Computers and Electronics in Agriculture 110 (2015) 91–102

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Behavioral classification of data from collars containing motion sensors in grazing cattle

L.A. González [a,e,*], G.J. Bishop-Hurley [b,e], R.N. Handcock [c,e], C. Crossman [d,e]

[a] CSIRO Animal, Food and Health Sciences, PMB PO, Aitkenvale, Queensland 4814, Australia
[b] CSIRO Animal, Food and Health Sciences, 306 Carmody Road, St. Lucia, Queensland 4067, Australia
[c] CSIRO Animal, Food and Health Sciences, Private Bag 5, Wembley 6913, Western Australia, Australia
[d] CSIRO Information and Communication Technologies Centre, Pullenvale 4069, Queensland, Australia
[e] CSIRO Sustainable Agriculture Flagship, Australia

A R T I C L E I N F O

Article history:
Received 28 March 2014
Received in revised form 29 September 2014
Accepted 26 October 2014

Keywords:
Accelerometer
GPS
Behavior
Livestock
Automatic
Wireless

A B S T R A C T

Remote monitoring of animal behavior offers great potential to improve livestock management however technologies able to collect data at high frequency and accurate data classification methods are required. The objective of this study was to develop a methodology capable of performing unsupervised behavioral classification of electronic data collected at high frequency from collar-mounted motion and GPS sensors in grazing cattle. Two independent trials were conducted, one for developing the classification algorithm (4 groups of 11 steers) and a second for its evaluation (14 steers). Each steer was fitted with a collar containing GPS and a 3-axis accelerometer that collected data at 4 and 10 Hz, respectively. Foraging, ruminating, traveling, resting and 'other active behaviors' (which included scratching against objects, head shaking, and grooming) were observed and recorded continuously at the nearest second in animals wearing collars. Collar data were aggregated to 10-s intervals through the mean (indicative of the position of the neck and travel speed) and standard deviation (SD; indicative of activity level) and then log-transformed for analysis. The histograms of travel speed showed 3 populations and observations revealed these populations represented stationary, slow and fast travel behaviors. The histograms of the accelerometer *X*-axis mean showed populations corresponding with behaviors of head down or head up. The histograms of the accelerometer *X*-axis SD showed 3 populations representing behaviors with high, medium and low activity levels. Mixture models were fitted to data from each animal in both trials to calculate threshold values corresponding to where behaviors transitioned between different states. These thresholds from the 3 sensor signatures were then used in a decision tree to classify all 10-s data where behaviors were unknown into 5 mutually exclusive behaviors. The algorithm correctly classified 85.5% and 90.5% of all data points in the development and evaluation datasets, respectively. Foraging showed the greatest sensitivity (93.7% and 98.4%) and specificity (94.6% and 99.4%) followed by ruminating (sensitivity 97% and 87%, and specificity 90% and 95%) for development and evaluation trials, respectively. Major advantages of mixture models include computational efficiency suitable for large data sets (e.g. >2 million data lines), minimal requirement for training datasets, and estimation of threshold values for individual animals under unknown and varying environmental conditions. The technology and methodology allows for the automatic and real-time monitoring of behavior with high spatial and temporal resolution which could benefit livestock industries beyond the research domain for improved animal and ecological management.

1. Introduction

Measuring animal location and behavior across different spatial and temporal scales can facilitate understanding of the factors that drive resource selection (Owen-Smith et al., 2013), growth, reproduction and survival (Gaillard et al., 2010), response to disease (González et al., 2008) and coping mechanisms with environmental conditions (Anderson et al., 2013). Therefore, monitoring behavior in near real-time can enable more accurate and timely management decisions to optimize animal performance, welfare and environmental outcomes. In grazing systems, Global Positioning Systems (GPS) and motion sensors (e.g. accelerometers) can monitor animal

* Corresponding author at: The University of Sydney, Centre for Carbon, Water and Food, Faculty of Agriculture and Environment, 380 Werombi Road, Camden, NSW 2570, Australia. Tel.: +61 2 4601 9834.
E-mail address: Luciano.Gonzalez@sydney.edu.au (L.A. González).

http://dx.doi.org/10.1016/j.compag.2014.10.018
0168-1699/

behavior in near real-time when integrated into wireless sensor networks (Wark et al., 2007; Handcock et al., 2009; Nadimi et al., 2012). However, the challenge in using sensor data is to automate the differentiation of behavioral activities. Several methodologies have previously been used to classify sensor data into behavioral states (Martiskainen et al., 2009; Ungar et al., 2005). However, those methodologies require a training dataset (e.g. direct observations) in every experiment or condition and do not account for differences among individual animals and devices. A methodology that is robust for use on data collected from different devices would also reduce the need to calibrate sensors and to fit collars with the same tension (Anderson et al., 2013). Mixture models address these constraints by allowing unsupervised classification of data using probability density functions (PDF; McLachlan and Peel, 2000; Tolkamp et al., 2000). Characterizing the structure of behavior using mixture models combined with observations in one experiment allows estimation of parameters describing the PDF without the need of direct observation in subsequent experiments.

The objective of the present study was to develop and evaluate a methodology to characterize the structure of electronically obtained data and classify such data into behavioral activities including foraging, resting, ruminating, traveling and 'other active behaviors' using GPS and motion sensor data from collars worn by steers. The goal was to characterize these behaviors from raw data once a dataset had been 'fingerprinted' based on observations and then apply the method to an independent dataset.

2. Materials and methods

All experimental procedures were approved by the institutional Animal Ethics Committee (Approval # A10/2010, A11/2010, and A8/2011). Two trials were conducted to collect data from electronic monitoring collars and direct visual observations of animal behavior. Data from one trial was used to develop an algorithm to classify collar data into behavioral categories that identify behaviors. Data from the remaining trial were used to evaluate the accuracy of the classification algorithm. Both trials were conducted at the Commonwealth Scientific and Industrial Research Organization (CSIRO) Lansdown Research Station near Townsville, Queensland, Australia (18°39′42″S and 146°51′12″E, elevation 63 m) using Brahman, Belmont Red Composites and crossbred Brahman steers. Paddocks contained tropical vegetation dominated by *Urocloa* spp., *Stylosanthes* spp., *Macroptilium* spp. and *Chloris* spp., and contained more than 2000 kg of DM/ha. Trees, woody vegetation and shrubs edible by cattle were not prevalent in trials of the present study. Steers had *ad libitum* access to one water trough in each paddock. The algorithm-development trial involved 4 groups of 11 steers with a mean initial body weight [BW] of 403 ± 30 kg and a mean average daily gain [ADG] of 0.37 ± 0.40 kg/day during 3 experimental weeks. Each steer was fitted with a CSIRO electronic cattle monitoring collar (Wark et al., 2007) for 21 days in October 2011 and each group of steers grazed a 7 ha flat treeless paddock. The evaluation trial was conducted in a 15 ha paddock and involved a single group of 14 steers (initial BW = 433 ± 32 kg; ADG = −0.63 ± 1.50 kg/day) fitted with collars for 10 days during November 2012.

2.1. Description of CSIRO cattle monitoring collars

The CSIRO collars (Wark et al., 2007) had a 20-channel GPS receiver chip (U-Blox, Thalwil, Switzerland), a GPS antennae, a microcontroller (Fleck™, CSIRO, Australia), 4 D-cell batteries in series (Duracell, Australia), a 4 GB micro Secure Digital card for data storage, a piezoelectric micro-electromechanical system (MEMS) chip containing a 3-axis accelerometer and a 3-axis magnetoresistive sensor (HMC6343 Honeywell, Plymouth, MN), and wireless network communication capability with 900 MHz radio antennae. All components except the GPS antenna were sealed in a plastic box and positioned on the animal such that the box remained below the neck when worn, and the GPS antenna remained on top of the neck to improve signal reception. The GPS measures animal location on the earth surface and calculates the speed the GPS unit is moving across the landscape on board of the collar. The accelerometer measures inertial acceleration in a 3 axis inertial and gravitational frame (fore-aft, right-left, up-down) with the *X*-axis detecting the vertical or up-down direction (tilt), the *Y* axis detecting the fore-aft and the *Z* axis the right-left direction. The collars were programmed to collect GPS data at 4 Hz (i.e. 345,000 data points/day) and accelerometer data at 10 Hz (i.e. 862,500 data points/day). Effective actual battery life was 12–14 days. After collar retrieval from the steers at the end of the trial, the memory storage cards were removed and the data downloaded.

