Paper:
Lush, L., Wilson, R., Holton, M., Hopkins, P., Marsden, K., Chadwick, D. & King, A. (2018). Classification of sheep urination events using accelerometers to aid improved measurements of livestock contributions to nitrous oxide emissions. *Computers and Electronics in Agriculture, 150*, 170-177.
http://dx.doi.org/10.1016/j.compag.2018.04.018

Title

Classification of sheep urination events using accelerometers to aid improved measurements of livestock contributions to nitrous oxide emissions

Authors

Lucy Lush[a], Rory P. Wilson[a], Mark D. Holton[b], Phil Hopkins[c], Karina A. Marsden[d], David R. Chadwick[d], Andrew J. King[a]

Affiliations

[a] Department of Biosciences, College of Science, Swansea University, Singleton Park, Swansea, SA2 8PP, UK

[b] College of Engineering, Swansea University, Bay Campus, Fabian Way, Swansea, SA1 8EN, UK

[c] College of Science, Swansea University, Singleton Park, Swansea, SA2 8PP, UK

[d] School of Environment, Natural Resources and Geography, Bangor University, Bangor, Gwynedd, LL57 2UW, UK

Corresponding author

Lucy Lush

Department of Biosciences, College of Science, Swansea University, Singleton Park, Swansea, SA2 8PP, UK

llush@hotmail.co.uk

Conflict of interest: None

Abbreviations

VeDBA: Vectorial Dynamic Body Acceleration; VeDBAs: Smoothed Vectorial Dynamic Body Acceleration; PSD: Power spectrum Density; StX, StY, StZ: Static acceleration on the X, Y, and Z axes; DyX, DyY, DyZ: Dynamic acceleration on the X, Y and Z axes; TP: True Positives; TN: True Negatives; FP: False Postives; FN: False Negatives

Abstract

Livestock emissions account for 74 % of nitrous oxide contributions to greenhouse gases in the UK. However, it remains uncertain how much is directly attributable to localised sheep urination events, which could generate nitrous oxide emission ‘hot spots’. Currently, IPCC emission factors are mainly extrapolated from lowland grazing systems and do not incorporate temporal or spatial factors related to sheep behaviour and movement. Being able to gather data that reliably measures when, where, and how much sheep urinate is necessary for accurate calculations and, to inform best management practices for reducing greenhouse gas emissions and minimizing emission-based climate change.

Animal-attached movement sensors have been shown to be effective in classifying different behaviours, albeit with varying classification accuracy depending on behaviour types. Previous studies have used accelerometers on cattle and sheep to assess active and non-active behaviours to help with grazing management, although no study has yet attempted to identify sheep urination events using this method.

We attached tri-axial accelerometer sensor tags to thirty Welsh Mountain ewes for thirty days to assess if we could identify urination events. We used random forest models using different sliding mean windows to classify behaviours. Urination had a distinctive pattern and could be identified from accelerometer data, with a 5 s window providing the best recall and a 10 s window giving the best precision. ‘State’ behaviours considered (foraging, walking, running, standing and lying down) were also identified with high recall and precision. This demonstrates the extent to which the identification of discrete ‘event’ behaviours may be sensitive to the window size used to calculate the summary statistics. The method shows promise for identifying urination in sheep and other livestock, being minimally invasive compared to other methods, and has clear potential to inform agricultural management practices and policies.

Keywords

Biologging, Climate change, Discrete behaviour, Greenhouse gas emissions, Sheep, Urination

1. Introduction

Agriculture contributes to 10 % of the total greenhouse gas emissions in the UK, with 74 % arising from nitrous oxide (N_2O) and 51 % from methane emissions (DEFRA, 2016). The latter is largely due to enteric fermentation by cattle and sheep (DEFRA, 2016), but N_2O is principally generated in the soil via nitrification and subsequent denitrification. Urine from livestock contains high concentrations of urea which can be hydrolysed in the soil to ammonium and subsequently nitrified. This means that urine patches can act as 'hot spots' for N_2O emissions (Hoogendoorn *et al.*, 2016; Marsden, Jones & Chadwick, 2016). There are uncertainties regarding the estimates of direct N_2O emission levels from urine and dung deposited by livestock, particularly from sheep and extensively grazed systems. Emission factors are currently extrapolated from cattle studies conducted in intensively managed systems (UNFCCC, 2016). The uncertainties surrounding N_2O emissions are also higher because precise measurements that incorporate spatial and temporal factors, along with animal behaviour and movement, are lacking (DEFRA, 2016). Being able to monitor when livestock urinate and understand any behavioural patterns that elucidate where and how often they urinate would help to reduce this uncertainty. Combining such data with other experimental studies to measure direct N_2O emissions released from soil due to urination in relation to edaphic factors, would enable more accurate calculations and better understanding of its contribution to climate change.

Previous studies have utilised thermistors in conjunction with GPS to determine the spatial distribution of urination events (Betteridge *et al.*, 2010). These have been modified to include a measure of urine volume and nitrogen content via refractive index (Betteridge *et al.*, 2013; Misselbrook *et al.*, 2016; Shepherd *et al.*, 2016). Flow meters in combination with data loggers have also been used to record cattle urine frequency and volume (Ravera *et al.*, 2015), but all these methods are quite invasive. The use of tri-axial accelerometers attached to a range of animals has proven to be a powerful method for determining animal behaviour (Shepard *et al.*, 2008; Nathan *et al.*, 2012; McClune *et al.*, 2014), although they have not yet been used to specifically detect urination events.

