Computers and Electronics in Agriculture 145 (2018) 298–301

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Monitoring drinking behavior in bucket-fed dairy calves using an ear-attached tri-axial accelerometer: A pilot study

Leonie Roland[a], Laura Lidauer[b], Georg Sattlecker[b], Florian Kickinger[b], Wolfgang Auer[b], Valentin Sturm[c], Dmitry Efrosinin[c], Marc Drillich[a], Michael Iwersen[a,*]

[a] *Clinical Unit for Herd Health Management in Ruminants, University Clinic for Ruminants, Department for Farm Animals and Veterinary Public Health, University of Veterinary Medicine, Vienna, Veterinärplatz 1, 1210 Vienna, Austria*
[b] *Smartbow GmbH, Jutogasse 3, 4675 Weibern, Austria*
[c] *Institute for Stochastics, Johannes Kepler University, Linz, Altenbergerstraße 69, 4040 Linz, Austria*

ARTICLE INFO

Keywords:
Calf
Drinking behavior
Accelerometer
Validation

ABSTRACT

Acceleration sensors allow a reduction of time-consuming visual observation. The drinking behavior of bucket-fed calves has not been monitored automatically yet, although sufficient milk intake is essential for calves' health and growth. The objectives of this pilot study were (1) to evaluate the technical and mathematical feasibility of using an acceleration sensor to detect drinking events in bucket-fed dairy calves, (2) to develop an algorithm for an acceleration sensor (**SMARTBOW** ear tag, Smartbow GmbH, Weibern, Austria) for monitoring drinking behavior in bucket-fed dairy calves, and (3) to validate the SMARTBOW sensor for monitoring drinking events in bucket-fed dairy calves to observations from video recordings. Three preweaned dairy calves were equipped with ear-tag accelerometers. Calves were housed in individual pens and fed milk from a teat-bucket twice a day. Acceleration data were collected and calf behavior was video-recorded for 5 d for 24 h d^{-1}. Based on a training data set, an algorithm was developed to predict drinking events. Further 15 d of data were generated by simulation. Video recordings were used to analyze whether drinking events (n = 174) were predicted correctly for the complete data set. Sensitivity (82.9%), specificity (96.9%), and accuracy (96.2%) were good, but precision (60.4%) was not yet optimal. Cohen's Kappa (0.68) indicated substantial agreement between sensor and video analysis. More research based on a larger number of animals with the aim to optimize the underlying algorithm and to further increase sensitivity and precision is planned.

1. Introduction

Wireless accelerometers are precision dairy farming tools that allow an automated real-time monitoring of animal health and welfare. Since the 1980s, a variety of acceleration sensors have been developed (Rutten et al., 2013) to identify and record activity (Darr and Epperson, 2009; Ledgerwood et al., 2010; Müller and Schrader, 2003; Robert et al., 2009), locomotion (De Passillé et al., 2010), estrus cycle (Brehme et al., 2008; Dolecheck et al., 2015), parturition (Krieger et al., 2016), and feeding behavior (Bikker et al., 2014; Burfeind et al., 2011; Scheibe and Gromann, 2006). In calves, previous studies have evaluated the use of data loggers for lying behavior (Bonk et al., 2013; Swartz et al., 2016), step activity (De Passillé et al., 2010; Swartz et al., 2016), gait patterns (De Passillé et al., 2010), rumination (Burfeind et al., 2011), and milk intake from a calf feeder (Breitenberger et al., 2015). The drinking behavior of bucket-fed calves has not been monitored automatically yet, although milk intake in calves is essential for their health and growth (Appleby et al., 2001; Miller-Cushon and DeVries, 2015). Early detection of a change in drinking behavior by an automated system allows for a prompt intervention, possibly reducing negative effects on the calves' health and weight gain.