2.2. Behavioral measurements

Direct visual behavioral observations (the gold standard method) were recorded using continuous sampling on random animals in each group for both experiments by recording the animal ID, time (to the nearest second) and type of activity at every occasion in which the steers changed from one activity to another (Altmann, 1974). Therefore, the number of animals from which data was collected and the length of the observation periods were variable for each behavioral activity. The percentage of data collected by electronic collars which had accompanying behavioral observations in the final datasets was 0.6%. Observations were made on days 2, 3, 5, 6, 7 and 8 after fitting the collars to the animals for the development trial with a total of 18 h of visual observations made during daylight hours from 1000 to 1800. The observations for the evaluation trial totaled 25 h over 5 days (days 1, 2, 3, 5 and 8 after fitting the collars). Animals got used to the collars very quickly and observations were started on the day of collar deployment. Five mutually exclusive activities (i.e. steers could only perform 1 activity at the time) were recorded: foraging, ruminating, resting, traveling and other active behaviors. Initially, grazing with the head down, browsing and searching for food were recorded separately however grazing occupied more than 95% of all foraging behaviors and it was decided to merge the 3 into the activity called 'foraging' for simplicity. Therefore, foraging was considered the act of searching for food while walking short distances with the head down without picking food up with the mouth, grazing with the head down while apprehending the forage with the mouth, browsing consisting of apprehending vegetation with the head held leveled with respect to the ground surface, and chewing either with the head down or the head leveled with regards to the ground surface. Ruminating was defined as chewing the cud while standing up or lying on the ground. 'Other active behaviors' was a category created to include vigorous head movement while standing with no forward movement of the body such as when rubbing or scratching their own body against an object (e.g. fence post), licking themselves or other herd mates with the mouth or tongue, or head shacking when attempting to get rid of insects. Traveling was defined as forward moving without foraging including walking or running and while the animal could be ruminating or not ruminating but not engaged in foraging activities. Resting was considered when the animals were stationary and not foraging, ruminating, traveling or performing other active behaviors (either in standing or lying down postures). The *WhatISee* smart phone application was used to register the activities (https://itunes.apple.com/us/app/whatisee/id332512569?mt=8) for iPhone, iPod or iPad (Apple, Cupertino, USA). Information to correct the time difference between *WhatISee* (local time) and the collars (GPS-system time) came from an online tool (www.leapsecond.com/java/gpsclock.htm).

2.3. Description of the data

Datasets collected from the 3-axis accelerometers fitted to each animal collar consisted of time and a reading of the X–Y–Z axis values (-4 to $4g \times 10^4$). The data recorded by the GPS included the spatial position of the animal on the earth-centered earth-fixed coordinate system (X-, Y-, and Z-axis coordinates in cm; Cai et al., 2011), the time as in the GPS system, and the 3-dimensional travel speed (cm/s) along with measures regarding the accuracy and quality of each measurement (e.g. number of satellites used to calculate the position, position dilution of precision and 3D accuracy).

2.4. Data processing

All data were processed using the SAS statistical software (v9.2, SAS Institute Inc., Cary, NC). No differential post-processing correction was done for the GPS data before analysis. The spatial location and travel speed datasets from the GPS chip were processed to determine the distance between consecutive GPS locations for each individual animal. A 2-dimensional plane considering longitude and latitude was assumed because the paddocks were virtually flat (height above the ellipsoid was disregarded). Erroneous values in the GPS data were identified and removed according to the following criteria where: the GPS fix was not 3-dimensional, the position dilution of precision was >6 (Hurn, 1993; cited by Ganskopp and Johnson, 2007), the accuracy of the 3-dimensional position was >2 m, travel speed was greater than 1.2 m/s (based on maximum speed of a Bos indicus plus 20%; Heglund and Taylor, 1988), travel speed accuracy was greater than 0.6 m/s, or if the distance between 2 consecutive GPS locations was greater than 3 m (based on analysis of unpublished data).

Data from the electronic collars were aggregated after downloading to a personal computer to reduce the amount of data to handle, the computing power required for developing the behavioral classification algorithms, and to obtain meaningful metrics to represent head position and activity level. Data from the accelerometers and GPS were aggregated by calculating the mean and standard deviation (SD) across 10 s intervals, generating 8 variables for analysis: 10-s mean and SD of the X-, Y- and Z-axis of the accelerometer and travel speed from GPS. The 10-s means indicate the mean motion and tilt states of the head and neck whereas the 10-s SD indicate the activity level or changes in acceleration of the neck. The 10-s time interval chosen for aggregation was based on a preliminary analysis which indicated that longer intervals of 20, 30 and 60 s reduced the accuracy of behavioral classification methods (data not shown). The entire experimental dataset was aggregated to produce 10-s means and SD's which was then used to generate a subset of sensor data matched with behavioral activities as recorded during observations (deleting all data with no recorded activity; $n = 8665$ for the development trial). Each 10-s data point from the collars was assigned a behavioral activity according to the start and end time of episodes of activity recorded during visual observations for each animal. Data were deleted if the episode of an activity was shorter than 30 s because of the uncertainty associated with whether a particular 10-s value was within the time frame. The first 10-s observation of an activity bout was not considered for analysis because the steer might have been performing the activity during only a fraction of the 10-s interval.

2.5. Statistical analysis and algorithm development

All electronic data obtained from the collars were transformed to the natural logarithm prior to analysis to normalize their distributions and homogenize variances (Tolkamp et al., 2000). Variables containing negative or zero values were made positive before log-transformation by adding positive values to all data because logarithm cannot be calculated for numbers equal or less than zero (i.e. 10,000 for 10-s means of the accelerometer data, and 1 for SD's and mean travel speed).

Briefly described, the procedure to develop the algorithm to classify each 10-s data point from the electronic collars into 1 of the 5 behavioral activities consisted of analysis on 2 datasets: the first dataset (A) was a subset of data where behaviors were identified from observations and it was used to determine differences among activities from values of the sensor data (step A1), to inspect the frequency distributions (histograms) of data with different activities (step A2), to select variables suitable for decision trees (step A3), and to construct conceptual decision trees (step A4). The second dataset (B) contained all data from the development trial where behaviors were unknown and it was used to fit PDF in mixture models (step B5; McLachlan and Peel, 2000; Tolkamp et al., 2000) which defined threshold values that separate populations of data points (step B6). Then, these threshold values formed part of the classification and decision trees as described in detail below. Finally, the proportion of daily time that individual animals spent in each activity was estimated.

The 4 steps (A1–A4) aimed at selecting the variables to form part of the classification and decision tree were applied as follows:

Step A1 consisted of determining which of the variables generated by the electronic sensors were significantly affected by behavioral category. To make this determination, data were first averaged on a per-animal and activity basis, and then mixed-effects regression analyses were conducted where activity was considered a fixed effect and animal as a random effect. Outliers were detected (studentized residual greater than 2.5) and removed from further analysis as appropriate, and differences among means obtained after Bonferroni's adjustment for multiple comparisons. Those variables which showed differences in the mean values among activities were then selected for further examination of frequency histograms for overlaps or breakpoints separating populations of data points suitable for unsupervised classification as described in step A2. Step A2 consisted of plotting frequency histograms of the selected variables for each behavioral activity overlaid on the same plot to visualize the number of populations of data points reflecting different behavioral activities. Frequency histograms also allow visualizing the presence of decision boundary values between populations suitable to fit mixture models for unsupervised classification. Step A3 consisted of further selection of the variables selected in step A1 by using Pearson correlations to select one from any 2 or more variables showing high correlations among each other. Step A4 consisted of the construction of conceptual classification and decision trees using the selected variables.

Four additional steps (B5–B9) were performed on both the entire experimental datasets ($n = 3{,}056{,}000$ and 662,602 data points for the development and evaluation trial, respectively) and on a per animal basis for both trials where the behavioral activities being performed by the animals were unknown. Step B5 consisted of fitting mixture distribution models to obtain threshold values that separate 2 populations of data points and are used for the classification of data points into 1 of the 5 behaviors. Mixture distributions arise from datasets containing 2 or more populations of data points which in data from electronic collars represent different positions and activity levels of the collar. The frequency distribution of such mixture datasets can allow visualization of the number of populations (e.g. multimodal distributions), shape (e.g. normal Gauss bell shape), skewedness, and degree of overlap between populations (McLachlan and Peel, 2000). Thus, each population within the mixture show best fit to a particular distribution according to its shape such as normal (N), lognormal (LN) or Weibull (W) distributions. Finite mixture models describe these mixture datasets using the sum of the probability density functions (PDF) by the mixing probabilities with k number of populations (e.g. $\text{PDF}_{\text{total}} = p_1 \times \text{PDF}_1 + p_2 \times \text{PDF}_2$ where p_1 and p_2 are the mixing probabilities or proportions of data points in population 1 and 2, respectively). Thus, using the mathematical expression of the PDF, mixture models allow calculating the probability of an event and to perform cluster analysis or unsupervised learning (i.e. calculate the probability of a data point to belong to one of the populations or a behavioral state, and assign a data point to the population or activity with greatest posterior probability). In the present study, the PDF of each population of data points identified in the mixture of distributions in step A2 was fitted to a 2-parameter N, LN, and W PDF with all possible combinations (e.g. N–N, N–LN, LN–N, LN–LN, N–W, LN–W, W–W, W–N, and W–LN for a bimodal distribution). The 2 parameters in the PDF describe the location (e.g. mean or median) and shape (e.g. skewedness or length of the tail). The combination of distributions with the best fit (i.e. lowest Bayesian Information Criteria, BIC) was selected. Fitting of PDFs was done using the NLMIXED procedure in SAS, applied to the data pooled across all animals and then to each animal–collar combination. Step B6 used the parameters describing each population's PDF to calculate the threshold values (i.e. decision boundaries) which separate populations of data points from each other and minimize the miss-assignment rate of data points to the wrong population (McLachlan and Peel, 2000; Tolkamp et al., 2000). Step B7 used the threshold values from each of the selected variables obtained in the previous step to form part of the conceptual trees from step A4 to classify each data point of the entire experimental dataset into 1 of 5 activities.

