

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system

Animal Biotelemetry (2015) 3:15

doi:10.1186/s40317-015-0045-8

Jorge A Vázquez Diosdado (javazqa@essex.ac.uk)
Zoe E Barker (Zoe.Barker@writtle.ac.uk)
Holly R Hodges (98347815@writtle.ac.uk)
Jonathan R Amory (Jonathan.Amory@writtle.ac.uk)
Darren P Croft (D.P.Croft@exeter.ac.uk)
Nick J Bell (njbell@rvc.ac.uk)
Edward A Codling (ecodling@essex.ac.uk)

ISSN 2050-3385

Article type Research

Submission date 17 February 2015

Acceptance date 25 May 2015

Article URL <http://dx.doi.org/10.1186/s40317-015-0045-8>

For information about publishing your research in BioMed Central journals, go to
<http://www.biomedcentral.com/info/authors/>

Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system

Jorge A Vázquez Diosdado^{1*}

* Corresponding author

Email: javazqa@essex.ac.uk

Zoe E Barker²

Email: Zoe.Barker@writtle.ac.uk

Holly R Hodges²

Email: 98347815@writtle.ac.uk

Jonathan R Amory²

Email: Jonathan.Amory@writtle.ac.uk

Darren P Croft³

Email: D.P.Croft@exeter.ac.uk

Nick J Bell⁴

Email: njbell@rvc.ac.uk

Edward A Codling¹

Email: ecodling@essex.ac.uk

¹ Department of Mathematical Sciences, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK

² Writtle College, Chelmsford, Essex CM1 3RR, UK

³ Center for Research in Animal Behaviour, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QG, UK

⁴ Royal Veterinary College, Hawkshead Lane, North Mymms, Hatfield AL9 7TA, UK

Abstract

Background

Advances in bio-telemetry technology have made it possible to automatically monitor and classify behavioural activities in many animals, including domesticated species such as dairy cows. Automated behavioural classification has the potential to improve health and welfare monitoring processes as part of a Precision Livestock Farming approach. Recent studies have used accelerometers and pedometers to classify behavioural activities in dairy cows, but such approaches often cannot discriminate accurately between biologically important behaviours

such as feeding, lying and standing or transition events between lying and standing. In this study we develop a decision-tree algorithm that uses tri-axial accelerometer data from a neck-mounted sensor to both classify biologically important behaviour in dairy cows and to detect transition events between lying and standing.

Results

Data were collected from six dairy cows that were monitored continuously for 36 h. Direct visual observations of each cow were used to validate the algorithm. Results show that the decision-tree algorithm is able to accurately classify three types of biologically relevant behaviours: lying (77.42 % sensitivity, 98.63 % precision), standing (88.00 % sensitivity, 55.00 % precision), and feeding (98.78 % sensitivity, 93.10 % precision). Transitions between standing and lying were also detected accurately with an average sensitivity of 96.45 % and an average precision of 87.50 %. The sensitivity and precision of the decision-tree algorithm matches the performance of more computationally intensive algorithms such as hidden Markov models and support vector machines.

Conclusions

Biologically important behavioural activities in housed dairy cows can be classified accurately using a simple decision-tree algorithm applied to data collected from a neck-mounted tri-axial accelerometer. The algorithm could form part of a real-time behavioural monitoring system in order to automatically detect dairy cow health and welfare status.

Keywords

Behavioural classification, Decision-tree algorithm, Precision livestock farming, Tri-axial accelerometer, Reality mining

Background

Over the past decade, there has been a huge increase in the use of remote monitoring devices such as global positioning (GPS) trackers, location sensors, proximity loggers and accelerometers for automated recording of both human and animal behaviour [1–16]. By necessity, this has led to the need for more efficient and accurate methods of analysing the vast amounts of movement and behavioural data that are being collected [17]. Data from accelerometers have frequently been used to monitor, classify and infer the behaviour of humans. For example, hidden Markov models (HMMs) have been used to classify human physical activity using data from accelerometers positioned at key points on the human body [3]. In addition, accelerometers have been used in wearable digital sensors that can detect falls in elderly patients [4]. In many cases, there is high crossover between the methodological approaches and objectives used to collect and classify behavioural data in humans and animals. This has led to calls for a more integrated approach for ‘reality mining’ of these data sets and for more cross-fertilisation of ideas between disciplines [17]. An example of this more integrated approach is a recent study by Banerjee et al. [18], who developed a method to detect jumps in laying hens based on some of the key features that are used to estimate forces during human vertical jumps [19].

Due to their small size and weight, low cost and their potential ability to record high resolution behavioural data for days or months at a time, bio-loggers and bio-telemetry devices are increasingly being used to monitor the entire populations of animals in order to infer both individual-level and social behaviour at a range of spatio-temporal scales [17, 20]. Using algorithms for reality mining of this type of individual and social behavioural data can provide new insights into dynamic processes such as disease transmission, as well as group structure and hierarchy, cooperation between individuals and other social behaviours [17]. One of the first studies that used accelerometers to identify and classify behavioural activities in free-ranging wild animals was undertaken by Yoda et al. [21]. Subsequently, studies on wild animals that specifically use tri-axial accelerometer data have been undertaken. For example, Nathan et al. [5] illustrated the general concepts and tools for using tri-axial accelerometer data in free-ranging vultures, while Resheff et al. [6] developed a free-access web application to classify such behaviours. Shepard et al. [15] used similar methods to identify a range of movement behaviour patterns across different wild animal species, and McClune et al. [22] specifically applied such techniques to quantify and automatically interpret behaviour in the Eurasian badger (*Meles meles*). The use of tri-axial accelerometers to determine behaviour has also been undertaken with domesticated animal species. Moreau et al. [7] used a threshold value approach with tri-axial accelerometer data to classify three different behaviours in goats. Martiskainen et al. [8] developed a method that uses accelerometer data and multiclass support vector machines (SVM) to automatically classify several behaviours in dairy cows. In a similar study, Robert et al. [9] implemented a decision-tree algorithm to classify different behaviours in cattles. Although these approaches all demonstrate the potential for this type of technology, there nevertheless remain a number of limitations to be overcome. For example, in [7] the true recognition, sensitivity was relatively low for some of the observed behaviours, specifically a sensitivity level of 68–93 % for resting and 20–90 % for walking. In [8], behavioural classification, results were poor for lying down (0 % sensitivity, 0 % precision) and standing up (71 % sensitivity, 29 % precision). In addition, the SVM algorithm used in [8] has a large computational cost. Finally, in [9], it was not possible to classify feeding behaviour due to the use of a leg-mounted accelerometer.

In general, studies that use accelerometers in order to infer animal behaviour collect and store data in one of two ways. Devices that store information internally for posterior acquisition are generally known as “bio-loggers” [17]. Such devices typically consume very little power, and hence, battery life is very rarely a problem over short to medium timescales. However, the fact that the accelerometer data is only stored internally (typically on a memory card within the sensor) means that the animal must be recaptured to recover the data; in addition, the amount of data that can be collected is limited by the size of the memory card within the device. Devices that transmit information to a central data receiver for subsequent processing are known as “bio-telemetry sensors” [17]. Bio-telemetry devices have the advantage that the animal does not need to be recaptured to access the accelerometer data and, as data do not need to be stored on the device, there is no limit (in principle) to the amount of data collected. However, a major issue with bio-telemetry devices is the power drain created as a result of sending and receiving data to the central receiver. This means that a bio-telemetry device will typically have a much shorter battery life than a bio-logger. One potential approach to overcome the issue of battery drain caused by sending and receiving large data sets is to undertake some form of preliminary processing of the accelerometer data on the bio-telemetry device itself. However, implementing such an approach in practice remains a major challenge due to limited available processing power and memory on the device and the additional drain on battery life caused by the processing of the data. Methods recently

proposed for automatic behavioural classification in animals are mainly based on different machine-learning algorithms such as decision-trees [6, 10, 22], k-means [11], SVM [8], and HMMs [23, 24]. SVM and HMMs come with large computational costs, which make implementation of such an algorithm inside a bio-telemetry device impractical. However, decision-trees have a much lower computational cost and can easily be implemented in real time. Hence, decision-trees may represent a good candidate for an algorithm to be implemented within a bio-telemetry device.