Methods used for analysing accelerometer data vary in terms of variables used to classify behaviours and the precise way the data are processed. Approaches used include template-matching (Walker *et al.*, 2015) and various clustering approaches (Sakamoto *et al.*, 2009; Nathan *et al.*, 2012), with accuracy depending on circumstance. In many clustering methods, the size of window used to summarise the data plays an important role in the accuracy with which the data can be classified (Gjoreski, Gams & Chorbev, 2010; McClune *et al.*, 2014). For example, Lush et al. (2015) used a 5 s window to classify brown hare (*Lepus europaeus*) behaviour resulting in high levels of classification accuracy for running, feeding and vigilance behaviours (> 90 %), but less than 50 % accuracy for resting, scratching and grooming. Similarly, McClune et al. (2014) used a 2 s window to analyse badger (*Meles meles*) behaviour and classified resting with nearly 100 % accuracy, but trotting, walking and snuffling was between 75 – 80 % accuracy, while Wang et al. (2015) also used a 2 s window to classify puma (*Puma concolor*) behaviour and achieved greater than 90 % classification accuracy for resting, walking, running and trotting, whilst feeding was 64 % and grooming was 0 %.

The variation in classification accuracies stem, in part, from the length of time over which a behaviour is expected to occur (Robert *et al.*, 2009). Behaviours, such as running, walking, feeding and resting that tend to occur over extended periods of minutes or longer and regarded as ‘state’ behaviours (Martin & Bateson, 1993), which facilitates their classification. In contrast, the short duration of many ‘event’ behaviours (Martin & Bateson, 1993), such as urination, makes them particularly sensitive to the window length used in the analysis (Robert *et al.*, 2009; Alvarenga *et al.*, 2015).

In this study, we used tri-axial accelerometers on Welsh Mountain ewes and then employed random forest models on the data using different sliding mean windows to assess if we could identify urination events. Accelerometers have been used previously on cattle and sheep to define active and non-active behaviours such as standing, lying down, feeding, walking and running using 3, 5, and 10 s windows (Martiskainen *et al.*, 2009; Robert *et al.*, 2009; Marais *et al.*, 2014; Alvarenga *et al.*, 2015).

However, this is the first study to attempt to use this approach to determine sheep urination events. Ewes exhibit a characteristic squat when they urinate, hence we hypothesised that a rear-mounted tri-axial accelerometer could reliably identify this behaviour. If successful it would provide a methodology that could improve the accuracy of N_2O emission estimates and help to define how much sheep contribute to greenhouse gas emissions.

2. Material and methods

The study was carried out in a semi-improved enclosed 11.5 ha upland pasture at Bangor University's Henfaes Research Centre, Abergwyngregyn, North Wales (53º13'13.75" N, 4º0'34.88" W). We attached a 'Daily Diary' tag (Wildbyte Technologies Ltd, UK) to each of 30 barren Welsh Mountain ewes for 30 d from 12th May – 16th June, 2016. Rear-mounted accelerometers were used since accelerometers mounted on a collar were not able to detect urination events. Average sheep weight was 36.8 kg (SD = 6.87 kg) and average age was 4.2 y (SD = 1.2). The work and methods used were approved by Swansea University's Animal Welfare and Ethical Review Group (Reference IP-1516-5) and by Bangor University's College of Natural Sciences Ethics Committee (Ethics approval code CNS2016DC01).

2.1 Daily Diary tags

The Daily Diaries' recorded accelerometer data at 40 Hz on each of the three orthogonal axes; X (surge), Y (sway), and Z (heave). The tags were powered by an A cell battery that was enclosed in a vacuform plastic housing and sealed using Poly Cement (Humbrol, Hornby Hobbies, UK) (Fig. 1). A small patch of wool was sheared from the rump of the sheep above their hips and the tags attached to the remaining shorter wool using a solvent free epoxy adhesive (Fig. 1). Positioning the tag at the rear of the sheep maximised the possibility of detecting the change in posture that occurs when sheep urinate. The tags weighed 50 g which was less than 0.002 % of their body weight, and therefore was likely to have minimal or no impact on sheep behaviour (Hobbs-Chell *et al.*, 2012).

