Logo

Description automatically generated A picture containing logo

Description automatically generated Text

Description automatically generated ****

**Final report**

**Developing and piloting a real-time monitoring system for sheep**

Project code: P.PSH.1205

Prepared by: Andrew Thompson

Murdoch University

Ian Harris and Mark Ferguson

neXtgen Agri

Date published: 15th April 2023

PUBLISHED BY

Meat & Livestock Australia Limited

PO Box 1961

NORTH SYDNEY NSW 2059

This is an MLA Donor Company funded project.

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

This publication is published by Meat & Livestock Australia Limited ABN 39 081 678 364 (MLA). Care is taken to ensure the accuracy of the information contained in this publication. However MLA cannot accept responsibility for the accuracy or completeness of the information or opinions contained in the publication. You should make your own enquiries before making decisions concerning your interests. Reproduction in whole or in part of this publication is prohibited without prior written consent of MLA.

# Abstract

The sheep industry has traditionally known very little about the productivity and wellbeing of individual sheep in a commercial flock. This makes both their individual management and the selection of the best individuals very difficult. Recent advances in the development of wearable accelerometer sensors together with new analytical techniques including machine learning have provided new opportunities for real-time quantitative assessments of animal behaviours related to their production and welfare. Consultation with producers has demonstrated that there is significant interest in the application of autonomous technology to sheep grazing systems. Producers participating in the project workshop identified key priorities for the design and applications for autonomous systems. Ideas included a tag-based system that could predict behaviours, such as time of parturition and feed intake, and that was able to be personalised with real-time information displayed for mobs or individual animals depending on the application. Alerts launched from the system to a phone were thought to be the most useful although the alert route could be determined based on its severity. Low level alerts could go to web or app whereas important and time critical alerts could go to phone.

Modelling completed by Farm Systems Analysis Service shows that the value proposition of this project for the Australian red meat industry has three components. The implementation of this technology would: (1) allow an improvement in profitability through enabling higher stocking rates to be managed through increasing pasture utilisation and enhancing labour use efficiency; (2) allow an improvement in productivity through more precise monitoring and selection; and (3) allow a reduction in production losses through early detection of disease, theft or predation. From the opportunities evaluated it was estimated that the successful implementation of sensors on farms across Australia would increase the gross value of production by $730m in total or $54,000/farm/year. This demonstrates the enormous potential of autonomous monitoring systems. The benefit-cost analysis completed around monitoring of lambing ewes demonstrated that the cost an autonomous system must be less than $5 per ewe to consistently provide a positive economic return.

The initial experimental phase of this project demonstrated across-flock average prediction accuracy for jaw mounted Actigraph sensors for the behaviours of grazing, ruminating and idle of 83%. This was achieved by converting the acceleration data to metrics and then training a relatively historic (Multi-Level Perceptron (MLP)) neural network. This technique has recently improved with the publication of some new metrics and has resulted in a model that predicts grazing, ruminating and idle behaviours with the accuracies of 93%, 88%, and 86% respectively. The last two years has seen the invention of transformer neural networks from the world’s three leading machine learning research organisations (Meta, Deepmind and Open AI). In December 2020 the first transformer designed to be used with multivariate time series data was published. As tri-axial acceleration is an example of a multivariate time series dataset this method looks like a promising addition to the metrics and MLP methods.

Experimentation for Phases 2 and 3 focused on developing and validating predictions for the time of parturition under synchronised and natural lambing scenarios. This used a range of analysis techniques of accelerometer data, observations of lambing behaviours and time of parturition from three experiments. The first method used, published by Smith *et al.* (2020), predicted the time of parturition with average errors ranging from 11 to 45 hours of the actual time of parturition when the duration of lambing was truncated to 10-days. This method produces variable results and does not seem to work so well when ewes are joined naturally and lamb over 4-5 weeks. Predictions over natural lambing periods will probably not be improved significantly on the back of ongoing technology improvements using this statistical approach. Work is on-going using MLPs and transformer neural network-based behaviour prediction models operating on accelerometer sensor data. This will improve with more time spent using more advanced and modern transformer orientated machine learning models.

The project moved to an image-based or machine vision approach for the real-time monitoring of natural lambing because it was the only system that met the requirements of the project. Custom-built cameras were used to test the feasibility of quantifying date of parturition in real-time over 42-days. The cameras contained three neural networks trained to recognise ewes, lambs and numbers branded onto the side of each. To date only a simple algorithm has been used to predict date of parturition from a single camera. The absolute average error for this first attempt using a very simple prediction algorithm was about 7 days, demonstrating proof of concept. The machine vision technique shows promise and should result in increasing accuracy as more time is spent developing algorithms that move away from accessing single images at a time to those that predict behaviour by tracking individual animals over a series of images i.e., ML based video analysis. It will also be improved by using larger datasets to train the constituent machine learning models that form part of the algorithm. Overall, the project has generated important information that will inform the future of autonomous sheep monitoring systems.

# Executive summary

## Background

The sheep industry has traditionally known very little about the productivity and wellbeing of individual sheep in a commercial flock. In addition to knowing little about individual animal performance, the monitoring of whole of flock health and welfare is a significant proportion of the labour use in a sheep production system. The automation of some of this monitoring is expected to both improve labour use efficiency as well as improve welfare through early detection of problems. With advances in available sensors and similar advances in analytical approaches using machine learning, the industry is now poised to make the most of these opportunities. This project undertook a range of research to provide the foundational datasets that will allow the industry to move forward into this exciting space.

## Objectives

1. Developed and verified normal behaviour algorithms on 50 ewes, on a minimum of five properties utilising different breeds of sheep under different nutritional conditions
2. Utilised behavioural algorithms developed in objective one to predict key behaviours over lambing for synchronised ewes on collaborating properties
3. Utilised outcomes from objectives one and two to support remote monitoring of a small number of ewes using real time sensors that will test the ability of the algorithms and processes developed to autonomously monitor lambing ewes and to provide remote alerts to the producer
4. Conducted a detailed market analysis, analysing the further development that would be required to enable the development and deployment of a system across the wider sheep industry. This will include:
   1. A detailed cost benefit analysis of the value of autonomous alerts at lambing supported by the experimental data outputs
   2. A review of the unit cost and proportion of the flock that would require sensors to derive the benefits of the technology
   3. A review of:
      1. Engagement with a commercial partner to investigate the development and release of a commercial version of the developments achieved
      2. A competitor analysis and value proposition-to a commercial partner-analysis
5. Engaged a minimum of 25 sheep producers in the project and surveyed their perceptions of the value of autonomous monitoring systems for their enterprises
6. Attended a minimum of two industry events to communicate the objective and outcomes of the project
7. Published a minimum of two producer and industry facing articles
8. Submitted a minimum of one peer-reviewed journal paper for review

## Methodology

* Farmer participation through a specific project workshop as well as presenting at industry days was used to ensure the project activities aligned with the interests and expectations of Australia Sheep Farmers
* A total of 5 different sites were established where 10 sheep that were fitted with a number of sensors were also videoed, these videos were later coded to identify the activity of sheep across 10 second intervals throughout the day
* A neural network was trained to convert sensor output into known sheep activity utilising the dataset generated as part of the project
* Lambing ewes that were wearing sensors were closely monitored over lambing to determine the exact time of birth
* A range of techniques were used to estimate the time of lambing from accelerometer output
* A commercial partner was identified to work towards a system that could identify lambing ewes and estimate time of birth

## Results/key findings

This project clearly identified that producers were very interested in the concept of autonomous monitoring of livestock and very willing to invest time and energy into seeing progress in this area of work. The project has developed the world’s largest data set of sheep behaviour linked to accelerometer output. The work conducted here will be fundamental in guiding future endeavours in automating aspects of livestock management. The algorithms that have been developed and refined throughout the process have demonstrated the power that machine learning will hold in the future of the grazing livestock industries. The autonomous monitoring of lambing ewes has provided several challenges. A range of aspects that the project team considered to be unique to lambing events turned out to be specific to individual sheep and not generic to all sheep. This made it difficult to define an accelerometer pattern that described when a ewe was due to lamb or had recently lambed. This prompted the project team to move toward a vision-based approach to lambing identification. This approach holds promise and has been a useful development of the project. While commercial application of these technologies to farmers still requires additional development, the project has demonstrated the significant potential that already exists to automate some aspects of livestock research and bring a new level of precision to livestock research.

## Benefits to industry

This project has thoroughly explored the sensor landscape as applicable to sheep and has built foundational data sets and techniques that can inform both future research efforts as well as commercial interests exploring this space. This work has been made publicly available so that technology developments in the future can start from a competitive advantage compared with where this project started. This project has developed a data set that is several fold larger than any database previously established. It has investigated the different algorithms that can be deployed to these types of data sets and found those that are most likely to deliver a successful outcome. The project has paved the way for more efficient and more accurate research in the future where sensors can be incorporated into research programs and animal activity can be predicted without the need to monitor them using research staff.

## Future research and recommendations

This project has set the foundation for a range of work using cameras and sensors in combination to automate several processes on Australian sheep farms. The work has demonstrated that while the technology holds significant promise and will eventually be a gamechanger for the industry, the foundational work is both difficult and risky. Future research efforts should be directed toward developing techniques that utilise camera-based approaches to ‘measure’ on-farm outcomes. The livestock industry sits at the edge of a technology revolution that will fundamentally change the way farms are operated. The transition to this new era will be considerably smoother if the likes of MLA continue to invest in technological change. The project has demonstrated that livestock research can significantly benefit from automation of data collection, and it is recommended that autonomous monitoring approaches become the norm in funded research projects.

Specifically, the project has demonstrated that:

* When utilising a neural network trained on a different data set, some labelling of data is required to ensure that predictions are relevant to a new data set
* Sensor systems based on cameras need to use video footage and animal tracking rather than still images
* The combination of metric-based behavioural models with time series transformers is the best method of prediction of time of lambing
* It is best to use both unsupervised learning and self-supervised learning to boost the effective size of the datasets
* Near real-time training of datasets on edge devices needs to be incorporated into future work to allow unsupervised learning and supervised learning to be merged in near real-time.

Table of contents

[Abstract 2](#_Toc114724968)

[Executive summary 4](#_Toc114724969)

[Background 4](#_Toc114724970)

[Objectives 4](#_Toc114724971)

[Methodology 5](#_Toc114724972)

[Results/key findings 5](#_Toc114724973)

[Benefits to industry 5](#_Toc114724974)

[Future research and recommendations 6](#_Toc114724975)

[1. Background 11](#_Toc114724976)

[1.1 Experimental objectives 11](#_Toc114724977)

[1.2 Market and strategic objectives 11](#_Toc114724978)

[1.3 Communication objectives 12](#_Toc114724979)

[2. Market analysis 12](#_Toc114724980)

[2.1 Challenges with on-animal sensors for sheep 12](#_Toc114724981)

[2.2 Potential benefits to the industry autonomous monitoring of sheep 13](#_Toc114724982)

[2.3 The current sensor landscape 14](#_Toc114724983)

[2.4 Viable options for the Australian sheep industry 15](#_Toc114724984)

[2.4.1 Digibale Smart Tag 15](#_Toc114724985)

[3.4.2 Genesmith 15](#_Toc114724986)

[3.5 Research devices versus commercial devices 16](#_Toc114724987)

[3.6 The value of early detection of dystocia 16](#_Toc114724988)

[3.7 Benefit-cost analysis of monitoring for dystocia 17](#_Toc114724989)

[3.7.1 Monitoring via daily mob checks 17](#_Toc114724990)

[3.7.2 Value of autonomous monitoring 20](#_Toc114724991)

[3.7.3 Relationship with price of implementation 22](#_Toc114724992)

[3. Methodology 23](#_Toc114724993)

[4.1 Farmer workshop 23](#_Toc114724994)

[4.2 Phase 1: Development and verification of normal behaviour algorithms 24](#_Toc114724995)

[4.2.1 Trial sites 24](#_Toc114724996)

[4.2.2 Experimental process 24](#_Toc114724997)

[4.2.3 In-flock machine-learning and predictions from Actigraph sensors 28](#_Toc114724998)

[4.2.4 Cross-flock machine-learning and predictions from Actigraph sensors 29](#_Toc114724999)

[4.3.3 Data from GPS Sensors 30](#_Toc114725000)

[4.3 Phase 2: Prediction of time of parturition when lambing is synchronised 31](#_Toc114725001)

[4.3.1 Data 31](#_Toc114725002)

[4.3.2 Predictions using statistical methods 32](#_Toc114725003)

[4.3.3 Predictions using metrics-based machine learning techniques 33](#_Toc114725004)

[4.3.4 Raw acceleration-based machine learning techniques for the prediction of time of parturition using Convolutional Neural Networks 36](#_Toc114725005)

[4.3.5 Raw acceleration-based machine learning techniques for the prediction of time of parturition using Long Short Term Memory (LSTM) 36](#_Toc114725006)