If an accurate behavioural monitoring system is in place, then information about individual and social behaviour (and potential changes in such behaviour) could subsequently be used as indicators of health, welfare and reproductive status. For example, acceleration data has been used in a self-learning classification model in order to predict oestrus status in dairy cows [25]. Similarly, the frequency of transitions between standing up and lying down has been suggested as a possible indicator of forthcoming calving [26]. In addition, several studies have found significant differences in lying, standing and feeding behaviour between healthy and diseased cows. For example, González et al. [27] observed changes in short term feeding behaviour during the onset of diseases such as ketosis and chronic lameness. Palmer et al. [28] observed that during lactation, cows that were severely lame ate fewer, larger meals and had shorter feeding times. Medrano-Galarza et al. [29] observed behavioural changes in lying behaviour and at milking times for cows with mild clinical mastitis. Blackie et al. [30] found significantly longer lying down and significantly shorter standing times for lame cows. Hence, by monitoring behaviour in real-time and observing changes in lying, standing and feeding, it may be possible to detect some of the most common diseases in cattle.

In this study, we develop a decision-tree algorithm that uses tri-axial accelerometer data from a neck-mounted sensor to classify biologically important behaviour in dairy cows such as lying, standing and feeding and to detect transition events between lying and standing. We show that the sensitivity and precision of the decision-tree algorithm matches the performance of more computationally intensive algorithms such as HMMs and SVMs. The algorithm functions in real-time and, given its simple structure, could feasibly be implemented directly in a remote sensor unlike more computationally intensive algorithms. We discuss how the algorithm could be extended to infer activity time budgets, behavioural bout duration and frequency of transitions. Finally, we discuss how this type of real-time behavioural monitoring could play a role in automated detection of dairy cow health and welfare status as part of a Precision Livestock Farming system.

Results

The tri-axial acceleration data were collected from six housed dairy cows wearing a neck collar with tag sensors from the Omnisense Series 500 Cluster Geolocation System [31] as shown in Fig. 1a, b. The sensors contain an accelerometer that records tri-axial acceleration continuously at 50 Hz. The acceleration data were collected from each cow continuously for 36 h. Direct visual observations of the cows were also recorded for a total of 33 h and 25 min in order to validate and quantify the sensitivity and precision of each algorithm.

Fig. 1 Sensor position, orientation and orientation changes when lying and standing. **a** Orientation and location of the neck collar mounted sensor on the neck of the cow. **b** Schematic figure of the coordinate frame of the sensor with X forwards, Y right and Z down according to the illustration. When a cow is wearing a neck collar with attached sensor, a

change in the acceleration in the x -axis corresponds to a sidewise movement to the left or to the right. A change in the acceleration in the y -axis measures the forward and backward movements while changes in the acceleration in the z -axis measure the sidewise rotation of the neck. **c** Example of the orientation of the sensor when a cow is observed standing. The component in the y -axis of the gravitation acceleration varies according to

$\vec{g}_y = g * \cos (180 - \beta)$ where β is the angle in degrees of the sensor relative to the horizontal. **d** Example of the orientation of the sensor when a cow is observed lying. The component \vec{g}_y will be different from standing as the angle α for lying is bigger than β

Summary of decision-tree algorithm performance

The decision-tree algorithm uses two thresholds to classify tri-axial acceleration data as either feeding (high activity) or lying or standing (both low activity). Fig. 2 shows the structure of the decision-tree, while further explanation of the algorithm and a systematic study of the effect of window size and threshold values used are given in Additional file 1.

Fig. 2 Decision-tree algorithm used for the classification of behaviours in dairy cows. Decision rules are evaluated downwards until the final behavioural class is assigned. The scheme contains the feature characteristic used as data input for the decision rule to partition the data. At each decision rule, data is partitioned into clusters with similar properties. The first decision rule in this algorithm discriminates between high and low energy expenditure activities using the overall dynamic body acceleration (VeDBA). High energy expenditure activities are classified as feeding. Low energy expenditure activities are further classified using a second decision rule which discriminates data by the static component of the acceleration in the y -axis (SCAY). Data with values above threshold B (-0.055 g) are classified as standing and data with values below this threshold are classified as lying

The classification performance of the decision-tree algorithm can be summarised in a confusion matrix, where each column represents the predicted behaviour from the algorithm and each row represents the ground truth observed behaviour. Table 1 shows the confusion matrix obtained by using the decision-tree algorithm with 1-min (2003 data points), 5-min (401 data points) and 10-min (200 data points) window sizes and with decision threshold values of 0.0413 g (threshold A) and -0.055 g (threshold B). With all window sizes, feeding is classified highly accurately by the decision-tree algorithm. However, it is clear that the decision-tree algorithm has a tendency to misclassify standing behaviour as lying (and vice versa). Hence, it seems clear that the behaviours that are most likely to be misclassified are those that have the most similarity in the relative neck movements of the cow (see Fig. 1). Note also that the number of standing events is significantly lower than the number of lying or feeding events.

Table 1 Confusion matrix obtained for the classification of dairy cow behaviour with the decision-tree classification algorithm. The results were obtained using a 1-min, 5-min and 10-min window size and with values of 0.0413 and -0.055 g for threshold A and B , respectively. The values given in the main part of the table correspond to the percentage of observations classified for each behaviour. The final column gives the total number of validated observations (data points) for each behaviour. Values marked in bold indicate the classification sensitivity for each behaviour

Observed behaviour	Predicted behaviours			Total number of observations
1-min window	Lying	Standing	Feeding	
Lying	74.09	22.27	3.64	988
Standing	8.96	82.08	8.96	279
Feeding	0.14	4.21	95.65	736
				2003
5-min window				
Lying	74.09	21.76	4.15	193
Standing	5.77	88.46	5.77	52
Feeding	0.0	2.56	97.44	156
				401
10-min window				
Lying	77.42	18.28	4.30	93
Standing	4.0	88.0	8.0	25
Feeding	0.0	1.22	98.78	82
				200

Comparative study of algorithm performance

To test the relative performance of our simple decision-tree algorithm, we directly compared its performance to alternative classification algorithms such as a k-means algorithm, a HMM and a SVM algorithm. The performance comparison was made using the same initial input data for all four algorithms and with 1-min, 5-min and 10-min window sizes. In the decision-tree algorithm, values of 0.0413 and -0.055 g were used for threshold A and B , respectively. The HMM also required initial values for the transition probability, initial distribution and emission distribution; the initialisation of these parameters and the selection of the training and testing sets for all the algorithms is further explained in Additional file 1.

Table 2 summarises the performance of the four different classification algorithms. For all window sizes, the highest overall sensitivity was obtained with the decision-tree algorithm (83.94 % for 1-min window, 86.66 % for 5-min window and 88.06 % for 10-min window). In contrast, the SVM algorithm achieved the highest overall precision for all window sizes (85.89 % for 1-min window, 87.72 % for 5-min window and 87.52 % for 10-min window). In general, the overall sensitivity and overall precision increased with window size for all the algorithms considered (the only exception being the precision of the k-means algorithm which decreased from 84.80 to 81.84 % when moving from a 5-min to a 10-min window). In general, the best classification performance for each behaviour was obtained using the decision-tree algorithm (sensitivity) or the SVM (precision). The HMM generally performed reasonably well but typically had lower sensitivity and precision than the decision-tree and SVM algorithms, except at the 10-min window size where it has the best performance for standing sensitivity (100 %) and lying precision (92 %). The k-means algorithm generally

had the worst overall performance, although it performed well for feeding sensitivity (99.36 %) and lying precision (98.10 %) at the 5-min window size. The decision-tree algorithm matched or exceeded the performance of the other (more computationally expensive) algorithms for classification sensitivity. The decision-tree algorithm did not perform as well as the SVM for classification precision, but was comparable to the k-means and HMM in this regard.