The objectives of this pilot study were (1) to evaluate if the detection of drinking events (**DE**) in bucket-fed dairy calves using an accelerometer is feasible from a technical and mathematical point of view, (2) to develop an algorithm for an acceleration sensor (**SMARTBOW** ear tag, Smartbow GmbH, Weibern, Austria) for monitoring DE in bucket-fed dairy calves, and (3) to validate the SMARTBOW sensor based on the algorithm developed under (2) for monitoring DE in bucket-fed dairy calves with observations from video.

Abbreviations: DE, drinking event; DE_a, actual drinking event (observed via video recordings); DE_p, predicted drinking event (calculated using the algorithm based on sensor data); FN, false negative; FP, false positive; TN, true negative; TP, true positive

* Corresponding author.
E-mail address: Michael.Iwersen@vetmeduni.ac.at (M. Iwersen).

https://doi.org/10.1016/j.compag.2018.01.008
Received 12 November 2017; Accepted 10 January 2018
0168-1699/

2. Material and methods

The study was approved by the institutional ethics committee of the University of Veterinary Medicine (ETK-03/09/2015), Vienna, Austria, as well as by the State Office of Agriculture, Food Safety and Fisheries Mecklenburg-Vorpommern, Germany (7221.3-2-028/15).

Data collection was performed on a commercial dairy farm in Mecklenburg-Vorpommern, Germany, in fall 2015. Three preweaned female Holstein Friesian calves (median age 15 d) were equipped with SMARTBOW sensors. The sensors were attached to the left ear 2 weeks prior to study begin. Calves were housed in a naturally ventilated stable in individual pens (1.37 × 2.00 m) with straw bedding. Twice a day (9–10 a.m. and 4–5p.m.) 7 L of whole milk were fed from a bucket (8 L hygienic bucket, Kerbl, Buchbach, Germany) with a rubber teat (product no. 1454, Kerbl, Buchbach, Germany). The SMARTBOW sensors recorded 10 three-dimensional acceleration values per second, which were transferred to an on-farm server via a wireless network. Acceleration data were collected continuously for 5 d for 24 h d^{-1}, resulting in 15 d of observed data in total. During the same time, calf behavior was recorded with infrared video cameras. One camera (IR Bullet Network Camera DS-2CD2632F-I(S), Hikvision, Hangzhou, China) was installed over each teat bucket and a fourth camera (Fisheye Network Camera DS-2CD6332FWD-I(V)(S), Hikvision) provided an overview over all 3 pens. Video material recorded by the 3 cameras installed over the buckets was analyzed by one observer with the video analysis software Mangold Interact (**Interact**, Mangold International, Arnstorf, Germany). Four behaviors were differentiated: milk intake, playing or sucking at the rubber teat without milk intake, no activity on the teat, and non-identifiable behavior. Viewing all videos recordings continuously in Interact, start and end time (in hh:mm:ss) of each behavior were observed visually and recorded with short-keys in a table in the Interact program. The tables were then converted into Microsoft Excel (**MS Excel**, Microsoft Excel for Mac, version 14.5.2, Microsoft Corporation, Redmont, WA) files. Intra-observer reliability was determined by having one observer watch the same video sequence twice (n = 1 calf, 24 h). It was calculated with MS Excel using Pearson's correlation and was found to be 0.99. Using the Interact software, intrarater reliability based on Cohen's Kappa (Cohen, 1960) was 0.96.

2.1. Algorithm development

Based on acceleration data, an algorithm was developed to predict the timing of DE. For this purpose, DE were defined as milk intake from the bucket including short gaps. They were regarded as separate actions if more than 300 s passed between subsequent events. Due to the fixation of the sensors it was not feasible to use the individual dimensions of the acceleration data, therefore the 3-dimensional acceleration data were transformed to form a single signal, the magnitude of the acceleration (absolute acceleration). The algorithm is based on a machine learning algorithm built from different features and consists of several steps.

