Automatic detection of feeding- and drinking-related agonistic behavior and dominance in dairy cows

B. Foris,¹ A. J. Thompson,² M. A. G. von Keyserlingk,² N. Melzer,^{1*} and D. M. Weary²

¹Institute of Genetics and Biometry, Leibniz Institute for Farm Animal Biology, 18196 Dummerstorf, Germany

²Animal Welfare Program, Faculty of Land and Food Systems, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada

ABSTRACT

Accurate assessments of social behavior and dominance relationships in cattle can be time consuming. We investigated whether replacements at the feed bunk and water trough—one type of agonistic interaction—can be used to automatically assess dominance relationships. Our study set out to (1) validate a replacement detection algorithm using combined data from electronic feed and water bins, and (2) investigate the applicability of this algorithm to identify individual dominance scores and group-level social hierarchy in freestall-housed dairy cows. We used 4 groups of lactating cows kept in different group sizes (11 to 20 cows) located at 2 research facilities. In both facilities, feed and water were provided via automated feeding systems. A trained observer recorded all agonistic interactions in the pen over multiple days using video. Data from the electronic feed and water bins for the same days were analyzed using an algorithm to detect replacements (i.e., visits where a receiver cow was competitively replaced by an actor cow). Most agonistic interactions at the feed bunk were replacements. These replacements were associated with a brief interval between the time the receiver cow left the bin and the actor cow took her place; the optimal threshold to detect these replacements varied from 22 to 27 s between groups, independent of stocking density. The recall and precision of an algorithm based upon this threshold was high (on average >0.8), comparable to that of trained human observers. We improved data preparation by controlling for detection errors and included filtering to reduce false positives. This resulted in a >20% decrease in false positives and an increase in precision of 0.043. The

dominance hierarchy based upon algorithm-detected replacements was similar to that based upon total agonistic interactions observed in the pen; the Spearman rank correlation coefficient between these hierarchies varied among the groups from 0.81 to 0.96. We conclude that data from electronic feed and water bins can accurately estimate agonistic behavior and dominance relationships among dairy cows.

Key words: social competition, precision livestock farming, welfare

INTRODUCTION

Automated methods are increasingly used to collect data on proximity and avoidance in cattle (Gygax et al., 2007; Boyland et al., 2013; Koene and Ipema, 2014) and these associations are sometimes used as a proxy for social interactions (Boyland et al., 2016). Detecting dyadic interactions could provide more specific information about social relationships (Foris et al., 2019). Existing work on social behavior in dairy cattle has relied on live observation or video (e.g., Wierenga, 1990; Val-Laillet et al., 2009; Tresoldi et al., 2015). Agonistic interactions occur mostly in areas where competition for resources is high, such as the feed bunk, drinker, lying stalls, mechanical brush, and automatic milking system (e.g., Val-Laillet et al., 2008a; Guzhva et al., 2016). Automated measures of attendance in these areas may also be used to estimate agonistic interactions (Guzhva et al., 2016, 2018; Huzzey et al., 2014; McDonald et al., 2019).

Previous work has assessed the use of an algorithm to detect agonistic replacements at the feed bunk using the data from an electronic feeding system (Huzzey et al., 2014; McDonald et al., 2019). A short time interval between visits by 2 different cows at the same bin was used to identify cases in which a receiver cow was competitively replaced by an actor cow. The optimal interval for identifying replacements was 26 s for feed bins (Huzzey et al., 2014) and 29 s for water bins (McDonald et al., 2019). Multiple studies have reported

Received March 27, 2019.

Accepted June 12, 2019.

*Corresponding author: melzer@fhn-dummerstorf.de

misidentification of cows and other technical issues (e.g., Chapinal et al., 2007; Ruuska et al., 2014; McDonald et al., 2019), but it is unclear how the performance of the algorithm is influenced by these issues. Moreover, the general applicability of a replacement detection algorithm to assess social relationships such as dominance has not been studied.

The objectives of the current study were to (1) evaluate the performance of a replacement detection algorithm using combined feed and water bin data, with a special focus on investigating the reasons for false-positive and false-negative detections; and (2) determine the efficacy of the algorithm for assessing individual- and group-level dominance within the group. To investigate the general applicability of the results, we addressed these objectives in different groups of dairy cows housed on different farms.