[4.3.6 Raw acceleration-based machine learning techniques for the prediction of time of parturition using Transformers 36](#_Toc114725007)

[4.3.7 Raw acceleration-based machine learning techniques for the prediction of onset of labour and parturition using combined grazing and lambing data sets 37](#_Toc114725008)

[4.4 Phase 3: Prediction of time of parturition during a natural 5-week lambing 38](#_Toc114725009)

[4.4.1 Prediction of time of parturition from Actigraph sensors 38](#_Toc114725010)

[4.4.2 Prediction of time of parturition from machine vision 38](#_Toc114725011)

[5. Results 44](#_Toc114725012)

[5.1 Farmer workshops 44](#_Toc114725013)

[5.1.1 Ideas generated for use of tag technology 44](#_Toc114725014)

[5.1.2 Fleshing out ideas in group discussion 44](#_Toc114725015)

[5.2 Phase 1: Development and verification of normal behaviour algorithms 45](#_Toc114725016)

[5.2.1 In-flock machine-learning and predictions from Actigraph sensors 45](#_Toc114725017)

[5.2.2 Cross-flock machine-learning and predictions from Actigraph sensors 46](#_Toc114725018)

[5.2.3 Data from GPS sensors 47](#_Toc114725019)

[5.3 Phase 2: Prediction of time of parturition when lambing is synchronised 48](#_Toc114725020)

[5.3.1 Predictions using statistical methods 48](#_Toc114725021)

[5.3.2 Predictions using metrics-based machine learning techniques 49](#_Toc114725022)

[5.3.3 Raw acceleration-based machine learning techniques for the prediction of time of parturition using Convolutional Neural Networks 49](#_Toc114725023)

[5.3.4 Raw acceleration-based machine learning techniques for the prediction of time of parturition using Long Short Term Memory (LSTM) 50](#_Toc114725024)

[5.3.5 Raw acceleration-based machine learning techniques for the prediction of time of parturition using Transformers 50](#_Toc114725025)

[5.3.6 Raw acceleration-based machine learning techniques for the prediction of onset of labour and parturition using combined grazing and lambing data sets 50](#_Toc114725026)

[5.4 Phase 3: Prediction of time of parturition during a natural 5-week lambing 52](#_Toc114725027)

[5.4.1 Prediction of time of parturition from Actigraph sensors 52](#_Toc114725028)

[5.4.2 Prediction of time of parturition from machine vision 52](#_Toc114725029)

[6. Discussion 54](#_Toc114725030)

[6.1 Phase 1: Development and verification of normal behaviour algorithms 54](#_Toc114725031)

[6.1.1 Machine-learning and predictions from Actigraph sensors 54](#_Toc114725032)

[6.1.2 Data from GPS Sensors 54](#_Toc114725033)

[6.2 Phase 2 and 3: Prediction of time of parturition 54](#_Toc114725034)

[6.2.1 Prediction of time of parturition from accelerometer data 54](#_Toc114725035)

[6.2.2 Prediction of time of parturition from machine vision 55](#_Toc114725036)

[7 Conclusion 55](#_Toc114725037)

[7.2 Key findings 56](#_Toc114725038)

[7.3 Benefits to industry 56](#_Toc114725039)

[8 Future research and recommendations 56](#_Toc114725040)

[9 References 57](#_Toc114725041)

[10 Appendix 59](#_Toc114725042)

[10.1. Machine Learning Metrics 59](#_Toc114725043)

[10.2. Selected metrics used for predicting time of parturition 61](#_Toc114725044)

[10.3. Published paper (Turner et al. 2022) 61](#_Toc114725045)

[10.4. Draft paper to be published (Turner et al. 2023) 62](#_Toc114725046)

# Background

The sheep industry has traditionally known very little about the productivity and wellbeing of individual sheep in a commercial flock. This makes both their individual management and the selection of the best individuals very difficult. Recent advances in technology have provided a new opportunity to bring a completely new level of individual assessment to sheep production systems. We now have available technology that can provide accelerometer and location/proximity data on individual sheep. These new sensor systems have been demonstrated to be extremely effective in matching lambs to their dams as well as predicting components of sheep behaviour. In addition to the developments in sensor technology there have also been major advances in analytical techniques through machine learning approaches.

Monitoring of various aspects of sheep production systems can make up a significant proportion of the labour use. The automation of some of this monitoring is expected to both improve labour use efficiency as well as improve welfare through early detection of problems. In order to realise this opportunity, data sets that link sensor output to actual outcome in the animal need to be established. In addition to these base datasets, there is also a need to establish the appropriate analytical technique to be used to convert the raw sensor data into an actionable output.

This project undertook a large range of field and sensor data collection and utilised the latest developments in machine learning for this project to develop and pilot a real-time monitoring system on Australian sheep farms. Further development will unlock enormous potential in the sheep industry and provide a platform for sophisticated business development. Objectives

## Experimental objectives

1. Develop and verify normal behaviour algorithms on 50 ewes, on a minimum of five properties utilising different breeds of sheep under different nutritional conditions
2. Utilised behavioural algorithms developed in objective one to predict key behaviours over lambing for synchronised ewes on collaborating properties
3. Utilised outcomes from objectives one and two to support remote monitoring of a small number of ewes using real time sensors that will test the ability of the algorithms and processes developed to autonomously monitor lambing ewes and to provide remote alerts to the producer

## Market and strategic objectives

1. Conducted a detailed market analysis, analysing the further development that would be required to enable the development and deployment of a system across the wider sheep industry. This will include:
   1. A detailed cost benefit analysis of the value of autonomous alerts at lambing supported by the experimental data outputs
   2. A review of the unit cost and proportion of the flock that would require sensors to derive the benefits of the technology
   3. A review of:
      1. Engagement with a commercial partner to investigate the development and release of a commercial version of the developments achieved
      2. A competitor analysis and value proposition-to a commercial partner-analysis

## Communication objectives

1. Engaged a minimum of 25 sheep producers in the project and surveyed their perceptions of the value of autonomous monitoring systems for their enterprises
   1. Attended a minimum of two industry events to communicate the objective and outcomes of the project
   2. Published a minimum of two producer and industry facing articles
   3. Submitted a minimum of one peer-reviewed journal paper for review

# Market analysis

This project is focussed on developing a library of accurate behaviour signatures that can be used by current or future sensor technologies to have a positive outcome on sheep welfare and sheep industry productivity. The project has focussed on both normal behaviours – grazing, walking, standing and ruminating as well as behaviours around lambing. The project has shown significant success around developing predictive algorithms for normal behaviours however indicators of lambing have proven to be more difficult.

## Challenges with on-animal sensors for sheep

The three greatest limitations to a viable on-animal sensor system for sheep are: power-use, real-time data transfer and individual sensor price point. It was anticipated that considering the rapid changes occurring in technology development, that challenges with on-sheep sensors would be overcome through the life of this project. This has proven not to be the case. Power use remains a significant challenge. The need for a sheep system on an ear-tag based system for widespread deployment presents a range of challenge around size and weight. This limits the amount of energy storage that is possible on each device. On device solar has shown to be effective in some scenarios however animals seeking shade reduces the reliability of this power source impacts considerably on the length of time the system can be deployed on an individual sheep and there is a need for whole of life deployment.

Real-time data transfer is critical to allow alert systems to be developed. This poses some challenges considering the terrain that a system would need to be deployed in. The size and form of a sheep-based system makes data transfer a challenge and there is a requirement to cover considerable distances across a sheep farm. This can either be achieved by having long range devices or relatively low-cost base stations that can be deployed in large numbers to provide a wide network. Advances in the availability of 3G and 4G networks as well as satellite do show some promise here.

The cost for on-animal devices is also a consideration for some use cases. For these cases, where individual animals each need to be fitted with a sensor, individual sensor cost remains a barrier to entry for the current technology types. For other use cases where it is only necessary to have a small number of sensors on a mob of sheep this is not an issue. In these scenarios the cost is more likely to be associated with the support network for the devices (base stations, connectivity etc) than the devices themselves.

## Potential benefits to the industry autonomous monitoring of sheep

****Modelling completed by John Young of Farm Systems Analysis Service shows that the value proposition of this project for the Australian red meat industry has three components. The implementation of this technology would: 1) allow an improvement in profitability through enabling higher stocking rates to be managed through increasing pasture utilisation and enhancing labour use efficiency; 2) allow an improvement in productivity through more precise monitoring and selection; and 3) allow a reduction in production losses through early detection of disease, theft or predation. From the opportunities evaluated it was estimated that the successful implementation of sensors on farms across Australia would increase the gross value of production by $730M in total or $54, 000/farm/year. This demonstrates the enormous potential of autonomous monitoring systems.

**Figure 1**. The industry value of autonomous monitoring to manage feed resources, improve labour efficiency, target feed to priority mobs, improve culling decisions based on the efficiency of individual ewes, improve the allocation of dry paddock feed during summer and autumn and thus reduce the costs of supplementary feeding, and improve animal performance through real-time welfare alerts.

Further details on the opportunities evaluated in Figure 1, and the magnitude of the benefits were:

1. Increasing the scale of the sheep enterprise by improving pasture production and utilisation (= $310M/year).
2. Improving labour efficiency also has a large payoff (=$250M/year). If sensor technologies could reduce by 50% the time spent monitoring livestock health & welfare status, a preliminary economic analysis suggests the average benefit could be an additional $45/ha for up to 50% of the sheep industry (Fig. 2; Farming Systems Analysis Service).
3. Improving the targeting of feed to high priority mobs (such as twin-bearing ewes) by providing feedback on animal intake levels (= $83M/year).
4. Improving culling decision for ewes in the flock using estimates of individual animal efficiency based on production and intake (= $41M/year).
5. Reducing cost of supplementary feeding through better allocation of dry paddock feed during summer and autumn = ($31M/year).
6. Improved animal performance through providing real-time animal welfare alerts (=$14M/year).



**Figure 2**. The increase in stocking rate and profit that labour-saving technology will allow.

## The current sensor landscape

A thorough analysis of current technologies in the market has been completed as part of the project. This has been reviewed based on advances in the last 12 months. This analysis was completed to investigate all of the commercial companies that are currently offering or developing sensor-based approaches to monitoring livestock. In addition to these companies there are a range of companies that are providing either human grade or research devices that are useful for proof of concept but are not able to be deployed in a commercial setting.

The companies that we identified that are developing commercial systems are:

1. **Allflex** – collar-based system for beef and dairy cattle well advanced but not sheep appropriate
2. **My Pocket Mate – Stock Keeper** – early stages of development, location tracking, no detail on individual animal data, likely to be focussed on mob data
3. **Cowlar** – development and form based on use in cows only, not suitable for sheep
4. **Foxteck** – Image based solution, no progress, not suitable for accelerometer data, focussed on cattle
5. **Smart Paddock** – GPS tracking only, cattle solution that has some early trial work being conducted in sheep
6. **Ceres tag** – location-based monitoring, animal health alerts possible, too large for implementation in sheep
7. **eShepherd by Agersens** – cattle-based solution with no adaptation available for sheep, virtual fencing, very large units
8. **Cainthus** – image analysis based, entirely focussed on dairy cows, not appropriate for accelerometer data and no work going on in sheep
9. **Herddogg** – Suitable capability, price point not suitable for individual deployment, retailing at approximately $60 individual unit
10. **AgriScan-** UHF RFID – potential to incorporate this capability but not at the right stage
11. **Vence** – cattle focussed and priced, virtual fencing plus animal monitoring, over specified for what is required for sheep industry, units not suitable for sheep
12. **Embedivet-** price point and ethical considerations for sheep rule this out
13. **Digibale Smart tag** by AWI – function and form closest to project requirements, the system has not been commercialised
14. **Genesmith –** Image analysis based, focussed on lambing time in sheep, not currently working on normal behaviours.

## Viable options for the Australian sheep industry

### Digibale Smart Tag

The Smart Tag being developed by AWI is the only tag that is currently under development specifically for sheep. It is also the only tag that is in a form that can be directly applied to sheep as an ear tag and can capture relevant information immediately. This tag system is likely to be sufficiently accurate to provide some welfare warnings and some location data but we have been unable to determine the level of accuracy that can be achieved by these tags. These tags have currently no algorithm to detect a lambing event either before or after it has occurred. A limitation with these tags is that the intention to take the tag to market is yet unclear. It is also unclear what the price point of these tags would be. We were expecting this tag to become commercially available over the life of this project but that has not occurred. However, of all the options that are currently available this tag and system is still considered the closest and most likely to deployed commercially.

