope data would improve all GRF measures due to GRF being correlated with ball release speed, and the gyroscope being the most important sensor when predicting ball release speed.[111,127] Future researchers and developers should consider the modest benefits in accuracy against the increased data volume, processing, and sensor requirements when incorporating a gyroscope.

It is hard to compare the results from the current study to Callaghan et al.(2020).[24] This is because the authors only compared overall force signatures to a force plate and did not predict GRFs from individual deliveries. It is also challenging to compare results against studies that have used IMUs and machine learning to predict GRF in other sports. This is because most studies display results in either MAE or RMSE,[35,38] which are generally proportional to the magnitude of GRF observed. Cricket has a relatively high average GRF and possibly greater inter-individual differences compared to other sports, likely resulting in higher MAE and RMSE scores. However, the results obtained were comparable to similar studies that displayed MAPE to estimate peak GRF in running and sidestepping drills (19.1–29.7% for vertical peak GRF, 21.8% for horizontal peak GRF).[36,37] Interestingly, these studies all used deep learning instead of the more conventional machine learning models used in the current study. Therefore, it is debatable whether more advanced machine learning techniques are warranted for estimating GRF using IMUs.

It is unclear why error rates are higher than other bowling workload metrics in cricket (i.e., ball release speed).[125] Cricket is a complex movement where the trunk experiences all three planes of motion during front foot contact. The high angular rotation has been associated with errors in acceleration data due to the crosstalk between sensing axes.[24,128] Furthermore, the IMUs are unlikely to be at the centre of mass during front foot contact, which was seen as a major limitation for achieving accurate GRF estimates.[39,40]

Practical applications

Although results were similar to other studies that have predicted GRF in running and sidestepping, it is unlikely that a measurement system with a MAPE of 22.1% for vertical and horizontal peak force would be useful for individual monitoring. Specifically, a previous study by McGrath et al. (2022)[127] showed that the difference in peak force between the high and low intensity deliveries was 0.67 and 0.46 BW in the vertical and horizontal directions, respectively. The corresponding RMSE of the 22.1%

error observed in this study is 1.29 and 0.71 BW, respectively. This means that such a system could not accurately differentiate GRFs across different zones of perceived bowling intensity. However, as there was a large range between participants for all GRFs, researchers may benefit from categorising values into levels (i.e., low and high). If a large amount of data is collected, this may help identify links among GRFs, injury and performance.

Limitations and future directions

There are several limitations that the reader needs to be aware of when interpreting these results. Firstly, although the sample size is considered large compared to similar studies, the generalisability of the models could be questionable. A model that has a high degree of generalisability means that it will work well on a range of cricketing populations. The study sample may not be representative of the wider fast bowling community. There is likely a larger variation in GRFs between academy fast bowlers compared to elite bowlers. Future studies should include a broader range of participants, such as juniors, females, and bowlers of varying abilities. This is important as these playing groups have similar injury rates to elite players.[93,94] Secondly, due to a technical issue with recording, approximately half the run-up was not captured from all bowlers. As bowling run-up velocity is linked with GRF,[43] more data relating to the run-up might improve model accuracy.

Although not a study limitation, researchers may consider using various other techniques to improve model performance. One option could be to train individualised models on each athlete or each type of bowling style (e.g., side-on, front-on, or mixed action). Individualised models may also help mitigate the inter-player variability evident when examining the horizontal impulse data. However, this method would require more individuals to get tested on a force plate which is one of the limitations this study was trying to eliminate. Lastly, researchers could look at other IMU locations (i.e., the front leg) or combine information from multiple IMUs. However, Hendry et al. (2020)[35] found that a single IMU on the sacrum was more accurate than IMUs located on five other sites (including the front leg and upper back) or a combination of data from all six IMUs. The locations in the current study were chosen due to useability. For example, an IMU positioned on the lower back may not be suitable in cricket due to players diving. Furthermore, as McGrath et al.(2021)[111] found, the current positions can accurately measure a range of bowling load parameters, so it would potentially be a barrier for use if athletes had to wear two or more IMUs.

Conclusion

This study determined whether an IMU can predict GRF in cricket fast bowling. Results showed a MAPE of 22.1% for vertical and horizontal peak force, 24.1% for vertical impulse, and 32.6% and 33.6% for vertical and horizontal loading rates, respectively. The LSVM model had the most consistent overall results. However, there was variability in model performance, with RF having the best results for measuring horizontal loading rate. Compared to just using accelerometer data alone, the results tended to show a small benefit when combining data from the accelerometer and gyroscope. Although results were similar to other papers that have estimated GRF, the error would likely prevent its use in individual monitoring. However, due to the large differences in raw GRFs between participants, researchers may be able to identify links between injury and performance by categorising values into levels (i.e., low and high).

Chapter 8 - General discussion

Fast bowlers are more than twice as likely to injure themselves during a season than any playing position in cricket.[1,2] Researchers have shown that certain acute and chronic bowling volumes are linked to injury.[1,4,8-16] However, players do not commonly record bowling volumes due to the manual, repetitive nature of the task.[5,17] In addition, as fast bowlers vary bowling intensity through a match and training session, bowling volume is not a complete measure of bowling workload.[5,95] Researchers have omitted bowling intensity from injury prevalence studies due to the lack of bowlers recording an intensity metric. This is because most bowling intensity metrics currently require expensive equipment or a purpose-built laboratory.

Inertial measurement units (IMUs) have shown promise in a range of sports for classifying movements and measuring variables. In cricket, they offer a new, cost-effective way of measuring bowling workload. As IMUs are integrated into most smart devices, they are assessable to a wide range of people from different social-economic backgrounds. Being accessible to many cricketers could allow researchers to collect data to better understand the link between bowling workload and injury. If bowling workload can be displayed in a user-friendly application, this could allow players and coaches to monitor loading and fatigue levels during a game, training session or throughout a season.

The aim of this thesis was to develop a cost-effective method to automatically predict bowling workload through bowling volume, ball release speed, perceived intensity, and ground reaction forces (GRFs). Although there were no agreed-upon methods of measuring bowling intensity,[5,19] the intensity metrics were chosen due to their previous use in other research papers,[17,24,26] and because they measured slightly different constructs.[26]

The thesis first explored past research in sports using IMUs to predict upper movement classification to learn what has been successful (Chapter 2). The chapter also provided an overview to sports scientists on the processes involved in developing these methods. Chapters 3 to 7 presented a progression of studies that designed and tested methods to predict bowling workload.

Research summary

Chapter 2 was a systematic literature review that explored past research in the classification and measurement of upper body movements in court and field-based sports using IMUs. It provided the sports science field with a non-technical description of how these systems work and the benefits and limitations that need to be considered when using IMUs. The review also informed future chapters by identifying equipment and techniques that worked and gaps in the literature. The main findings were that a machine learning approach was best for sports involving complex biomechanical movements. Since cricket is a complex movement involving the whole body, machine learning was used for the remaining chapters to estimate bowling workload. Also noted was a lack of research detailing the best IMU specification, data processing technique, or machine learning algorithm for any given sport. Therefore, subsequent studies compared different hardware, pre-processing methods, and machine learning models.