Once data points were classified into a behavioral activity, the probability to belong to the actual predicted activity according to the sensor data was calculated (i.e. its location in the continuum PDF). Data points with low probability (<0.7) to belong to the predicted behavior ('borderline' data points) were then 'revised' and imputed with the most frequent activity for the period of the minute preceding and following the actual data point. This process accounted for the fact that animals perform behaviors in bouts and the accuracy of classification could be improved if the predominant behavior surrounding the actual data point is considered to predict the actual behavior. To achieve this revision, the probability of every data point to belong to each of the 5 activities (i.e. probability of class membership) was calculated using the cumulative distribution function (CDF) of observed values for mean travel speed, and mean and SD of the accelerometer X axis. For example, data from the X axis accelerometer ranged from 8 to 10 and showed 2 populations of data points, one with low values indicating head down position and another with high values indicating head up position, and a threshold value of 9.20 separating head down from head up positions (i.e. values lower than this threshold are classified with head down while values greater than 9.20 would be classified with head up). A data point with a value of 9.15 would have lower probability to belong to head down position compared to a data point with a value of 8.6. The closer a data point is to the threshold value then the lower will be the probability to belong to the assigned population while values equal to the threshold values have the same probability to belong to either population. In step B8, the CDF from the decision boundary value was calculated using the fitted PDF's parameters according to Eqs. (1)–(7). During step B9, activities classified with a probability lower than 0.7 were imputed with the activity that yielded the greatest (i.e. most frequent) probability during the time frame of 1 min before and 1 min after the actual data point, i.e. with the greatest 1-min centered mean posterior probability.

$$P_{\text{Foraging}} = (1 - \text{CDF}_{\text{P1 mean X-axis accelerometer}}) \times \text{CDF}_{\text{P3 SD X-axis accelerometer}} \quad (1)$$

$$P_{\text{Resting}} = (1 - \text{CDF}_{\text{P1 SD X-axis accelerometer}}) \quad (2)$$

$$P_{\text{Ruminating}} = \text{CDF}_{\text{P2 SD X-axis accelerometer}} \quad \text{if } \text{CDF}_{\text{P2 SD X-axis accelerometer}} < 0.5 \quad (3)$$

$$P_{\text{Ruminating}} = (1 - \text{CDF}_{\text{P2 SD X-axis accelerometer}}) \quad \text{if } 0.5 < \text{CDF}_{\text{P2 SD X-axis accelerometer}} < 0.998 \quad (4)$$

$$P_{\text{Ruminating}} = 0 \quad \text{if } \text{CDF}_{\text{P2 SD X-axis accelerometer}} > 0.998 \quad (5)$$

$$P_{\text{Travel}} = \text{CDF}_{\text{P3 mean Speed}} \times \text{CDF}_{\text{P3 SD X-axis accelerometer}} \quad (6)$$

$$P_{\text{Other active behaviors}} = \text{CDF}_{\text{P2 mean X-axis accelerometer}} \times \text{CDF}_{\text{P3 SD X-axis accelerometer}} \quad (7)$$

where P is the probability of membership of a data point to the behavioral activity, and $\text{CDF}_{\text{P1–P2–P3}}$ are the cumulative distribution functions of the first (i.e. lowest values), second or third (i.e. largest values) population of data points for each variable, respectively.

Once the final loop was performed (step B9), the proportion of daily time spent in each of the predicted activities was calculated for each animal, disregarding days with more than 10% of data points lost in regards to the expected number of data points (8640 data points/animal/d) due to deletion of data as a result of the aforementioned data processing method or to malfunction in collars.

The entire experimental datasets were classified into 1 of 5 behavioral activities according to each decision tree described in A4 and the procedure described in B5–B9. Then, the best algorithm (tree) was selected based on the following priority of statistical measures: (a) the greatest sensitivity and specificity for detecting foraging, (b) the greatest sensitivity and specificity for detecting ruminating, and (c) the greatest overall coefficient of agreement Kappa (Cohen, 1960) between actual (from visual observations) and predicted activities (from the algorithm). Statistical measures were calculated using the number of data points that are true positive (TP; e.g. a data point classified as foraging by the algorithm when observations of the animal confirmed that it was truly foraging), true negative (TN; e.g. a data point classified as not foraging when observations also confirmed that the animal was performing other behavior than foraging), false positive (FP; e.g. a data point classified as foraging when observations indicated that the animal was not foraging) and false negative (FN; e.g. a data point classified as not foraging when the animal was foraging). Thus, each activity had an associated precision or sensitivity (TP/[TP + FN]), specificity (TN/[TN + FP]) and statistical concordance (the proportion of predicted data points in relation to actual (observed) data points for a given activity.

3. Results

3.1. Relating the sensor signatures from the collars to the corresponding behavior

In this section, when the following terms are used they refer to 10-s mean and SD of the data from the specified sensors after log-transformation: mean and SD of X-, Y- and Z-axis of the accelerometers and travel speed. Each behavioral activity showed a characteristic sensor signature coming from the motion sensors. The mean of the accelerometer X-axis values were lower during foraging compared to the values when the steers were ruminating, resting, traveling or performing other active behaviors ($P < 0.01$; Table 1) as a result of the head position down while grazing. The activity level of the neck, as measured through the SD of the accelerometer X-axis, was greatest during foraging and traveling, intermediate during other active behaviors, and lowest during ruminating and resting ($P < 0.05$). The SD of the accelerometer X-axis was more sensitive in detecting differences among behaviors compared to the Y- and Z-axis. The SD of the X-axis was able to separate ruminating from resting ($P < 0.05$), and foraging from other active behaviors ($P < 0.05$). Travel was fastest during traveling, intermediate during foraging and other active behaviors, and slowest during ruminating and resting ($P < 0.05$; Table 1). Pearson correlation coefficients indicated high correlations among the SD's for all 3 axes of the accelerometer ($r \geqslant 0.95$; data not shown) because high variability exists in all 3 axes when animals are involved in behaviors in which the neck and head is moving, and low when steers are inactive as during resting. Correlation coefficients between 0.6 and 0.7 (data not shown) were found between mean travel speed and the SD of most accelerometer-related variables as a result of high movement of the neck in all directions while traveling.

3.2. Frequency distributions and probability density functions

Based on its ability to capture head position and activity level, and considering the high correlation with other variables, the mean and SD of the accelerometer X-axis data were selected to explore whether these variables present different populations of data points that represent different behaviors. Mean travel speed was also selected because it was the only variable that could be used to differentiate forward movement from all other behaviors. Frequency distributions of these 3 selected variables for each known behavior are shown in Fig. 1. For the 10-s mean of the accelerometer X-axis, the population of data points corresponding to foraging is well separated from the rest of the activities (Fig. 1A). However, high overlap exists between the frequency distributions of the other 4 activities measured, which indicate that these activities present more of a challenge to separate on the basis of this variable. Visual inspection of Fig. 1A suggests that a value of just over 9 log-units of the mean of the accelerometer X-axis data could be a suitable criterion for separation of head down from head up positions while minimizing the miss-assignment of data points to the wrong population.

The frequency distribution of mean travel speed was characterized by 3 populations of data points, although the first population (representing stationary states dominated by ruminating and resting) overlapped with the second population (representing slow movement dominated by foraging), whereas traveling (forward movement) was separated and characterized by high rates of travel greater than 36 cm/s or 1.28 km/h (Fig. 1B). Histograms for the SD of the accelerometer X-axis data showed 3 clearly separated and distinct populations (Fig. 1C), with the first population (<4.6 log-units) formed by activities with very low activity levels of the neck representing inactive states (i.e. resting), the second population (between 4.6 and 6.1 log-units) representing a medium activity level of the neck (i.e. ruminating), and the third and largest population (>6.1 log-units) representing activities where the neck has a high activity level (i.e. foraging and traveling).