### Genesmith

Genesmith is an autonomous system that uses machine vision rather than on animal sensors to monitor animals. To meet the objectives of this project we require a system that can be deployed to monitor lambing ewes. This system meets those requirements. The Genesmith system uses solar powered cameras that can run machine learning algorithms on the device. The devices rotate 360 degrees and scan the paddock. When a sheep is in range of the system the devices can detect its presence, determining its identity through automated recognition of facial features or side brand number and then determining its status and the presence or otherwise of a lamb. It is currently developed to determine the date of birth of a lamb but has not yet been trained to identify when a ewe is in the process of lambing. This system has become available during the life of the project and is the most advanced of any of the technologies we have been following. The Genesmith team are focussed on commercially releasing the product to accurately determine date of birth and the match between lamb and dam on a commercial scale. Genesmith is an off shoot of neXtgen Agri.

## Research devices versus commercial devices

This comprehensive search of commercial companies developing capability has demonstrated that there is very limited commercial activity for on-sheep devices. There are many reasons for this, the most significant one being that the individual unit cost and weight can be significantly higher on devices deployed on cattle which makes it easier to overcome some of the limitations in current technology. There are currently a range of devices that can be used for research purposes, and these have been utilised in this and associated research work. These devices and the new algorithms developed in this project have significant value for research purposes and can have a major influence on the efficiency and accuracy of research data collection. In addition, for scenarios where only a small number of animals per mob need to be monitored, there is potential for some of the systems being developed to be used and add value.

The lack of commercial device development of on-sheep devices has resulted in the project re-thinking the way sheep can be autonomously monitored. In addition, the benefit-cost analysis completed around monitoring of lambing ewes demonstrated that the cost per ewe of a system must be less than $5 per ewe. Over the life of the project the capabilities and commercial activity in machine vision have presented opportunities for the project to explore. The benefit of machine vision systems over tag-based systems is that the cost is in the individual units (which can handle hundreds of animals) rather than on a per head basis. In the last 12 months, Genesmith cameras have become available and provide the potential of monitoring ewes in large numbers without the per animal cost and are therefore well suited to monitoring lambing ewes. The lambing monitoring cameras provided by Genesmith for the natural lambing trial present the most likely commercially deployable option currently.

## The value of early detection of dystocia

The cost of dystocia to the Australian sheep industry is estimated at $780 million (Bruce *et al.* 2021). On an individual enterprise basis, the cost of ewe and lamb mortality from dystocia ranges between $2 and $24 per lambing ewe depending on the prevailing markets and frequency of dystocia in the flock. Figure 3 demonstrates the total loss in production from dystocia assuming a lamb price of $80 and ewe price of $250 across a range of enterprise scales and levels of dystocia.



**Figure 3.** Total cost to the farm of lamb and ewe mortality from dystocia assuming a ewe value of $250 and lamb value of $80.

## Benefit-cost analysis of monitoring for dystocia

The total value to an enterprise of monitoring ewes for dystocia is highly variable depending on a range of factors. The value of the practice is most sensitive to enterprise scale, prevailing lamb and ewe values and the level of dystocia normally occurring in the flock. A variety of scenarios were tested across these three factors. The net benefit of monitoring ewes for dystocia and assisting any ewes that require help was calculated using a range of assumptions. Labour was assumed to be $30/hour and it was assumed it would take 1 hour to treat each ewe with dystocia (which includes travel time to the paddock). Vehicle costs were assumed to be $0.75/km.

### Monitoring via daily mob checks

Enterprises that do currently monitor their flocks for dystocia do so by manually checking each lambing mob on a daily basis. Within this scenario, due to the extended time between checks, the survival rate of ewes with dystocia that required assistance was assumed to be 75% and survival rate of assisted lambs assumed to be 50%. Because each mob is disrupted daily within this method the increase in lamb mortality due to starvation, mismothering and exposure was assumed to be 3% (3% higher than would occur without any disruption).



**Figure 4.** Net value to the farm of monitoring for dystocia assuming a ewe value of $150 and lamb value of $60.



**Figure 5.** Net value to the farm of monitoring for dystocia assuming a ewe value of $200 and lamb value of $70.



**Figure 6.** Net value to the farm of monitoring for dystocia assuming a ewe value of $250 and lamb value of $80.



**Figure 7.** Net value to the farm of monitoring for dystocia assuming a ewe value of $300 and lamb value of $90.

### Value of autonomous monitoring

The value of autonomous monitoring was calculated assuming that labour was available on farm to attend to ewes when alerted. The cost of the labour and vehicle was still included in the analysis but only accounted for the time and cost of attending to each dystocia event not the down time in between. Within this scenario, it was assumed that it would be possible to attend to ewes in a timely manner. The survival rate of ewes with dystocia that required assistance was assumed to be 90% and survival rate of assisted lambs assumed to be 80%. Because each mob is disrupted only when a dystocia event has been detected, within this method the increase in lamb mortality due to Starvation, mismothering and exposure was assumed to be 1.5% (1.5% higher than would occur without any disruption).



**Figure 8.** Net value to the farm of autonomous monitoring for dystocia assuming a ewe value of $150 and lamb value of $60.



**Figure 9.** Net value to the farm of autonomous monitoring for dystocia assuming a ewe value of $200 and lamb value of $70.



**Figure 10.** Net value to the farm of autonomous monitoring for dystocia assuming a ewe value of $250 and lamb value of $80.



**Figure 11.** Net value to the farm of autonomous monitoring for dystocia assuming a ewe value of $300 and lamb value of $90.

### Relationship with price of implementation

The benefit-cost analysis is obviously dependent on the cost per ewe to implement a monitoring system (Table 1). To calculate these ratios, the assumption was made that dystocia (to the point that both the ewe and lamb will die without intervention) was set at 3% of ewes. These calculations do include the associated cost of intervening to increase the chance of ewe and lamb survival. These calculations do not consider the lower-level dystocia cases where lambs are eventually born but lamb survival is reduced due to a slow birth. The monitoring system would alert these slow births and allow some intervention which would increase lamb survival and the value of monitoring, but this has not been considered in this analysis.

**Table 1. The benefit-cost ratio of a dystocia monitoring system across a range of implementation costs and ewe and lamb prices.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cost per ewe of implementation | Lamb and Ewe prices | | | |
| $60 & $150 | $70 & $200 | $80 & $250 | $90 & $300 |
| $1.00 | $6.21 | $7.92 | $9.63 | $11.34 |
| $2.00 | $3.11 | $3.96 | $4.82 | $5.67 |
| $3.00 | $2.07 | $2.64 | $3.21 | $3.78 |
| $4.00 | $1.55 | $1.98 | $2.41 | $2.84 |
| $5.00 | $1.24 | $1.58 | $1.93 | $2.27 |
| $6.00 | $1.04 | $1.32 | $1.61 | $1.89 |
| $7.00 | $0.89 | $1.13 | $1.38 | $1.62 |
| $8.00 | $0.78 | $0.99 | $1.20 | $1.42 |
| $9.00 | $0.69 | $0.88 | $1.07 | $1.26 |
| $10.00 | $0.62 | $0.79 | $0.96 | $1.13 |

This analysis shows there is a significant cost of dystocia to a sheep enterprise however daily monitoring of lambing ewes can result in a negative outcome for profitability, but a positive outcome for animal welfare. An autonomous system that provides an alert when there is a problem and allows a much more targeted use of labour and can provide a positive economic outcome as well as a better welfare outcome with more ewes and lambs saved as a result of more timely intervention. The benefit-cost analysis revealed that the per ewe cost of implementation would need to be less than $5 per ewe to consistently provide a positive economic return. This doesn’t consider the other benefits of the monitoring system. However, what it does show that a lambing monitoring system must be implemented at less than $5/ewe.

# Methodology

## Farmer workshop

A farmer workshop was held at the Muresk Institute near Northam in Western Australia in August 2019 to gather further insights on how technology could be utilised on-farm. It was attended by 25 producers and industry professionals. The workshop followed a format to get as much group participation as possible:

1. *Setting Expectations:* Participants were asked to outline their name, their farm and why they came along.
2. *Technology Explanation:* The current and potential technology capabilities were then outlined as a presentation to allow participants to see what is possible.
3. *Capture of group ideas:* Participants were then asked to come up with two ideas they would use this new technology for if they had it on their farm. They were asked to share these ideas with the person next to them and then eventually with the entire group.
4. *Project explanation:* The project was then explained to the group, including the relative pay-offs that are likely to be achieved from the various potential applications of the technology.
5. *Group think:* Participants were then divided into two groups to develop up a concept of using the technology for a particular solution. The focus of these discussions was determined by the ideas that had been put forward earlier in the workshop.
6. *Presentation of ideas:* The two groups then presented their ideas back to the wider group.
7. *Summary and close:* There was then a period of general discussion before closing the workshop.

In addition to the workshop, the project was outlined at a series of industry events that were attended by over 300 participants during 2019, including:

* Muresk Institute Farm Smart Showcase near Northam, Western Australia (40 participants)
* Techspo Presentation at Wagin, Western Australia (60 participants)
* Sheeps Back Easy Sheep Day at Moora, Western Australia (150 participants)
* Southwest Prime Lamb Group at Hamilton, Victoria (30 participants)
* The Sheep’s Back Producer Advisor panel at Perth, Western Australia (15 participants)
* PlanFarm Professional Development Day at Perth, Western Australia (30 participants)

## Phase 1: Development and verification of normal behaviour algorithms

### Trial sites

Trial sites were established in Western Australia between 2018 and 2020 to collect data and video footage across five different scenarios. These scenarios were:

1. Merino ewe hoggets grazing dry pasture at the Muresk Institute near Northam
2. Adult Suffolk ewes grazing short green pasture at the Murdoch University Farm in Perth
3. Merino ewe hoggets grazing stubble at the Muresk Institute near Northam
4. Adult Merino ewes grazing ate winter pasture during lambing at Katanning
5. Adult Merino wethers grazing a vegetative barley crop at the Muresk Institute near Northam

### Experimental process

**Sensors fitted**

On each of the five sites, 10 sheep were fitted with three different sensors in three different locations on the head:

1. An Actigraph accelerometer (ActiGraph device GT3X-BT) were fitted on a halter and attached near the jaw o. The sensor is the size of a small bike light (46 mm x 33 mm x 15 mm).
2. A second smaller (23 mm x 32 mm x 7.6 mm) accelerometer sensor were mounted on a larger cattle tag and attached to the ear
3. A i-gotU GT-600 USB GPS Travel & Sports Data Logger (46 mm x 41.5 x 14 mm; 37 g) that logged location at 1-minute intervals were attached to the sheep using a neck collar

The sensors were programmed to collect data over the period of the experiment. They ran 24hrs while video collection and therefore coding could only be completed during daylight hours. Actigraph sensors sampled at 30Hz and Axivity sensors sampled at 25Hz. Examples of the sensors are shown in Figure 12.

|  |  |  |
| --- | --- | --- |
| A close-up of a race car  Description automatically generated with low confidence | Text  Description automatically generated | A picture containing indoor, iPod, white  Description automatically generated |

**Figure 12.** An ActiGraph device GT3X-BT (left), Axivity accelerometer sensor (middle) and i-gotU GT-600 USB GPS Travel & Sports Data Logger (right) used to collect behavioural data from sheep

**Video collection and coding**

Each of the 10 sheep in each experiment were numbered, by painting numbers at several sites on each side. Over 3-4 days, observers then used 3-4 separate video cameras to collect footage of the animals, rotating the individual animal that they were tracking. At the same time the observers recorded the activity that the individual sheep was undertaking while the video was being captured. This information was used to balance both the individual animal being recorded as well as the activity being recorded where possible. The duration of studies and number of observations made are outlined in Table 2. After the experiment, all the sensors were removed, and the data was downloaded. The videos had a time stamp added so that sensor output could be matched with animal behaviour.

A subset of each of the videos were then coded by separating the video into ten second blocks and assigning the block one or more behaviours. The behaviours used were sitting, standing, walking, grazing, and ruminating. Additional observations were also made within the ten second epochs: number of steps taken, time spent walking, time of transition between standing and sitting, and time of transition between sitting and standing. An observation may have a single behaviour e.g., Sitting, or a combination of behaviours e.g., Standing, Ruminating. Sometimes the combination of behaviours were concurrent (Standing, Grazing) and others were sequential (Sitting, Walking). In a prior study (Keebles 2016) each observation was coded as a single category, which resulted in lots of different behaviours as the study attempted to codify most of the combinations described in Table 3. This resulted in some confusion between the participants involved in the coding. Having several binary observation types (Sitting, Standing, Walking, Grazing, Ruminating), from which, one or more could be selected for each 10 second epoch, was a cleaner and more explicit way of describing the behaviours.