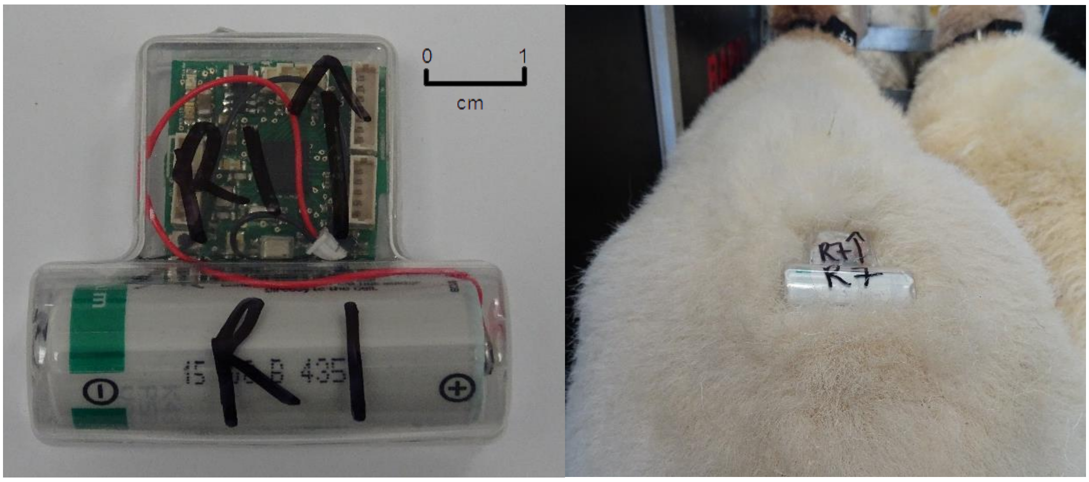

Fig. 1: Rear tag consisting of a Daily Diary and an A cell battery and a tag in position on the rear of the sheep.

2.2 Behavioural measurements

Twenty of the tagged sheep were filmed using a Panasonic HC-W570 full HD camcorder (Panasonic UK & Ireland) over four separate filming sessions to record the different types of sheep behaviour. Not all thirty sheep were filmed due to difficulties of observing all of them within the field. Sheep (n = 20) were filmed for 5 min at a time unless they moved out of view. A total of 335 min of behaviour from the video footage was logged, representing 15.9 ± 11.7 mins per sheep. Using the timestamp, the logged behaviours were synchronised to the accelerometer data to create a labelled behaviour file. An ethogram was produced of the main behaviours (Table 1). Six main behaviours were used to label the accelerometer data and in subsequent analysis. Infrequently observed behaviours were omitted. Urination events created a distinctive pattern within the acceleration trace that was identified using the observed dataset (Fig. 2). Filmed urination events had an average duration of 7 s (SD = 4.9 seconds).

Behaviour	Description	Sample (seconds)
Foraging	**Feeding with head down, small movements of head side to side and small steps forward**	**7595**
Walking	**Moving at slow pace**	**2170**
Running	**Moving at fast pace**	**126**
Standing	**Stationary with head raised**	**1653**
Lying	**Lying down with head raised or lowered**	**8345**
Urinating	**Rear of sheep lowers in a squatting position**	**127**
Scratching	Using the back leg to scratch body or head	64
Grooming	Bending head to lick leg	8
Interaction	Physical interaction between two sheep such as head butting	8

Table 1: Ethogram of sheep behaviour and number of seconds of observed behaviour logged (335 min) from video footage of 20 sheep. Behaviours in bold are those used for further analysis.

Fig. 2: Example time series of raw acceleration of the X, Y and Z axes from 40 Hz sampling rate showing a single urination event of (a) 11 s duration, and (b) 5 s identified from the observed behaviour (bounded in black box). The shaded rectangle represents a 3 s window. Urination is associated with a sharp increase in the acceleration of the X axis combined with a decrease in acceleration along the Z axis, and the Y axis generally remaining low, unless the sheep turns its head.

2.3 Random Forest model

Random Forests are machine learning models that test large numbers of regression or classification trees on a training dataset to identify the best ensemble model. R (version 3.2.5), RandomForest package (Liaw & Wiener, 2002) and RATTLE (R Analytical Tool To Learn Easily, Williams 2007) were used for analysis. Previous studies have shown the merits of using random forest as a robust method to classify behaviour from accelerometer data that also allows classification accuracy to be measured for individual behaviours (Nathan *et al.*, 2012; Lush *et al.*, 2015; Fehlmann *et al.*, 2017).

A series of descriptive statistics were calculated using a 3, 5 and 10 s sliding windows on the accelerometer data for the labelled behaviour dataset. These window sizes were chosen to allow comparison with other behaviours and other studies that used the same window sizes. The variables calculated were the static and dynamic acceleration (for each axis), the pitch, sway, Vectorial Dynamic Body Acceleration (VeDBA), smoothed VeDBA with the mean, standard deviation, minimum and maximum for all variables calculated. In addition, the maximum Power spectrum Density (PSD) and associated frequency and second maximum PSD and frequency for each axis (Wang *et al.*, 2015; Pagano *et al.*, 2017) were also calculated (Table 2, see Fehlmann *et al.*, (2017) for example R code). This gave 52 variables to be used in the initial model. 75 % of the labelled dataset was used as the training data to create the random forest model, with the remaining 25 % used to validate the model's accuracy (how well the model classified the behaviours). 500 trees were grown with 5 splits at each node. The mean decrease in accuracy was used to improve the model (Cutler *et* *al.*, 2007) and resulted in VeDBA, dynamic acceleration, and frequency variables being removed, reducing the number of variables used in the final models to 30 (Table 2). A random forest model was created for each of the time windows to assess how window size affected the accuracy with which each of the main behaviours could be classified. We were particularly interested in how well the model could classify urination events.