Chapter 3 was the first of five studies to develop a system that could predict bowling workload. A standard IMU (accelerometer ±16 g) located on the upper back and five different machine learning models were used to estimate bowling volume in a training setting. When using a 250 Hz IMU, utilising data from all three phases (pre-delivery, delivery, and post-delivery), the F-score from two models were 1.0. The same method was also tested on lower sampling frequencies and data from just the delivery phase. While down-sampling did not completely simulate different IMU devices, it did suggest that this method could work on a smaller amount of data with sampling rates as low as 25 Hz (F-score = 0.97 from two models). No single machine learning model outperformed the others. Although this method was not tested in a game, non-delivery activities included the worst possible case where a fielding and throwing drill closely resembled a bowl. It was concluded that a standard IMU can accurately predict bowling volume. This was an essential step in developing a system to predict bowling workload, as an algorithm would not be able to predict any intensity metrics accurately without distinguishing a bowl from other random events.

Chapter 4 quantified intensity through ball release speed and the perceived intensity zone using a standard IMU (accelerometer ±16 g) located on the upper back during a training session. To the authors' knowledge, this was the first study to record these metrics using an IMU in cricket. Again, data were down-sampled to determine if accuracy was reduced in each of the four models used. Gradient boosting (XGB), a type of machine learning model, was the most consistent across all sampling frequencies at measuring ball release speed (mean absolute error (MAE) = 3.61 km/h at 25 Hz) and the perceived intensity zone (F-score = 0.88 at 25 Hz). Its success is thought to be due to its scalability in all scenarios like classification and regression problems.[109] The first key finding was that there was no significant difference between models that were trained from all phases compared to the delivery phase. This was surprising as run-up speed is a strong indicator of ball release speed.[42] Furthermore, as arm and wrist speed are good indicators of ball release speed,[121] it was also unexpected to receive such accurate results from an IMU located on the upper back. Lastly, there was no significant difference in results between sampling frequencies for the two most consistent models – XGB and random forest (RF). This shows that sampling frequency is not as important as initially thought. This may mean that a range of IMUs can predict these parameters, including consumer-grade wearables and smart devices that might have fluctuating sampling rates.

Chapter 5 was conceived by the prediction that an IMU located on the wrist should improve the accuracy of the results obtained in Chapter 4 by providing machine learning models with more relevant data pertaining to the bowling wrist. This is because ball release speed is highly correlated to arm speed just before ball release.[43] It was unknown whether the high g-forces that the bowling wrist is subjected to, upwards of 70 g near ball release, would affect the accuracy of the models when using a standard IMU (where the accelerometer is capped at ± 32 g). This proposed limitation was tested by comparing an IMU with a high threshold accelerometer (SABELSense, ± 100 g) with an Apple Watch (± 32 g). Both devices were placed on the bowling and non-bowling wrists because some bowlers opposed wearing a watch on their bowling wrist. The authors also wanted to test the performance of a consumer-grade product as they were not stand-alone IMU devices, meaning they simultaneously performed numerous functions. Therefore, less priority may be given to the IMU. The same models and outcome variables from the previous study were used with the addition of bowling volume. XGB models again had the best results across all measures. There was only a slight improvement compared to the previous study (bowling volume: F-score = 1.0; ball release speed: MAE = 2.76 km/h; the perceived intensity zone: F-score = 0.92). There was no significant difference between the SABELSense and Apple Watch; however, a significant improvement in classifying the perceived intensity zone was observed for IMUs located on the dominant wrist.

Chapters 6 and 7 determined if an IMU and a machine learning approach could predict GRF. Measuring GRF in bowling currently requires a laboratory environment, and the associated cost means only a few locations in the world offer this capability. This also means that limited longitudinal data has been collected to determine a link with injury. However, it has been proposed that exposure to repeated high magnitude ground impacts, combined with spinal rotation, may be a significant cause of injury, especially in the lower body.[116-118]

Chapter 6 firstly investigated the association between GRF, ball release speed, and perceived intensity. If this association were not present, the prediction of GRF would be less relevant as it would not be a measure of bowling intensity. The study found that all GRF measures in the horizontal axis increased significantly across low, medium and high intensity zones. Both peak force and loading rate were significantly different among all three perceived intensity zones in the vertical axis. An increase in ball release speed was associated with increases in peak GRF and loading rate on both the horizontal and vertical axis. Lastly, moving from high to medium intensity, or medium to low intensity, was associated with a larger relative decrease in GRF on the horizontal axis compared to the relative reduction in ball release speed. For example, a drop from high to medium intensity zones resulted in a 7–17% decrease in the horizontal axis compared to a 5% decrease in ball release speed. This could influence bowlers' strategies during an unlimited overs match as they could conserve energy and reduce the load through their body with only a small reduction in ball release speed.

Chapter 7 was the final study determining whether an IMU combined with machine learning can predict GRF. Two IMUs were placed on the upper back and the bowling wrist. As previous similar studies have omitted data from the gyroscope,[35-38] this study also investigated whether accuracy improves with the addition of the gyroscope. Three machine learning models were tested – RF, LSVM and XGB. A mean absolute percentage error (MAPE) of 22.1% was recorded for vertical and horizontal peak force, 24.1% for vertical impulse, and 32.6% and 33.6% for vertical and horizontal loading rates, respectively. The LSVM model had the most consistent overall results. However, there was variability in model performance, with RF having the best results for measuring horizontal loading rate. There were small trends toward using the upper back IMU and data from the accelerometer and gyroscope. Although results were similar to other papers that have estimated GRF, the error would likely prevent its use in individual monitoring. However, due to the large differences in raw GRFs between participants, researchers may be able to help identify links among GRF, injury, and performance by categorising GRFs into levels (i.e., low and high).

Significance of findings

This body of work makes several novel contributions to measuring bowling workload using an IMU combined with machine learning. Each of these contributions and their implications are discussed below.

The adaptivity of general machine learning techniques

The machine learning models used in this PhD were general models that have been successfully used in a range of situations involving sports. Throughout this thesis, the models were tested under different IMU sampling frequencies. In Chapters 3 and 4, the sampling rate was reduced from 250 Hz to 25 Hz to simulate other IMU devices. Overall, it will appear that these machine learning models will perform well regardless of the sampling frequency. Accuracy was also not affected by the location where the IMUs were placed (i.e., the upper back or the wrist of the bowling and non-bowling arm) or when data from the run-up or follow-through phases were omitted. This was surprising given that run-up speed highly correlates with ball release speed.[43] Pre-processing techniques – event detection, filtering and windowing – were all common techniques that have been used in similar movement classification and metric estimation studies. The features used were similar throughout the studies and closely resembled a volleyball study that classified different shots.[68] Due to their robust nature, sport and data scientists should consider these general techniques and models when classifying movement patterns or measuring outputs in similar sports. It also bodes well for similar techniques to work on smart devices, greatly increasing accessibility.

An automated measure of bowling workload

Chapter 5 was the first study to use a smart device (Apple Watch) to predict components of bowling workload – bowling volume, ball release speed and the perceived intensity zone. The results obtained from the Apple watches were similar to the research-grade IMUs. Although there was a greater fluctuation in the sampling rate for the Apple Watch due to the processor having to run other hardware and software, the algorithms used were robust enough to deal with this. There were also only slight differences between wearing the watch on the bowling or non-bowling wrist. This was a novel finding and addressed the issue of certain bowlers not wanting to wear a watch on their bowling wrist.