**Table 2. Study duration and number of observations for Phase 1 trial sites.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Murdoch -green pasture** | **Muresk - dry pasture** | **Muresk - stubble** | **Katanning -green pasture** | **Muresk -barley** |
| Duration | 3 days 6 hours | 3 days 21 hours | 3 days 21 hours | 22 days | 23 days |
| Number of observations | 28,730 | 31,041 | 24,406 | 37,065 | 20,993 |

**Table 3. Breakdown of observation combinations from Phase 1 trial sites.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Behaviours** | **Murdoch -green pasture** | **Muresk - dry pasture** | **Muresk - stubble** | **Katanning - green pasture** | **Muresk -barley** |
| Standing | 2,031 | 6,505 | 6,392 | 6,659 | 1,863 |
| Sitting | 6,984 | 10,626 | 2,469 | 6,919 | 9,997 |
| Walking | 476 | 1,002 | 1,715 | 501 | 104 |
| Standing, Grazing | 5,615 | 7,069 | 5,253 | 6,427 | 1,026 |
| Standing, Ruminating | 1,175 | 2,184 | 1,684 | 4,513 | 231 |
| Standing, Walking | 593 | 397 | 1,470 | 1,503 | 507 |
| Sitting, Grazing, | 298 |  | 5 | 2 | 19 |
| Sitting, Ruminating | 6,021 | 1,538 | 836 | 3,905 | 3,718 |
| Sitting, Walking | 47 | 4 | 11 | 10 | 16 |
| Sitting, Standing | 138 |  | 19 | 48 | 60 |
| Standing, Grazing, Walking | 4,871 | 929 | 4,218 | 5,866 | 3,271 |
| Standing, Walking, Ruminating | 136 | 39 | 89 | 549 | 39 |
| Sitting, Standing, Walking | 51 |  | 3 | 25 | 46 |

**Machine-learning – Actigraph and Axivity sensors**

***Categories***

Observations were split into three final behaviours: Ruminating, Grazing and Idle. Ideally “Walking” would also be included. However, sheep do not tend to spend much of their time walking, so there were not enough observations of Walking to be included as a separate final category at this stage. Additionally, it is not straightforward to define walking, because there are times when sheep are grazing when they take several steps during a ten second epoch. Perhaps two models are required, a classifier for Grazing, Ruminating, Idle and a regressor that predicts the number of steps taken.

As previously mentioned, each ten second epoch could be assigned one or more of the basic behaviours: Sitting, Standing, Walking, Grazing, Ruminating. This method was chosen as a simple and explicit way of describing the behaviour during the ten second epochs. When splitting the observations into just three final behaviours: (Ruminating, Grazing and Idle) some decisions must be made about the observations with more than one basic behaviour. For observations with just a single behaviour there is no issue i.e., observations with just one of (Sitting, Standing, Walking, Grazing, Ruminating) get mapped as follows: Sitting, Standing, Walking end up as Idle, whereas Grazing and Ruminating are mapped into their own final behaviours.

The epochs with compound behaviours still must be mapped into a single final behaviour. In the previous report, observations coded as just Walking were ignored and removed from the dataset. As were compound observations with both Ruminating and Grazing. This makes the dataset cleaner and makes the accuracy results higher as some of the observations that can be confused between final behaviours have been removed. As the final purpose of a machine learning behaviour model is to predict a full set of behaviours, all combinations of behaviours were used in the training of models for this report.

The logic used to split the observations into final behaviours was:

* Set all observations to Idle
* Set all observations with Ruminating as Ruminating
* Set All observations with Grazing as Grazing

Each subsequent line in the logic can overwrite the results of the previous line. This logic results in the mappings outlined in Table 4.

**Table 4. Observation behaviour mappings.**

|  |  |
| --- | --- |
| **Behaviour** | **Final Behaviour** |
| Standing | Idle |
| Sitting | Idle |
| Walking | Idle |
| Standing, Grazing | Grazing |
| Standing, Ruminating | Ruminating |
| Standing, Walking | Idle |
| Sitting, Grazing, | Grazing |
| Sitting, Ruminating | Ruminating |
| Sitting, Walking | Idle |
| Sitting, Standing | Idle |
| Standing, Grazing, Walking | Grazing |
| Standing, Walking, Ruminating | Ruminating |
| Sitting, Standing, Walking | Idle |

**Metrics**

Machine learning (ML) algorithms can use any sort of data to train a model. That data can be raw acceleration traces, or some metrics derived from the acceleration traces. The derived metrics can be created from the separate x, y or z axes of the tri-axial accelerometer or calculated from a combination of the three axes. Most (as of 2022) recent published studies of applying machine learning to predict sheep behaviour have used metrics derived from the acceleration traces rather than the raw acceleration traces. All the metrics published in (Emily Walton, 2018) (Jamie Barwick, 2018) (Solomon Petrus le Roux, 2017) (L. Riaboff, 2022) were used (Appendix A).

**The process of training a machine learning model with metrics**

The method of generating the ML model using data from Actigraph sensors involved the following steps:

1. Convert Actigraph gt3x files to acceleration csv files with x, y, z acceleration values and a timestamp per row
2. Convert acceleration csv files into 10 second epoch acceleration csv files with 10 seconds of x, y, z acceleration data and a timestamp on each row
3. Merge the 10 second epoch acceleration csv files with coded observation data with 10 seconds of x, y, z acceleration data, a timestamp, and a coded observation on each row
4. Convert the 10 seconds of acceleration data into a list of calculated metrics, a timestamp, and a coded observation on each row
5. Train a Random Forest Classifier or an Extra Tree Classifier on all the metrics for the sole purpose of getting an ordered list of the effective and relevant metrics
6. Train a neural network using the best metrics and the coded observations

***Axivity Ear Mounted Accelerometers***

Machine learning models trained from data collected from the ear mounted Axivity sensors were generally at least 4% less accurate than the jaw mounted Actigraph sensors. The probable cause of the loss of accuracy is due to the sheep’s ears being more susceptible to movements not being caused by their behaviour, for example wind conditions. One extreme case were the results collected from, Murdock Green Pasture. The resulting ML model had very poor accuracy whereas the corresponding jaw mounted Actigraph trained model had similar accuracy as models trained on other flocks. The period of the Murdock Green Pasture was matched to a period of very windy conditions. The conclusion being that the Axivity sensors were buffeted by high winds during the Murdock Green Pasture study to the extent that it made detecting the behaviour from those sensors difficult. Due to this effective loss of the Murdock Green Pasture Axivity dataset we focused on the jaw mounted data.

Despite the ear mounted sensors being subject to more spurious noise as compared to the jaw mounted sensor, they are still a promising avenue of research for commercial sensors.

### In-flock machine-learning and predictions from Actigraph sensors

To train a ML model there should be a the very least a training dataset and a validation dataset. The training set is used to train the ML model and the validation set is used to measure the performance by producing several performance metrics: accuracy, precision recall, F1 and confusion matrices.

(L. Riaboff, 2022) states that the following methods were used to split the acceleration datasets into a training set and a validation set.

1. Random splits (used in 63% of studies) where the partitioning into training and validation sets is done randomly.
2. Time-based splits (used in 5% of studies) where certain observation periods are used for the training set and other time periods are used for the validation sets.
3. Animal-based splits (used in 32% of studies) where some of the flock (~80%) is used for the training set and the remaining animals are used for the validation set (~20%)

None of these methods are practically good, and the first two are useless. We have done some animal-based splits which has been described as in-flock testing so that we can compare our accuracy and other performance metrics with other studies that have used method 3. The first two methods train and test on data from the same animals so the potential to generalise to unseen sheep is not measured. Method 3. will generalise across sheep, but not across different environments because all the data was collected from one flock. For any of these methods to be optimal the splits should have been three-way instead of two-way, so that a training, validation, and test set were collected. The three sets are used as follows:

1. Training set: used to train the ML model
2. Validation set: used to check that the trained ML model generalises outside of its training set
3. Test set: used to check that the trained and validated model generalises outside of its hyper-parameters (used to optimise results during validation)

None of the published studies validate and test this way, so we implemented a method of animal-based splits that is a version of k-fold cross validation and named it in-flock testing. This is effectively the method described in “Animal-based splits” and allows the us to compare our results with some of the published studies.

The Muresk Barley flock of 9 sheep were used with k-fold cross-validation where each fold involves the data from a single sheep being used as the validation set. Then each of the nine sheep are cycled through as the validation set and the results are aggregated where the weight of the results from each sheep are proportional to the number of samples for each sheep. A three-class behaviour MLP classifier (Grazing, Ruminating, Idle) was trained.

### Cross-flock machine-learning and predictions from Actigraph sensors

In-flock/animal-based validation is not a valid way to access the performance of an ML model designed to be used on flocks that were not involved with the training of the model. The scenario of a ML model being trained on several different flocks representing a spread of environments and types of sheep, and then being used to do some predictions on an unseen flock, is a far more useful scenario than those described in other publications. Consequently, our main protocol for accessing the performance of ML behaviour models is to train the model on several flocks and then train on the flock that was left out of the training i.e., cross-flock validation.

The five flocks that make up the dataset:

1. Murdoch green pasture
2. Muresk dry pasture
3. Muresk stubble
4. Katanning green pasture
5. Muresk barley

To date the most effective model trained to classify behaviour from ten second epochs of acceleration data have been small Multi-Level Perceptrons (MLP) whose features have been selected from the collection of calculated metrics listed in Appendix A. by an Extra Tree Classifier. It should be noted that the previous use of a Random Forest Classifier or Extra Tree Classifier to pre-select the metrics prior to training the MLP resulted in a local peak accuracy in the low range of 6 – 15 metrics. On extending the metric selection to a larger number it was found that the peak in the low range was in fact a local maxima and that the global peak was at a larger number of metrics (18 – 28).

The MLP neural network used was implemented in Keras and Tensorflow. It consisted of just two dense layers and a single dropout layer to restrict overfitting.

**Table 5. MLP neural network structure**

|  |  |  |
| --- | --- | --- |
| **Layer Type** | **Layer Description** | **Activation** |
| Dense | 16 neurons | relu |
| Dense | 32 neurons | relu |
| Dropout | 0.2 |  |
| Dense | 3 neurons for 3 classes (Grazing, Ruminating, Idle) | softmax |

Hyperparameters:

* Adam optimizer
* Sparse Categorical Cross-entropy loss function
* Batch size: 64

### 4.3.3 Data from GPS Sensors

To determine the error inherent in the data collected from the GPS loggers on the sheep, a GPS receiver was left in a stationary location at Kensington, Western Australia, for three days to get an idea of the noise in the GPS receiver data. Figure 13. displays the variation in latitude over the three days, where a significant proportion of the samples have errors greater than 10 meters. Variations in the longitude are of the same order.

A screenshot of a cell phone

Description automatically generated

**Figure 13.** Variation in measured latitude of a stationary receiver for i-gotU GT-600 USB GPS Travel & Sports Data Logger GPS located in Kensington, Western Australia.

## Phase 2: Prediction of time of parturition when lambing is synchronised

### Data

Two datasets have been used to develop and validate predictions of time of parturition. Firstly, data from the ‘Lambing Density’ project collected in 2016 at the University of Western Australia Research Farm near Pingelly in Western Australia and reported by Lockwood *et al*. (2018). In this experiment, 360 twin bearing Merino ewes from two replicates of high mob size (*n* = 130 ewes) and low mob size (*n* = 50 ewes) were fitted with two jaw-mounted Actigraph tri-axial accelerometers. One on the left side configured as a beacon and one on the right side configured as a receiver, as described by Sohi *et al.* (2017). All ewes were continuously monitored between 730 and 1600 hours for 11 days commencing 147-148 days after the first of 4 days of artificial insemination. A range of ewe and lamb behaviours and the time of lambing was recorded for 91 ewes as described by Lockwood *et al.* (2018), and the lambing event itself was captured on video for 56 of these ewes. Video recording were used to code various ewe behaviours over five second periods from a few minutes before until a few minutes after the birth of the first lamb.

A second data set was collected in June 2020 from lambing ewes at the Muresk Institute Smart Farm, near Northam in Western Australia. About 350 single and twin bearing ewes were artificially inseminated on the 16th and 17th of January 2020 as part of a AMSEA Sire Evaluation Trial. The ewes were fitted with a single Actigraph mounted on a halter 140 days from artificial insemination and allocated to four lambing paddocks. Forty-eight ewes were also be fitted with a GPS unit mounted to a collar. All ewes were side-branded with a unique number for clear identification of the individual ewe from a distance. The ewes were monitored continuously during daylight hours over 10 days and where possible lambing events were recorded by video. Date of lambing was recorded for all ewes. The video recording has been coded as described for the lambing density dataset. The data analysed across the two data sets is summarised in Table 6.