Variable	**Label**	**Definition**
Raw acceleration	Raw X, Y, Z	Raw output of each acceleration channel
Static acceleration*	StX, StY, StZ	$StX = \frac{1}{n}\sum_{i=0}^{n-1} RawX - i$
Dynamic acceleration	DyX, DyY, DyZ	$DyX = StX - RawX$
Vectorial Dynamic Body Acceleration	VeDBA	$\sqrt{DyX^2 + DyY^2 + DyZ^2}$
Smoothed VeDBA*	VeDBAs	VeDBA calculated over sliding mean of 3, 5 or 10 s
Pitch*	Pitch	Asin(*StZ*)
Sway*	Sway	Asin(*StY*)
Power Spectrum Density* (PSD) and Frequency	PSD1X, PSD1Y, PSD1Z, PSD2X, PSD2Y, PSD2Z,	Fast Fourier analysis to calculate dominant frequencies, and respective strengths for windows of 3, 5 or 10 s for *DyX*, *DyY* and *DyZ*. Values used were the maximum and second maximum PSD and associated frequency calculated for each axis.

Table 2: Calculated variables from the raw X, Y, and Z acceleration axes used in the models. * indicates those variables used in the final models.

2.4 Comparisons between models

To assess model performance for classifying the six behaviours, a confusion matrix was created based on the number of true positives (TP), which was the number of events correctly classified, the true negative (TN), which was those events correctly identified as being a different behaviour, the false positive (FP), where behaviours were incorrectly classified as the behaviour, and false negative (FN), where the behaviour was incorrectly classified as another behaviour (Martiskainen *et al.*, 2009; Alvarenga *et al.*, 2015). This allowed us to calculate the precision (TP / (TP+FP)) and recall/sensitivity (TP / (TP+FN)) for each time window generated from the validation data.

The Kappa statistic (Kappa = (observed accuracy – expected accuracy) / (1 – expected accuracy)), was also calculated to compare models and evaluate the classifiers by comparing the observed accuracy with the expected accuracy against random chance (Cutler *et al.*, 2007; Martiskainen *et al.*, 2009; Alvarenga *et al.*, 2015).

3 Results

3.1 Model fitting

The mean static acceleration of the Z axis was the most useful variable for classifying behaviours from our acceleration data across all three different time windows (3, 5 and 10 s models; Fig. 3). Static acceleration (Z and Y axis), pitch and smoothed VeDBA were also important for distinguishing among behaviours performed by the sheep for each of our models, but the mean smoothed VeDBA, minimum static acceleration of the Y axis (Min stY) and standard deviation of the static acceleration of the X axis (SD stX) had higher importance in the 10 s model compared to both the 3 and 5 s models.