The algorithms created from this PhD can be used as the framework for a future application on a smartwatch that automatically estimates bowling workload throughout a training session, match, week, or season. An automated measurement device stops the monotonous tasks of manually recording bowling volume, which was a significant barrier to consistent monitoring.[5,17] The template images below show what a potential application on an Apple Watch may look like using the algorithms created in this thesis.

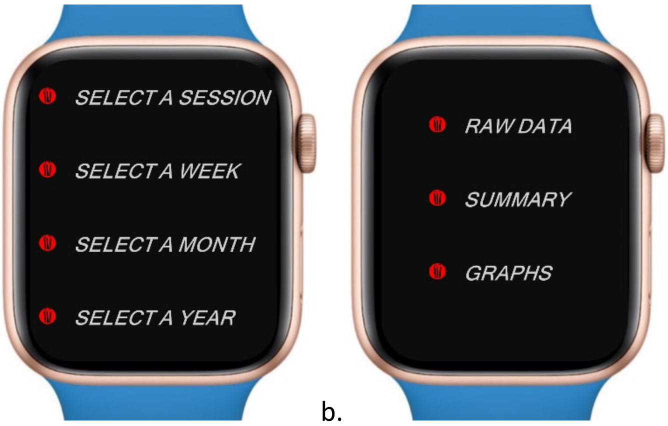
Figure 8-1: Opening screen.

The Apple Watch application would have an option to record a live session or view metrics from past sessions.

Figure 8-2: Real-time summary for a delivery.

Figure 8-2 shows a real-time summary of an individual delivery for each metric. It will also automatically count the number of deliveries.

Figure 8-3: a) Viewing history main page. b) History option screen.



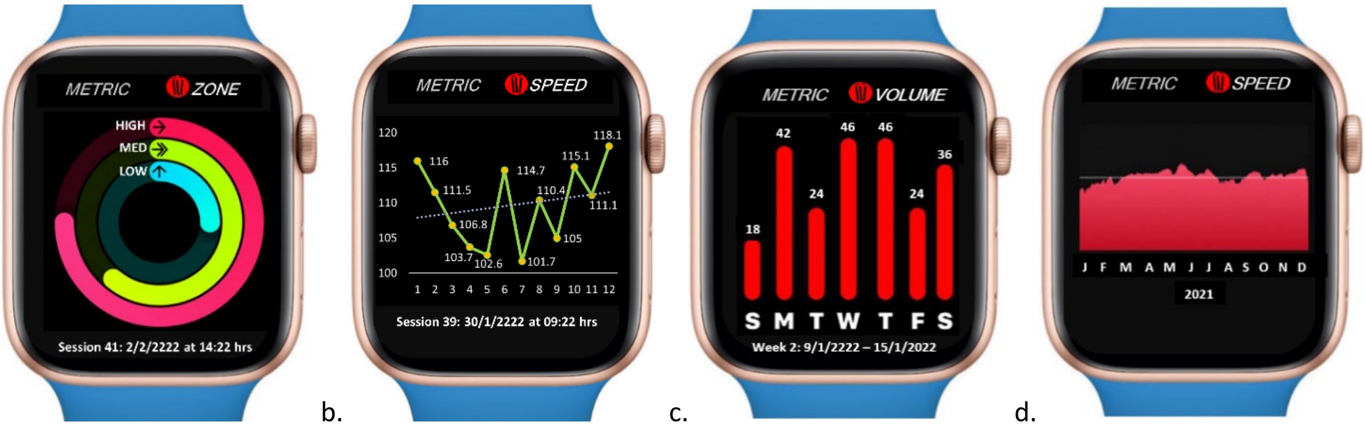
a. b.

Figure 8-4: a) Raw data from a session. b) Session summary.

a. b.

When viewing the history (Figure 8-3-a), users could have the option (Figure 8-3-b) to select raw data (seen in Figure 8-4-a) or a summary (seen in Figure 8-4-b) from an individual session or sessions over the week, month or year. Users would have the option of selecting what metric they want to view: speed, zone, or GRF.

Figure 8-5: Examples of viewing summaries across a session (a and b), week (c) and year (d).



a. b. c. d.

Users will also have the option of viewing summaries in a graphical form. Figure 8-5 show examples from different metrics across a session, week, and year.

Potential to assist researchers in determining more accurate links to injury.

As previously mentioned, due to the cost of the specialised equipment and expertise needed to measure bowling workload parameters, researchers have had a lack of intensity data to draw possible links with injury. Furthermore, bowling workload data collected previously, mainly bowling volume and the rate of perceived exertion rating (RPE), has had its reliability questioned.[5,17] This is primarily due to bowlers failing to fill out logbooks.[17] If this application comes to fruition on a smart device, researchers could collect more reliable data[5] from many participants. Surveys or follow-up appointments could be arranged to ask the participant whether they are injury-free or not. With this information, researchers could look at different combinations of parameters to determine what type of load correlates with injury. This could be displayed as a risk meter which lets bowlers know whether they are at a high, medium or low risk of injury (Figure 8-6-a). An insight could also inform the bowler on what to do to move into a lower risk category (Figure 8-6-b).

Figure 8-6: Personal insights into workload management.

A more personalised approach to injury prevalence may be achieved if further information is collected from participants – i.e., height, weight, age, and bowling style (front on, side on, or mixed action). Lastly, the application could gather injury prevalence information from players from developing nations. Little is known about injuries rates for fast bowlers in countries like India, Sri Lanka, Pakistan, Bangladesh and Afghanistan.

A coaching tool to improve performance and monitor fatigue.

An application that predicts multiple metrics of bowling workload could be used to improve performance and monitor fatigue levels. For example, ball release speed could be tracked over time to see if improvements occur from changing technique or improving physical strength. An individual who increases their bowling speed is likely to be more successful due to the reduction in reaction time for the batsman.[28,121] In addition, if the perceived intensity rate decreases, but the ball release speed stays the same, this may indicate an improvement in technique or muscular power.

If an application can automatically track bowling workload over long periods, players and coaches can use this data to assess fatigue levels. In particular, ball release speeds can be monitored for decreases despite similar perceived intensity levels. This may indicate fatigue, overreaching, or technique issues and can aid a bowler's or coach's decision to reduce workload or analyse technique. Strength and conditioning coaches could analyse ball release speed over a bowling spell to determine if bowlers hold speeds for long periods. If a sudden drop off is seen after a few overs, this may indicate that further work is needed to improve fitness and muscular endurance levels. Information from different bowling workload components could also aid team selection decisions by letting coaches know which bowler is at their peak or bowling faster for more extended periods.

Study limitations and future directions

In addition to the limitations discussed within each research chapter, several broader limitations must be noted. These, along with future directions, will be addressed in the themes below.

