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

L. Roland[1], G. Sattlecker[2], F. Kickinger[2], L. Lidauer[2], A. Berger[2], W. Auer[2], V. Stum[3], D. Efrosinin[3], M. Drillich[1], M. Iwersen[1]

[1]Department for Farm Animals and Veterinary Public Health, University Clinic for Ruminants, Clinical Unit for Herd Health Management in Ruminants, University of Veterinary Medicine, Vienna, Austria

[2]Smartbow GmbH, Research and Development, Weibern, Upper Austria, Austria

[3]Institute for Stochastics, Johannes Kepler University, Linz, Altenbergerstraße 69, 4040 Linz, Austria

A paper from the Proceedings of the
14th International Conference on Precision Agriculture
June 24 – June 27, 2018
Montreal, Quebec, Canada

Abstract. *Accelerometers support the farmer with collecting information about animal behavior and thus allow a reduction in visual observation time. The milk intake of calves fed by teat-buckets has not been monitored automatically on commercial farms so far, although it is crucial for the calves' development. This pilot study was based on bucket-fed dairy calves and intended (1) to evaluate the technical feasibility of using an ear-attached accelerometer (SMARTBOW, Smartbow GmbH, Weibern, Austria) to identify drinking events, (2) to develop an algorithm to detect milk intake, and (3) to validate the SMARTBOW sensor incorporating the algorithm developed under (2) for identifying drinking events against observations from video. The acceleration data used in this study were generated from three sensors attached to the ears of three preweaned calves. Sensor data were recorded for 5 d for 24 h/d and calf behavior was video camera-recorded during the same time period . Based on a training data set, an algorithm was developed to identify drinking events. In addition, a mathematical data simulation was performed which generated further 15 d of data. The complete data set was compared with video recordings to analyze whether drinking events (n = 174) were detected correctly. Sensitivity (82.9 %), specificity (96.9 %), and accuracy (96.2 %) were good, but precision (60.4 %) was not yet satisfactory. Cohen's kappa (0.68) indicated a substantial agreement between sensor and video analysis. Additional work with a larger number of animals is planned to further improve the algorithm.*

Keywords. *Calf, drinking behavior, accelerometer, validation.*

Introduction

Wireless accelerometers are precision dairy farming devices that allow an automated real-time monitoring of animal behavior. Starting in the 1980s, different accelerometers have been developed (Rutten et al. 2013) to 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. 2017), rumination (Reiter et al. 2018), and feeding behavior (Bikker et al. 2014; Burfeind et al. 2011; Scheibe and Gromann 2006). Previous studies in calves have evaluated the use of automated devices for activity (Hill et al. 2017), lying behavior (Bonk et al. 2013; Swartz et al. 2016), step activity (Swartz et al. 2016; De Passillé et al. 2010), gait patterns (De Passillé et al. 2010), eating (Hill et al. 2017), rumination (Burfeind et al. 2011; Hill et al. 2017), and milk intake from a calf feeder (Breitenberger et al. 2015). Milk intake by bucket-fed calves has not been monitored automatically yet, although it is crucial for a healthy development (Appleby et al. 2001; Miller-Cushon and DeVries 2015). The early recognition of a difference in a calf's drinking pattern by an automated system allows an early intervention, which is likely to decrease adverse effects on a calf's health and weight gain.

The objectives of this pilot study were (1) to evaluate if the identification of drinking events (DE) in bucket-fed calves using an acceleration sensor is feasible under technical and mathematical points of view, (2) to develop an algorithm for an accelerometer (SMARTBOW Eartag, Smartbow GmbH, Weibern, Austria) for detecting DE in bucket-fed dairy calves, and (3) to evaluate the SMARTBOW sensor incorporating the algorithm for monitoring DE in bucket-fed dairy calves against observations from video.

Material and methods

Data was collected on a commercial dairy farm in Mecklenburg-Vorpommern, Germany. SMARTBOW Eartags were applied to three preweaned female Holstein-Friesian calves with a median age of 15 d. The sensors were attached to the left ear two weeks prior to the beginning of the study. Calves were housed in individual pens (1.37 x 2.00 m) with straw bedding. Seven liters 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) twice a day. The SMARTBOW sensors recorded 10 Hz of three-dimensional acceleration values. Acceleration data were collected continuously for 5 d for 24 h/d.