Table 2 Performance comparison of four different machine-learning algorithms using the same input data set. Performance measures (sensitivity and precision) were obtained using 1-min, 5-min and 10-min windows. HMM refers to the hidden Markov model, and SVM refers to the support vector machine algorithm. Overall sensitivity is calculated as the arithmetic mean sensitivity for the three behaviours. Overall precision is calculated in a similar manner. Values marked in bold indicate the best performing algorithm for each behaviour classification

Behaviour	Performance metric	Classifier algorithm			
1-min window		Decision-tree	K-means	HMM	SVM
Lying	Sensitivity	74.09	85.93	90.17	92.91
	Precision	96.57	91.88	85.41	89.65
Standing	Sensitivity	82.08	59.50	38.35	51.65
	Precision	47.01	29.28	37.28	77.01
Feeding	Sensitivity	95.65	59.92	83.83	98.01
	Precision	92.03	86.13	91.54	91.01
Overall	Sensitivity	83.94	68.45	70.78	80.85
	Precision	78.53	69.09	71.41	85.89
5-min window					
Lying	Sensitivity	74.09	55.37	80.31	92.91
	Precision	97.95	98.10	94.51	91.66
Standing	Sensitivity	88.46	69.23	76.92	60.89
	Precision	47.92	69.23	54.05	79.15
Feeding	Sensitivity	97.44	99.36	97.44	98.29
	Precision	93.25	87.08	93.25	92.36
Overall	Sensitivity	86.66	74.65	84.89	84.03
	Precision	79.71	84.80	80.60	87.72
10-min window					
Lying	Sensitivity	77.42	80.65	70.97	89.60
	Precision	98.63	96.15	100.00	93.35
Standing	Sensitivity	88.00	76.00	92.00	68.00
	Precision	55.00	59.38	50	76.04
Feeding	Sensitivity	98.78	98.78	100	100.00
	Precision	93.10	90.00	93.18	93.18
Overall	Sensitivity	88.06	85.14	87.65	85.86
	Precision	82.24	81.84	81.06	87.52

Decision-tree algorithm classification at the individual-level

Performance of the decision-tree algorithm was also analysed at the level of the individual cow (Table 3). For this analysis, a 1-min window size was used to avoid having too few data

points for each individual cow; values of 0.0413 and -0.055 g were used for threshold A and B , respectively. Cow 1b (i.e. cow 1 on day 2) was not observed standing at any point during the observation period. In general, classification of feeding showed the smallest variation in sensitivity across individual cows (sensitivity 78.49–100 % and precision 27.59–100 %). Classification of lying showed wider variation for sensitivity (21.82–100 %) but less variation for precision (89.91–100 %). The widest variation, in both sensitivity and precision, for the three different behaviours was obtained for standing (sensitivity and precision 0–100 %). These results match with the previous analysis (Table 1), suggesting that standing is the behaviour most likely to be misclassified. Comparing individual cows on successive days, it seems likely that the decision-tree algorithm consistently performed better with particular cows. For example, comparing the lying sensitivity of cow 4 (100 % day 1, 85.59 % day 2) or cow 6 (85 % day 1, 100 % day 2) with cow 3 (21.82 % day 1, 60.27 % day 2), it is clear that there may be a consistent misclassification of this behaviour in cow 3. There are also day to day variations in the sensitivity, which could be due to individual differences in how the accelerometer sensor was positioned on the cow or due to different individual cow behaviour (e.g. if the cow does not raise its neck as high as other cows when standing or feeding).

Table 3 Performance of the decision-tree classification algorithm across individual cows. Results were obtained using a 1-min window and values of 0.0413 and -0.055 g for threshold A and B , respectively. Characters a (day 1) and b (day 2) after the cow number correspond to the day the observations were recorded. For all of the cows, observations were recorded over two successive days. For cow number 2, no accelerometer records were collected in the second day due to a battery failure in the sensor. No observations of standing behaviour were recorded for cow 1 on day 2

Cow	Lying		Standing		Feeding	
	Sensitivity	Precision	Sensitivity	Precision	Sensitivity	Precision
1a	91.59	89.91	80.65	84.75	91.30	93.33
1b	69.62	98.21	No obs	0	100	75.38
2	71.17	94.05	68.75	68.75	92.54	92.54
3a	21.82	100	100	27.59	100	27.59
3b	60.27	100	91.67	40.74	98.61	81.61
4a	100	95	95	100	93.75	100
4b	85.59	100	100	23.26	100	71.05
5a	66.67	100	0.00	0	88.89	53.33
5b	47.83	100	39.39	15.85	78.49	51.41
6a	85.00	100	100	85.19	100	84.62
6b	100	94.12	89.04	85.53	84.62	80.00

Transition events between lying and standing

Transitions from standing-to-lying or lying-to-standing were relatively infrequent throughout the observation period: only 23 transition events were observed of which 13 were lying-down events and 10 were standing-up events. These 23 transition events were used to test the performance of the transition detection algorithm. In the first step of the algorithm, transition events are detected without distinguishing between lying down and standing up. Subsequently, when a transition event has been detected, the decision-tree classification algorithm described in the previous sections is used to classify the behaviour either side of

the transition and hence discriminate between lying-down events and standing-up events. Further details can be found in Methods section and in Additional file 1.

From Table 4, it is clear that the transition detection algorithm can accurately detect transition events (95.45 % sensitivity, 87.5 % precision) without distinguishing between the different types of transitions. However, it performs less well when trying to classify the transition events as lying down (66.67 % sensitivity, 83.33 % precision) or standing up (63.64 % sensitivity, 70 % precision).

Table 4 Performance of the detection algorithm for transitions between standing and lying. The non-specific transition refers to a behavioural transition that is detected but without distinguishing between lying down and standing up

Type of transition	Sensitivity	Precision
Non-specific	95.45	87.50
Standing up	63.64	70
Lying down	66.67	83.33

Discussion

Analysis of behaviour and behavioural changes has been suggested as a potential way to indirectly monitor health and welfare of dairy cows [27–30], and several automated systems have been proposed to identify different biologically important behaviours [8, 9]. In this study, we have developed a simple decision-tree classification algorithm that uses tri-axial accelerometer data from a neck-mounted sensor to accurately classify biologically relevant behaviours such as lying (77.42 % sensitivity, 98.63 % precision), standing (88.00 % sensitivity, 55.00 % precision) and feeding (98.78 % sensitivity, 93.10 % precision). A further algorithm can detect transition events between lying and standing or vice versa (95.45 % sensitivity and 87.50 % precision when transition events are not classified as lying down or standing up specifically). The main decision-tree classification algorithm performs at least as well as more complex algorithms, such as HMMs or SVMs but is much simpler and less computationally expensive than these approaches and hence may be suited for direct incorporation in the sensor itself. The decision-tree algorithms use intuitive and easy to interpret characteristics of the biomechanics of behaviour based on the static component of the acceleration in the y -axis (SCAY) or the overall vectorial dynamic body acceleration (VeDBA). The parameters used in the algorithms (window size and threshold values) were explored using a single data set (see Additional file 1 for full details), but the approach could be adapted to construct similar algorithms in different contexts or for different data sets. The output of the behavioural classification and transition detection algorithms can be extended to infer activity budgets, behavioural bout duration and frequency of transitions. No behavioural classification algorithm will ever be free from error, but our simple decision-tree algorithm performs relatively accurately (in terms of sensitivity and precision). The tri-axial accelerometers used in this study (Omnisense Series 500 Cluster Geolocation System [31]) are one element of a more general wireless location sensor network that can accurately track spatial position of each cow. Although this feature was not used in this study, it may be possible to combine accelerometer data with spatial location data to more accurately determine real-time behaviour and behavioural changes as part of an automated detection system for dairy cow health and welfare status within a Precision Livestock Farming approach.