The frequency distribution of the mean and SD of the accelerometer X-axis and of the mean travel speed in the entire experimental datasets where behaviors were unknown were also plotted to determine if these variables followed similar patterns to those in the subset of data where behaviors were known. If histograms with the entire datasets also show different populations of data points then this would be an indication that unsupervised classification of individual data points into the correct population (head position and activity level) is possible and this would allow unsupervised classification of data points into behavioral states. Interestingly, the distinct populations of data points found in the data subset with matching behavioral states from visual observations (Fig. 1) were also evident in the entire experimental dataset where behavioral activities were not known (Figs. 2 and 3). This confirms that finite mixture distributions (i.e. a distribution made up of two or more component populations) exist in data obtained from motion sensors because animals spent most of their time in definite positions or activities, including head down–head up, low–medium–high activity levels, and low–medium–high travel speeds. The parameters in the mathematical equations describing each population's PDF (i.e. mixing proportions, location and shape or scale parameters), the value in the X-axis where the PDF of individual populations intersect with each other, and the goodness-of-fit statistics (BIC) were obtained by fitting finite mixture models to the entire experimental dataset. Overlays of both the observed and fitted (or modeled PDF) frequency distributions are shown in Fig. 2 for the data pooled across animals to show the overall shape of the mixture distributions although this fitting technique was also made on a per animal basis to account for differences among animals and collars.

For the mean of the accelerometer X-axis data, the best fit was a mixture of lognormal and Weibull distributions to separate the head down from head up positions (best fit in 53% of all animals). The SD of the accelerometer X-axis data resulted in the best fit to a

Table 1
Differences among behaviors based on 10-s mean and standard deviation (SD) of data obtained from 3-axis accelerometers and GPS sensors in collars worn by grazing steers (classification algorithm development trial).*

		Foraging	Ruminating	Resting	Travel	Other active behaviors†	SEM	P-value
Accelerometer	Mean							
	X	8.87[c]	9.43[a]	9.37[a,b]	9.32[a,b]	9.29[b]	0.032	<0.001
	Y	8.76[a]	8.69[a]	8.67[a]	8.63[a]	8.71[a]	0.067	0.15
	Z	9.71[a]	9.71[a]	9.70[a]	9.73[a]	9.70[a]	0.010	0.43
	SD							
	X	7.04[a]	5.34[c]	4.89[d]	6.90[a,b]	6.65[b]	0.085	<0.001
	Y	6.51[a]	4.96[b]	4.65[b]	6.55[a]	6.24[a]	0.088	<0.001
	Z	6.51[a]	4.96[b]	4.65[b]	6.55[a]	6.21[a]	0.094	<0.001
Speed	Mean	2.65[b]	2.06[c]	2.06[c]	4.11[a]	2.59[b]	0.082	<0.001
	SD	2.18[b]	1.67[c]	1.76[c]	2.72[a]	2.23[b]	0.069	<0.001
n‡	–	23	21	19	10	17	–	–

* All variables were log-transformed after the 10-s mean and SD were obtained. Speed is presented as Log (1 + cm/s) and accelerometer data as Log ($10{,}000 + g \times 10^4$).
† Other active behaviors included scratching, self- and cross-grooming, and head shacking to get rid of insects which involved head/neck movement or shaking.
‡ Number of animals from which behavioral data were collected.
a,b,c,d Mean without a common superscript differ ($P < 0.05$).

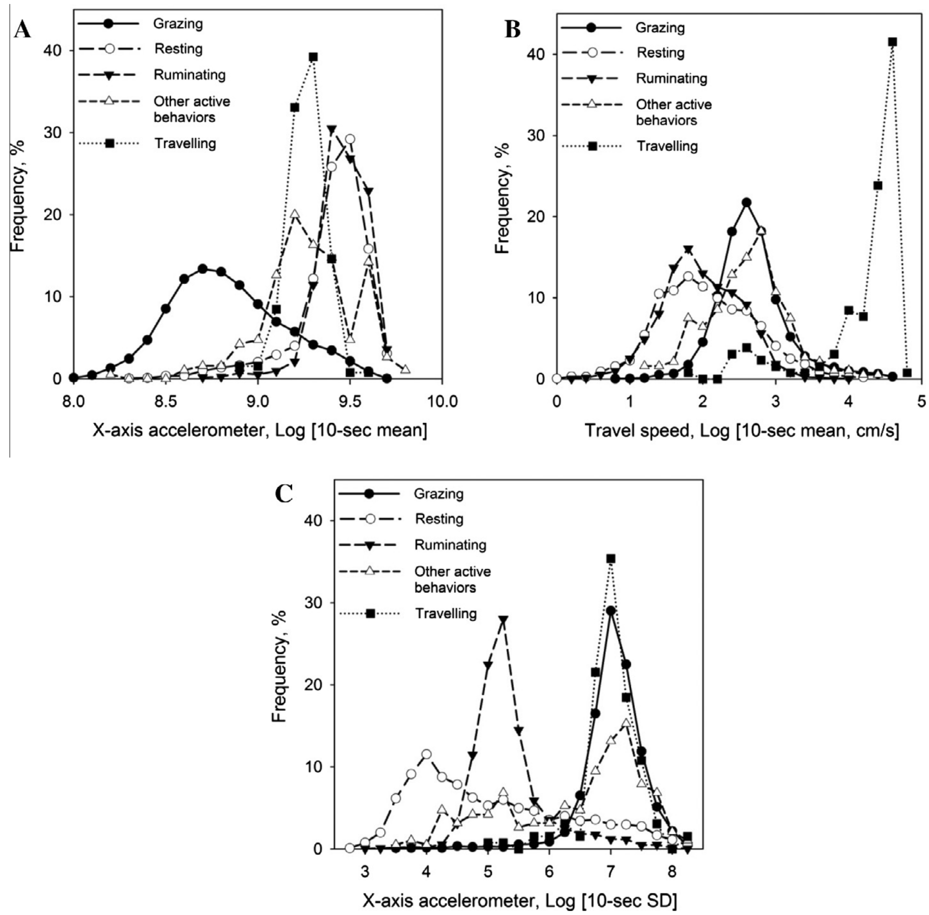

Fig. 1. Frequency distribution of 10-s mean of the accelerometer *X*-axis (A) and travel speed (B), and 10-s SD of the accelerometer *X*-axis (C) from electronic collars worn by steers to monitor foraging, ruminating, resting, traveling and performing other active behaviors ($n = 8665$). Other active behaviors included those with vigorous head movement while standing with no forward movement of the body and not engaged in foraging activities. Behavior was recorded by observing steers. All variables were log-transformed after the 10-s mean and SD were obtained. Speed is presented as Log (1 + cm/s) and accelerometer data as Log ($10{,}000 + g \times 10^4$).

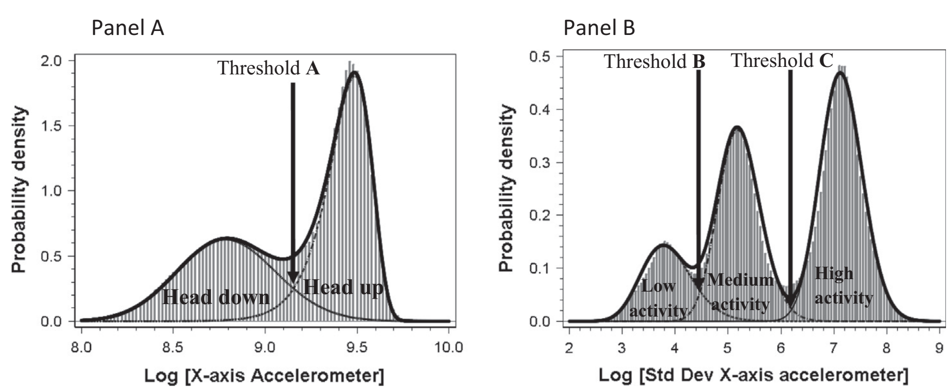

Fig. 2. Frequency distribution of observed data (gray bars) and fitted probability density function (PDF; continuous lines) of selected variables obtained from sensors in collars worn by grazing cattle ($n = 42$). Variables plotted are 10-s mean (panel A) and SD (panel B) of the accelerometer *X*-axis. The thin gray lines represent the PDF of each component population in the total PDF (thick black lines). The point where individual populations intersect determines the threshold value (threshold A, B, and C) used to separate individual populations such that miss-assignment of data points to the wrong population is minimized. Figures presented are from data pooled across all individual animals in the development trial. Raw data (10 Hz; $n = 327{,}230{,}000$) was aggregated for 10-s intervals by their means and SDs ($n = 3{,}310{,}988$ data points from 42 steers). All variables were log-transformed after the 10-s mean and SD were obtained.

mixture of 3 component lognormal distributions in 60% of the steers allowing to separate low, medium and high activity states (data not shown). A model with a mixture of normal–lognormal–Weibull distributions was the best fit for travel speed data (48% of animals) reflecting behavioral states characterized as stationary, slow and fast. Parameters describing the model fitted PDFs to both the data from each animal and to the entire experimental dataset, are shown in Table 2. In the development trial, the first distinct population, corresponding to head-down positions, characterized 46% of all electronic data across animals (Table 2 and Fig. 2A). However, the proportion of data points in the first population ranged from 36% to 60% across animals (Table 2) suggesting that a