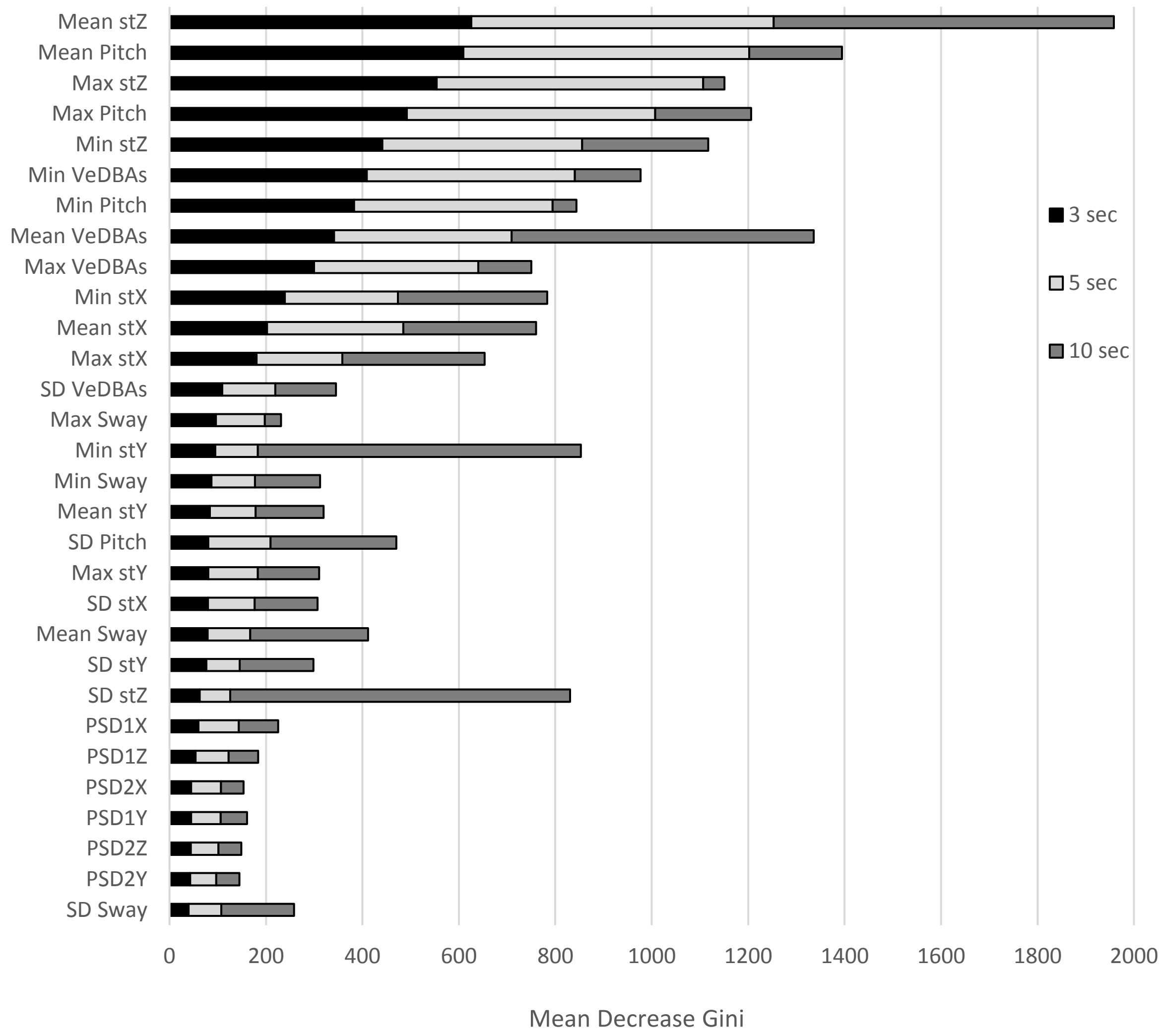

Fig. 3: Variable importance for the 3, 5 and 10 s window models. For terms see Table 2.

The 3 s window model classified most behaviours with the lowest error rate, for both the training (Table 3) and the validation data (Table 4). Foraging was an exception to this, being classified with lower error on the 10 s window for the training data (2.6 %), as was urination, which was classified with lower error on the 5 s window validation data (28.0 %), although the training data error was much higher (54.3 %).

Behaviour	**Class Error (%)**		
	3 s window	**5 s window**	**10 s window**
Foraging	3.1	3.1	2.6
Walking	9.9	13.4	19.0
Running	16.6	18.8	27.9
Standing	21.5	23.5	23.0
Lying	0.2	0.3	0.4
Urinating	31.5	54.3	67.4
OOB estimate of error rate (%)	**4.38**	**5.22**	**5.88**

Table 3: Class errors (amount of classification error) for each behaviour using the training data to create the Random Forest model for each time window.

Observed behaviour (%)	**Predicted behaviour (%)**						
	Foraging	Walking	Running	Standing	Lying	Urinating	**Class Error**
3 s window	**Overall error = 4 %, average class error = 13 %**						
Foraging	**97.7**	2.2	0.1	0.06	0.0	0.0	2.0
Walking	7.7	**91.3**	0.2	0.2	0.6	0.0	9.0
Running	10	0.0	**90.0**	0.0	0.0	0.0	10.0
Standing	17.5	1.3	0.0	**80.5**	0.3	0.5	20.0
Lying	0.0	0.0	0.0	0.0	**100**	0.0	0.0
Urinating	16.0	8.0	0.0	4.0	8.0	**64**	36.0
Performance	**Kappa = 0.945**						**Mean %**
Precision	93.9	90.3	90.0	99.1	99.7	88.9	93.7
Recall/Sensitivity	97.7	91.3	90.0	80.5	99.9	64.0	87.2
5 s window	**Overall error = 5 %, average class error = 15 %**						**Class Error**
Foraging	**97.0**	2.8	0.1	0.1	0.0	0.1	3.0
Walking	12.3	**86.5**	0.4	0.5	0.4	0.0	14.0
Running	10.0	8.3	**81.7**	0.0	0.0	0.0	18.0
Standing	19.0	4.9	0.0	**73.6**	0.0	0.3	24.0
Lying	0.0	0.0	0.1	0.0	**99.9**	0.0	0.0
Urinating	16.7	0.0	0.0	0.0	11.1	**72.2**	28.0
Performance	**Kappa = 0.927**						**Mean %**
Precision	92.6	86.6	92.5	98.2	99.8	81.3	91.8
Recall/Sensitivity	97.0	86.5	81.7	75.7	99.9	72.2	85.5
10 s window	**Overall error = 5 %, average class error = 20 %**						**Class Error**
Foraging	**97.4**	1.9	0.1	0.7	0.0	0.0	3.0
Walking	15.2	**84.4**	0.0	0.2	0.2	0.0	16.0
Running	13.2	13.2	**71.7**	0.0	1.9	0.0	28.0
Standing	19.5	1.9	0.0	**78.1**	0.3	0.3	22.0
Lying	0.1	0.0	0.0	0.0	**99.9**	0.0	0.0
Urinating	5.6	0.0	0.0	27.8	16.7	**50.0**	50.0
Performance	**Kappa = 0.926**						**Mean %**
Precision	92.1	90.0	97.4	93.8	99.7	90.0	93.8
Recall/Sensitivity	97.4	84.4	71.7	78.1	99.9	50.0	80.3
Mean precision	92.9	89.0	93.3	97.0	99.7	86.7	
Mean recall	97.3	87.4	81.1	78.1	99.9	62.1	

Table 4: Confusion matrix of the validation datasets and the performance of the Random Forest model in classifying six sheep behaviours using three different mean sliding time windows (3, 5 and 10 s). The numbers in bold are the correct classifications. (Values are percentages)

3.2 Model accuracy and performance

Overall, the 3 s window model performed the best for most of the behaviours, with the highest kappa statistic (Table 4). In fact, the kappa statistic was very high across all three models and, according to Landis and Koch's (1977) criteria, was almost perfect (0.81 – 1.00). Running was predicted with the highest precision in the 10 s window model, whereas, urination had the highest precision in the 10 s window and the highest recall in the 5 s window.