The behaviours investigated in this study (lying, standing and feeding) have been suggested as indicators of health and welfare in dairy cows [27–30]. Using a neck-mounted sensor, we are able to include feeding behaviour in the repertoire of behaviours, something that is not usually possible in studies that use leg-mounted sensors. The position of the sensor on the body of the animal determines the behaviours that can be discriminated and multiple sensors could potentially be used to improve the behavioural classification [14]. The position of the sensor can also affect the performance of the classification, as illustrated by Moreau et al. [7] who deployed sensors at different positions on the body of a goat when classifying grazing behaviour. The counterweights in the neck collars used in this study help to reduce positional changes that can affect the performance of the classification.

Sensitivity and precision were used as statistical measures of the performance of the algorithm. Both performance measures were validated and quantified through direct visual observations of the cows. Performance of any classification algorithm can depend on a range of factors as discussed in [32]. In our case, the performance of the decision-tree algorithm was explored in relation to the choice of window size and the selection of threshold values within the decision-tree. Window sizes below 60 s showed a low overall sensitivity, particularly for feeding behaviour (Additional file 1: Figure S1). At small window sizes it may not be easy to perceive the regular up and down movements of the cow's neck while eating, which will result in apparently low activity values (VeDBA) and hence lead to misclassification. Classification performance for lying and standing were very similar for all window sizes but the best overall accuracies were found above 60 s. This result, along with the fact that visual observations of bouts of behaviour of less than 60 s were rarely recorded, means that a window size above 60 s represents the most appropriate choice. Tables 1 and 2 illustrate that a small increase in the decision-tree classification algorithm performance is obtained at the largest window size of 10 min. Note also that the low values for precision are likely related to the difficulty of distinguishing standing from lying behaviour and also to the fact that there were significantly less observations of standing behaviour than lying behaviour for the cows in this study (Table 1). A similar analysis was undertaken to explore the effect of the threshold value used at each step of the decision-tree (see Additional file 1), and values with the best overall performance were selected (0.0413 and -0.055 g for thresholds *A* and *B* respectively).

In addition to the parameter choice used within the algorithm, behavioural variation across individual cows could also have an effect on the classification performance [33], see Table 3. This behavioural variation might explain the differences in the performance when applying the algorithm at the individual level. For example, some cows may lie down or stand in different positions, causing the algorithm to misclassify these two behaviours. Similarly, a cow strongly moving its head while standing might be misclassified as feeding. In addition to the behavioural variation, low numbers of behavioural observations can also explain some of the low sensitivity and precision values obtained at the individual level (e.g. cow 1 on day 2 was not observed standing at any point during the observation period). Further investigations of these variations should be undertaken if the decision-tree algorithm is to be used for longitudinal studies in a larger number of animals. In principle, and if long enough time series of data are available, it should be possible to train the decision-tree at the individual level (so that each cow would have different values for threshold *A* and *B*) and relative to their underlying behavioural characteristics.

In an earlier study, Martiskainen et al. [8] used a SVM algorithm to classify eight different behaviours in cattle (feeding, lying, standing, transitions between lying and standing, plus

two walking behaviours and ruminating behaviour), while Robert et al. [9] used a generalised mixed linear model (GMLM) to classify only three behaviours (lying, standing, walking). Values on sensitivity and precision were reported by Martiskainen et al. [8], while Robert et al. [9] only reported the sensitivity (called the ‘agreement’). Classification of standing (88 % sensitivity) and feeding (98.78 % sensitivity) in our decision-tree classification algorithm compares well to the figures reported by Martiskainen et al. [8] for their SVM (80 % sensitivity for standing, 75 % sensitivity for feeding), although it should be noted that when more behaviours are considered (as with the eight behaviours considered in [8]) the individual classification accuracy for each behaviour is likely to be lower. Sensitivity for lying and standing was lower for the decision-tree classification algorithm (77.42 and 88 %, respectively), when compared to the GMLM reported in [9] (99.2 and 98 %, respectively), although the decision-tree is a much simpler algorithm. Walking was not included in our study, while in [9] it was considered in the behaviours; conversely, feeding behaviour was not included in the GMLM algorithm in [9] since data was collected using a leg-mounted sensor. Despite some advantages in terms of classification performance when using SVM and GMLM algorithms, they remain difficult to implement and require much more computational power than a simple decision-tree algorithm. Simplicity in our decision-tree comes from not only the algorithm structure but also from the small number of feature characteristics (VeDBA and SCAY). These are based on parameters that are easy to use and to interpret biologically.

The additional detection algorithm for transition events between lying and standing (or vice-versa) provided a satisfactory sensitivity (95.45 %) and precision (87.50 %) when detecting transitions without considering the type (lying down or standing up). When the transition type was considered, both sensitivity and precision decreased significantly (standing up, 63.64 % sensitivity, 70 % precision; lying-down, 66.67 % sensitivity, 83.33 % precision). In a previous study, Martiskainen et al. [8] reported a performance of 71 % sensitivity, 29 % precision for standing up and 66.67 % sensitivity, 83.33 % precision for lying-down. Due to the limited observation time period, only a small number of transitions were observed in the present study. In future work, the algorithm could be refined and validated by collecting observational data over longer periods of time.

Using the output of the decision-tree algorithms described in this study, statistical measures of the activity budget, bout duration and frequency of transitions for each cow (or the herd as a whole) can be computed (Table 1, Figs. 3 and 4). Such measures can provide relevant information about the behaviour and behavioural changes of cattle over time and can potentially be used as indirect indicators of the health and welfare of dairy cows as part of a Precision Livestock Farming approach [27–30].

Fig. 3 Example time series of raw tri-axial accelerometer and its component outputs for lying, standing and feeding. **a** Example time series of the raw tri-axial accelerometer output for observed periods of lying, standing and feeding for a single cow. The *x*-, *y*- and *z*-axis correspond to the blue, green and red lines, respectively. When a cow is lying or standing, little change in the acceleration is registered because these two behaviours exhibit little overall movement. The shifts in the acceleration observed when the cow is feeding are caused by the cow moving its head up and down. **b** Output readings of the running mean of the acceleration in the *y*-axis and vectorial dynamic body acceleration (VeDBA) values under the three different behaviours. These two parameters correspond to the static and dynamic components of the acceleration. There is a clear difference in the VeDBA outputs between feeding and lying or standing. There is also a difference in the running mean between

standing and lying which is caused by a difference in the component in the y -axis of the gravity field (see also Fig. 1c, d)

Fig. 4 Examples of lying and standing transitions and results for their detection. **a** Example time series of the raw tri-axial accelerometer output for standing up and lying down transitions. A rapid change in acceleration for all three axes can be observed. **b** Output of the results for the transition detection algorithm. Values of the range of the y -axis above a predefined threshold determine if a transition has occurred. Visual observations are displayed in *green* and prediction by the algorithm in *red* (“up” corresponds to standing up and “down” to lying down)

Conclusion

Our results show that a simple decision-tree classification algorithm that uses data from a neck-mounted tri-axial accelerometer can classify, with a high level of accuracy, biologically relevant behaviours in cattle such as feeding, lying and standing. The decision-tree classification algorithm matched the performance of other more computationally intensive machine-learning algorithms. The detection algorithm which proposed to distinguish between lying-down and standing-up events also showed satisfactory performance but needs further refinement to improve accuracy. The decision-tree algorithm has great potential for use directly within a sensor for real-time calculations and monitoring of animal behaviour. By extension, it would be feasible to determine activity time budgets, bout durations and frequency of transitions. Such a system could offer a new potential technology for the automated detection of health and welfare problems in dairy cows. The specific decision-tree algorithm we describe here could possibly be adapted to work with other similar housed animal species such as pigs. More generally, simple behavioural classification algorithms can play a key role in automated behavioural detection within Precision Livestock Farming.