**Table 6. Total number of instances and number of sheep observed exhibiting behaviours.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Behaviour Description** | **Lambing Density 2016** | | **Muresk Lambing 2020** | |
|  | **Instances** | **Sheep exhibiting behaviour** | **Instances** | **Sheep exhibiting behaviour** |
| Walking but not grazing | 2390 | 56 | 3218 | 54 |
| Grazing with head down | 2297 | 53 | 1634 | 48 |
| Lying idle | 1246 | 36 | 2236 | 41 |
| Standing no push | 3497 | 56 | 8819 | 54 |
| Agitated Behaviour | 341 | 31 | 316 | 20 |
| Pawing Ground | 149 | 17 | 336 | 28 |
| Pushing standing | 339 | 29 | 1236 | 31 |
| Pushing lying | 1269 | 35 | 2384 | 42 |
| Bending neck back - skyward | 474 | 26 | 963 | 35 |
| Bending neck back - lateral | 245 | 24 | 415 | 19 |

### Predictions using statistical methods

(Daniel Smith, 2020) predicts the time of parturition using various statistical methods applied to activity values derived from the acceleration data. This method does not involve machine learning and therefore does not require any behaviour observations but does require the date and time of birth to be known to verify the accuracy of the methods. The method that resulted in the smallest mean error for the Lambing Density 2016 dataset was the Earth Mover Distance (EMD) matrix method.

The EMD matrix method involves several steps:

1. Approximately remove the effects of gravity from the accelerometer data by subtracting 30 second averages from each axis
2. Convert the tri-axial accelerometer data in activity values using the formula:
3. Each of the ten days of the study were divided into six four hours blocks: 0-4am, 4am-8am, 8am-12pm, 12pm-4pm, 4pm-8pm, 8pm-12am, resulting in 60 four-hour blocks for the whole study
4. Accumulate all the activity values (30 per second) into four-hour blocks and form a histogram of the activity values for each block
5. Sheep have a diurnal activity pattern so comparing the activity histogram for one time of day against another time of day will not work because normal diurnal variations will obscure any changes caused by a lambing event. So, four-hour blocks of time can only be compared to other four-hours blocks from different days from the same time of day.
6. The EMD algorithm allows a distance (i.e., difference) to be calculated between the activity histograms (probability distributions) of the four-hour blocks that are all from the same time of day. This results in a 10 by 6 (10 days by 6 blocks per day) matrix of EMD distances.
7. Select the maximum distance from the EMD distance matrix to indicate the four-hour block in which the parturition is most likely to have occurred.
8. Select the middle of the select four-hour block as the date and time of parturition of the first lamb.

The EMD matrix method was applied to three datasets. In each case the datasets had their time truncated to 10 days so that prediction errors could be compared directly. This involved removing some sheep whose birth dates could not be squeezed into the 10 period.

**Table 7.** **Comparison of lambing prediction errors using Smith et al Earth Mover calculations.**

|  |  |  |
| --- | --- | --- |
| **Study** | **Duration (days)** | **Mean Absolute Error (hours)** |
| Lambing Density 2016 | 10 | 10.7 |
| Muresk Lambing 2020 | 10 | 45 |
| Katanning Green Pasture 2019 | 10 | 18 |

The best result obtained by (Daniel Smith, 2020) was a mean absolute error of 5.3 hours for a study period of seven days. The shorter the study period the easier the prediction problem becomes and the lower the expected error. If Smith’s study was 10 days, then the reported error may have been in the region of ~7 hours. This seems comparable with the result we obtained for Lambing Density 2016 i.e., 10.7 hours. However, the errors for Muresk Lambing 2020 and Katanning Green Pasture 2019 were considerable larger, to the extent that the result for Muresk Lambing 2020 although better than random, is not useful. One explanation is that Smith et. al. were lucky with their dataset as we were with Lambing Density 2016. Another limitation of this method is that it assumes that a lamb is going to be born in the period of the study, be that 7 or 10 days. It doesn’t take account of the situation where the ewe does not give birth in the period of the study. Some sort of threshold technique would need to be added to the method to indicate as to whether a birth had occurred or not. As the results were quite varied across the three datasets, we did not add the threshold set and decided not to progress with this method.

### Predictions using metrics-based machine learning techniques

Machine learning can be used to train a classifier to be able to predict when a certain behaviour is happening during a study. It is theoretically possible for a neural network (NN) to be trained to recognise the actual birth event. However, the chances of being able to train a NN to recognise the birth event with just 110 (56 plus 54 from two datasets) examples is extremely low. Another option is to find a behaviour that all sheep exhibit around the time of birth and then look for the greatest concentration of that behaviour to pinpoint a date and time of parturition. The behaviour that looks to be most promising is ‘Licking after birth’ because most sheep exhibited this behaviour after parturition and because there are numerous observations of this behaviour.

Using a metrics-based ML model to predict time of parturition involved the following steps:

1. Convert Actigraph gt3x files to acceleration csv files with x, y, z acceleration values and a timestamp per row
2. Convert acceleration csv files into 5 second epoch acceleration csv files with 5 seconds of x, y, z acceleration data and a timestamp on each row
3. Merge the 5 second epoch acceleration csv files with coded observation data (licking, not-licking) resulting in 5 seconds of x, y, z acceleration data, a timestamp, and a coded observation on each row
4. Convert the 5 seconds of acceleration data into a list of calculated metrics, a timestamp, and a coded observation on each row
5. Train an Extra Tree Classifier on all the metrics for the sole purpose of getting an ordered list of the effective and relevant metrics (Appendix b.)
6. Train a neural network using the best metrics and the coded observations
7. Run the neural network over the full datasets of sheep not used training so that every five second epoch across the study has a prediction of whether the sheep was licking or not
8. Count the number of times that licking is predicted within each hour of the study
9. Plot the hourly counts of licking behaviour for each sheep for the whole study duration and look for a peak in licking behaviour and compare with the known lambing time

There were noticeable peaks in licking behaviour coinciding with the actual parturition time in about one third of the sheep in the Lambing Density 2016 study. Examples are shown in Figure 14.

**Chart

Description automatically generated**

**Chart, histogram

Description automatically generated**

**Chart, histogram

Description automatically generated**

**Figure 14.** Examples of a peak in predicted licking behaviour coinciding with the parturition event

For the remaining two thirds of sheep in the Lambing Density 2016 study, there was no discernible peak in licking behaviour as shown in Figure 15.

Chart

Description automatically generated

***A picture containing histogram

Description automatically generated***

***Chart, histogram

Description automatically generated***

**Figure 15.** Examples of an absence in a peak of predicted licking behaviour coinciding with the parturition event

### Raw acceleration-based machine learning techniques for the prediction of time of parturition using Convolutional Neural Networks

Convolutional Neural Networks (CNN) were designed specifically to classify, segment and object detect in image data. They can be adapted to classify multivariate time series data. The following variations of CNN were used to classify raw acceleration data. That involved creating matrices with dimensions of 3 x 150 out of five seconds of triaxial data. The CNNs were then trained to classify the acceleration data using the observations from the Lambing Density dataset. (Fawaz, Forestier, Idoumghar, & Muller, 2019) detailed several CNNs that had been adapted for time series classification and their publication came with a companion GIT repo. Their code was used to test the various CNN configurations on the Lambing Density 2016 dataset.

The types of CNN used:

* Fully Convolutional Neural Network (FCN)
* Residual Network (ResNet)
* Encoder: hybrid deep CNN inspired by FCN
* Convolutional Neural Network (CNN)
* Multi-Channel Deep Convolutional Neural Network (MCDCNN)

### Raw acceleration-based machine learning techniques for the prediction of time of parturition using Long Short Term Memory (LSTM)

A LSTM was coded in and was trained on the Muresk Lambing 2020 dataset for 10,000 epochs. The maximum test accuracy attained prior to overfitting was ≈70% which occurred around epoch 310 (Figure 16).

A picture containing arrow

Description automatically generated

**Figure 16.** The accuracy of the training and test sets for the LSTM on the first 310 epochs on the Muresk Lambing 2020 dataset for Phase 2 of the project

### Raw acceleration-based machine learning techniques for the prediction of time of parturition using Transformers

During the last couple of years, a lot of progress has been made with a new form of machine learning model called transformers. Primarily transformers have been created for very large language models allowing for almost human level creation of text and the creation of images. These models have been created by the leading organisations in the machine learning space (Meta, Deepmind, Google, and Open AI). Late in 2020 a version of a transformer was published that could be used for multi-variate time series data (George Zerveas, 2020). Tri-axial accelerometer data is a multivariate time series. The publication was paired with code repository that allowed the transformer to be used almost out of the box. All that was required was the acceleration data from the Actigraph and any observation data was merged into a format called sktime (Franz Király, 2019).

In the five flocks’ behavioural data it was clear that the Katanning flock was different. When used as a test set, the accuracy was ~20% lower than the other flocks. One reason for the difference could be because the Katanning flock were lambing. Perhaps pregnant sheep move differently. These led to the idea of creating a huge, labelled dataset.

Between the Lambing Density 2016 and Muresk Lambing 2020 datasets, there are over 600 sheep that have a known date for the birth of their first lamb. If a classifier wants to learn pregnant versus lambed behaviours, then every sample prior to birth is classified as pregnant and every sample after birth is classified as lamb. If the area of interest is ten days each side of the first lamb, then the dataset has around hundred million samples of five second acceleration epochs and their associated behaviour classification (pregnant or lambed).

The process to train a transformer on both the Lambing Density and Muresk Lambing dataset:

1. Convert the Actigraph gt3x files into acceleration files with a sample per row
2. Convert acceleration files into 5 second epoch files
3. Merge the five second epoch files with a pregnant or lambed classification using date of birth spreadsheets for Lambing Density and Muresk Lambing
4. Convert the five second epoch files into sktime format
5. Train and validate the performance of a multivariate time series transformer using the code base supplied by (George Zerveas, 2020)

As the entire dataset, of ~100 million samples, is too large to be used with our computer facilities, one percent of the dataset was selected resulting in a 1 million sample training set. A PC with 128MB RAM and a RTX 3060 GPU can load the dataset into memory and training a transformer within a few hours.

### **Raw acceleration-based machine learning techniques for the prediction of onset of labour and parturition using combined grazing and lambing data sets**

Real-time monitoring of pregnant ewes at the time of labour, would allow intervention with prolonged labour. To undertake this task requires the ability to detect the labour and licking behaviours to identify the start of labour, and the time of parturition. The observations for the Lambing Density 2016 and Muresk Lambing 2020 datasets focused on the parturition period. Therefore, a combined dataset utilising the grazing behaviours and the lambing behaviours was created.

As noted, there is an issue of generalising results across sheep and flocks. To address this issue with the combined dataset, fine-tuning was performed on individual sheep to create a specialised model for individual ewes. Six ewes from the Lambing Density 2016 dataset were chosen as their first birth fell within the observation period.

A CNN-LSTM was trained on the combined dataset, addressing imbalance in the data using Synthetic Minority Oversampling Technique as defined in (Kirk E. Turner, 2022). Labels (*ruminating, labour, licking, grazing, walking, idle)* were allocated to the samples by priority order, promoting the dataset specific behaviours (*ruminating* for the grazing behaviours, *labour* and *licking* for the lambing behaviours dataset), then selecting the smaller number of samples, where they overlapped within the same sample.

**Table 8**. **Mapping of labels to lambing and grazing behaviours.**

|  |  |  |
| --- | --- | --- |
| **Label** | **Lambing Behaviours** | **Grazing Behaviours** |
| ruminating |  | ruminating |
| labour | agitated, pawing ground, pushing while standing, pushing while lying, pushing neck back skyward, pushing neck back laterally, lamb born |  |
| licking | licking after birth |  |
| grazing | grazing head down | grazing |
| walking | walking not grazing | walking |
| idle | standing not pushing, lying idle | standing, sitting |

The process to train the LSTM model:

1. Starting with the coded behaviours for the Lambing Density 2016 and Muresk Lambing 2020 datasets from 4.3.4, the 5s epochs were combined into 10s epochs where the time periods were continuous.
2. The lambing behaviours were labelled with a single behaviour for each 10s epoch.
3. The grazing behaviours were labelled with a single behaviour for each 10s epoch.
4. The datasets were combined and split into 5-fold cross-validation sets ensuring all samples had an opportunity to be excluded from the training set. The training sets were then augmented with SMOTE data.
5. Train and validation of performance using the CNN-LSTM.
6. The above steps were repeated, excluding the selected ewes for the fine-tuning, creating a training and test set for the individual ewes.
7. Train and validation of the fine-tuning were performed using a warmup phase of 100 epochs, freezing the early stages of the network, before unfreezing all the layers and performing fine-tuning for 20 epochs.

## 4.4 Phase 3: Prediction of time of parturition during a natural 5-week lambing

### 4.4.1 Prediction of time of parturition from Actigraph sensors

Data were collected in June 2019 from 250 lambing ewes at the DPIRD Research Facility near Katanning in Western Australia. In this experiment, all ewes were fitted with colour mounted Actigraph tri-axial accelerometers for a 5-week lambing period. Ewes wearing neck tags with their individual identification were observed continuously during day light hours to record time and day of lambing. Actual time of lambing was recorded for 66 ewes. Time of parturition was predicted using the various statistical methods of Smith *et al* (2020) that have been previously described in section 4.3.2.