Generalisability

No outputs from the research chapters are generalisable to the whole cricketing population. A device for all cricketers would require machine learning models to train on a diverse range of bowlers and abilities. Although the authors tried to measure a wide variety of players, the algorithms developed might not be useable for females and juniors. Future studies should address these issues as female cricket is a rapidly growing sport, and both females and junior fast bowlers are just as prone to overuse injuries as male adults.[93,94]

As all studies were conducted in a training situation, there may be generalisability issues in a match setting. Although the authors tested participants against worst-case scenarios to try and confuse machine learning models, it may not accurately mimic a game where there are random events. Further studies should investigate this, as two studies have shown conflicting results on whether model performance is affected.[62,98]

The measurement of GRF

There was a considerable error when predicting GRFs. Although results were similar to other studies that predicted GRF in running and sidestepping, such a system could not accurately differentiate GRFs across different zones of perceived bowling intensity. It is unclear why error rates were higher than other bowling workload metrics in cricket (i.e., ball release speed). What was evident was the variety in accuracy among players, which is probably due to differences in technique. A more stable technique – which is often associated with elite players – would likely give more consistent IMU and GRF measurements, which would help the performance of the machine learning models. As the cohort from Chapter 7 was from a similar playing level – club cricket – future researchers could look at whether results improve when models are trained on elite players. Furthermore, researchers could look at an individualised approach where models are trained on the individual instead of a large group of players.

Slower deliveries

There were no slower deliveries analysed in any of the research chapters. Slower deliveries are now common in twenty20 and one-day limited-overs cricket. The authors could only find one study that included slower deliveries when correlating ball release speed to perceived exertion.[26] It concluded that ball release speed was related to perceived exertion. However, the number of slower bowls analysed was small, with only four slower deliveries from 48 deliveries. The different types of slower balls (e.g., knuckleball, off-cutter, back-of-the-hand) should be explored in future studies. Although accuracy might not be affected when estimating the perceived intensity zone and GRF, it is unlikely that an accurate ball release speed calculation could be obtained. This is because slower deliveries can be bowled at maximum intensity with the same arm speed.[129] The delivery is slower due to the late changes applied before ball release, commonly by adding or reducing rotation to the ball.

Methodology

Some participants had issues consistently bowling at the required perceived effort in the three separate data collection events. Estimating the perceived intensity zone may become compromised as models will learn from data that does not accurately represent each submaximal zone. However, the authors tried to mitigate the effects of this by including larger sample sizes (above the recommended sample size), providing a familiarisation bowl before testing, and using consistent verbal instructions. Issues with bowling at submaximal intensities have also been highlighted in a recent study,[26] with some participants showing differences between the prescribed intensity, perceived intensity, and actual intensity (measured as a percentage of maximum ball release speed). However, it should be noted that using the percentage of maximum ball release speed as the actual intensity metric does not consider small changes in technique and timing, which can alter speed without changing the perceived effort score.[27]

Smartwatch application

The models created in this thesis have yet to be implemented in a smart device application. Whether the models created using the Apple Watch will work on other smartwatches is unknown. Although the authors believe that algorithms could be integrated into an application, there may be reliability issues regarding the IMU inside the Apple Watch. The Apple Watch is not a standalone IMU; therefore, priority is not given from a system-level to this unit. In this thesis, SensorLog, the Apple Watch application used in Chapter 5 to collect raw data from the IMU, had problems recording data after a software update. This meant that the Apple Watch could not be used in Chapter 7. Because of this, it is unknown whether accuracy decreases when using the Apple Watch to estimate GRFs. If the algorithms developed are put into an application, it would require considerable resources to keep the application working. Due to international cricket council (ICC) match-fixing regulations, it is also unlikely that elite players will be permitted to use a smartwatch in a televised game. This might not be an issue, as broadcasters often record bowling volume and ball release speed.

Conclusion

An IMU can predict bowling workload through bowling volume, ball release speed, perceived intensity, and GRFs. This thesis represents an original contribution to the current body of knowledge in wearable sensors, machine learning in sports, and cricket bowling workload. The results show that general machine learning models can predict bowling workload with varying amounts of data using IMUs positioned at different locations with a range of sampling frequencies and threshold limits. This offers promise for using similar techniques to create an application on a smart device that predicts bowling workload in real-time. Such a system may also be accessible to most of the cricketing population. It will allow fast bowlers to track bowling workload automatically without manually recording data or using expensive equipment, normally out of reach for most players. Measurements can also be used to improve and track performance, monitor fatigue levels, and aid coaches in decision making. Lastly, it will allow researchers to capture more bowling workload data from several bowling workload measures to forge better links with injury and performance. A better understanding of the links between injury and workload may stimulate a more individualised approach to bowling workload management, hopefully reducing bowling-related injuries.

References

1. Orchard, J. W., Kountouris, A., & Sims, K. (2016). Incidence and prevalence of elite male cricket injuries using updated consensus definitions. *Open Access Journal of Sports Medicine, 7*, 187. doi:http://dx.doi.org/10.2147/OAJSM.S117497

2. Orchard, J., Kountouris, A., Sims, K., Orchard, J., Beakley, D., & Brukner, P. (2014). Feature injury report 2014: Injury report. *Sports Medicine Australia, 32*(4), 14-18. doi:http://dx.doi.org/10.3316/informit.037642732357295

3. Soomro, N., Redrup, D., Evens, C., Strasiotto, L. P., Singh, S., Lyle, D., Singh, H., Ferdinands, R. E. D., & Sanders, R. (2018). Injury rate and patterns of sydney grade cricketers: A prospective study of injuries in 408 cricketers. *Postgrad Med J, 94*(1114), 425-431. doi:http://dx.doi.org/10.1136/postgradmedj-2018-135861

4. Orchard, J., Blanch, P., Paoloni, J., Kountouris, A., Sims, K., Orchard, J. J., & Brukner, P. (2015). Cricket fast bowling workload patterns as risk factors for tendon, muscle, bone and joint injuries. *British Journal of Sports Medicine, 49*(16), 1064-1068. doi:http://dx.doi.org/10.1136/bjsports-2014-093683

5. Perrett, C., Lamb, P., & Bussey, M. (2020). Is there an association between external workload and lower-back injuries in cricket fast bowlers? A systematic review. *Physical Therapy in Sport, 41*, 71-79. doi:https://doi.org/10.1016/j.ptsp.2019.11.007

6. Portus, M., Mason, B. R., Elliott, B. C., Pfitzner, M. C., & Done, R. P. (2004). Technique factors related to ball release speed and trunk injuries in high performance cricket fast bowlers. *Sports Biomech, 3*(2), 263-284. doi:10.1080/14763140408522845

7. Herbert, A. J., Williams, A. G., Hennis, P. J., Erskine, R. M., Sale, C., Day, S. H., & Stebbings, G. K. (2019). The interactions of physical activity, exercise and genetics and their associations with bone mineral density: Implications for injury risk in elite athletes. *European journal of applied physiology, 119*(1), 29-47. doi:https://doi.org/10.1007/s00421-018-4007-8

8. Dennis, R., Farhart, P., Goumas, C., & Orchard, J. (2003). Bowling workload and the risk of injury in elite cricket fast bowlers. *J Sci Med Sport, 6*(3), 359-367. doi:http://dx.doi.org/10.1016/s1440-2440(03)80031-2