Methods

Instruments

The acceleration data were collected using a wireless sensor system (Ominsense Series 500 Cluster Geolocation System [31]; <http://www.omnisense.co.uk/>) that includes an embedded tri-axial accelerometer (Xtrinsic MMA8451Q 3-Axis, 14-bit/8-bit Digital Accelerometer with a sensitivity between -8 and $+8$ g). Accelerometer data was collected at 50 Hz which allowed for effective battery life of approximately 2 days. The wireless sensors contain a 2.4 GHz, IEE 802.15.4a transmitter module to remotely send messages to the CLS-504 location server. The Series 500 sensors can be used to form a wireless mesh sensor-node network that is able to compute relative spatial locations of the sensor nodes using the arrival time of periodic messages sent from each node to its neighbours. In principle, acceleration data could be processed on the sensor in real-time and outputs sent across the network as part of a more general monitoring system. However, in this study, only data from the tri-axial accelerometer were recorded using a 4 GB micro SD flash memory card for posterior data analysis. The sensors were fixed in the same orientation on the right hand side of a neck collar worn by the cows (Fig. 1a). Counterweights (0.5 kg) were used on the neck collars to ensure a stable position of the sensor on the body of the animal. The sensor weighs approximately 0.25 kg in total (including batteries), half the weight of the counterweight. The coordinate frame of the sensors corresponds to X forwards, Y right and Z down as shown in Fig. 1b). At the end of the

study, the SD card was removed from the sensor and the accelerometer data was converted from its hexadecimal format to g units ($g = 9.81 \text{ ms}^{-2}$).

Study site, animals and observation of behavioural activities

The data collection was carried out on a commercial farm of Holstein dairy cattle located in Essex, UK. The cows were loose housed in a cubicle shed. The herd was milked three times a day at approximately 5 a.m (morning), 1 p.m (afternoon) and 9 p.m (evening). The duration of milking time for each individual cow varied between 1 and 1 ½ h. The herd mean 305-day milk yield was 11,000 litres per cow. Cows were fed a commercial total mixed ration. A total of six cows that had not shown signs of severe lameness, or other disease that might affect their behavioural repertoire, were selected for this study. Cows were selected and collared during morning or afternoon milking and were wearing the collar for a maximum of 2 days (since battery of the sensors could not be guaranteed after this point). Cows were monitored between milking periods; during milking, no visual observations were recorded.

Cow behavioural activities were recorded by observers (ZB and HH) performing a visual focal tracking on each individual cow that was wearing a sensor collar according to the following criteria for each behavioural activity:

1. *Feeding (state)*: cows located at the feeding zone, ingesting food;
2. *Lying (state)*: cows located in a cubicle in a lying down position;
3. *Standing (state)*: cows standing on their four legs;
4. *Lying down (transition)*: cows changing from a standing state to a lying state;
5. *Standing up (transition)*: cows rising from a lying state to a standing state

Drinking, brushing and walking activities were observed less frequently and for short durations and therefore not considered for classification in the algorithm. It should be stressed that these rarer activities and events may still be biologically important in the context of detecting health and welfare status. Hence, although we do not try to classify them here, future studies should also consider methods for detecting these rarer behaviours.

From the data set of visual observations, only the activities of interest for this study were selected to validate the classifier algorithm. The new data set used for validation contains the following observational data:

Cow 1: a total of 8 h and 2 min extracted from observation on 28 August 2014, between 08:00–18:00; 29 August 2014, between 08:00–12:40

Cow 2: a total of 3 h and 6 min extracted from observations on 28 August 2014, between 08:00–18:00

Cow 3: a total of 5 h and 40 min extracted from observations on 3 September 2014, between 15:00–17:30; 4 September 2014, between 08:00–18:00

Cow 4: a total of 5 h and 5 min extracted from observations on 15 September 2014, between 15:30–18:30; 16 September 2014, between 07:40–17:40

Cow 5: a total of 6 h and 5 min extracted from observations on 3 September 2014, between 15:10–17:10; 4 September 2014, between 08:00–18:10

Cow 6: a total of 5 h and 22 min 15 September 2014, between 15:30–18:10; 16 September 2014, between 07:30–17:50

In total, direct visual observations of the cows were completed for 33 h and 20 min, of which 15 h and 30 min were lying, 4 h and 10 min were standing and 13 h and 40 min were feeding. All behavioural observations were entered into a spreadsheet with the start and stop time of every activity and identification of the corresponding cow. Observer and sensor watches were synchronised at the start of the observation period so that observation data could be accurately aligned with the tri-axial accelerometer data retrieved from the sensors in a single database.

Algorithms for behavioural state classification

Raw acceleration data

Figure 3a illustrates example time series of the raw tri-axial accelerometer output for observed periods of lying, standing and feeding behaviour for a single cow. It is clear that there is very little qualitative difference in the acceleration output for the lying and standing behaviours, since for both these behaviours the cow exhibits very little overall movement. When the cow is feeding there is a clear regular shift in the acceleration in the y and z axes that corresponds to the cow moving its head up and down. Figure 3 is only a representative example but similar qualitative patterns in the acceleration output were observed for the other cows in the study. These qualitative observations offer a useful intuitive starting point for determining the most appropriate feature characteristics to include in the classification algorithm.

Feature characteristics

Machine-learning algorithms use feature characteristics (also called summary statistics) calculated from the input data (e.g. the raw accelerometer data) to classify different states (e.g. feeding, lying or standing). The algorithms in this study have been developed using two intuitive and easy to interpret characteristic features based on the biomechanics of the movement behaviour of the cows. These two feature characteristics consist of two different components of the raw acceleration data: a static component caused by the gravity field (SCAY) and a dynamic component caused by the movement of the animal (vectorial dynamic body acceleration, VeDBA [34, 35]). Other studies have used a far larger number of feature characteristics (e.g. 30 or even higher) [5, 6, 8]. In our study, the use of only two features was motivated by the need to reduce computational time and complexity and also to allow more intuitive biological interpretation of the results.

Figure 3b illustrates a typical example time series of running mean in the y -axis and VeDBA output for observed periods of lying, standing and feeding behaviour for a single cow. Low VeDBA output values for lying and standing are caused by the low movement exhibited by cows during these behaviours. In contrast, high VeDBA values obtained for feeding are caused by the upward and downward head movement cows perform during this behaviour. In this figure, it is also possible to observe a small difference in the SCAY outputs between lying and standing. Since the running mean in the y -axis represents the static component caused by the gravity field, output values obtained for this parameter correspond to the orientation of the sensors during the behaviour as seen in Fig. 1c, d. Figure 1c shows an example of the orientation of the sensor when the cow was observed standing, while Fig. 1d shows the orientation of the sensor when the cow was observed lying. The component in the y -axis of the gravity field is given by $\overline{g}_y = g * \cos(180 - \beta)$. Using this expression, a

preselected threshold of -0.055 g for the static component in the y -axis corresponds to an angle of $\beta = 86.84^\circ$ (where an angle of $\beta = 90^\circ$ can be interpreted as the cow having its neck aligned horizontally). Therefore, the decision-tree classifies standing and lying behaviour if the neck (and therefore sensor) is above or below this threshold. Figure 1c, d are only representative examples, but similar patterns in the static component were found for other cows in this study.

The VeDBA and SCAY feature characteristics are calculated as a mean over a given moving window size centred at the time point of interest (see Additional file 1). This requires a moving window size to be specified before any algorithm is run. A range of moving window sizes was tested for each algorithm and we report results for sizes of 1, 5 and 10 min (Table 2). Results for other moving window sizes are explored for the decision-tree algorithm in Additional file 1.