### Prediction of time of parturition from machine vision

The commercial landscape for sheep-ready sensors has not matured significantly in the period this project has been running. The only sensor system that is suitable for deployment on sheep currently is the AWI developed system. Unfortunately, this is yet to be commercialised and there is no indication of a price point for the system. A benefit-cost analysis for autonomous monitoring over lambing showed that a system would need to be deployed at a cost of less than $5 per ewe in most scenarios. To reach this price point, systems that work per mob rather than per individual are required. Machine vision is developing rapidly and providing a new opportunity for mob monitoring for important time periods like lambing. These systems can monitor large numbers of animals and can provide real time analysis of animal welfare. Hence, the project moved to an image-based approach for the real-time monitoring of a natural lambing because it was the only system that met the requirements of the project.

Custom built cameras were installed in four paddocks at the Muresk Institute Farm, near Northam in WA, to test the feasibility of quantifying date of lambing in real-time for ewes. The trial used Merino ewes that had been joined for 35 days in 2021. Single-bearing and twin-bearing ewes were side-branded before lambing with unique three-digit brands on their sides and were each allocated into two mobs for lambing. Within pregnancy status, mobs were supplementary fed during lambing with either a self-feeder or via trail feeding. Subsequently, this resulted in three different scenarios for observation via machine vision: feeding from a self-feeder, feeding from a trail, or grazing in a paddock. Cameras were positioned at the location of self-feeders or the feed trail. In addition, one of the paddocks with a self-feeder also had cameras located away from the feeder that rotated to view various parts of the paddock (Table 8). The cameras ran for the entire duration of lambing.

**Table 9**. **Paddock, ewe and camera details for the machine vision trial which used Merino ewes during lambing at Muresk in 2021.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Paddock number** | **Number of ewes and pregnancy status** | **Camera number** | **Camera location** | **Camera rotation** |
| 14 | 129 single-bearing ewes | 14 | Feed trail | Scanning 45° |
| 15 | Feed trail | Scanning 45° |
| 15 | 71 twin-bearing ewes | 13 | Feed trail | Scanning 45° |
| 17 | Feed trail | Scanning 45° |
| 16 | 137 single-bearing ewes | 4 | Self-feeder | Static |
| 5 | Paddock | Rotating 360° |
| 6 | Paddock | Rotating 360° |
| 10 | Paddock | Rotating 360° |
| 101 | Paddock | Rotating 360° |
| 103 | Self-feeder | Static |
| 26 | 70 twin-bearing ewes | 2 | Self-feeder | Static |
| 102 | Self-feeder | Static |

***Camera hardware***

The cameras were designed to take photographs, detect ewes, and lambs and the numbers branded on the ewes, and relate the appearance of a lamb/s with a ewe as the time of birth. To achieve these tasks the camera needed a high-resolution camera with a long lens, the ability to pan and enough computing power to run three neural networks. The camera had to be solar-powered and needed the ability to connect to a network. The camera included two computers; one to do the machine learning (ML) computing and a low power watchdog computer that runs 24/7 that wakes up the ML computer during daylight hours (7am to 5:30pm). Both computers need access to real-time clocks (RTC) so that the watchdog computer knows what time to wake up the ML computer and the ML computer knows what time a picture was taken.

Icon

Description automatically generated with medium confidence

**Figure 17.** Camera enclosure.

A picture containing electronics

Description automatically generated

**Figure 18.** Camera internal components.

***Camera software***

The Nvidia Jetson Nano machine learning (ML) computer contains a graphics processing unit (GPU) that is capable for running neural networks much faster than a normal computer containing just a central processing unit (CPU). Python software was written on the Jetson Nano to predict lambing datetimes for the ewes photographed. An approximate sequence of steps conducted by the software are:

1. Take a photo
2. Use a neural network-based object detector to find instances of sheep and lambs in the photo
3. Crop the instances of sheep from the original photo and use a neural network-based object detector to find a string of digits comprising the brand on the side of the ewe.
4. Crop the instances of string of digits from the cropped sheep image and use a neural network-based object detector to find digits in the string
5. Determine if the ewe with a recognised brand is standing near a lamb(s)
6. Record the number of the ewe and the information about the ewe’s proximity to lambs
7. Uses statistics to determine the birth timestamp of a ewes’ lamb(s) by calculating the transition time from when an ewe was seen without lamb(s) to when a ewe was seen with lamb(s)

The cameras experienced several hardware issues due to the short development time frame, resulting in just five cameras collecting decent datasets. Additional real time clock (RTC) problems resulted in issues with the image timestamps. The timestamps were corrected by correlating weather in images from the camera with a working RTC (camera 2), with the weather in the images from the other cameras. Having looked at some of photos of lambs that appear to be a few hours old and comparing the corrected datetimes with their known birth dates, the corrected datetimes are sometimes one day out.

A group of animals stand in a grassy field

Description automatically generated

**Figure 19. Example photo of trail from paddock 15 including camera #13 taken by camera #17.**

***Analysis of lambs per image***

Three neural networks where trained to recognise objects in the images collected by the cameras. Those objects being ewes, lambs, and numbers. The neural networks were trained with datasets of images of ewes, lambs, and numbers where the location of those items are known within the image. We hand labelled some of the images collected in this study to enhance the accuracy of the neural network object detectors. As a sanity check the object detectors were run through all the images collected by the cameras and the numbers of sheep and lambs found per image were plotted with respect to time. We would expect to see the numbers of sheep per image to stay the same while the numbers of lamb per image to increase over the duration of recording. The number of lambs trend upwards in Figure 20 confirming the expected increase in numbers of lambs being recorded.

Chart, line chart, histogram

Description automatically generated

Chart, line chart, histogram

Description automatically generated

Chart, line chart, histogram

Description automatically generated

Chart, line chart, histogram

Description automatically generatedChart, line chart, histogram

Description automatically generated

**Figure 20.** Ewe and Lamb counts per image as found by the YOLO convolutional neural network object detector for the duration of the study.

# 5. Results

## Farmer workshops

### Ideas generated for use of tag technology

The ideas that people came up with that the thought this new technology could be used for included:

1. Matching lambs to their dams and therefore yielding mapping sheep flocks
2. Birth weight (from outside the paddock)
3. Determining when ewes are having lambing difficulty
4. Defining the energy balance of ewes based on behaviour
5. Complete structural assessment of animals by monitoring their gait
6. Determining feed use efficiency in a grazing scenario
7. Measure and improve pasture utilisation
8. Determine the time to shift sheep between paddocks with a lot more accuracy
9. Determine the appropriate amount and timing of supplementation
10. Estimate feed-on-offer based on grazing behaviour
11. Monitoring of sheep normal behaviour and alert when something changes eg.
    1. Illness
    2. Flystrike
    3. Predation
    4. Theft
12. Integration with Walk-over weighing integrated to see trends between behaviour and weights
13. Use behaviour information as well as performance information from the tag to assist with animal culling
14. When animals being maintained on self-feeders, determine the amount of time each animal is on the feeder.
15. Understand animals grazing behaviour and determine how far away different animals travel away from the feeder.
16. Use the system to determine animals that are shy feeders and not getting sufficient access to the self-feeder

### Fleshing out ideas in group discussion

The groups then explored the options that would be required in a system that could be deployed on farm and investigated alert systems that would be required. These alerts would be set for any change inconsistent with ‘normal’ behaviour.

Key data to be accessed from sensor information:

1. Predicted time of lambing based on mating behaviours
2. Birth weight – if could be achieved?
3. The proportion of the mob that has lambed
4. Dam to lamb matching
5. Birth weight to weaning, define growth rates
6. Mapping the lambing areas on a farm and the relative densities of each
7. Assess joining length, self-draft those animals that don’t return to heat
8. Feed intake assessment and daily intake information, enabling supplementary feeding to be matched

The participants thought that it would be important that the application that was running a tag-based system also had the capability to integrate with other information sources, either automatically or manually. Integration with weather forecasts was one of these applications. A link to pastures from space was another similar thought. The participants also thought that it would be important to manually enter information like a change in forage type or supplementation so that the appropriate algorithms could be applied. Participants also thought it would be good to integrate with any system that is developed that can automatically assess weight or condition score. To have a useful and deployable system, participants thought that it would be important to have a personalised dashboard that had real time information displayed on either mobs or individual animals depending on the application. Alerts launched from the system to a phone were thought to be the most useful although the alert route could be determined based on its severity. Low level alerts could go to web or app whereas important and time critical alerts could go to phone.

Overall, there was a strong engagement of producers and a lot of interest in the application of this technology to sheep grazing systems.

## Phase 1: Development and verification of normal behaviour algorithms

### In-flock machine-learning and predictions from Actigraph sensors

A MLP was trained in-flock on the nine sheep in the Muresk Barley flock.

*Muresk Barley Flock In-Flock Accuracy: 90%*

Figure 21. displays a summary of prediction accuracies from recent papers in the field. Note that 68% of these accuracies were generated using validation methods that provide overly optimistic results. Achieving 90% accuracy for the Muresk Barley flock is comparable with the top 40% of publications which is a good result considering we have used the more rigorous validation technique.

Chart, histogram

Description automatically generated

**Figure 21.** Distribution of ML model prediction accuracies from (L. Riaboff, 2022).

### Cross-flock machine-learning and predictions from Actigraph sensors

The Katanning green pasture flock was significantly different from the other four flocks due to its sheep being pregnant. When the ML model was trained on the other four flocks and tested on Katanning, the prediction accuracy was low at 59% indicating that a model trained on non-pregnant sheep did not fare well when used on pregnant sheep. To access the accuracy of the technique the Katanning flock was dropped. The process became:

* Train on three flocks
* Validated on the flock left out
* Cycle each flock as the left-out flock

Chart, waterfall chart

Description automatically generated

Chart, waterfall chart

Description automatically generated

Chart

Description automatically generatedChart, waterfall chart

Description automatically generated  
**Figure 22.** Confusion matrices for where each flock is the left-out test flock.

If the ML model is trained on the above four flocks and is, then validated on the pregnant ewes in the Katanning green pasture the result is poor as is shown in Figure 23.

Chart

Description automatically generated with medium confidence

**Figure 23.** Confusion matrix for when the model is trained on four flocks and tested on Katanning.

Chart, waterfall chart

Description automatically generated

**Figure 24.** Confusion matrix where each of four flocks is used as the validation flock and the results are aggregated.

### Data from GPS sensors

Data collected from the GPS sensors found that the GPS trace spent considerable time outside the paddock boundary. This confirms the results of the stationary receiver log locations over 2 meters away from it’s know. A possible method of removing the noise from the GPS is to run a filter over the results logged by the GPS. The GPS path was found to be predominately within the paddock boundary upon doing this. However, given that some parts of the trace are still outside the paddock boundary it appears that the GPS trace would need to be averaged over a greater number of samples to realise a sensible path. As 13 samples corresponds to an approximate duration of 20 minutes, the amount of detail being lost is significant. In conclusion, the idea of simply filtering the GPS trace to remove the noise appears to be a non-starter.

One idea of removed the GPS noise is to use a Kalman filter. A Kalman filter is an algorithm which uses a series of measurements and estimates predictions over time to remove noise from a time series of measurements. Kalman filters are often used in situations requiring trajectory optimisation. There are some published accounts of using a combination of GPS measurements and accelerometer data to remove the noise from the GPS (Amin, Bin Ibne Reaz, & Arif Sobhan Bhuiyan, 2014).

Usually, the accelerometers are in a known and static orientation with respect to their vehicle. This is not the case when mounted on a sheep, which means integrating the acceleration traces into a measure of distance is not possible.

The behaviour observation data includes the number of steps taken, which will provide a measurement of the distance travelled in each ten second epoch. Perhaps combining the distance travelled and the GPS data into a Kalman filter will result in the removal of a lot of the noise inherent in the GPS measurements.

To calculate the distance travelled for each ten second epoch, a steps predictor needs to be trained from the acceleration data and the steps observations.

The number of steps listed in all the observations ranged from 0 to 23 steps in a ten second epoch. The initial attempt to train a regressor to predict the number of steps failed with a prediction accuracy of a random result (~33%). The next attempt involved dividing the steps in three buckets:

1. 0 steps
2. 1 – 4 steps
3. 5 – 23 steps

A classifier was then trained to predict which steps bucket an epoch was in and resulted in an accuracy of 61%. This value is too low for it to be integrated by a Kalman filter with the GPS data.

## Phase 2: Prediction of time of parturition when lambing is synchronised

### Predictions using statistical methods

The average error between the predicted and actual time of parturition using this method was 10.7 hours using data from the Lambing Density 2016 project. Due to the distribution of the actual birth dates and times across the 240-hour study, random guesses would have resulted in a mean error of 75 hours. This result is comparable to the 5.3-hour error reported by Smith *et al.* (2020) and would seem to validate the approach in a different lambing environment under commercial conditions. Likewise, both studies were predicting the time of parturition within 12 hours of the actual time for 81 to 84% of ewes, and the predicted error was less than 24 hours for more than 90% of the ewes in our data (Figure 25). This level of accuracy is sufficient to identify day or birth with confidence which hence contribute to increasing the accuracy of estimating the genetic merit of early-life traits such as weaning weight.