9. Dennis, R., Farhart, P., Clements, M., & Ledwidge, H. (2004). The relationship between fast bowling workload and injury in first-class cricketers: A pilot study. *Journal of Science and Medicine in Sport, 7*(2), 232-236. doi:http://dx.doi.org/10.1016/s1440-2440(04)80014-8

10. Hulin, B. T., Gabbett, T. J., Blanch, P., Chapman, P., Bailey, D., & Orchard, J. W. (2013). Spikes in acute workload are associated with increased injury risk in elite cricket fast bowlers. *British Journal of Sports Medicine*, bjsports-2013-092524. doi:http://dx.doi.org/10.1136/bjsports-2013-092524

11. Dennis, R., Finch, C. F., & Farhart, P. (2005). Is bowling workload a risk factor for injury to australian junior cricket fast bowlers? *British Journal of Sports Medicine, 39*(11), 843-846. doi:http://dx.doi.org/10.1136/bjsm.2005.018515

12. Orchard, J., James, T., Portus, M., Kountouris, A., & Dennis, R. (2009). Fast bowlers in cricket demonstrate up to 3-to 4-week delay between high workloads and increased risk of injury. *The American Journal of Sports Medicine, 37*(6), 1186-1192. doi:https://doi.org/10.1177/0363546509332430

13. Orchard, J., Blanch, P., Paoloni, J., Kountouris, A., Sims, K., Orchard, J. J., & Brukner, P. (2015). Fast bowling match workloads over 5–26 days and risk of injury in the following month. *Journal of Science and Medicine in Sport, 18*(1), 26-30. doi:https://doi.org/10.1016/j.jsams.2014.09.002Get

14. Olivier, B., Taljaard, T., Burger, E., Brukner, P., Orchard, J., Gray, J., Botha, N., Stewart, A., & Mckinon, W. (2015). Which extrinsic and intrinsic factors are associated with non-contact injuries in adult cricket fast bowlers? *Sports Medicine, 46*(1), 79-101. doi:http://dx.doi.org/10.11124/jbisrir-2015-1387, 2015

15. Stretch, R. A. (2014). Junior cricketers are not a smaller version of adult cricketers: A 5-year investigation of injuries in elite junior cricketers. *South African Journal of Sports Medicine, 26*(4), 123-127. doi:http://dx.doi.org/10.7196/SAJSM.543

16. Gabbett, T. J., Hulin, B., Blanch, P., Chapman, P., & Bailey, D. (2019). To couple or not to couple? For acute: Chronic workload ratios and injury risk, does it really matter? *International Journal of Sports Medicine, 40*(09), 597-600. doi:https://doi.org/10.1055/a-0955-5589

17. Dennis, R. J., Finch, C. F., McIntosh, A. S., & Elliott, B. C. (2008). Use of field-based tests to identify risk factors for injury to fast bowlers in cricket. *British Journal of Sports Medicine, 42*(6), 477-482. doi:http://dx.doi.org/10.1136/bjsm.2008.046698

18. Alway, P., Brooke-Wavell, K., Langley, B., King, M., & Peirce, N. (2019). Incidence and prevalence of lumbar stress fracture in english county cricket fast bowlers, association with bowling workload and seasonal variation. *BMJ open sport & exercise medicine, 5*(1), e000529. doi:http://dx.doi.org/10.1136/bmjsem-2019-000529

19. Constable, M., Wundersitz, D., Bini, R., & Kingsley, M. (2021). Quantification of the demands of cricket bowling and the relationship to injury risk: A systematic review. *BMC sports science, medicine and rehabilitation, 13*(1), 1-12. doi:https://doi.org/10.1186/s13102-021-00335-8

20. McNamara, D. J., Gabbett, T. J., Chapman, P., Naughton, G., & Farhart, P. (2015). Variability of playerload, bowling velocity, and performance execution in fast bowlers across repeated bowling spells. *Int J Sports Physiol Perform, 10*(8), 1009-1014. doi:http://dx.doi.org/10.1123/ijspp.2014-0497

21. Senington, B. (2019). *An investigation into the spinal kinematics and lower limb impacts during cricket fast bowling and their association with lower back pain.* (Doctoral Thesis), Bournemouth University.

22. King, M. A., Worthington, P. J., & Ranson, C. A. (2016). Does maximising ball speed in cricket fast bowling necessitate higher ground reaction forces? *Journal of Sports Sciences, 34*(8), 707-712. doi:https://doi.org/10.1080/02640414.2015.1069375

23. Tallent, J., Higgins, M., Parker, N., Waldron, M., Bradford, E., Keenan, J., O'Neill B, V., & Bell, P. G. (2019). Quantification of bowling workload and changes in cognitive function in elite fast bowlers in training compared with twenty20 cricket. *J Sports Med Phys Fitness, 59*(1), 35-41. doi:http://dx.doi.org/10.23736/S0022-4707.17.07940-3

24. Callaghan, S., Lockie, R., Andrews, W., Chipchase, R., & Nimphius, S. (2020). The relationship between inertial measurement unit-derived 'force signatures' and ground reaction forces during cricket pace bowling. *Sports Biomechanics, 19*(3), 307-321. doi:https://doi.org/10.1080/14763141.2018.1465581

25. Ahmun, R., McCaig, S., Tallent, J., Williams, S., & Gabbett, T. (2019). Association of daily workload, wellness, and injury and illness during tours in international cricketers. *Int J Sports Physiol Perform, 14*(3), 369-377. doi:http://dx.doi.org/10.1123/ijspp.2018-0315

26. Feros, S. A., Bednarski, D. A., & Kremer, P. J. (2021). The relationship between prescribed, perceived, and actual delivery intensity in cricket pace bowling. *International journal of sports physiology and performance, 1*(aop), 1-4.

27. Kiely, N., Pickering Rodriguez, L., Watsford, M., Reddin, T., Hardy, S., & Duffield, R. (2021). The influence of technique and physical capacity on ball release speed in cricket fast-bowling. *Journal of Sports Sciences*, 1-9. doi:https://doi.org/10.1080/02640414.2021.1933349

28. Malhotra, A., & Krishna, S. (2018). Release velocities and bowler performance in cricket. *Journal of Applied Statistics, 45*(9), 1616-1627. doi:https://doi.org/10.1080/02664763.2017.1386772

29. Middleton, K. J., Mills, P. M., Elliott, B. C., & Alderson, J. A. (2016). The association between lower limb biomechanics and ball release speed in cricket fast bowlers: A comparison of high-performance and amateur competitors. *Sports Biomechanics, 15*(3), 357-369. doi:https://doi.org/10.1080/14763141.2016.1163413

30. Callaghan, S., Lockie, R., Yu, W., Andrews, W., Chipchase, R., & Nimphius, S. (2018). Delivery length and front foot kinetics of cricket fast bowling: Potential impact on fast bowling workload. *Journal of Science and Medicine in Sport, 21*, S58-S59. doi:https://doi.org/10.1016/j.jsams.2018.09.133

31. Felton, P. J., Yeadon, M. R., & King, M. A. (2020). Optimising the front foot contact phase of the cricket fast bowling action. *J Sports Sci, 38*(18), 2054-2062. doi:https://doi.org/10.1080/02640414.2020.1770407