Machine-learning algorithms

There are a range of different machine-learning algorithms that could be used to classify different animal behaviours. These algorithms can be described as either *supervised* or *unsupervised* approaches. A supervised learning algorithm is formed by two processes: training and testing. A supervised learning algorithm uses a known data set to construct a model (training process) that is then used for making predictions on a new data set (testing process). Unsupervised machine-learning algorithms explore the data to find hidden patterns or to cluster the data input in classes with similar statistical properties. In this study, the three following unsupervised algorithms for the classification of the dairy cow behaviours were used: decision-tree, k-means and a HMM. The decision-tree was selected based on its simple structure and low computational cost, making it feasible to be implemented directly in a remote sensor. The selection of the k-means algorithm was based also on the simplicity of its structure and the possibility to compare the decision-tree to methods with similar levels of simplicity (although the k-means may have high computational costs due to a recursive component in the algorithm). The HMM was chosen in order to compare the decision-tree performance with a more sophisticated statistical model that is often used to classify animal behavioural states [23, 24]. Finally, a supervised SVM algorithm was also chosen in order to compare the decision-tree performance to a more complex algorithm that has been used for the classification of accelerometer data to distinguish between different behaviours in dairy cows [8]. The decision-tree and k-means algorithms were custom written by the authors in Matlab [36]. The HMM was applied using the Matlab toolbox for HMMs developed in [37]. The SVM was applied using the machine-learning toolbox provided in [38].

Decision-tree

A full description of the decision-tree algorithm used in this study is available in Additional file 1. We summarise the key features of the algorithm here. The decision-tree algorithm uses two rules with associated thresholds to classify tri-axial acceleration data as either feeding (high activity) or lying or standing (both low activity). The first rule in the decision-tree uses the mean of the VeDBA values and a predefined threshold A to discriminate between cases with high and low energy expenditure activities. Those cases resulting in a high energy expenditure activity are labelled as feeding, and those with low energy expenditure activities are used in the second step of the decision-tree (Fig. 2). The second decision rule of the tree compares the running mean of the acceleration in the y -axis (SCAY) to a predefined threshold B value in order to partition the data into two clusters (mean of static component in

the y -axis above or below the threshold value). Cases resulting in values below the threshold are labelled as lying, and those with values above are labelled as standing (Fig. 2). A range of different predefined threshold values were considered (see Additional file 1), and values of $A = 0.0413$ g and $B = 0.055$ g were found to give the best performance with this data set. Similarly, to explore the effect of the choice of window size, the performance of the algorithm was investigated using windows ranging from 1–600 s (window sizes above 600 s resulted in too few data points for a fair comparison of performance) and full details are given in Additional file 1.

k-means

Observations for the k -means algorithm are given by the 2-dimensional feature characteristics. The first dimension is represented by the mean of the VeDBA values over the window size, whereas the second dimension is represented by the mean of the acceleration in the y -axis (SCAY). The k -means algorithm discriminates between the observations in one step using both feature characteristics at the same time. This represents a key difference between the decision-tree and the k -means, since the former uses one feature characteristic at each decision rule. A full description of the k -means algorithm is given in Additional file 1.

Hidden Markov model

A sequence of behaviours in dairy cows can be modelled as a first-order HMM with a finite number of hidden states (behaviours) where each activity can be observed through a set of characteristic features (observations). The observations for the HMM correspond to the same characteristic features used for the decision-tree, i.e. mean of VeDBA over the window size and running mean of the acceleration in the y -axis over the window size (SCAY). The hidden Markov model was applied using the Matlab toolbox for hidden Markov models developed in [37]. This toolbox randomly generates an initial transition probability matrix A and an initial probability π . The emission probability distribution B is initialised using a static Gaussian Mixture model. Since the results can depend on the initialisation parameters, we run a total of 100 random initialisations to select the highest scoring model. Further details of the implementation, use and application of the Baum-Welch, the Viterbi and the forward-backward algorithms for HMMs can be found in [39], and further details are given in Additional file 1.

Support vector machines

SVMs are a supervised learning algorithm requiring training and testing processes. In this study, training was performed using $k-1$ folds and tested in the fold left out. We used a 3-fold cross validation for the implementation of the SVM algorithm. Further details of the SVM algorithm are provided in Additional file 1 and can also be found in [38, 40–42].

Performance of the classification algorithm

Comparison of algorithm classification performance

The performance of the decision-tree classification algorithm was compared across a range of values for the algorithm parameters (window size, thresholds A and B); for details see Additional file 1. The performance of the algorithm was directly compared to alternative

classification algorithms such as k-means, HMM and SVM using the same input data set (Table 2). The performance of an automated behavioural classification algorithm can often vary across individuals or breeds of the same species [33]. Hence, we also considered the performance of the decision-tree algorithm at the level of the individual cow. In order to do this, we computed the performance metrics for each individual cow at a window size of 1 min (Table 3). The 1 min window was selected in this context to avoid having only a small number of samples for each individual cow (which can occur at larger window sizes).

Sensitivity and precision

When comparing algorithm classification performance, we considered two performance metrics: the sensitivity of classification and the precision of classification. In standard statistical process control, the sensitivity (Sen) and precision (Pre) are defined as:

$$\text{Sen} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$
$$\text{Pre} = \frac{\text{TP}}{\text{TP} + \text{FP}}.$$

Here, TP (true positive) is the number of instances where the behavioural state of interest that was correctly classified by the algorithm after validation by the visual observer. FN (false negative) is the number of instances where the behavioural state of interest was visually observed in reality but was incorrectly classified as some other behaviour by the algorithm. FP (false positive) is the number of times the behavioural state of interest was (incorrectly) classified by the algorithm but not observed in reality. TN (true negative) is the number of instances where the behavioural state of interest was (correctly) classified as not being observed.

An algorithm for detection of transitions between lying and standing

A further two-step algorithm was developed to detect the transitions between lying and standing (Table 4). The first step of the algorithm (non-specific) uses a threshold over the range of the acceleration in the y -axis to determine if a transition occurs or not. Range in the y -axis represents a good candidate for the threshold due to the biomechanics of the rapid change in this axis when cows exhibit a transition between lying and standing or vice-versa. As described by Martinskainen et al. [8], a cow that lies down bends one front leg, lowers its forequarters then its hindquarters until it settles into a lying position. When a cow stands up, it lunges forward, lifts its hindquarters, then rises to stand up on its four legs. According to this definition and the orientation of the sensors in Fig. 1a, b, a transition movement implies a significant change in the orientation of the sensor in the y -axis (Fig. 4a).

The second step of the transition detection algorithm is performed by applying the decision-tree classification algorithm described previously to infer the anterior and posterior behaviour on either side of the transition and hence discriminate between standing up and lying down. Further details of the transition detection algorithm are given in Additional file 1.

Availability of supporting data

The data collected as part of this study is available in Additional file 2.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JA, EC, JVD, ZB, NB and HH participated in the design of the field trial. ZB and HH conducted the fieldwork and data collection. JVD and EC designed the algorithms and comparative classification study. JVD wrote the Matlab code and performed the data analyses. JVD and EC wrote the manuscript. JA, ZB, HH, DC and NB helped to draft the manuscript. All the authors read and approved the final manuscript.

Acknowledgements

This study is supported by the UK Biotechnology and Biological Sciences Research Council through grants BB/K002562/1 (EC & JVD), BB/K002376/1 (JA & ZB), BB/K001302/1 (DC) and BB/K003070/1 (NB). The study was reviewed and approved by the Royal Veterinary College Ethics and Welfare Committee under the unique reference number 2012 1223. HH was supported by the Colin Spedding Memorial Research Studentship awarded by The Farm Animal Welfare Trust.