Chart, bar chart

Description automatically generated

**Figure 25**. Difference between the time of parturition predicted using the Smith *et al*. (2020) method and the actual time of parturition observed for 56 ewes for the Lambing Density 2016 dataset for Phase 2 of the project.

Accelerometer data from the Muresk Lambing 2020 dataset comprised of 16 days of accelerometer data as compared with 7 days for Smith’s dataset and 10 days for Lambing Density 2016. However, all the births were within a 10-day period, so the 16 days was truncated down to the 10-day period encompassing the births. Smith’s Earth Mover Distance (EMD) matrix method was then applied and the lambing datetime predictions had an average error of over 60 hours. This average error could be reduced to 45 hours by changing some of the parameters of Smith’s technique, e.g., reducing the data blocks from 4 hours to 3 and changing the profile of the activity histograms. Changing the parameters is not a useful method because the best parameters are never known prior to using the technique. It appears that the Smith technique does not produce uniform results across datasets.

### Predictions using metrics-based machine learning techniques

The third of sheep where the dominant peak in the licking predictions coincide with the parturition timestamp are promising. The binary licking classifier trained on the metrics had an accuracy of 72% which is 22% above a random result. Without further testing it’s not possible to know what affect the accuracy of the classifier has on the ability to identify the obvious peak in licking activity. Perhaps the current classifier could be degraded to a lower accuracy and then accessed as to the number of licking peaks that coincide with lambing. The drop in number of correlations might provide an indication as to what an increase in classifier accuracy would result in with respect to the number of accurately predicted parturition events.

### Raw acceleration-based machine learning techniques for the prediction of time of parturition using Convolutional Neural Networks

Table 10 lists the architectures and the prediction accuracy. The better TSC results (65.7 – 66.9%) are close to the value gained by the metrics-based technique (66%).

**Table 10. CNN Time Series Classifier Results on the Lambing Density 2016 dataset for Phase 2 of the project**

|  |  |
| --- | --- |
| Deep Neural Network Architecture Name | Accuracy |
| Fully Convolutional Neural Network (FCN) | 61.5% |
| Residual Network (ResNet) (Kirk E. Turner, 2022) | 61.0% |
| Encoder: hybrid deep CNN inspired by FCN | 65.7% |
| Convolutional Neural Network (CNN) | 65.5% |
| Multi-Channel Deep Convolutional Neural Network (MCDCNN) | 66.9% |

### Raw acceleration-based machine learning techniques for the prediction of time of parturition using Long Short Term Memory (LSTM)

The LSTM has an accuracy of 70% for classifying the ‘Licking after birth’ behaviour. It is at least 3 percent better than metrics-based classifiers and CNNs adapted for time series classification. There is further potential with LSTMs through filtering and splitting the acceleration data into its gravitation and activity components. (Kirk E. Turner, 2022) has shown that there is promise in using LSTM to classify sheep behaviour (see Appendix 10.3).

### Raw acceleration-based machine learning techniques for the prediction of time of parturition using Transformers

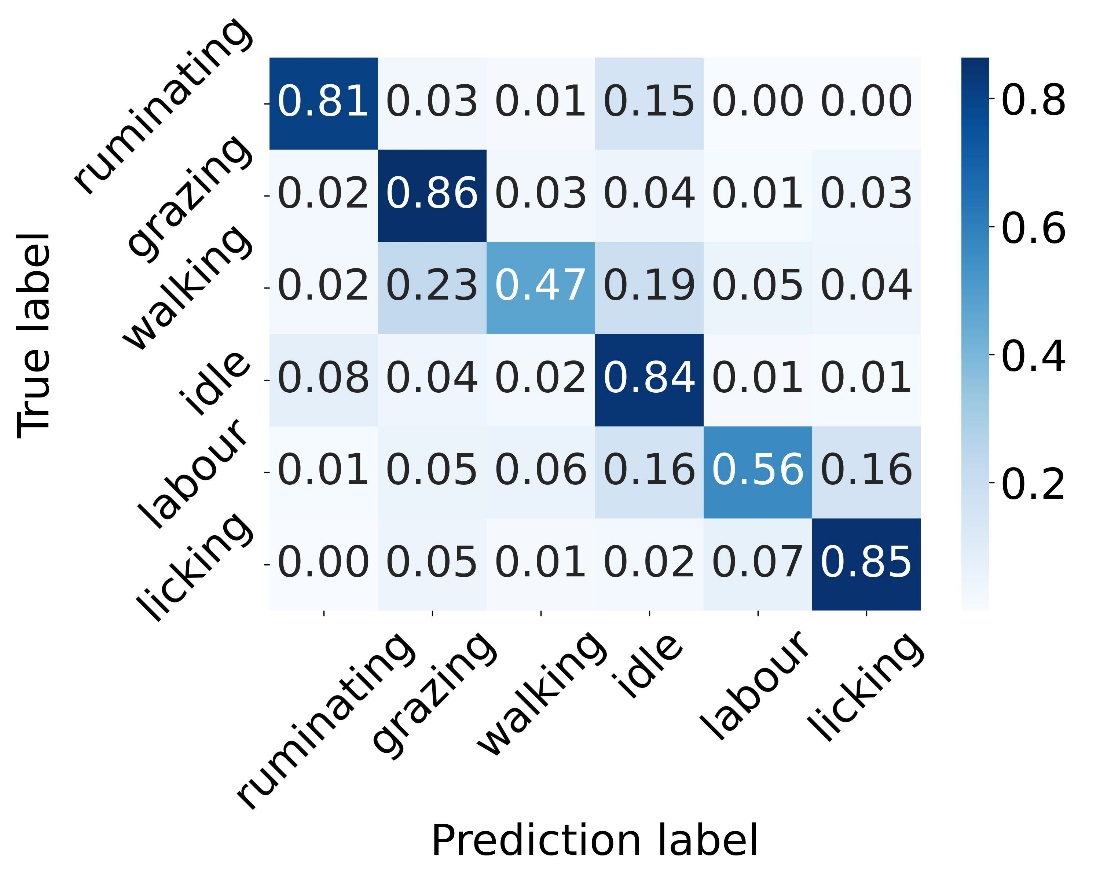
A transformer was trained on 1% of the combined Lambing Density 2016 and Muresk Lambing 2020 datasets.

*The transformer trained on 1/100 of the dataset resulted in an accuracy: 61%*

Given that only 1% per of the dataset was used, this is a promising result. A dataset was created that was 5% of the available rows, but this caused the training computer to hang. At this stage the cause is not known. It might be a memory limitation or a bug in the software from the publisher. This method shows a lot of promise, especially if the accuracy increases with dataset size.

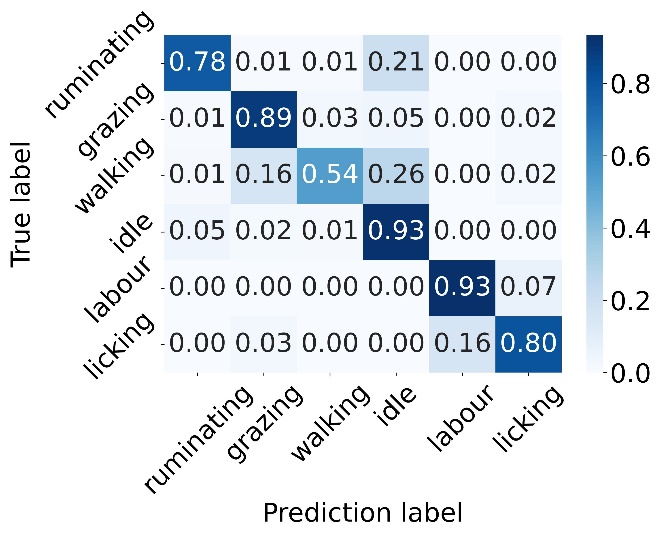
### **Raw acceleration-based machine learning techniques for the prediction of onset of labour and parturition using combined grazing and lambing data sets**

When combining the grazing and lambing behaviour datasets, an accuracy of 80.9% was achieved, with a recall of 0.85 for the *licking* behaviour. The model struggled to identify the *labour* behaviours, showing confusion with both the *licking* and *idle* behaviours (Figure 26). Full draft paper provided as Appendix 10.4.



**Figure 26**. Normalised confusion matrix for the combined grazing and lambing behaviours.

However, fine-tuning on the individual sheep increased the accuracy to an average of 86.3%, promoting the detection of the *labour* behaviours. In individual cases this resulted in a loss of performance in the *licking* behaviour recall (Figure 26), but this was not a universal result (Figure 27). Overall, fine-tuning increased the reliability of detection of the *labour* and *licking* behaviours providing the basis for further work in detecting the onset of birth, and the time of parturition.

Text

Description automatically generated with medium confidence

**Figure 27**. Normalised confusion matrix for the combined grazing and lambing behaviours, fine-tuned on individual sheep.

## Phase 3: Prediction of time of parturition during a natural 5-week lambing

### Prediction of time of parturition from Actigraph sensors

When the data was truncated to a duration of ten days, so that the prediction errors could be easily compared to the previous two datasets, the average error between the predicted and actual time of parturition for 54 ewes using this method was 18 hours. For most sheep the error was low, but a few sheep with large errors increased the average error (Figure 28).

Chart

Description automatically generated

**Figure 28**. Difference between the time of parturition predicted using the Smith *et al*. (2020) method and the actual time of parturition observed for 54 ewes at the Katanning Research Facility in 2019 for Phase 3 of the project.

### Prediction of time of parturition from machine vision

***Statistical prediction of the date of parturition of lambs from individual ewes***

The initial algorithm to predict the date of birth for a specific ewe uses the steps:

1. Find all the ewes and lambs in a picture
2. Select only the ewes where the number can be read and is valid
3. If a lamb is with 2.5 body lengths of the ewe, assume that the lamb belongs to the ewe
4. Take the day of the first instance of a lamb belonging to an ewe as the “date of parturition”

The following fig shows the error in days between what camera 6’s (paddock 16) predicted birthdates and the real birthdates. The error being negative means that the camera’s predicted date is after the true birth date.

Chart, histogram

Description automatically generated

**Figure 29.** Lambing prediction error from simple algorithm run on images from camera 6.

This is the first attempt at a very simple birth date prediction algorithm. Ideally the errors would display a large peak near 0. Negative errors mean that the system predicted the birth after the event, whereas positive errors mean that the system predicted the birth before it happened, which must be caused by another ewe’s lamb being seen near the subject ewe or the system reading the ewe’s number incorrectly. The absolute average error is ~7 days.

***Machine Learning prediction of the date of parturition of lambs from individual ewes***

The statistical prediction of parturition date relied on the camera identifying an ewe on multiple occasions, near lambs to predict that parturition had occurred. Another option is to train a ML model to detect an instance of a ewe and lamb in an image where it is highly likely that they are lamb and dam. The benefit of have a ML object detector that can recognise when it is looking at a lamb-dam pairing is that only a single image where ewe identification is successful would be required to conclude that a particular ewe has lambed.

A dataset of pictures taken by the cameras were labelled by an expert in sheep behaviour. The expert labelled the dataset by putting boxes around an ewe and lamb(s) where she was sure they were together. Ten thousand images where labelled and a YOLO version 5 object detector was trained on the labelled dataset. Unfortunately, the accuracy of the model was poor i.e., it would miss some lamb to dam pairings and would also some false positives.

*Yolo Lamb-Dam Object detector: precision: 72%, recall 62%*

At this stage we don’t know what size training set would be required for an object detector to learn the subtle differences discernable by a human expert.

# Discussion

## Phase 1: Development and verification of normal behaviour algorithms

### Machine-learning and predictions from Actigraph sensors

The process of converting accelerometer traces into metrics from which machine learning (ML) models are trained is the most common method used in recent publications regarding predicting sheep behaviour. The additional step of using a Random Forest Classifier (RFC) to select the best metrics (Jamie Barwick, 2018) prior to training a neural network has turned out to be an important step for increasing the accuracy of predictions across the five trial flocks. This process has resulted in cross-flock prediction accuracy for jaw-mounted Actigraph sensors for the behaviours of grazing, ruminating and idle approaching 90%. The cross-flock accuracy for ear-mounted Axivity sensors is 71%. However, the Axivity dataset has not been processed with the newest metrics, so its expected that its accuracy will approach 80%.

The data from these five experiments now provides a unique dataset that can create ML models to predict behaviours on other datasets which do not have any observations. This provides the opportunity of other experiments being mined for conclusions.