Calf behavior was recorded with infrared video cameras during the same time. One camera (IR Bullet Network Camera DS-2CD2632F-I(S), Hikvision, Hangzhou, China) was installed over each bucket and one camera (Fish-eye Network Camera DS-2CD6332FWD-I(V)(S), Hikvision) recorded an overview over all three pens. Video material was analyzed by one observer with Mangold Interact (Interact, Mangold International, Arnstorf, Germany). Four behaviors were observed: milk intake, playing or sucking at the rubber teat without milk intake, no activity on the teat, and non-identifiable behavior. With Interact, start and end time of each behavior were determined and converted into Microsoft Excel (MS Excel, Microsoft Excel for Mac, version 14.5.2, Microsoft Corporation, Redmont, WA) files. Intra-observer reliability based on Cohen's kappa (κ) (Cohen 1960) was calculated with Interact as 0.96.

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).

An algorithm was developed to detect the timing of DE based on acceleration data generated by the SMARTBOW sensor. DE were defined as milk intake from the bucket including short gaps. If more than 300 s passed between subsequent DE they were regarded as separate events. The algorithm was based on a machine learning algorithm built from different features and was developed in several steps: First, the complete acceleration data set (3 calves, 5 d of observation per calf, resulting in 15 d in total) was split up in a training set (9 d) and a test set (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. A more detailed review of the mathematical process is covered by Sturm et al. (2017).

In a simulation study, 15 additional data sets (five days of simulation per calf) were generated to expand the experimental data. They were incorporated into building each of the classifiers outlined previously.

All of the video recordings (120 h/calf) were used to analyze whether the algorithm identified DE correctly. Number and duration of DE observed via video recordings (DE_v) and DE calculated using the algorithms based on sensor acceleration data (DE_a) were calculated using MS Excel. 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) events were determined. Sensitivity, specificity (Lundorff Jensen and Kjelgaard-Hansen 2011), accuracy, precision (Powers 2011), negative predictive value (Lundorff Jensen and Kjelgaard-Hansen 2011), F1 score (Powers 2011), κ (Cohen 1960), and the Youden index (Youden 1950) were calculated for both experimental and simulated data. Interpretation of κ was made according to McHugh (2012).

In a second step, video material was analyzed again to determine what behavior calves were involved in during 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-identifiable.

Results and Discussion

During the observation period of 120 h per calf 174 DE_v and 170 DE_a were identified, 133 of which overlap in time. The average daily number of DE_v per calf was calculated as 11.6 ± 4.1 (mean ± standard deviation) and the daily number of DE_a per calf was 11.3 ± 2.8. The mean duration of DE_v per day amounted to 01:15:47 h, whereas the mean duration of DE_a was 01:46:00 h. An average day (24 h) comprised 01:03:59 h (4.4 %) of TP, 22:00:50 h (91.7 %) of TN, 42:01 min (2.9 %) of FP, and 13:10 min (0.9 %) of FN events. A sample segment of acceleration data showing DE_v versus DE_a is presented in Figure 1.