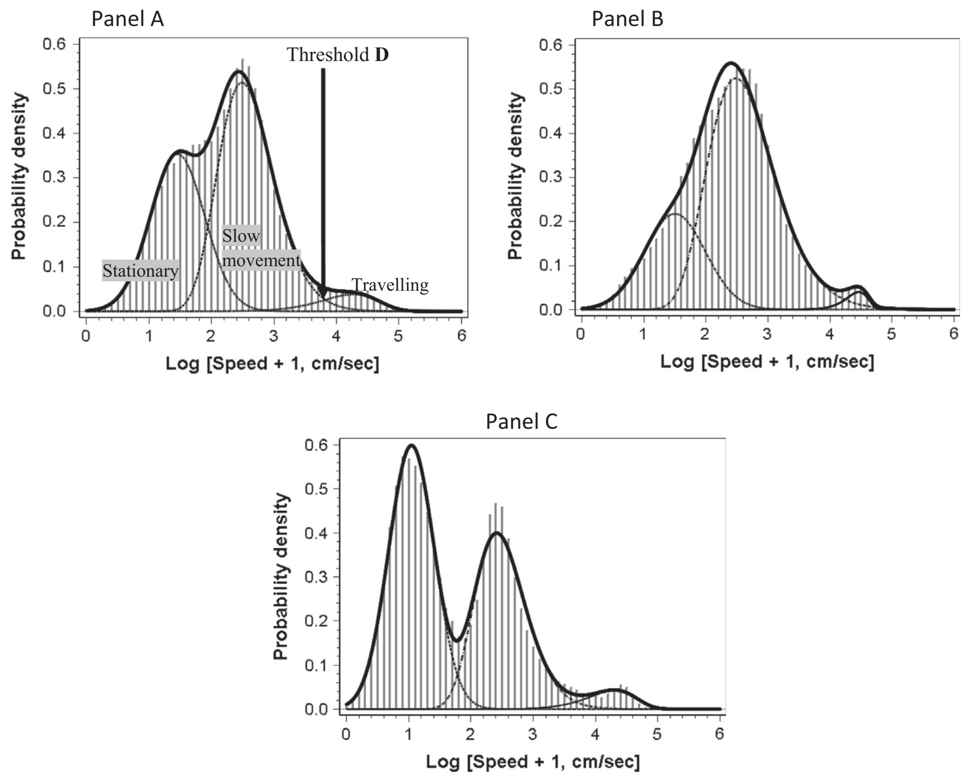

Fig. 3. Frequency distribution of observed data (gray bars) and predicted probability density functions (PDF, continuous lines) of travel speed pooled across 42 steers (panel A), and for 2 individual steers using GPS collars selected to show the most (panel B) and the least (panel C) overlap between the first two populations. The short dash line represents the first population of data points with low speed or stationary states (resting or ruminating), the long dash line represents the second population dominated by foraging, the thin black line is the third population dominated by traveling, and the thick solid line is the sum of all individual PDF. Speed is presented as Log (1 + cm/s). Threshold D is used to separate the population of intervals belonging to traveling behavior from the rest of the data.

large variation exists between individual steers and collars and that such variation should be accounted for by fitting mixture models to individual animals.

The threshold value A separating the head down (P1) from the head up (P2) population of data points was 9.21 log-units, i.e. the intersection between P1 $\times$ P2 which is used for the classification of data points into head down or up positions and then used for behavioral classification (Table 2 and Fig. 2A). Data from the SD of the accelerometer *X*-axis indicated that on average, animals had 16% of their data points (or spent 16% of their time) in inactive states, 36% with medium activity level of the neck (ruminating) and 48% of their time in active behaviors involving high activity level of the neck such as foraging, walking and other active behaviors (Fig. 2B). These 3 populations can be separated by the threshold B and C which are the values of the SD of the accelerometer *X*-axis where the 1st and 2nd PDF, and 2nd and 3rd PDF intersect, respectively (Table 2; Fig. 2C). Such threshold values are required in classification trees to assign a given data point to a specific activity according to whether it falls above or below the threshold.

Steers moved at medium speed for 56.4% of the day, were stationary for 39.4% of the day and spent 4.2% of the time traveling at high speed (Fig. 3A; Table 2). Travel speed showed the largest variability across individuals for the parameters of the PDF and the threshold values with CV ranging from 17% to 25% for the mixing proportions, 2.5 to 15% for the medians, and 5% of CV across animals/collars for threshold E which separates medium from high travel rate populations (Table 2). In addition, the shape and degree of overlap among populations for travel speed also showed large variation among animals (Fig. 3B and C). For instance, some individuals have clearly displayed all 3 populations of data points (Fig. 3C) while other animals showed high overlap between the first 2 populations (Fig. 3B) indicating that a large proportion of data points could belong to both the first and second population. Nevertheless, the third population representing 'directed walking or high-speed traveling' was evident in every animal (data not shown) although the threshold value D differed among steers or collars, or both. The threshold value D for the data pooled across animals was 3.83 log-units of travel speed (Table 3 and Fig. 3A) whereas individual animals could have higher or lower values for threshold D (4.24 and 3.71 log-units for the animals depicted in Fig. 3B and C, respectively). Application of the threshold estimate based on pooled data would assign almost double (90% more) the number of observations to high-speed traveling than the individual estimate for the animal in Fig. 3B. In contrast, application of the threshold estimate on pooled data for the animal in Fig. 3C would assign 24% lesser number of observations to high-speed traveling compared to the individual estimate.

3.3. Classification algorithm

Fig. 4 depicts the decision tree of the best algorithm obtained from the search of combinations of variables which resulted in the highest sensitivity and specificity for foraging and ruminating and Table 3 presents the classification results. The algorithm used 2 sensors (i.e. accelerometer and GPS) and 3 variables (mean and SD from the accelerometer *X*-axis, and mean travel speed). However, it is important to notice that similar results are obtained when using other variables highly correlated to those used in the decision tree of Fig. 4. For example, similar results were obtained using the SD of the accelerometer *Y*- or *Z*-axis instead of the *X*-axis of the accelerometer because of the high correlation between SD of

Table 2
Proportion of data points (P) and median (M) of each component population of finite mixture distributions of selected variables obtained from motion and GPS sensors in collars worn by grazing steers (development dataset). The probability density functions were fitted to the data for each individual animal and also the data pooled across all animals in the trial.[a]

Variable	Minimum	Mean	Maximum	CV	Pooled data[b]
Mean X-axis accelerometer					
P 1st population[c]	0.361	0.462	0.598	10.66	0.424
P 2nd population[c]	0.402	0.538	0.639	9.16	0.576
M 1st population[d]	8.551	8.830	9.134	1.19	8.793
M 2nd population[d]	9.370	9.458	9.584	0.60	9.447
Threshold A[e]	8.975	9.214	9.399	0.94	9.157
SD X-axis accelerometer					
P 1st population[c]	0.099	0.162	0.202	14.96	0.169
P 2nd population[c]	0.296	0.362	0.399	7.00	0.360
P 3rd population[c]	0.428	0.476	0.520	4.62	0.470
M 1st population[d]	3.640	3.820	4.055	2.45	3.834
M 2nd population[d]	4.948	5.199	5.474	2.63	5.212
M 3rd population[d]	6.945	7.144	7.396	1.45	7.142
Threshold B[e]	4.120	4.479	4.840	3.99	4.480
Threshold C[e]	5.890	6.188	6.530	2.60	6.210
Travel speed					
P 1st population[c]	0.228	0.394	0.537	24.39	0.383
P 2nd population[c]	0.420	0.564	0.740	17.55	0.572
P 3rd population[c]	0.019	0.042	0.062	25.03	0.044
M 1st population[d]	1.151	1.540	2.100	15.01	1.455
M 2nd population[d]	1.081	2.548	2.872	10.44	2.565
M 3rd population[d]	4.035	4.214	4.465	2.49	3.056
Threshold E[e]	3.320	3.842	4.245	4.97	3.830

[a] $n = 42$ animals. Total number of data points in finite mixture models was 3,056,728 each being 10-s long. All variables were log-transformed after the 10-s mean and SD were obtained. Speed is presented as Log (1 + cm/s) and accelerometer data as Log ($10{,}000 + g \times 10^4$).
[b] Data obtained from all experimental animals was pooled and fitted to finite mixture models.
[c] Proportion of data points belonging to the first, second and third population.
[d] Median of the first, second or third population of data points.
[e] Value at which the probability density functions of the first and second, or second and third populations of data points intersected. These are the threshold values that separate those populations with minimal miss-assignment of data points to the wrong population.

the accelerometer *X*, *Y* and *Z* axis, or the distance traveled between consecutive points instead of travel speed (data not shown). This might be important with devices that do not provide travel speed by the chip on board of the collar but where distance can be calculated from the GPS positions, or devices with different data format or sensor design (e.g. uni-axial accelerometers) or if data is missing for some variables.