32. Glazier, P. S., & Worthington, P. J. (2014). The impact of centre of mass kinematics and ground reaction forces on ball release speeds in cricket fast bowling. *Sports Technology, 7*(1-2), 4-11. doi:https://doi.org/10.1080/19346182.2014.893351

33. Pew Research Center. (2019). *Mobile connectivity in emerging economies*. Retrieved from https://www.pewresearch.org/internet/2019/03/07/use-of-smartphones-and-social-media-is-common-across-most-emerging-economies/

34. Bullock, G. S., Panagodage-Perera, N. K., Murray, A., Arden, N. K., & Filbay, S. R. (2019). Relationship between cricket participation, health and well-being: Scoping review protocol. *BMJ open, 9*(11), e032070. doi:http://dx.doi.org/10.1136/bmjopen-2019-032070

35. Hendry, D., Leadbetter, R., McKee, K., Hopper, L., Wild, C., O'Sullivan, P., Straker, L., & Campbell, A. (2020). An exploration of machine-learning estimation of ground reaction force from wearable sensor data. *Sensors (Basel), 20*(3), 740. doi:10.3390/s20030740

36. Johnson, W. R., Mian, A., Robinson, M. A., Verheul, J., Lloyd, D. G., & Alderson, J. A. (2019, 31 July–4 August 2019). *Multidimensional ground reaction forces and moments from a single sacrum mounted accelerometer via deep learning.* Paper presented at the In Proceedings of the International Society of Biomechanics/American Society of Biomechanics Calgary Annual Conference, Calgary, AB, Canada.

37. Stetter, B. J., Ringhof, S., Krafft, F. C., Sell, S., & Stein, T. (2019). Estimation of knee joint forces in sport movements using wearable sensors and machine learning. *Sensors, 19*(17), 3690.

38. Wouda, F. J., Giuberti, M., Bellusci, G., Maartens, E., Reenalda, J., Van Beijnum, B.-J. F., & Veltink, P. H. (2018). Estimation of vertical ground reaction forces and sagittal knee kinematics during running using three inertial sensors. *Frontiers in physiology, 9*, 218. doi:https://doi.org/10.3389/fphys.2018.00218

39. Gurchiek, R. D., McGinnis, R. S., Needle, A. R., McBride, J. M., & van Werkhoven, H. (2017). The use of a single inertial sensor to estimate 3-dimensional ground reaction force during accelerative running tasks. *Journal of biomechanics, 61*, 263-268. doi:https://doi.org/10.1016/j.jbiomech.2017.07.035

40. Pogson, M., Verheul, J., Robinson, M. A., Vanrenterghem, J., & Lisboa, P. (2020). A neural network method to predict task- and step-specific ground reaction force magnitudes from trunk accelerations during running activities. *Med Eng Phys, 78*, 82-89. doi:10.1016/j.medengphy.2020.02.002

41. Portus, M. R., Mason, B. R., Elliott, B. C., Pfitzner, M. C., & Done, R. P. (2004). Cricket: Technique factors related to ball release speed and trunk injuries in high performance cricket fast bowlers. *Sports Biomechanics, 3*(2), 263-284.

42. Worthington, P. J., King, M. A., & Ranson, C. A. (2013). Relationships between fast bowling technique and ball release speed in cricket. *Journal of applied biomechanics, 29*(1), 78-84. doi:https://doi.org/10.1123/jab.29.1.78

43. Salter, C. W., Sinclair, P. J., & Portus, M. R. (2007). The associations between fast bowling technique and ball release speed: A pilot study of the within-bowler and between-bowler approaches. *Journal of Sports Sciences, 25*(11), 1279-1285. doi:https://doi.org/10.1080/02640410601096822

44. 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. *J Sports Sci, 37*(5), 568-600. doi:http://dx.doi.org/10.1080/02640414.2018.1521769

45. Bulling, A., Blanke, U., & Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys (CSUR), 46*(3), 33. doi:https://doi.org/10.1145/2499621

46. Schuldhaus, D., Zwick, C., Koerger, H., Dorschky, E., Kirk, R., & Eskofier, B. M. (2015). *Inertial sensor-based approach for shot/pass classification during a soccer match.* Paper presented at the KDD Workshop on Large-Scale Sports Analytics 2015.

47. Carling, C., Reilly, T., & Williams, A. M. (2008). *Performance assessment for field sports*. Abingdon, Oxon: Routledge.

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

49. Nguyen, L. N. N., Rodríguez-Martín, D., Català, A., Pérez-López, C., Samà, A., & Cavallaro, A. (2015). *Basketball activity recognition using wearable inertial measurement units.* Paper presented at the Proceedings of the XVI International Conference on Human Computer Interaction. doi:https://doi.org/10.1145/2829875.2829930

50. Qaisar, S. B., Imtiaz, S., Faruq, F., Jamal, A., Iqbal, W., Glazier, P., & Lee, S. (2013). A hidden markov model for detection & classification of arm action in cricket using wearable sensors. *J. Mobile Multimedia, 9*(1&2), 128-144.

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

52. Bai, L., Efstratiou, C., & Ang, C. S. (2016). *Wesport: Utilising wrist-band sensing to detect player activities in basketball games.* Paper presented at the Pervasive Computing and Communication Workshops (PerCom Workshops), 2016 IEEE International Conference on. doi:https://doi.org/10.1109/PERCOMW.2016.7457167

53. Stöggl, T., Holst, A., Jonasson, A., Andersson, E., Wunsch, T., Norström, C., & Holmberg, H.-C. (2014). Automatic classification of the sub-techniques (gears) used in cross-country ski skating employing a mobile phone. *Sensors, 14*(11), 20589-20601. doi:https://doi.org/10.3390/s141120589

54. Ericsson. (2018). Ericsson mobility report 2018. Retrieved from https://www.ericsson.com/assets/local/mobility-report/documents/2018/ericsson-mobility-report-june-2018.pdf

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

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

57. Kos, M., Ženko, J., Vlaj, D., & Kramberger, I. (2016). *Tennis stroke detection and classification using miniature wearable imu device.* Paper presented at the Systems, Signals and Image Processing (IWSSIP), 2016 International Conference on Systems, Signals and Image Processing (IWSSIP). doi:https://doi.org/10.1109/IWSSIP.2016.7502764

58. Mlakar, M., & Luštrek, M. (2017). *Analyzing tennis game through sensor data with machine learning and multi-objective optimization.* Paper presented at the Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers. doi:https://doi.org/10.1145/3123024.3123163

59. 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.* Paper presented at the SENSORS, 2015 IEEE. doi:https://doi.org/10.1109/ICSENS.2015.7370311

60. 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*, 1-20. doi:https://doi.org/10.1123/ijspp.2016-0683

61. Wang, Z., Guo, M., & Zhao, C. (2016). Badminton stroke recognition based on body sensor networks. *IEEE Transactions on Human-Machine Systems, 46*(5), 769-775. doi:https://doi.org/10.1109/THMS.2016.2571265

62. McNamara, D. J., 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. *Int J Sports Physiol Perform, 10*(1), 71-75. doi:https://doi.org/10.1123/ijspp.2014-0062

63. 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. doi:https://doi.org/10.1080/02640414.2018.1553270

64. 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. doi:https://doi.org/10.1007/s10618-017-0511-4

65. Khan, A., Nicholson, J., & Plötz, T. (2017). Activity recognition for quality assessment of batting shots in cricket using a hierarchical representation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 1*(3), 62. doi:https://doi.org/10.1145/3130927

66. Murray, N. B., Black, G. M., Whiteley, R. J., Gahan, P., Cole, M. H., Utting, A., & Gabbett, T. J. (2016). Automatic detection of pitching and throwing events in baseball with inertial measurement sensors. *International Journal of Sports Physiology and Performance*, 1-18.