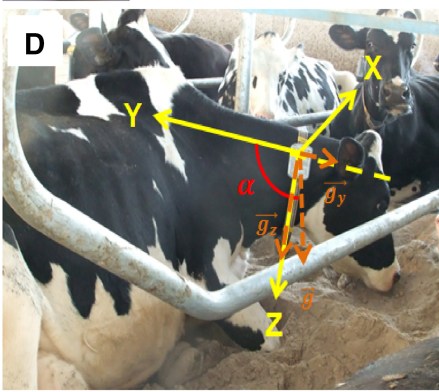
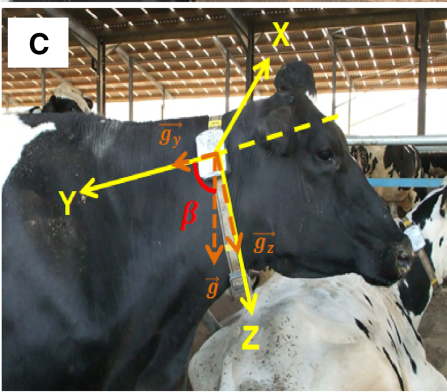
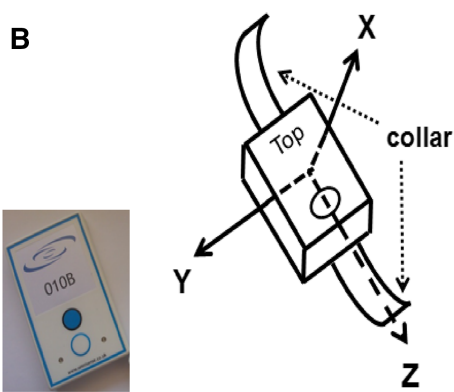
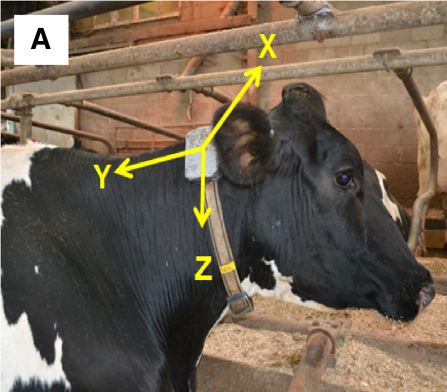
References

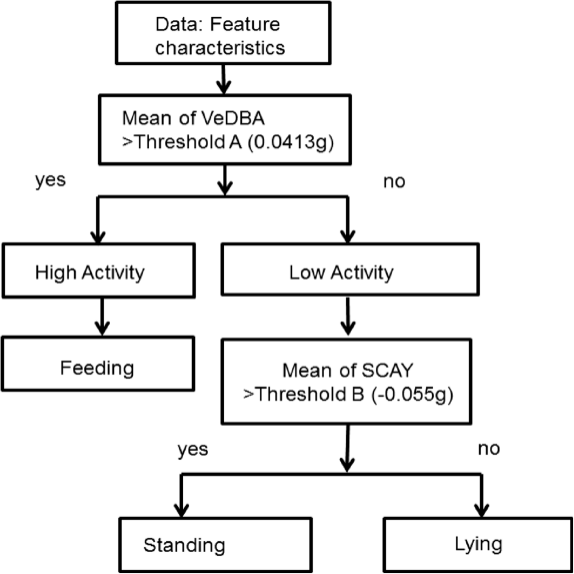
1. Tomkiewicz SM, Fuller MR, Kie JG, Bates KK. Global positioning system and associated technologies in animal behaviour and ecological research. *Philos Trans R Soc B*. 2010;365:2163–76.
2. Baratchi M, Meratnia N, Havinga PJM, Skidmore AK, Toxopeus BAG. Sensing solutions for collecting spatio-temporal data for wildlife monitoring applications: a review. *Sensors*. 2013;13:6054–88.
3. Mannini A, Sabatini AM. Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors*. 2010;10:1154–75.
4. Bourke AK, O'Brien JV, Lyons GM. Evaluation of threshold-based tri-axial accelerometer fall detection algorithm. *Gait Posture*. 2007;26:194–9.
5. Nathan R, Spiegel O, Formann-Roe S, Harel R, Wikelski M, Getz WM. Using tri-axial acceleration data to identify behavioural modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. *J Exp Biol*. 2012;215:986–96.
6. Resheff YS, Rotics S, Harel R, Spiegel O, Nathan R. AcceleRater: a web application for supervised learning of behavioural modes from acceleration measurements. *Mov Ecol*. 2014;2:27. doi:10.1186/s40462-014-0027-0.
7. Moreau M, Siebert S, Buerket A, Schlecht E. Use of tri-axial accelerometer for automated recording and classification of goats' grazing behaviour. *Appl Anim Behav Sci*. 2009;119:158–70.

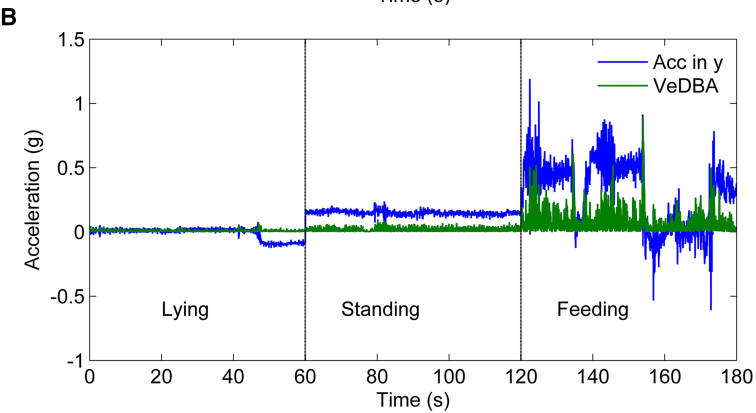
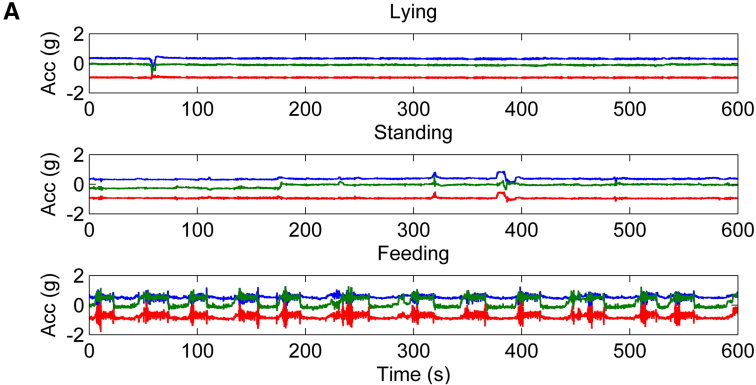
8. Martiskainen P, Järvinen M, Skön JP, Tiirikainen J, Kolehmainen M, Mononen J. Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Appl Anim Behav Sci.* 2009;119:32–8.
9. Robert B, White BJ, Renter DG, Larson RL. Evaluation of three-dimensional accelerometers to monitor and classify behaviour patterns in cattle. *Compu Electron Agric.* 2009;67:80–4.
10. Shamoun-Baranes J, Bom R, van Loon EE, Ens BJ, Oosterbeek K, Bouten W. From sensor data to animal behaviour: an oystercatcher example. *PLoS One.* 2012;7, e37997. doi:10.1371/journal.pone.0037997.
11. Bidder OR, Campbell HA, Gómez-Laich A, Urgé P, Walker J, Cai Y, et al. Love thy neighbour: automatic animal behavioural classification of acceleration data using the k-nearest neighbour algorithm. *PLoS One.* 2014;9, e88609. doi:10.1371/journal.pone.0088609.
12. Chen KY, Bassett Jr DR. The technology of accelerometry-based activity monitors: current and future. *Med Sci Sports Exer.* 2005;37:490–500.
13. Sakamoto KQ, Sato K, Ishizuka M, Watanuki Y, Takahashi A, Daunt F, et al. Can ethograms be automatically generated using body acceleration data from free-ranging birds? *PLoS One.* 2009;4, e5379. doi:10.1371/journal.pone.0005379.
14. Brown DD, Kays R, Wikelski M, Wilson R, Klimley AP. Observing the unwatchable through acceleration logging of animal behaviour. *Animal Biotelem.* 2013;1(1):20.
15. Shepard ELC, Wilson RP, Quintana F, Gomez LA, Liebsch N, Albareda DA, et al. Identification of animal movement patterns using tri-axial accelerometry. *Endang Species Res.* 2010;10:47–60.
16. Yoda K, Naito Y, Sato K, Takahashi A, Nishikawa J, Ropert-Coudert Y, et al. A new technique for monitoring the behaviours of free-ranging Adélie penguins. *J Exp Biol.* 2001;204:685–90.
17. Krause J, Krause S, Arlinghaus R, Psorakis I, Roberts S, Rutz C. Reality mining of animal social systems. *Trends Ecol Evol.* 2013;28:541–51.
18. Banarjee D, Diagle CL, Dong B, Wurtz K, Newberry RC, Siegford JM, et al. Detection of jumping and landing force in laying hens using wireless wearable sensors. *Poult Sci.* 2014;93:2724–33.
19. Liikavainio T, Bragfge T, Hakkarainen M, Jurvelin JS, Karjalainen PA, Arokoski JP. Reproducibility of loading measurements with skin-mounted accelerometers during walking. *Arch Phys Med Rehabil.* 2007;88:907–15.
20. Chen S, Amiyaal L, White BJ, Sanderson MW, Lanzas C. Spatial-temporal dynamics of high-resolution animal social networks: what we can learn from domestic animals. q-bio PE. 2014. arXiv: 1407.6074.