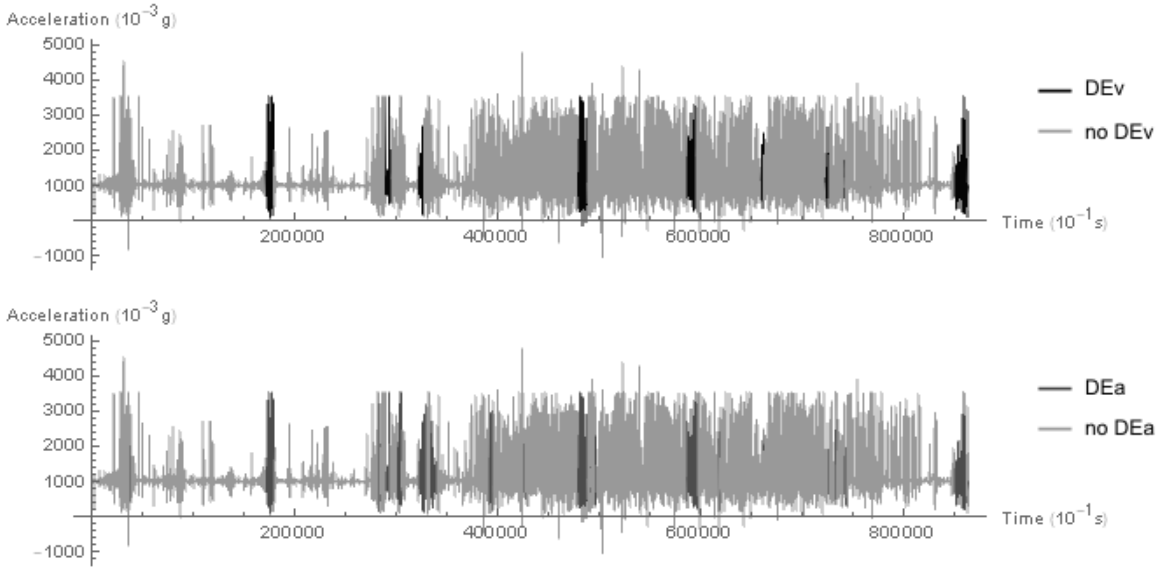

Fig. 1. Comparison of drinking events identified by video analysis (top) vs. algorithm based on acceleration data (bottom) drinking events.

The test characteristics for the calculated DE_a and the simulated results (DE_s) are presented in Table 1.

Table 1. Test characteristics of actual (video) vs. calculated (based on sensor data, DE_a) and vs. simulated drinking events (DE_s)

Parameter	DE_a	DE_s
True positive (%)	4.4	4.3
True negative (%)	91.7	92.4
False positive (%)	2.9	2.2
False negative (%)	0.9	1.1
Sensitivity (%)	82.9	79.9
Specificity (%)	96.9	97.7
Accuracy (%)	96.2	96.7
Precision (%)	60.4	65.9
Negative predictive value (%)	99.0	98.8
F1 Score (%)	69.9	72.1
Cohen's Kappa (κ)	0.68	0.70
Youden Index	0.80	0.77

To our knowledge, this is the first study evaluating the use of an accelerometer to identify milk intake in bucket-fed calves. The SMARTBOW accelerometer was overall successful in detecting DE. Sensitivity (82.9 %), specificity (96.9 %), and accuracy (96.2 %) were good. Cohen's κ was calculated as 0.68, which indicates a substantial agreement between sensor data and video analysis. However, because calves were only drinking for short time periods during the day, the validity of the specificity and accuracy is limited. When developing the algorithm, it was particularly challenging to obtain a high sensitivity without causing a concurrent decrease in precision. The accurate identification of the duration of DE, the percentage of FP results, and the precision should be further improved. The moderate precision was possibly caused by the small number of animals enrolled in this pilot study and differences in drinking behavior between individuals. This complicated the transfer of movement pattern calculations to different animals. The results from the simulation study were comparable to the results obtained from the observed data. Selected video sequences were analyzed a second time to identify what activities the animals were engaged in during FP events. The analysis indicated that during approximately a third of the time calves were involved in active behaviors (e.g. licking, feed intake) that the algorithm might have categorized as DE incorrectly.

Prior studies evaluating the use of data loggers for calf feeding or drinking behavior were either based on different technology (e.g. Burfeind et al. (2011) studied a logger based on audio recordings to evaluate rumination in calves) or presented different test characteristics. Hill et al. (2017) evaluated a different ear-attached accelerometer (SensOor, CowManager BV, Harmelen, The Netherlands) to describe activities in calves. They concluded that the SensOor was a suitable tool to detect eating, but described difficulties in identifying drinking behavior. Bikker et al. (2014) validated the SensOor for cows and reported a κ for 'eating' of 0.77. The authors suggest that behavior involving simple and repetitive movements might be detected more easily. Adult cows are likely to be less active than calves, have larger extension of jaw movements, and a higher degree of standardization and repetition in their movement patterns. This could explain the slightly lower κ calculated for milk intake in this study. Another reason could be that feeding might be easier to be detected by an ear-attached accelerometer than drinking behavior.