The 3 s model had the highest mean recall across all six behaviours (Table 4). All behaviours except urination had high mean precision and recall (> 75 %) across all models. Urination had high mean precision (86.7 %) but the mean recall was lower at 62.1 %.

4. Discussion

4.1 Behaviour identification in sheep

Overall, the random forest approach identified the behaviours well, with the 3 s window model performing the best for classifying ‘state’ behaviour (e.g. foraging, walking and lying) and relatively well for the ‘event’ behaviour we were interested in; that is, urination, for both precision and recall. Unsurprisingly, our ability to detect state behaviours were little affected by the size of window used, because the duration of the window was great enough to incorporate multiples of any repetitive frequency within the behaviour, while only being a small fraction of the likely length of any bout of the behaviour. However, longer time windows have been found to perform less well, as found in a study on cattle behaviour (Robert et al. 2009).

Conversely, urination, a discrete event behaviour, was the least well classified out of all the behaviours, with the degree of success depending greatly on window size. In fact, although the 5 s window model classified urination with the highest classification accuracy on the validation data the classification accuracy for the training model was only 54 %. High training data error and low validation error is indicative of a poorly fitting model (Sujatha, Prabhakar & Devi, 2013). Ideally, the validation error should be low, and the training error marginally higher. Therefore, the 3 s model, with a training error of 31.5 % and validation error of 36 %, indicates a better model fit. Model precision for urination was relatively high across all models. However, it was the recall, critical for showing how good a classification model is at correctly identifying the behaviour, which varied greatly. This could be because the window may miss either the start and/or the end of urination events, which are defined by the change in pitch (and the value of smoothed acceleration X and Z) as the sheep squats and returns to standing (Fig. 2), interspaced with lower VeDBA, because sheep remain stationary whilst urinating. Therefore, the interplay between window size and the duration of the urination event may modulate the classification error overall. In addition, the sample size of urination events was one of the lowest of our selected behaviours, as it was difficult to film, resulting in a reduced training dataset to inform the model.

Urination had a visually very distinctive pattern within the raw acceleration data (Fig. 2), which arises from the time-separated ‘squat’, ‘hold’ and ‘return-to-standing’ sequence. Such readily identifiable patterns in the accelerometer trace may be better dealt with by an algorithm that accurately defines the time-based order of important variables in sequence, as done by template matching (Walker *et al.*, 2015), for example. The immediate difficulty here, is coping with variable durations within such event behaviours. It may also be more difficult for identifying behaviours that occur simultaneously within state behaviours.

Despite the issues associated with identifying infrequent and transient behaviours like urination, this study has nonetheless identified urination events from accelerometer data. This approach, therefore, provides valuable information about urination frequency and duration. When combined with high-resolution GPS data (e.g. Haddadi et al., 2011) it can provide spatial and temporal information on urine emissions (Fig. A1). This method of using rear-mounted tags to identify urination events would not be suitable to detect urination events of rams, as they do not exhibit the characteristic squat movement that is used for ewes. However, the number of rams grazing compared to breeding ewes would be negligible and therefore would not have as much impact on greenhouse gas emissions. Given that sheep movement is not random (Harris & O'Connor, 1980) their patterns of urination are not expected to be either. In fact previous work over a six-day trial estimated that sheep deposit about 30 % of their urine over only 7.5 % of the pasture area used for grazing (Betteridge et al. 2010). This heterogeneity of urine deposition to pasture soils could create highly concentrated 'hot spot' areas that potentially release N_2O through nitrification and subsequent denitrification. By combining information on where and when sheep urinate with data on N_2O emissions from urine patches on different soil types and under different environmental conditions, could improve greenhouse gas estimates from grazed pastures.

4.2 Conclusions

We suggest that our method of using a rear-mounted tri-axial accelerometer may provide a non-invasive method to record urination events in sheep and other livestock to estimate urination patterns (frequency and duration). This would provide important information to measure livestock urination contributions to greenhouse gas emissions and to inform better agricultural management practices and policies.

Acknowledgements

We thank the Uplands N_2O Project team, and especially Mick Whelan, David Jones and Heiko Balzter for discussion and comments on this manuscript. Thanks to Jon Holmberg, Emma Archibald, Richard Lewis, Agustina di Virgilio and Ines Fürtbauer for their assistance with data collection and Gaelle Fehlmann for help with part of the analysis. Also, thanks to Mark Hughes, Llinos Hughes and team at the Henfaes Research Centre for assistance in the field. This work was funded by the Natural Environment Research Council [grant number NE/M015351/1].

References

Alvarenga, F.A.P., Borges, I., Palkovič, L., Rodina, J., Oddy, V.H. & Dobos, R.C. (2015). Using a three-axis accelerometer to identify and classify sheep behaviour at pasture. *Appl. Anim. Behav. Sci.* **181**, 91–99.

Betteridge, K., Costall, D.A., Li, F.Y., Luo, D. & Ganesh, S. (2013). Why we need to know what and where cows are urinating – a urine sensor to improve nitrogen models. *Proc. New Zeal. Grassl. Assoc. 75* **75**, 33–38.