In the development trial, the algorithm correctly classified 85.5% of all data points into the correct activity (sensitivity), with foraging and ruminating having the greatest sensitivity whereas resting and other active behaviors showed the least sensitivity (Table 3). The statistical concordance indicated that foraging was predicted with an underestimation of 1.3% (98.7% of observed; statistical concordance below 100% indicates underestimation of time spent in a particular behavior whereas above 100% indicates overestimation by the algorithm in relation to the time actually observed). However, ruminating was over-estimated (131% of observed) because of the lower specificity and higher misclassification rate of resting as ruminating (Table 3). The correlation between predicted and observed time spent performing the different activities from visual observations for the subset of data where activities were known resulted in an R^2 of 94.6%.

The dataset from each individual animal in the evaluation trial (unknown behaviors) were fitted with the PDF, then the threshold values obtained (data not shown) and finally the binary decision tree obtained with the development trial applied to the evaluation trial dataset. Then, a subset of data were created with those data points where behavioral activities were known from observations, for the purpose of evaluating the accuracy and precision of the classification method as done with the development trial dataset. Overall precision (90.5% vs. 85.6%), Kappa (0.86 vs. 0.77) and determination (94.5% vs. 99.6%) coefficients were higher in the evaluation trial compared to the development trial (Table 3). Foraging and resting showed greater sensitivity in the evaluation trial whereas sensitivity was lower for ruminating, traveling and other active behaviors. Statistical concordance was better for all activities but other active behaviors were severely overestimated and traveling was underestimated. In regards to daily time spent in each behavior, the algorithm predicted that animals spent 37.2% and 42.6% of the day foraging during the development and evaluation trial, respectively (Table 3). However, ruminating time was longer in the development compared to the evaluation trial.

4. Discussion

Behavioral classification of data from collars containing GPS and captive-bolt sensors have previously been reported for grazing cattle (Ungar et al., 2005; Augustine and Derner, 2013). Differences between those studies and the present study are the type of sensor used (MEMS sensor in the present study vs. captive bolt), greater frequency of data collection, the greater number of activities classified and the proposed analyses methodology. A high frequency rate of data collection seems important to accurately classify behavioral activities that resemble each other in terms of position and movement of the neck such as ruminating from resting, and foraging from traveling. Indeed, Anderson et al. (2013) argues in favor of high fix rate of GPS data for cattle behavioral studies to 'change the industry norm' and to improve classification of traveling and foraging activities. In addition, behavioral data obtained at a high frequency is important for several research and commercial applications. For example, GPS data obtained at a high frequency is critical for determining whether animals grazed on small patches of vegetation (e.g. weeds) or are just traveling through them without eating them (Swain et al., 2008). Differentiating between these 2 behaviors could give insight into factors driving diet selection and other relevant plant-animal interactions (Owen-Smith et al., 2013). High GPS fix rates may also become critical to accurately measure distance traveled per day, especially in small paddocks where animals do not cover long distances in straight lines but change direction frequently as they move in a forward direction and encounter fences. Another example of the potential benefits for recording high frequency behavioral data is in the study of transitions among key behaviors, such as the segue between foraging and travel, in order to understand underlying mechanisms producing these transitions that may be related to energy expenditure (travel) vs. energy harvest (foraging; Owen-Smith et al., 2013). Frequent data acquisition is especially important for behaviors lasting only a short duration such as traveling that may alternate with foraging according to varying vegetation conditions (e.g. sward height, density, availability, and quality). A third example of the benefit of measuring foraging activity with high frequency is to study underlying mechanisms in the regulation of feed intake as shown under intensive dairying (Tolkamp et al., 2000). Finally, high frequency data are also needed for some commercial applications such as for virtual fencing where animals have to be cued with high spatial and temporal accuracy when approaching the virtual fence (Ruiz-Mirazo et al., 2011).

Previous methodologies used for behavioral classification of data from cattle collars include clustering, discriminant, classification and regression trees, and support vector machine analysis (Ungar et al., 2005; Martiskainen et al., 2009; Augustine and Derner, 2013). However, these methodologies have one or more

Table 3
Number of data points classified into 1 of 5 activities (foraging, resting, ruminating, traveling and other active behaviors) by a decision tree (predicted activity) using data from cattle electronic monitoring collars containing accelerometers and GPS sensors compared to the actual activity based on observations for both algorithm-development and evaluation trials.[a]

Observed activity	Predicted activity					
	Foraging	Resting	Ruminating	Other active behaviors	Traveling	Total
Development trial						
Foraging	4185	66	88	65	64	4468
Resting	133	1121	559	23	12	1848
Ruminating	2	52	1968	8	0	2030
Other active behaviors[b]	84	44	35	24	2	189
Traveling	7	4	8	2	109	130
Total	4411	1287	2658	122	187	8665
Sensitivity or precision (%)	93.7	60.7	96.9	12.7	83.8	85.48[d]
Specificity (%)	94.6	97.6	89.6	98.8	99.1	0.77[e]
Statistical concordance (%)	98.7	69.6	130.9	64.6	143.8	94.55[f]
Predicted (% daily)[c]	37.16	21.66	36.36	2.51	2.31	–
Evaluation trial						
Foraging	3351	9	4	35	8	3407
Resting	13	2022	263	41	3	2342
Ruminating	8	352	3151	96	3	3610
Other active behaviors[b]	0	0	5	0	0	5
Traveling	14	11	14	22	20	81
Total	3386	2394	3437	194	34	9445
Sensitivity or precision (%)	98.4	86.3	87.3	0.0	24.7	90.46[d]
Specificity (%)	99.4	94.8	95.1	97.9	99.9	0.87[e]
Statistical concordance (%)	99.4	102.2	95.2	3880.0	42.0	99.61[f]
Predicted (% daily)[c]	38.99	19.21	35,64	3.30	2.86	–

[a] Each data point is equivalent 10-s intervals of time. The classification tree used the mean and SD of the accelerometer *X*-axis and the mean travel speed (Fig. 4). Threshold values for the decision tree were obtained after fitting probability density functions to finite mixture distributions for each individual animal.
[b] Other active behaviors included those with vigorous head movement while standing with no forward movement of the body and not engaged in foraging activities.
[c] Proportion of daily time that steers spent performing each behavioral category calculated from the entire experimental datasets.
[d] Overall precision of the classification algorithm (% of all data points correctly classified).
[e] Kappa coefficient used as overall coefficient of agreement ($P < 0.001$).
[f] Correlation coefficient between observed and predicted times spent in each activity.

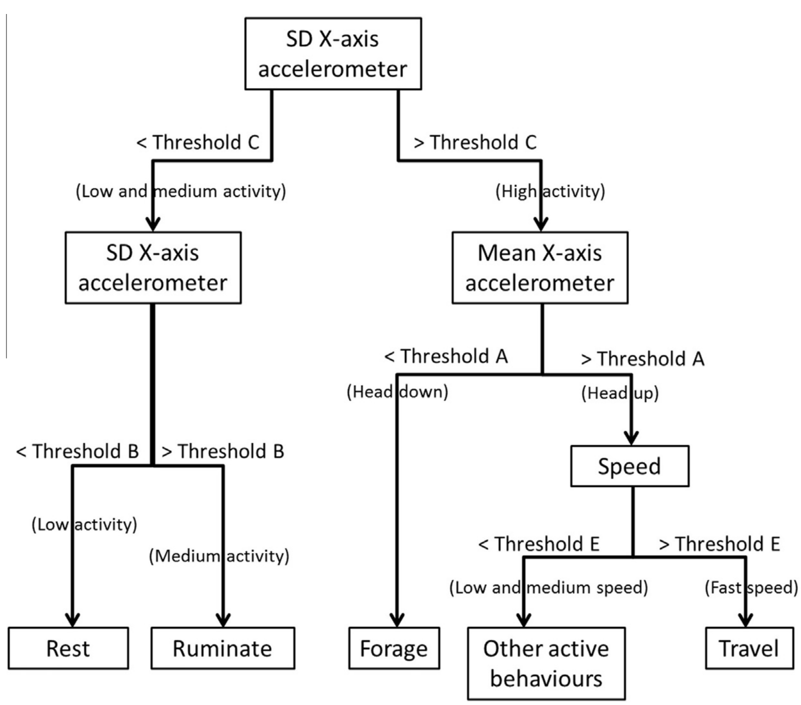

Fig. 4. Decision tree used to classify data obtained from cattle monitoring collars into 1 of 5 activities (foraging, ruminating, resting and other active behaviors) according to threshold values A, B, C and D obtained from fitting probability density functions using finite mixture models methodology. Other active behaviors included those with vigorous head movement while standing with no forward movement of the body and not engaged in foraging activities.

of the following disadvantages. First, previously used methodologies require high computing power which limits their use for large datasets. Second, observations are essential in every experiment where autonomous data are collected and evaluated to obtain both the classification tree and the threshold values. Augustine and Derner (2013) clearly specified that classification of cattle grazing behavior requires calibration specific to the environment and vegetation, and that sample size of the training data is also critical. In addition to being time consuming and subject to human error, this also limits the possibility of accounting for differences between individual animals and collars because observations are rarely obtained for large number of animals. Establishing the minimum number of experimental units is a challenge in behavioral studies with tracking devices (Anderson et al., 2013) however further research with technologies and analytical methodologies such as those presented herein could provide valuable information to determine sample size based on variation amongst individuals and temporal frequency of sampling (e.g. fix rate).