### 6.1.2 Data from GPS Sensors

The steps bucket classifier needs to be improved beyond its current accuracy of 61% before attempting to de-noise the GPS data with a Kalman filter. Perhaps a transformer could be used as the classifier instead of a metrics based MLP. This should be pursued because it has the chance of making the GPS data useable and because it could be used as the basis of a ‘walking’ classifier i.e., a certain number of steps per ten seconds could be classified as walking.

## Phase 2 and 3: Prediction of time of parturition

### 6.2.1 Prediction of time of parturition from accelerometer data

The method recently published by Smith *et al.* (2020) predicted the time of parturition within 11 to 60 hours of the actual time of parturition when the duration of lambing was truncated to 10-days. The precise reasons for the variable results across different lambing flocks using the Smith technique are not known. However, even an error of 2-3 days would still be acceptable for genetic analysis where adjustment for date of birth is required e.g., when generating breeding values for weaning weight. A more significant limitation of the Smith technique appears to be that the error was much greater when used to predict time of parturition over a normal lambing period of 4-5 weeks. This will probably not be improved significantly on the back of ongoing technology improvements.

The current limitations of this technique are:

* The method always assumes that a birth has taken place during the period being examined. At this stage, the method cannot determine if a birth has happened or not within a period. It just finds the four-hour block of time that a birth is most likely to have happen. The method needs to be extended with a thresholding function to also indicate the possibility that a birth has not occurred.
* The method’s most effective granularity is four hours. So, it finds the four-hour block of time that is most likely to have contained the time of birth. This limits the accuracy.

The primary goal of the machine learning based techniques is to increase the accuracy beyond what is possible with the statistical techniques developed by Smith et al. (2020). To date, this has not been successful. More work needs to be done on two aspects using neural networks to predict the time of parturition. Currently, the maximum accuracy of any of the DNNs at predicting ‘Licking after birth’ is the 66.9% achieved by the MCDNN variant of a Time Series Classifier (TSC). Further work needs to be done to increase this accuracy. This might involve using an ensemble of networks that vote for the most likely prediction outcome or perhaps using a totally different type of DNN that has also shown promise in classifying time series e.g., Long Short Term Memory (LSTM) DNN. Additionally, more work needs to be done in perfecting the technique in using a DNN licking classifier of limited accuracy to find a single dominant cluster of licking instances within the entire period of the study. Perhaps combining the Smith et al method to narrow down the lambing time prior to employing the TSC may improve performance.

### Prediction of time of parturition from machine vision

The machine vision technique using cameras in the lambing paddocks to predict date of parturition clearly shows promise and accuracy should be increased as more time is spent developing algorithms. It will also be improved by using larger datasets to train the constituent machine learning models that form part of the algorithm. The next step is to use more complicated algorithms to reduce the prediction error. The current algorithm measures the distance between a lamb and ewe in the images and decides if the lamb belongs to the ewe if the lamb is within a certain distance. The date of the first instance of a lamb being seen next to an ewe is used as the predicted birthdate. This algorithm is using the output of the sheep, lamb, and number machine learning detectors in conjunction with some logic regarding distances between sheep and ewes. This algorithm could be made more nuanced and complicated to minimise the difference between the observed and predicted birthdates. Another approach is to swap the distance related logic with another machine learning model. The model would look for instances in images of an ewe with a lamb i.e., the resulting object detector would draw boundary boxes around instances of an ewe with a lamb. This maybe the best approach because it is often obvious from the ewes’ and lambs’ body position and relative locations, that they are together. The “obviousness” is quite difficult to implement with logic, but maybe far more effectively implemented in a neural network. Recent improvements in the cameras used (improved zoom) dramatically increases the likelihood of success of this technique.

# Conclusion

The commercial landscape for sheep-ready sensors has not matured significantly in the period this project has been running. The only sensor system that is suitable for deployment on sheep currently are the AWI-developed Smart Tags. Unfortunately, this is yet to be commercialised, there is no indication of a price point for the system The current project moved to an image-based approach for the real-time monitoring of a natural lambing because it was the only system that met the requirements of the project. The machine vision technique shows promise and should result in increasing accuracy as more time is spent developing algorithms and using larger datasets to train the constituent machine learning models that form part of the algorithm. Overall, the project has generated important information that will inform the future of autonomous sheep monitoring systems.

## Key findings

* Neural networks trained on metrics derived from accelerometer data continue to improve with behaviour prediction accuracies approaching 90%.
* Neural networks will struggle to predict behaviours on datasets the are significantly different to the training dataset which implies that large diverse training datasets are required to train neural networks intended to work across diverse environments.
* Neural networks trained on raw acceleration data are improving rapidly thanks to the industries’ emphasis on multi-model networks like transformers.
* Computer vision tasks like identifying time of lambing will most likely need neural networks trained on video rather than individual images.

## Benefits to industry

This project has thoroughly explored the sensor landscape as applicable to sheep and has built foundational data sets and techniques that can inform both future research efforts as well as commercial interests exploring this space. This work has been made publicly available so that technology developments in the future can start from a competitive advantage compared with where this project started. This project has developed a data set that is several fold larger than any database previously established. It has investigated the different algorithms that can be deployed to these types of data sets and found those that are most likely to deliver a successful outcome. The potential to reduce the cost of future livestock research projects by automating some of the monitoring is an important outcome of this project. This has a completely different cost:benefit calculation than the deployment of sensors on commercial farms. Further automation of the processes used in this project can pave the way for more efficient animal research in the future.

# Future research and recommendations

* All accelerometer datasets collected to predict some behaviour using an existing neural network need to collect a small set of observations for a test set so that the model’s accuracy can be verified.
* Computer vision systems need to process video rather than individual images to implement ML trackers so that a sheep’s identity and behaviour can be linked over extended periods of time e.g., a sheep in the process of lambing will often need to be tracked to allow it to be identified.
* Combine the metric-based behavioural models with time series transformers to predict lambing i.e., feed the aggregated output of a metric-based behaviour model into a time series transformer to try and find long time-based patterns involved with lambing
* Use both unsupervised learning and self-supervised learning to boost the effective size of the datasets
* Consider near real-time training of datasets on edge devices to allow unsupervised learning and supervised learning to be merged in near real-time.
* Invest in further automation of sensor systems to allow them to be used by a range of research groups allowing enhanced efficiency of livestock research.
* Throughout the course of this project, machine learning techniques have continued to evolve, future improvements are likely to shed new light on the data set that has been collected through this project.

# References

Amin, M. S., Bin Ibne Reaz, M., & Arif Sobhan Bhuiyan, M. (2014, June). Kalman filtered GPS accelerometerbased accident detection and location system: A low-cost approach. *Current Science, 106*(11), 1548-1554

Bruce, M., Young,J., Masters,D.G., Refshauge, G., Thompson, A.N., Kenyon, PR., Behrendt, R., Lockwood, A., Miller, DW. and Jacobson, C (2021). The impact of lamb and ewe mortality associated with dystocia on the Australian and New Zealand sheep farms: a systematic review, meta-analysis and bio-economic model. Preventive Veterinary Medicine, 196, [105478]. https://doi.org/10.1016/j.prevetmed.2021.105478

Daniel Smith, J. M. (2020). Automatic detection of partuation in pregnant ewes using a three-axis accelerometer. *Computers and Electronics in Agriculture, 173*, 105392.

Emily Walton, C. C.-D. (2018, February 7). Evaluation of sampling frequency, window size and sensor position for classification of sheep behaviour. *Royal Society Open Science*.

Fawaz, H. I., Forestier, G., Idoumghar, L., & Muller, O.-A. (2019). Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery*(33), 917-963.

Franz Király, M. L. (2019, April). sktime. Retrieved from sktime.org

George Zerveas, S. J. (2020, December 8). A Transformer-based Framework for Multivariate Time Series Representation Learning. USA: arxiv. Retrieved from https://arxiv.org/pdf/2010.02803.pdf

Jamie Barwick, D. L. (2018, January 11). Predicting Lameness in Sheep Activity Using Tri-Axial Acceleration Signals. *Animals (Basel)*.

Kirk E. Turner, A. T. (2022, April 18). Deep learning based classification of sheep behaviour from accelerometer data with imbalance. *Information Processing in Agriculture*. Retrieved from https://www.sciencedirect.com/science/article/pii/S2214317322000415?via%3Dihub

L. Riaboff, L. S. (2022, January). Predicting livestock behaviour using accelerometers: A systematic review of processing techniques for ruminant behaviour prediction from raw accelerometer data. *Computers and Electronics in Agriculture, 192*.

Solomon Petrus le Roux, J. M. (2017). Animal-borne behaviour classification for sheep (Dohne Merino) and Rhinoceros (Ceratotherium simum and Diceros bicornis). *Animal Biotelemetry*.

# Appendix

## 10.1. Machine Learning Metrics

Triaxial acceleration was converted to the following metrics when used to train multi-level perceptron neural networks.

|  |  |
| --- | --- |
| **Metric** | **Equation(s) or Details1** |
| Average Signal Magnitude |  |
| Average Signal Magnitude with approximate gravity removed |  |
| Maximum Value for x, y, z |  |
| Minimum Value for x, y, z |  |
| Mean () | , , |
| Standard Deviation ( ) | , , |
| Variance for x, y, z | , , |
| Skewness for x, y, z | , , |
| Kurtosis for x, y, z | , , |
| Energy for x, y, z | , , |
| Spectral Entropy for x, y, z | , , |
| Pairwise correlation between the axes | , |
| Acceleration difference in 75th and 25th percentile |  |
| Acceleration slope difference in 75th and 25th percentile |  |
| Kurtosis of acceleration signal magnitude |  |
| Kurtosis of the slope of the acceleration signal magnitude |  |
| Standard deviation of the acceleration signal magnitude |  |
| Standard deviation of the slope of the acceleration signal magnitude |  |
| Minimum of the acceleration signal magnitude |  |
| Minimum of the slope of the acceleration signal magnitude |  |
| Maximum of the acceleration signal magnitude |  |
| Maximum of the slope of the acceleration signal magnitude |  |
| Mean of the slope of the acceleration signal magnitude |  |
| Acceleration signal zero crossing count |  |
| Acceleration signal slope zero crossing count |  |
| Dominant frequency of acceleration signal | The frequency from derived from an FFT with the largest amplitude |
| Dominant frequency of slope of the acceleration signal | The frequency from derived from an FFT with the largest amplitude |
| Spectral entropy of the acceleration signal |  |
| Spectral entropy of the slope of the acceleration signal |  |
| The area under the acceleration signal |  |
| The area under the magnitude of the acceleration signal |  |
| Movement Variation |  |
| Acceleration Entropy |  |
| Roll (Version 1) Minimum | Roll calculated using two different formulae.  Version 1: |
| Roll (Version 2) Minimum | Version 2: |
| Roll (Version 1) Maximum |  |
| Roll (Version 2) Maximum |  |
| Roll (Version 1) 25th Percentile |  |
| Roll (Version 2) 25th Percentile |  |
| Roll (Version 1) 75th Percentile |  |
| Roll (Version 2) 75th Percentile |  |
| Pitch (Version 1) Minimum | Pitch calculated using two different formulae.  Version 1: |
| Pitch (Version 2) Minimum | Version 2: |
| Pitch (version 1) Maximum |  |
| Pitch (version 2) Maximum |  |

* + - 1. P is the normalized power spectrum. FFT is fast Fourier transform.

## 10.2. Selected metrics used for predicting time of parturition

Twenty eight metrics selected by an Extra Tree Classifier listing in order of priority:

|  |  |
| --- | --- |
| **Name** | **Formula if not previously defined** |
| Y theta average |  |
| Acceleration Entropy |  |
| Movement Variation |  |
| Average Signal Magnitude with approximate gravity removed |  |
| Roll (Version 2) Minimum |  |
| Pitch (Version 1) Minimum |  |
| Standard deviation of the slope of the acceleration signal magnitude |  |
| Minimum Value for x |  |
| Roll (Version 1) Minimum |  |
| Acceleration slope difference in 75th and 25th percentile |  |
| Average Signal Magnitude |  |
| Pitch (version 1) Maximum |  |
| Acceleration signal slope zero crossing count |  |
| Acceleration magnitude 25th Percentile |  |
| Minimum Value for y |  |
| Roll (Version 2) 25th Percentile |  |
| Pitch (Version 2) Maximum |  |
| Roll (Version 1) Maximum |  |
| Pitch (Version 2) Minimum |  |
| Roll (Version 2) Maximum |  |
| Energy for z |  |
| Maximum Value for x |  |
| Acceleration magnitude 75th Percentile |  |
| Minimum Value for z |  |
| Maximum of the slope of the acceleration signal magnitude |  |
| Average dynamic body acceleration for z | TODO formula |
| Z theta average |  |
| Roll (Version 2) 75th Percentile |  |

## 10.3. Published paper (Turner et al. 2022)



## 10.4. Draft paper to be published (Turner et al. 2023)