First, the acceleration data set of 15 d (3 calves, 5 d of observation per calf) was split up in two parts: a training set, consisting of 9 d, and a test set, consisting of 6 d. Second, the members of the training set were divided into non-overlapping subintervals of equal length (60 s). In these intervals, three different measures (variance, skewness, and kurtosis) of the absolute acceleration were calculated. Based on the training data, a lower and upper bound for each measure were defined and all intervals for which at least one of these statistics laid outside the boundaries were excluded from further consideration as a possible DE. This led to elimination of more than 60% of all intervals in the training set, while keeping about 90% of the intervals overlapping with a drinking phase. Third, a total of 40 features based on the absolute acceleration data in each interval were calculated. The features included various statistical quantities, such as the sample central moments up to order 4, the autocorrelation up to lag 5, features based on the estimated power spectral density, and parameters of modeling the acceleration with moving-average and autoregressive processes (Brockwell and Davis, 1991).

Prior to further use of these features in the machine learning algorithm, a preliminary feature selection based on the correlation between the different features was implemented (Hall, 1999; Jain et al., 1996). This procedure was employed to define a measure for importance and to eliminate unimportant features for further consideration. With the remaining features, the Mathematica software (version 11.0, Wolfram Research, Inc., Champaign, Illinois) was used to build up different ensembles of multilayer perceptron learning algorithms with two hidden layers (Jain et al., 1996) as follows: The training data were split up into every combination of 7 training days and 2 validation days (36 possibilities). For each combination, a classifier function out of the 7 training days was generated, the model was applied to the remaining 2 validation days, and the overall average accuracy and sensitivity were calculated. By changing the parameters of the underlying learning algorithm it was tried to maximize the average accuracy on the validation sets while maintaining an average sensitivity of at least 80%. Due to the severe imbalance in the data (only approximately 5% of the observation time accounted for DE), the feature vectors belonging to the drinking events in each calculation of a new model were randomly oversampled to approximately reach a balanced data base. Once appropriate parameters were found, all 36 models voted for the intervals in the training set and it was specified that an interval is chosen as a drinking phase if a least amount of models individually voted for it to be one. This minimum amount was chosen to maximize the sum of accuracy and sensitivity on the training set while keeping a sensitivity of at least 90%. Next, the intervals which were shorter than a predefined length were eliminated, and the remaining ones were merged to bigger drinking phases if they did not lie more than 300 s apart. Finally, the same procedure was used on the test set.

2.2. Simulation study

A simulation study was conducted to expand the experimental data set. The training data was used to produce 15 additional data sets (5 days of simulation for each calf). The length of DE and the time between two such events was estimated. Moreover it was tried to find stochastic processes that rebuild the properties of the given data. The new simulated datasets were incorporated into building every one of the 36 classifiers as mentioned above.

2.3. Validation

The complete video data set (i.e. 120 h per calf) was used to validate whether the algorithms predicted DE correctly. Number and duration of actual (**DE_a**, observed via video recordings) and predicted (**DE_p**, calculated using the algorithms based on sensor data) drinking events were calculated using MS Excel. Based on the timing of DE, the duration and percentage of true positive (**TP**, correctly predicted time period of milk intake), true negative (**TN**, correctly predicted time period of no milk intake), false positive (**FP**, falsely predicted time period of milk intake), and false negative (**FN**, falsely predicted time period of no milk intake) DE was calculated. On this basis, the following test characteristics were calculated: sensitivity and specificity (Lundorff Jensen and Kjelgaard-Hansen, 2011), accuracy and precision (Powers, 2011), negative predictive value (Lundorff Jensen and Kjelgaard-Hansen, 2011), F1 score (Powers, 2011), Cohen's Kappa (Cohen, 1960), and the Youden index (Youden, 1950). The simulated data was processed analogously.

In order to understand what behavior the calves expressed during FP events, video footage recorded by the camera covering the whole pen area was analyzed with the Interact software for FP episodes. Behavior was categorized into licking (self grooming, licking at objects, playing with or sucking at the rubber teat without milk intake), roughage intake, neutral (no active muzzle movement), and non-