67. Rawashdeh, S. A., Rafeldt, D. A., & Uhl, T. L. (2016). Wearable imu for shoulder injury prevention in overhead sports. *Sensors (Basel), 16*(11), 1847. doi:https://doi.org/10.3390/s16111847

68. 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*, 1-28. doi:https://doi.org/10.1007/s10618-017-0495-0

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

70. Mangiarotti, M., Ferrise, F., Graziosi, S., Tamburrino, F., & Bordegoni, M. (2019). A wearable device to detect in real-time bimanual gestures of basketball players during training sessions. *Journal of Computing and Information Science in Engineering, 19*(1), 011004. doi:https://doi.org/10.1115/1.4041704

71. 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.* Paper presented at the HEALTHINF. doi:https://doi.org/10.5220/0006148503190325

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

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

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

75. Chen, C., Jafari, R., & Kehtarnavaz, N. (2015). *Utd-mhad: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor.* Paper presented at the Image Processing (ICIP), 2015 IEEE International Conference on. doi:https://doi.org/10.1109/ICIP.2015.7350781

76. Ahmadi, A., Mitchell, E., Destelle, F., Gowing, M., O'Connor, N. E., Richter, C., & Moran, K. (2014). *Automatic activity classification and movement assessment during a sports training session using wearable inertial sensors.* Paper presented at the Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on. doi:https://doi.org/10.1109/BSN.2014.29

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

78. Yazdansepas, D., Niazi, A. H., Gay, J. L., Maier, F. W., Ramaswamy, L., Rasheed, K., & Buman, M. P. (2016). *A multi-featured approach for wearable sensor-based human activity recognition.* Paper presented at the Healthcare Informatics (ICHI), 2016 IEEE International Conference on. doi:https://doi.org/10.1109/ICHI.2016.81

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

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

81. Laguna, J. O., Olaya, A. G., & Borrajo, D. (2011). A dynamic sliding window approach for activity recognition. *UMAP, 11*, 219-230. doi:https://doi.org/10.1007/978-3-642-22362-4_19

82. Lara, O. D., & Labrador, M. A. (2013). A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys and Tutorials, 15*(3), 1192-1209. doi:https://doi.org/10.1109/SURV.2012.110112.00192

83. Khalid, S., Khalil, T., & Nasreen, S. (2014). *A survey of feature selection and feature extraction techniques in machine learning.* Paper presented at the Science and Information Conference (SAI), 2014. doi:https://doi.org/10.1109/SAI.2014.6918213

84. Sze, V., Chen, Y.-H., Yang, T.-J., & Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE, 105*(12), 2295-2329. doi:https://doi.org/10.1109/JPROC.2017.2761740

85. Gómez, D., & Rojas, A. (2016). An empirical overview of the no free lunch theorem and its effect on real-world machine learning classification. *Neural Computation, 28*(1), 216-228. doi:https://doi.org/10.1162/NECO_a_00793

86. Wundersitz, D. W., 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. doi:https://doi.org/10.1016/j.jbiomech.2015.09.015

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

88. Hall, M. A. (1999). *Correlation-based feature selection for machine learning.* (Doctor of Philosophy), The University of Waikato. New Zealand.

89. Chapman, A., Vicenzino, B., Blanch, P., & Hodges, P. (2009). Do differences in muscle recruitment between novice and elite cyclists reflect different movement patterns or less skilled muscle recruitment? *Journal of Science and Medicine in Sport, 12*(1), 31-34. doi:https://doi.org/10.1016/j.jsams.2007.08.012

90. Petersen, C. J., Pyne, D., Dawson, B., Portus, M., & Kellett, A. (2010). Movement patterns in cricket vary by both position and game format. *Journal of Sports Sciences, 28*(1), 45-52. doi:https://doi.org/10.1080/02640410903348665

91. di Prampero, P. E., Fusi, S., Sepulcri, L., Morin, J. B., Belli, A., & Antonutto, G. (2005). Sprint running: A new energetic approach. *J Exp Biol, 208*(Pt 14), 2809-2816. doi:https://doi.org/10.1242/jeb.01700

92. 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. doi:https://doi.org/10.1145/1882471.1882479

93. Forrest, M. R., Hebert, J. J., Scott, B. R., Brini, S., & Dempsey, A. R. (2017). Risk factors for non-contact injury in adolescent cricket pace bowlers: A systematic review. *Sports Medicine, 47*(12), 2603-2619. doi:https://doi.org/10.1007/s40279-017-0778-z

94. Jacobs, J., Olivier, B., Dawood, M., & NK, P. P. (2021). Prevalence and incidence of injuries among female cricket players: A systematic review and meta-analyses. *JBI evidence synthesis*. doi:https://doi.org/10.11124/JBIES-21-00120

95. McNamara, D. J., Gabbett, T. J., & Naughton, G. (2017). Assessment of workload and its effects on performance and injury in elite cricket fast bowlers. *Sports medicine, 47*(3), 503-515.

96. Orchard, J. W., Kountouris, A., Sims, K. J., Orchard, J., Beakley, D. T., & Brukner, P. D. (2015). Change to injury profile of elite male cricketers in the t20 era. *New Zealand Journal of Sports Medicine, 42*(1).

97. McGrath, J., Neville, J., Stewart, T., & Cronin, J. (2020). Upper body activity classification using an inertial measurement unit in court and field-based sports: A systematic review. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology, 235*(2), 1754337120959754. doi:https://doi.org/10.1177/1754337120959754

98. Jowitt, H. K., Durussel, J., Brandon, R., & King, M. (2020). Auto detecting deliveries in elite cricket fast bowlers using microsensors and machine learning. *Journal of Sports Sciences, 38*(7), 1-6. doi:https://doi.org/10.1080/02640414.2020.1734308

99. McNamara, D. J., Gabbett, T. J., Blanch, P., & Kelly, L. (2017). The relationship between wearable microtechnology device variables and cricket fast bowling intensity. *International Journal of*

Sports Physiology and Performance, 13(2), 1-20. doi:https://doi.org/10.1123/ijspp.2016-0540

100. Bredt, S., Chagas, M. H., Peixoto, G. H., Menzel, H. J., & de Andrade, A. G. P. (2020). Understanding player load: Meanings and limitations. *J Hum Kinet, 71*(1), 5-9. doi:https://doi.org/10.2478/hukin-2019-0072

101. Nyoni, B., Nleya, M., & Mtunzi, B. (2018). *A training utility for estimating the bowling speed of a cricketer using accelerometer data.* Paper presented at the 2018 International Conference on Intelligent and Innovative Computing Applications (ICONIC). doi: https://doi.org/10.1109/ICONIC.2018.8601232