Betteridge, K., Hoogendoorn, C., Costall, D., Carter, M. & Griffiths, W. (2010). Sensors for detecting and logging spatial distribution of urine patches of grazing female sheep and cattle. *Comput. Electron. Agric.* **73**, 66–73.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J. & Lawler, J.J. (2007). Random forests for classification in ecology. *Ecology* **88**, 2783–92.

DEFRA. (2016). Agricultural Statistics and Climate Change. *Dep. Environ. Food Rural Aff.* 1–109.

Fehlmann, G., O'Riain, M.J., Hopkins, P.W., O'Sullivan, J., Holton, M.D., Shepard, E.L.C. & King, A.J. (2017). Identification of behaviours from accelerometer data in a wild social primate. *Anim. Biotelemetry* **5**, 6.

Gjoreski, H., Gams, M. & Chorbev, I. (2010). 3-Axial Accelerometers Activity Recognition. *ICT Innov.* 51–58.

Haddadi, H., King, A.J., Wills, A.P., Fay, D., Lowe, J., Morton, A.J., Hailes, S. & Wilson, A.M. (2011). Determining association networks in social animals: Choosing spatial-temporal criteria and sampling rates. *Behav. Ecol. Sociobiol.* **65**, 1659–1668.

Harris, P.S. & O'Connor, K.F. (1980). The Grazing Behaviour of Sheep (Ovis aries) on a High-Country Summer Range in Canterbury, New Zealand. *N. Z. J. Ecol.* **3**, 85–96.

Hobbs-Chell, H., King, A.J., Sharratt, H., Haddadi, H., Rudiger, S.R., Hailes, S., Morton, A.J. &

Wilson, A.M. (2012). Data-loggers carried on a harness do not adversely affect sheep locomotion. *Res. Vet. Sci.* **93**, 549–552.

Hoogendoorn, C.J., Luo, J., Lloyd-west, C.M., Devantier, B.P., Lindsey, S.B., Sun, S., Pacheco, D., Li, Y., Theobald, P.W. & Judge, A. (2016). Agriculture , Ecosystems and Environment Nitrous oxide emission factors for urine from sheep and cattle fed forage rape (Brassica napus L .) or perennial ryegrass / white clover pasture (Lolium perenne L ./ Trifolium repens). *Agric. Ecosyst. Environ.* **227**, 11–23.

Landis, J.R. & Koch, G.G. (1977). The Measurement of Observer Agreement for Categorical Data. *Int. Biometric Soc.* **33**, 159–174.

Liaw, A. & Wiener, M. (2002). Classification and Regression by randomForest. *R News* **2**, 18–22.

Lush, L., Ellwood, S., Markham, A., Ward, A.I. & Wheeler, P. (2015). Use of tri-axial accelerometers to assess terrestrial mammal behaviour in the wild. *J. Zool.* **298**, 257–265.

Marais, J., Petrus, S., Roux, L., Wolhuter, R. & Niesler, T. (2014). Automatic classification of sheep behaviour using 3-axis accelerometer data. In *Pattern Recognition Association of South Africa*: 1–6. Puttkammer, M. & Eiselen, R. (Eds). . Cape Town, South Africa.

Marsden, K.A., Jones, D.L. & Chadwick, D.R. (2016). The urine patch diffusional area: An important N2O source? *Soil Biol. Biochem.* **92**, 161–170.

Martin, P. & Bateson, P. (1993). *Measuring behavior: An introductory guide.* Second edi. Cambridge, England: Cambridge University Press.

Martiskainen, P., Järvinen, M., Skön, J.P., Tiirikainen, J., Kolehmainen, M. & Mononen, J. (2009). Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Appl. Anim. Behav. Sci.* **119**, 32–38.

McClune, D.W., Marks, N.J., Wilson, R.P., Houghton, J., Montgomery, I.W., Mcgowan, N.E., Gormley, E. & Scantlebury, M. (2014). Tri-axial accelerometers quantify behaviour in the Eurasian badger (Meles meles): towards an automated interpretation of field data. *Anim.*

Biotelemetry **2**, 1–6.

Misselbrook, T., Fleming, H., Camp, V., Umstatter, C., Duthie, C.A., Nicoll, L. & Waterhouse, T. (2016). Automated monitoring of urination events from grazing cattle. *Agric. Ecosyst. Environ.* **230**, 191–198.

Nathan, R., Spiegel, O., Fortmann-Roe, S., Harel, R., Wikelski, M. & Getz, W.M. (2012). Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. *J. Exp. Biol.* **215**, 986–96.

Pagano, A.M., Rode, K.D., Cutting, A., Owen, M.A., Jensen, S., Ware, J.V., Robbins, C.T., Durner, G.M., Atwood, T.C., Obbard, M.E., Middel, K.R., Thiemann, G.W. & Williams, T.M. (2017). Tri-axial accelerometers remotely identify wild polar bear behaviors. *Endanger. Species Res.* **32**, 19–33.

Ravera, B.L., Bryant, R.H., Cameron, K.C., Di, H.J., Edwards, G.R. & Smith, N. (2015). Use of a urine meter to detect variation in urination behaviour of dairy cows on winter crops. *Proc. New Zeal. Soc. Anim. Prod.* **75**, 84–88.