A follow-up study with a larger number of calves is planned to further improve the algorithm. Another objective of that study will be to assess if the SMARTBOW sensor can be used to quantify milk intake. Once a suitable algorithm has been developed, it should be evaluated if changes in milk intake detected by acceleration sensors can be used to identify discomfort or pathological states in calves.

Conclusion

Overall, the SMARTBOW sensor was successful in predicting drinking events. The technical feasibility of using an accelerometer to detect milk intake in calves was proven. Sensitivity, specificity, and accuracy were good, but precision is subject to further improvement. Kappa indicated a substantial agreement between sensor data and video material. More research based on a larger number of animals is planned with the objective to increase the precision of the algorithm, to quantify milk intake, and to identify diseased animals.

Acknowledgements

We thank Vanessa Schweinzer and Peter Kanz from the University of Veterinary Medicine, Vienna, Austria, and the staff from Rinderzucht Peters KG, Körchow, Germany for their technical assistance.

References

Appleby, M. C., Weary, D. M., & Chua, B. (2001). Performance and feeding behaviour of calves on ad libitum milk from artificial teats. *Applied Animal Behaviour Science,* 74(3), 191-201.

Bikker, J. P., van Laar, H., Rump, P., Doorenbos, J., van Meurs, K., Griffioen, G. M., et al. (2014). Technical note: Evaluation of an ear-attached movement sensor to record cow feeding behavior and activity. *Journal of Dairy Science,* 97(5), 2974-2979, doi: http://dx.doi.org/10.3168/jds.2013-7560.

Bonk, S., Burfeind, O., Suthar, V., & Heuwieser, W. (2013). Technical note: Evaluation of data loggers for measuring lying behavior in dairy calves. *Journal of Dairy Science,* 96(5), 3265-3271, doi: http://dx.doi.org/10.3168/jds.2012-6003.

Brehme, U., Stollberg, U., Holz, R., & Schleusener, T. (2008). ALT pedometer - New sensor-aided measurement system for improvement in oestrus detection. *Computers and Electronics in Agriculture,* 62(1), 73-80, doi:10.1016/j.compag.2007.08.014.

Breitenberger, S., Efrosinin, D., Auer, W., Deininger, A., & Waßmuth, R. (2015). Change point detection in piecewise stationary time series for farm animal behavior analysis. In K. F. Doerner, I. Ljubk, G. Pflug, & G. Tragler (Eds.), *Operations Research Proceedings, Vienna, 2015* (pp. 369-375). Vienna: Springer

Brockwell, P. J., & Davis, R. A. (1991). In *Time series: theory and methods* (2 ed.). New York: Springer.

Burfeind, O., Schirmann, K., von Keyserlingk, M. A. G., Veira, D. M., Weary, D. M., & Heuwieser, W. (2011). Technical note: Evaluation of a system for monitoring rumination in heifers and calves. *Journal of Dairy Science,* 94(1), 426-430, doi: http://dx.doi.org/10.3168/jds.2010-3239.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement,* 20(1), 37-46, doi: http://dx.doi.org/10.1177/001316446002000104.

Darr, M., & Epperson, W. (2009). Embedded sensor technology for real time determination of animal lying time. *Computers and Electronics in Agriculture,* 66(1), 106-111, doi:10.1016/j.compag.2009.01.004.

De Passillé, A. M., Jensen, M. B., Chapinal, N., & Rushen, J. (2010). Technical note: Use of accelerometers to describe gait patterns in dairy calves. *Journal of Dairy Science,* 93(7), 3287-3293, doi: http://dx.doi.org/10.3168/jds.2009-2758.

Dolecheck, K., Silvia, W., Heersche, G., Chang, Y., Ray, D., Stone, A., et al. (2015). Behavioral and physiological changes around estrus events identified using multiple automated monitoring technologies. *Journal of Dairy Science,* 98(12), 8723-8731, doi:10.3168/jds.2015-9645.