In contrast, finite mixture models used in the present study are computationally efficient and do not require behavioral observations for each experiment or environmental condition once the classification tree was obtained using a training dataset as it was demonstrated with the evaluation trial of the present study. Indeed, the threshold values can be obtained from finite mixture models from the entire datasets as long as the structure of data is similar (in terms of the number and shape of the populations) to that data from where the tree was obtained. The presence of mixture distributions in the entire dataset of individual animals under different environmental conditions of both trials of the present study demonstrated that unsupervised classification of sensor data may be possible under a wider range of conditions (e.g. pastures and animals). This would be an important advantage if monitoring systems are to be adopted across commercial production systems where farmers are unlikely to perform behavioral observations.

In addition, finite mixture models allow accounting for differences among individual devices or animals after fitting the PDFs, which also reduces the need of calibrating sensors and mounting all collars with the same tension as it would be the case when using a unique threshold value across animals and collars (Anderson et al., 2013). The proposed methodology also allows quality control of the data by permitting inspection of the frequency histograms to ensure that the populations of data points reflecting characteristic behavioral patterns of grazing cattle are present such as head down and head up, or high, medium and low activity levels. The lack of such populations in the present study has been shown to indicate that the collar has fallen off the animal, the sensors are not operational or the collar was fitted in the wrong position.

The large differences among animals and collars observed in the fitted parameters of the PDF, the threshold values and structure of the frequency distributions (e.g. overlap of populations) clearly demonstrate the need for accounting for such differences. However, more research is needed to determine the proportion of the observed variation between experimental units that are a result of differences in movement patterns between animals, differences in 'fitness' of collars (loose vs. tight fitting) and differences in the measurements between sensors. Information from fitting the PDFs to data from individual animals allowed obtaining the parameters that describe each population representing a specific head position and activity level, and travel speed. These parameters include the proportion of data points that belong to each population (i.e. head down and up), the mean (location or scale parameter) and the shape of the PDF. This information could be valuable to study animal behavior without further classification. For example, the proportion of data points within the population representing head down position is equivalent to the proportion of time spent in such a position, whereas the median of this population may quantify how low the head was. In addition, parameters from the PDF have also been used to respond to biological questions including mechanisms regulating satiety and hunger (Tolkamp et al., 2000). Furthermore, parameters from the PDF and visual inspection of frequency histograms allowed characterization of individual profiles. However, the high degree of overlap between the first and second population for travel speed in some individuals indicated that travel speed may not be reliable in some animals to differentiate foraging from resting and ruminating as it is the case for accelerometer data. Ungar et al. (2005) also reported that differences among animals were evident for distance traveled between 2 successive GPS points. Such differences among animals or collars highlight the need to obtain threshold values adjusted to each animal and collar combination for the classification algorithm. Anderson et al. (2012) characterized 3 populations of data points in the frequency histogram of travel speed from 4 cows with data collected at 1 s intervals and classified these data into stationary, foraging and walking behaviors. We agree with these authors that producing mean speeds for periods longer than 30 s will flatten the third peak corresponding to traveling.

In the present study, foraging was characterized by a head down position, high activity level of the neck and medium travel speed. Traveling was characterized by high travel speed, head up position and high activity level of the neck. Ruminating was characterized by head up position with medium activity level, and low travel speed, whereas resting was characterized by a head up position, with low activity level and low travel speed. Other active behaviors resulted in head positions intermediate between up and down, as well as high activity level and intermediate travel speed. Therefore, all behaviors differentiated from each other by at least 1 of the 3 variables used in the classification algorithm despite the fact that 2 behaviors might have similar electronic signature in one of the variables, which would make them not possible to differentiate from each other (e.g. high activity level with high SD in the accelerometer *X*-axis for both foraging and traveling but the later showing high speed of forward movement).

The 10-s mean values of the accelerometer *X*-axis was the most suitable to quantify the head down position while foraging which agrees with other authors who used the *Y* axis sensor count (Augustine and Derner, 2013) or the pitch angle (Nadimi and Søgaard, 2009) to determine head down. However, it is important to note that the position of the head while foraging might change depending on the height and type of the vegetation being grazed, e.g. prevalence of shrubs or trees. Furthermore, the position of the head and the threshold value separating head down from head up could change over time as a result of changes in vegetation height as the available forage is depleted over time. Browsing occupied a small proportion of the total observation time in the present study and therefore, it was not possible to differentiate browsing from grazing. However, browsing (9.22 ± 0.036 log-units) showed higher values in the accelerometer *X*-axis compared to grazing (8.83 ± 0.036 log units; $P < 0.01$). More research is needed to determine if browsing can be separated from grazing accurately as this could have an important application for research and grazing management. These activities may be captured in the other active behaviors category of the present algorithm which considers high activity of the neck performed with the head up and at medium to low travel speed (data not shown). It would be interesting to test the present methodology in pastures where tall, woody shrub vegetation prevails such as in leucaena plantations. However, other active behaviors might also be important under tropical conditions because animals may spend long periods per day trying to avoid flies, insects or ticks. Nevertheless, it is not possible to drawn firm conclusions about browsing and other active behaviors from the present study because these occurred at low frequency and parasites were not quantified. Traveling could also occupy a larger proportion of their time under certain conditions, such as in extensive rangelands where animals need to travel long distances searching for water or forage. Traveling and other active behaviors could be 'lumped' into foraging because of similar activity level if these are not classified separately with the aid of speed and head position data requiring both GPS and accelerometer sensors. It is therefore important to classify these behaviors separately from grazing although this requires the use of 3 variables in the decision tree.

The SD of the accelerometer *X*-axis was the most valuable variable to quantify the level of activity or mobility of the neck. Interestingly, results from the present study showed 3 clear populations or activity levels which were not evident in the sensor count data from Ungar et al. (2005) and Augustine and Derner (2013). This suggests that the distribution of the data may depend on the type and configuration of the sensor, and the aggregation window of the data. The captive bolt sensor in the collars used by Ungar et al. (2005) and Augustine and Derner (2013) are activated by tilt and require setting up a threshold above which counts of the number of 'hits' per unit of time. Therefore, those commercial collars do not provide a quantification of the acceleration but the number of times the captive bolt was above certain threshold which may not allow identification of the 3 populations of data points as in the present study. Furthermore, simulations with our data indicated that increasing the aggregation time window for calculation of the SD of the accelerometer *X* axis from 10 s to 20 min flattened the frequency distributions and abolished the ability to separate populations of data points from activities alike such as resting from ruminating (data not shown). Interestingly, Ungar et al. (2005) found that the sensor count was useful to differentiate foraging from other activities. However, Augustine and Derner (2013) reported that sensor count was of minimal value compared to 'head down' position which was predicted from the count of up-down movement of the neck along the vertical axis (*Y* axis). The later authors attributed such a difference with Ungar et al. (2005) to the fact that their animals were grazing short and homogeneous vegetation where head movement was not as relevant while grazing.

Long distances between 2 consecutive GPS measurements were associated with traveling by Ungar et al. (2005) and Augustine and Derner (2013). However, Ungar et al. (2005) concluded that traveling was difficult to differentiate from grazing if it did not cover more than 50 m because of the long time between GPS fixes (5 and 20 min) in that study. It is unknown however if differential correction of GPS logs could improve the classification accuracy of traveling with the commercial collars used by the aforementioned studies however this would make it difficult for real time classification of behavior on-board of the collar before wireless data transfer. In contrast, the present study used the mean of 40 measurements of speed of traveling calculated on-board by the GPS chip, which may increase the accuracy of the speed measurement and reduce the need for differential correction. Compared to the present study, Ungar et al. (2005) reported similar sensitivity and lower specificity for foraging (93.7% and 82.7%, respectively) and traveling (78.3% and 97.6%, respectively) using both discriminant and regression tree analysis in 2 studies with collars containing GPS and a captive bolt sensor in beef cattle under extensive conditions. Ungar et al. (2005) recommended caution when interpreting grazing hot-spots in spatial data with the methodology and collar technology they used because of the lower specificity of grazing. However, these studies used longer sample aggregation periods and classified only 3 activities (overall precision of 87.2%). Similar results were reported by Augustine and Derner (2013) with 83.6% overall precision using similar collars at 5 min intervals to detect 4 activities (grazing, resting, traveling and 'mixed' activities).