Robert, B., White, B.J., Renter, D.G. & Larson, R.L. (2009). Evaluation of three-dimensional accelerometers to monitor and classify behavior patterns in cattle. *Comput. Electron. Agric.* **67**, 80–84.

Sakamoto, K.Q., Sato, K., Ishizuka, M., Watanuki, Y., Takahashi, A., Daunt, F. & Wanless, S. (2009). Can ethograms be automatically generated using body acceleration data from free-ranging birds? *PLoS One* **4**, e5379.

Shepard, E., Wilson, R., Quintana, F., Gómez Laich, a, Liebsch, N., Albareda, D., Halsey, L., Gleiss, a, Morgan, D., Myers, A., Newman, C. & McDonald, D. (2008). Identification of animal movement patterns using tri-axial accelerometry. *Endanger. Species Res.* **10**, 47–60.

Shepherd, M.A., Welten, B.G., Costall, D., Cosgrove, G.P., Pirie, M. & Betteridge, K. (2016). Evaluation of refractive index for measuring urinary nitrogen concentration in a sensor worn by grazing female cattle. *New Zeal. J. Agric. Res.* **60**, 23–31.

Sujatha, M., Prabhakar, S. & Devi, G. (2013). A Survey of Classification Techniques in Data Mining. *Int. J. Innov. Eng. Technol.* **2**, 86–92.

UNFCCC. (2016). UK Greenhouse Gas Inventory, 1990 to 2014: Annual report for submission under the framework convention on climate change. *Dep. Energy Clim. Chang.* 1–569.

Walker, J.S., Jones, M.W., Laramee, R.S., Holton, M.D., Shepard, E.L., Williams, H.J., Scantlebury, D.M., Marks, N.J., Magowan, E. a, Maguire, I.E., Bidder, O.R., Di Virgilio, A. & Wilson, R.P. (2015). Prying into the intimate secrets of animal lives; software beyond hardware for comprehensive annotation in “Daily Diary” tags. *Mov. Ecol.* **3**, 1–16.

Wang, Y., Nickel, B., Rutishauser, M., Bryce, C., Williams, T., Elkaim, G. & Wilmers, C. (2015). Movement, resting, and attack behaviors of wild pumas are revealed by tri-axial accelerometer measurements. *Mov. Ecol.* **3**, 2.

Williams, G. (2007). Rattle: A graphical user interface for data mining in R using GTK. *R J.* **1**, 45–55.

Appendix

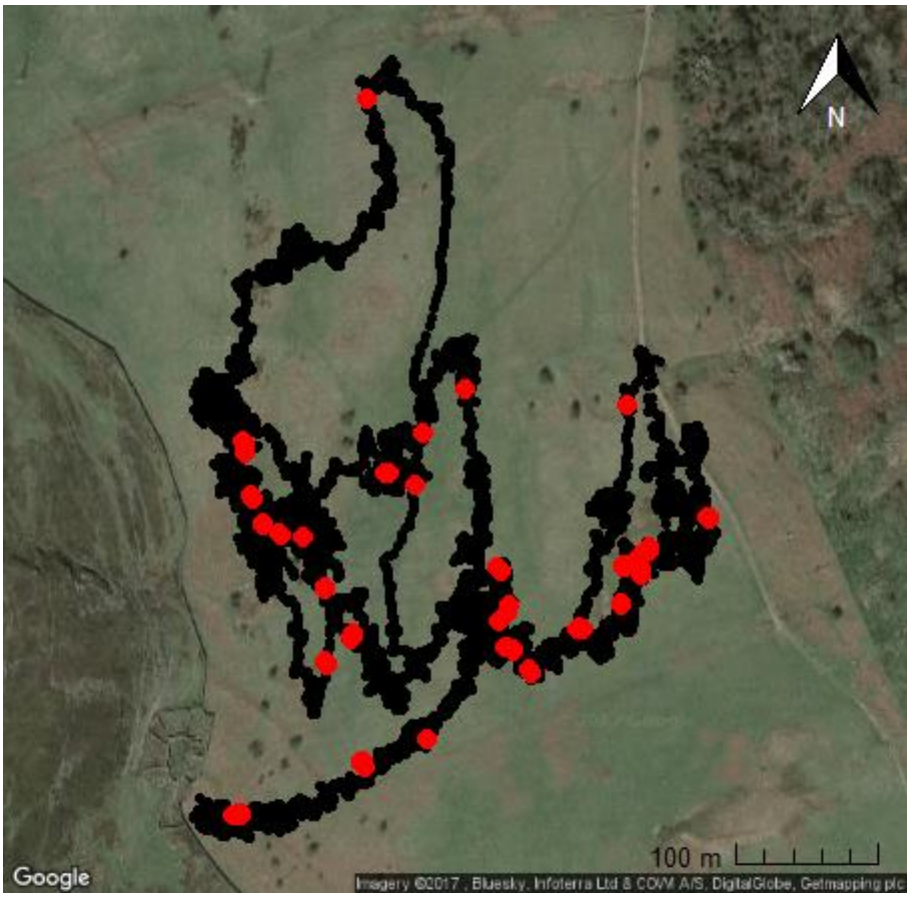

Fig. A1: Movement of 1 sheep over the duration of a day plotted on the study site. Red dots are urination events.