Hall, M. A. (1999). *Correlation-based feature selection for machine learning*. Doctoral dissertation, The University of Waikato, Hamilton, New Zealand.

Hill, T., Suarez-Mena, F., Hu, W., Dennis, T., Schlotterbeck, R., Timms, L., et al. (2017). Technicel note: Evaluation of an ear-attached movement sensor to record rumination, eating, and activity behaviors in 1-month-old calves. *The Professional Animal Scientist,* 33(6), 743-747, doi: https://doi.org/10.15232/pas.2017-01623.

Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer,* 29(3), 31-44.

Krieger, S., Sattlecker, G., Kickinger, F., Auer, W., Drillich, M., & Iwersen, M. (2017). Prediction of calving in dairy cows using a tail-mounted tri-axial accelerometer: A pilot study. *Biosystems Engineering*, doi: https://doi.org/10.1016/j.biosystemseng.2017.11.010.

Ledgerwood, D. N., Winckler, C., & Tucker, C. B. (2010). Evaluation of data loggers, sampling intervals, and editing techniques for measuring the lying behavior of dairy cattle. *Journal of Dairy Science,* 93(11), 5129-5139, doi: 10.3168/jds.2009-2945.

Lundorff Jensen, A., & Kjelgaard-Hansen, M. (2011). Diagnostic test validation. In D. J. Weiss, & K. J. Wardrop (Eds.), *Schalm's Veterinary Hematology* (6th ed., pp. 1027-1033). Ames: John Wiley & Sons, Ltd.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica,* 22(3), 276-282.

Miller-Cushon, E. K., & DeVries, T. J. (2015). Invited review: Development and expression of dairy calf feeding behaviour. *Canadian Journal of Animal Science,* 95(3), 341-350, doi: http://doi.org/10.4141/CJAS-2014-163.

Müller, R., & Schrader, L. (2003). A new method to measure behavioural activity levels in dairy cows. *Applied Animal Behaviour Science,* 83(4), 247-258, doi: 10.1016/S0168-1591(03)00141-2.

Powers, D. M. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies,* 2(1), 37-63.

Reiter, S., Sattlecker, G., Lidauer, L., Kickinger, F., Öhlschuster, M., Auer, W., et al. (2018). Evaluation of an ear-tag-based accelerometer for monitoring rumination in dairy cows. *Journal of Dairy Science,* 101, 1-14, doi: https://doi.org/10.3168/jds.2017-12686.

Robert, B., White, B. J., Renter, D. G., & Larson, R. L. (2009). Evaluation of three-dimensional accelerometers to monitor and classify behavior patterns in cattle. *Computers and Electronics in Agriculture,* 67(1), 80-84, doi: 10.1016/j.compag.2009.03.002.

Rutten, C. J., Velthuis, A. G. J., Steeneveld, W., & Hogeveen, H. (2013). Invited review: Sensors to support health management on dairy farms. *Journal of Dairy Science,* 96(4), 1928-1952, doi: http://dx.doi.org/10.3168/jds.2012-6107.

Scheibe, K. M., & Gromann, C. (2006). Application testing of a new three-dimensional acceleration measuring system with wireless data transfer (WAS) for behavior analysis. *Behavior Research Methods,* 38(3), 427-433.

Sturm, V., Efrosinin, D., Efrosinina, N., Roland, L., Iwersen, M., Drillich, M., et al. (2017). Automated classification of a calf's feeding state based on data collected by active sensors with 3D-accelerometer. In *International conference on distributed computer and communication networks, Moscow, Russia, 25-29 September, 2017 2017* (pp. 120-134). Cham: Springer

Swartz, T., McGilliard, M., & Petersson-Wolfe, C. (2016). Technical note: The use of an accelerometer for measuring step activity and lying behaviors in dairy calves. *Journal of Dairy Science,* 99, 1-5, doi: http://dx.doi.org/10.3168/jds.2016-11297.

Youden, W. J. (1950). Index for rating diagnostic tests. Cancer, 3(1), 32-35.