Nadimi and Søgaard (2009) obtained 87.2% overall classification rate of 2 activities (grazing and non-grazing) in cows using pitch angle data in K-means and multiple-model adaptive estimation. Therefore, the use of both head position and activity level, and the higher frequency of data collection in our study seem to improve the precision (by 2–13% units) and sensitivity (by 10–16% units) of classification of foraging behavior compared to previous studies which could make the methodology and electronic technology developed in the present study more appropriate to identify and interpret foraging hotspots. Ueda et al. (2011) achieved 94.5% for both precision and specificity for grazing vs. 'all other behaviors' using discriminant analysis of uni-axial accelerometer data in dairy cows grazing small paddocks with temperate pastures. Martiskainen et al. (2009) classified 6 activities using support vector machines and data from collars containing 3-axis accelerometers in dairy cows housed in a barn. Frequency of data collection was similar to the present study and they summarized 10-s time windows to yield 28 variables used for the classification methodology. They found lower overall algorithm precision (77%) and Kappa (0.68), lower precision for feeding (81%), greater for traveling (79%) and similar for resting (lying = 65%; standing = 83%) compared to the present study. In agreement with our findings, Martiskainen et al. (2009) also highlighted that misclassification occurred most often with similar behaviors such as ruminating, standing and lying, or between feeding, traveling and standing. However, the complexity and variety of behaviors is expected to be higher in extensive tropical environments of the present study because vegetation is highly heterogeneous in terms of height, density, and palatability. Finally, an important outcome of the present study is the high statistical concordance (close to 100%) for foraging, indicating no over- or under-estimation. In addition, predictions of daily durations of each behavior are similar to dairy cows grazing tropical vegetation as measured using vibracorders for 24-h (Stobbs, 1970).

In the present study, the lowest precision of classification was found in 'other active behaviors', which consisted of high activity with the head up while traveling at low or medium speed. Caution should be taken when interpreting this finding because of the very low number of observations for this behavior. In addition, other active behaviors comprised a range of activities with similar motion patterns such as head shaking, rubbing against objects, grooming and scratching as a result of an insect burden, and social interactions. Also, animals may perform vigorous head movements or 'shaking' for short durations and we may have failed to record these accurately through visual observations. Therefore, the high misclassification rate of other active behaviors may be a result of the observer's inability to accurately capture every head movement in a timely fashion. Further research should register active behaviors with greater detail to allow separating them more accurately. However, this may require video recording because behaviors of short duration and with rapid transition are difficult to register by direct observation in the paddock.

5. Conclusion

The present study showed that information collected at high frequency (10 and 4 Hz) by accelerometers and GPS sensors embedded in cattle collars have the ability to capture fine scale spatio-temporal differences in the position and activity level of the neck while cattle are engaged in various behaviors in the paddock. This ability to capture behavior at fine spatio-temporal scales may help improving the accuracy of behavioral classification methods and to develop a consistently reliable means to remotely access data from the collars in real-time for virtual fencing and management applications. However, it is important to notice that the increased precision achieved with 10-s data comes at the cost of substantial increases in the size of databases collected, battery needs (especially for the GPS chip), and computing power required for managing such large databases. Finite mixture models developed for the classification of behavior in the present study may provide advantages over previously published ones (e.g. classification and regression trees, and support vector machine) because it reduces the need of training datasets for every experiment, environmental conditions and animals, and allow accounting for differences among individual animals and devices. This would allow greatly reducing the time and cost involved in field personnel to collect observational data. Furthermore, the approach could extend behavior classification beyond the realm of research into practical applications for livestock producers for commercial gain and ecological management purposes. However, further research is needed to confirm that the methodology has the ability to measure grazing cattle behavior under varying conditions ranging from sparse desert shrub land vegetation to rangelands and wet tropical vegetation. Wireless transfer in real time of collar data is now also possible. Such methodologies that remotely and automatically monitor cattle behavior will improve our understanding of ecological drivers of behavior and provide tools to improve their management in extensive grazing conditions.

Acknowledgements

The authors want to thank Sharon McGavin and Ombeline Chapolard for helping with field work, electronic collars and behavioral observations, Wayne Flintham and Kevin Barrett for their help with cattle work, and Ed Charmley, Dave Henry, Dave Swain and Bert Tolkamp for their useful comments of the manuscript.

References

Altmann, J., 1974. Observational study of behavior: sampling methods. Behaviour 49, 227–267.
Anderson, D.M., Winters, C., Estell, R.E., Fredrickson, E.L., Doniec, M., Detweiler, C., Ru, D., James, D., Nolen, B., 2012. Characterising the spatial and temporal activities of free-ranging cows from GPS data. Rangeland J. 34, 149–161.
Anderson, D.M., Estell, R.E., Cibils, A.F., 2013. Spatiotemporal cattle data—a plea for protocol standardization. Positioning 4, 115–136.
Augustine, D.J., Derner, J.D., 2013. Assessing herbivore foraging behavior with GPS collars in a semiarid grassland. Sensors 13, 3711–3723.
Cai, G., Chen, B.M., Lee, T.H., 2011. Unmanned rotorcraft systems. In: Advances in Industrial Control. Springer-Verlag London Limited.
Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20, 37–46.
Gaillard, J.M., Hebblewhite, M., Loison, A., Fuller, M., Powell, R., Basille, M., Van Moorter, B., 2010. Habitat–performance relationships: finding the right metric at a given spatial scale. Philos. Trans. R. Soc. B 365, 2255–2265.
Ganskopp, D.C., Johnson, D.D., 2007. GPS error in studies addressing animal movements and activities. Rangeland Ecol. Manage. 60, 350–358.
González, L.A., Tolkamp, B.J., Coffey, M.P., Ferret, A., Kyriazakis, I., 2008. Changes in feeding behavior as possible indicators for the automatic monitoring of health disorders in dairy cows. J. Dairy Sci. 91, 1017–1028.
Handcock, R.N., Swain, D.L., Bishop-Hurley, G.J., Patison, K.P., Wark, T., Valencia, P., Corke, P., O'Neill, C.J., 2009. Monitoring animal behaviour and environmental interactions using wireless sensor networks, GPS collars and satellite remote sensing. Sensors 9, 3586–3603.
Heglund, N.C., Taylor, C.R., 1988. Speed, stride frequency and energy cost per stride: how do they change with body size and gait? J. Exp. Biol. 138, 301–318.
McLachlan, G.J., Peel, D., 2000. Finite Mixture Models. John Wiley & Sons Publishers.
Martiskainen, P., Jarvinen, M., Skon, J.P., Tiirikainen, J., Kolehmainen, M., Mononen, J., 2009. Cow behavior pattern recognition using a three-dimensional accelerometer and support vector machines. Appl. Anim. Behav. Sci. 119, 32–38.
Nadimi, V., Søgaard, H.T., 2009. Observer Kalman filter identification and multiple-model adaptive estimation technique for classifying animal behaviour using wireless sensor networks. Comput. Electr. Agric. 68, 9–17.
Nadimi, E.S., Jørgensen, R.N., Blanes-Vidal, V., Christensen, S., 2012. Monitoring and classifying animal behavior using ZigBee-based mobile ad hoc wireless sensor networks and artificial neural networks. Comput. Electr. Agric. 82, 44–54.
Owen-Smith, N., Fryxell, J.M., Merrill, E.H., 2013. Foraging theory upscaled: the behavioural ecology of herbivore movement. Philos. Trans. R. Soc. B 365, 2267–2278.
Ruiz-Mirazo, J., Bishop-Hurley, G.J., Swain, D.L., 2011. Automated animal control: can discontinuous monitoring and aversive stimulation modify cattle grazing behaviour? Rangeland Ecol. Manage. 64, 240–248.
Stobbs, T.H., 1970. Automatic measurement of grazing time by dairy cows on tropical grass and legume pastures. Trop. Grasslands 4, 237–244.
Swain, D.L., Wark, T., Bishop-Hurley, G.J., 2008. Using high fix rate GPS data to determine the relationships between fix rate, prediction errors and patch selection. Ecol. Model. 212, 273–279.
Tolkamp, B.J., Schweitzer, D.P., Kyriazakis, I., 2000. The biologically relevant unit for the analysis of short-term feeding behavior of dairy cows. J. Dairy Sci. 83, 2057–2068.
Ueda, Y., Akiyama, F., Asakuma, S., Watanabe, N., 2011. Technical note: the use of a physical activity monitor to estimate the eating time of cows in pasture. J. Dairy Sci. 94, 3498–3503.
Ungar, E.D., Henkin, Z., Gutman, M., Dolev, A., Genizi, A., Ganskopp, D., 2005. Inference of animal activity from GPS collar data on free-ranging cattle. Rangeland Ecol. Manage. 58, 256–266.
Wark, T., Corke, P., Sikka, P., Klingbeil, L., Guo, Y., Crossman, C., Valencia, P., Swain, D., Bishop-Hurley, G., 2007. Transforming agriculture through pervasive wireless sensor networks. IEEE Pervasive Comput. 6, 50–57.