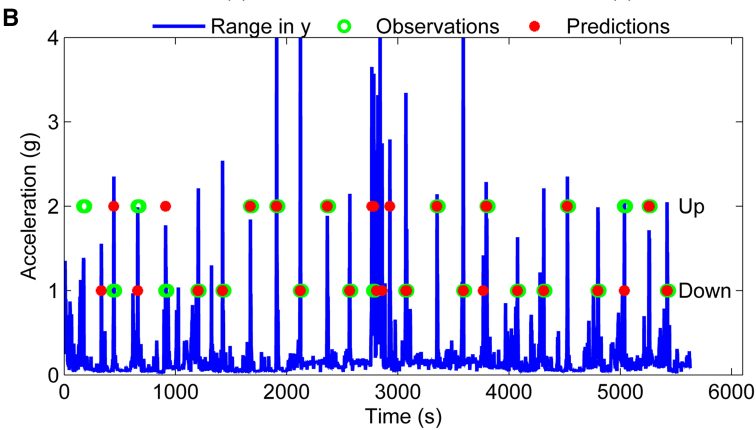
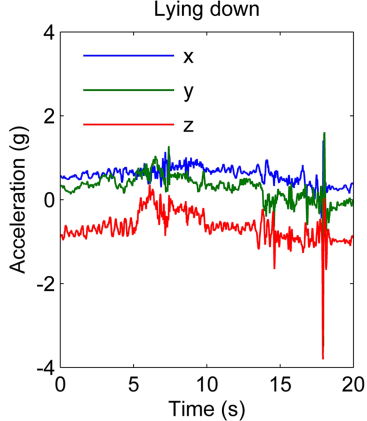
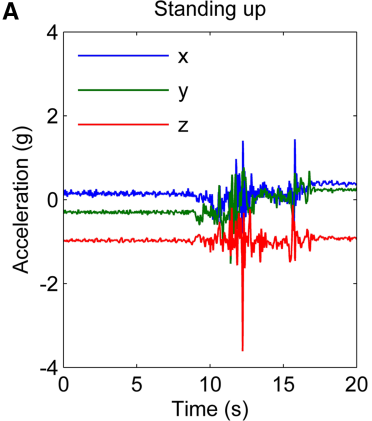
21. Yoda K, Sato K, Niizuma Y, Kurita M, Bost CA, Le Maho Y, et al. Precise monitoring of proposing behaviour of Adélie penguins determined using acceleration data loggers. *J Exp Biol.* 1999;202:3121–6.
22. Mclune DW, Marks NJ, Wilson RP, Houghton JDR, Montgomery IW, McGowan NE, et al. Tri-axial accelerometers quantify behaviour in the Eurasian badger (*Meles meles*): towards an automated interpretation of field data. *Animal Biotelem.* 2014;2:5.
23. Langrock R, King R, Mathiopoulos J, Thomas L, Fortin D, Morales JM. Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology.* 2012;93:2336–42.
24. Guo Y, Poulton G, Corke P, Bishop-Hurley GJ, Wark T, Swain DL. Using accelerometer, high sample rate GPS and magnetometer data to develop a cattle movement and behaviour model. *Ecol Model.* 2009;220:2068–75.
25. Yin L, Hong T, Liu C. Estrus detection in dairy cows from acceleration data using self-learning classification models. *J Comput.* 2013;8:2590–7.
26. Morttram T. Wireless technologies for managing grazing animals. In. Profitable and sustainable grazing systems—moving forward with science. Worcester, UK. 2013.
27. González LA, Tolkamp BJ, Coffrey MP, Ferret A, Kyriazakis I. Changes in feeding behaviour for the automatic monitoring of health disorders in dairy cows. *J Dairy Sci.* 2008;91:1017–28.
28. Palmer MA, Law R, O’Connell NE. Relationships between lameness and feeding behaviour in cubicle-housed Holstein-Friesian dairy cows. *Appl Anim Behav Sci.* 2012;140:121–7.
29. Medrano-Galarza C, Gibbons J, Wagner S, de Passillé AM, Rushen J. Behavioural changes in dairy cows with mastitis. *J Dairy Sci.* 2012;95:6994–7002.
30. Blackie N, Amory JR, Bleach ECL, Scafie J. The effect of lameness on lying behaviour of zero grazed Holstein dairy cattle. *Appl Anim Behav Sci.* 2011;14:124–37.
31. Spence AJ, Tan H, Wilson A. Accuracy of the TurfTrax racing data system for determination of equine speed and position. *Equine Vet J.* 2010;40:680–3.
32. Homburger H, Schneider MK, Hilfiker S, Lüscher A. Inferring behavioral states of grazing livestock from high-frequency position data alone. *PLoS One.* 2014;9, e114522. doi:10.1371/journal.pone.0114522.
33. Bailey DW, Van Wagoner HC, Weinmeister R. Individual animal selection has the potential to improve uniformity of grazing on foothill rangeland. *Rangl Ecol Manag.* 2006;59:351–8.
34. Qasem L, Cardew A, Wilson A, Griffiths I, Halsey LG, Shepard ELC, et al. Tri-axial dynamic acceleration as a proxy for animal energy expenditure; should we summing values or calculating the vector? *PLoS One.* 2012;7, e31187. doi:10.1371/journal.pone.0031187.

35. Gleiss AC, Wilson RP, Shepard ELC. Making overall dynamic body acceleration work: on the theory of acceleration as a proxy for energy expenditure. *Methods Ecol Evol.* 2011;2:23–33.
36. Matlab release 2014a. Natick, Massachusetts: The MathWorks, Inc.; 2014
37. Rezek I, Roberts SJ. Ensemble hidden Markov models for biosignal analysis. In. 14th International Conference on Digital Signal Processing. Santori, Greece. 2002.
38. Murphy KP. Machine learning: a probabilistic perspective. Cambridge, Massachusetts: MIT Press; 2012.
39. Rabiner L. Tutorial on hidden Markov models and selected applications on speech recognition. *Proc IEEE.* 1989;77:257–86.
40. Cortes C, Vapnik V. Support-vector networks. *Mach Learn.* 1995;20:273–97.
41. Duan KB, Keerthi SS. Which is the best multiclass SVM method? *Lect Notes Comput Sc.* 2005;3541:278–85.
42. Hsu CW, Lin CJ. A comparison of methods of multiclass support vector machines. *IEEE Trans Neural Netw.* 2002;13:415–25.









Additional files provided with this submission:

Additional file 1. Methods to develop an algorithm for behavioral state classification. Further details of the decision-tree, k-means, HMM and SVM algorithms, results from a systematic exploration of the decision-tree parameters (window size and threshold values), and further details of the transition detection algorithm (693kb)

<http://www.animalbiotelemetry.com/content/supplementary/s40317-015-0045-8-s1.pdf>

Additional file 2. Feature characteristics for classification. The feature characteristics used in this study to classify the three different behaviours using a 1-min, 5-min and 10-min window size (data from 6 dairy cows over a 3-day period) (107kb)

<http://www.animalbiotelemetry.com/content/supplementary/s40317-015-0045-8-s2.txt>