102. Lopez, G., Abe, S., Hashimoto, K., & Yokokubo, A. (2019). *On-site personal sport skill improvement support using only a smartwatch.* Paper presented at the 2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). doi: https://doi.org/10.1109/PERCOMW.2019.8730681

103. Van den Tillaar, R., Bhandurge, S., & Stewart, T. (2020). Detection of different throw types and ball velocity with imus and machine learning in team handball. *ISBS Proceedings Archive, 38*(1), 184. doi: http://hdl.handle.net/10292/13633

104. Gençoğlu, C., & Gümüş, H. (2020). Standing handball throwing velocity estimation with a single wrist-mounted inertial sensor. *Annals of Applied Sport Science*, 0-0. doi:https://doi.org/10.29252/aassjournal.893

105. Skejø, S. D., Bencke, J., Møller, M., & Sørensen, H. (2020). A novel method for estimating throwing speed in handball using a wearable device. doi:https://doi.org/10.3390/s20174925

106. Skejø, S., Møller, M., Bencke, J., & Sørensen, H. (2018). Predicting throwing velocity using accelerometers.

107. Kuhn, M. (2008). Building predictive models in r using the caret package. *Journal of Statistical Software, 28*(5), 1-26. doi:https://doi.org/10.18637/jss.v028.i05

108. Narayanan, A., Desai, F., Stewart, T., Duncan, S., & Mackay, L. (2020). Application of raw accelerometer data and machine-learning techniques to characterize human movement behavior: A systematic scoping review. *J Phys Act Health, 17*(3), 360-383. doi:https://doi.org/10.1123/jpah.2019-0088

109. Chen, T., & Guestrin, C. (2016). *Xgboost: A scalable tree boosting system.* Paper presented at the Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. doi:https://doi.org/10.1145/2939672.2939785

110. Van Hooren, B., Goudsmit, J., Restrepo, J., & Vos, S. (2020). Real-time feedback by wearables in running: Current approaches, challenges and suggestions for improvements. *J Sports Sci, 38*(2), 214-230. doi:https://doi.org/10.1080/02640414.2019.1690960

111. McGrath, J., Neville, J., Stewart, T., Clinning, H., & Cronin, J. (2021). Can an inertial measurement unit (imu) in combination with machine learning measure fast bowling speed and perceived intensity in cricket? *Journal of Sports Sciences, 39*(12), 1402-1409. doi:https://doi.org/10.1080/02640414.2021.1876312

112. Felton, P. J., & King, M. A. (2016). The effect of elbow hyperextension on ball speed in cricket fast bowling. *J Sports Sci, 34*(18), 1752-1758. doi:https://doi.org/10.1080/02640414.2015.1137340

113. Poushter, J. (2016). Smartphone ownership and internet usage continues to climb in emerging economies. *Pew Research Center, 22*.

114. Hulin, B. T., Gabbett, T. J., Blanch, P., Chapman, P., Bailey, D., & Orchard, J. W. (2014). Spikes in acute workload are associated with increased injury risk in elite cricket fast bowlers. *Br J Sports Med, 48*(8), 708-712. doi:10.1136/bjsports-2013-092524

115. Senington, B., Lee, R. Y., & Williams, J. M. (2020). Biomechanical risk factors of lower back pain in cricket fast bowlers using inertial measurement units: A prospective and retrospective investigation. *BMJ open sport & exercise medicine, 6*(1), e000818. doi:http://dx.doi.org/10.1136/bmjsem-2020-000818

116. Hurrion, P. D., Dyson, R., & Hale, T. (2000). Simultaneous measurement of back and front foot ground reaction forces during the same delivery stride of the fast-medium bowler. *J Sports Sci, 18*(12), 993-997. doi:https://doi.org/10.1080/026404100446793

117. Bartlett, R. M., Stockill, N. P., Elliott, B. C., & Burnett, A. F. (1996). The biomechanics of fast bowling in men's cricket: A review. *J Sports Sci, 14*(5), 403-424. doi:https://doi.org/10.1080/02640419608727727

118. Crewe, H., Campbell, A., Elliott, B., & Alderson, J. (2013). Lumbo-pelvic loading during fast bowling in adolescent cricketers: The influence of bowling speed and technique. *J Sports Sci, 31*(10), 1082-1090. doi:https://doi.org/10.1080/02640414.2012.762601

119. Perrett, C., Bussey, M., & Lamb, P. (2021). The relationship between release speed, heart rate and rate of perceived exertion across maximal and submaximal intensities in fast bowlers. *The Journal of Sport and Exercise Science, 5*(2), 114-120. doi:10.36905/jses.2021.02.04

120. Feros, S. A., Bednarski, D. A., & Kremer, P. J. (2021). The relationship between prescribed, perceived, and actual delivery intensity in cricket pace bowling. *International Journal of Sports Physiology and Performance, 16*(5), 731-734. doi:https://doi.org/10.1123/ijspp.2020-0558

121. Ranson, C., King, M., Burnett, A., Worthington, P., & Shine, K. (2009). The effect of coaching intervention on elite fast bowling technique over a two year period. *Sports Biomech, 8*(4), 261-274. doi:https://doi.org/10.1080/14763140903469908

122. Senington, B., Lee, R. Y., & Williams, J. M. (2018). Ground reaction force, spinal kinematics and their relationship to lower back pain and injury in cricket fast bowling: A review. *Journal of back and musculoskeletal rehabilitation, 31*(4), 671-683. doi:10.3233/BMR-170851

123. Worthington, P., King, M., & Ranson, C. (2013). The influence of cricket fast bowlers' front leg technique on peak ground reaction forces. *Journal of Sports Sciences, 31*(4), 434-441. doi:https://doi.org/10.1080/02640414.2012.736628

124. Callaghan, S. J., Govus, A. D., Lockie, R. G., Middleton, K. J., & Nimphius, S. (2021). Not as simple as it seems: Front foot contact kinetics, muscle function and ball release speed in cricket pace bowlers. *J Sports Sci, 39*(16), 1807-1815. doi:10.1080/02640414.2021.1898192

125. McGrath, J. W., Neville, J., Stewart, T., Clinning, H., Thomas, B., & Cronin, J. (2022). Quantifying cricket fast bowling volume, speed and perceived intensity zone using an apple watch and machine learning. *J Sports Sci, 40*(3), 323-330. doi:10.1080/02640414.2021.1993640

126. McNamara, D. J., Gabbett, T. J., & Naughton, G. (2017). Assessment of workload and its effects on performance and injury in elite cricket fast bowlers. *Sports Med, 47*(3), 503-515. doi:https://doi.org/10.1007/s40279-016-0588-8

127. McGrath, J. W., Neville, J., Stewart, T., Lamb, M., Alway, P., King, M., & Cronin, J. (2022). The relationship between bowling intensity and ground reaction force in cricket pace bowlers. *Journal of Sports Sciences, 40*(14), 1602-1608. doi:https://doi.org/10.1080/02640414.2022.2094561

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

129. Perrett, C., Bussey, M., & Lamb, P. (2021). External workload intensity in cricket fast bowlers across maximal and submaximal intensities: Modifying playerload and imu location. *Journal of Sports Sciences*, 1-7. doi:https://doi.org/10.1080/02640414.2021.2003570