MATERIALS AND METHODS

Animals and Housing

We followed 4 groups of lactating Holstein cows (see Supplemental Table S1 for age, parity, and DIM; <https://doi.org/10.3168/jds.2019-16697>). Two groups of 20 cows were at the University of British Columbia Dairy Education and Research Centre in Agassiz, BC, Canada (**UBC** 1 and 2), housed in pens containing 24 lying stalls, 12 electronic feed bins, and 2 electronic water bins (see Neave et al., 2017 for a full description of housing conditions). Two other groups were at the Leibniz Institute for Farm Animal Biology in Dummerstorf, Germany (**FBN** 1: 14 cows, **FBN** 2: 11 cows), housed in pens containing 15 lying stalls, 10 electronic feed bins, and 2 electronic water bins (see Foris et al., 2019 for a full description of housing conditions). Thus, the groups differed in the stocking ratio of cows to feed bins, water bins, and lying stalls. All groups used Insentec feed and water bins (Roughage Intake Control System; Insentec, Marknesse, the Netherlands; validated by Chapinal et al., 2007 for intake and attendance measures) to provide ad libitum access to a TMR and water. To facilitate the identification of cows from video, cows were marked with unique alphanumeric symbols. Animals were cared for following the guidelines of the Canadian Council on Animal Care (CCAC, 2009) and the Association for the Study of Animal Behaviour/Animal Behavior Society (ASAB/ABS) *Guidelines for the Use of Animals in Research* (Anonymous, 2016). Data collection was approved by the UBC Animal Ethics Committee (protocol A10-0163) and by the Committee for Animal Use and Care of the Ministry of Agriculture, Environment and Consumer Protection of

the federal state of Mecklenburg-Western Pomerania, Germany (Reference number: 7221.3-2-033/15).

Video Recording and Analysis

In both facilities, multiple days of video recording were used based on previous work (Foris et al., 2019). The 2 UBC groups were video-recorded continuously for 2 consecutive days in summer 2013 using 4 Panasonic WVCW504SP video cameras (Panasonic Corp., Osaka, Japan) mounted approximately 6 m above the experimental pen. Cameras were connected to a digital video surveillance system (GeoVision, GeoVision Inc., Corona, CA). Red lights (100 W) were hung next to the cameras to facilitate video recording at night. The FBN groups were video-recorded in autumn 2016, over 3 consecutive days. Each group was recorded with 2 cameras (Sony YC 3189, Sony Corp., Tokyo, Japan; including infrared light) and a digital recorder (EDR HD-2H14/4H4, EverFocus Electronics Corp., New Taipei City, Taiwan).

All videos were analyzed using Mangold Interact V15 (Mangold International GmbH, Arnstorf, Germany) following Foris et al. (2019). All agonistic interactions in the entire pen were identified continuously. We excluded milking time, which began when the first cow left the pen before milking and ended when the last cow entered the pen after milking, and times when management activities disturbed the cows, such as cleaning the lying stalls. Agonistic interactions were defined as (1) displacements at the bins (aggressive contact; i.e., butt or push from the actor that resulted in the receiver leaving the bin); replacements were those displacements where the actor took occupancy of the bin within 60 s after the receiver left; and (2) displacements at the lying stall or the walking alley (aggressive contact that resulted in the receiver leaving the stall or taking a minimum of 2 steps in the walking alley). We only considered aggressive interactions that included contact between the actor and the receiver and where the outcome was clearly identifiable. We did not include more subtle agonistic encounters with no clear outcomes to avoid confounding the actor-receiver matrix when calculating dominance.

Replacement Detection Algorithm

All data preparation and analyses were performed using R version 3.4.2 (R Core Team, 2017).

Data Preparation. Data were downloaded from the electronic feed and water bins recording the start and end time of each visit and the identity of the cow. Data were combined for feed and water bins in each

group and summarized by day (visits could not overlap days because all bins closed shortly before midnight and reopened just after midnight). A technical issue associated with the water bins in group UBC 1 during the first day resulted in these data not being used. As a first step, data were screened to exclude any cow identification numbers from cows not housed within the group. Also, any times from when the bins were in the closed position (i.e., “0” transponder entries) were removed. As a second step, we checked for technical errors by visually assessing the plausibility of visit durations using boxplots to identify outliers; feeder visits longer than 30 min and drinker visits longer than 10 min were individually validated using video observation. We also quantified bouts when a cow was registered at 2 bins simultaneously and determined their cause based on video. As a third step, for each bin, the time-ordered data were used to determine the replacements by calculating the time difference between the exit time of one cow and the entry time of another (replacement duration); a replacement was recorded when this interval was less than or equal to a set threshold. Replacements were discarded when (1) the replacement was recorded during milking times or other management activities that were excluded from the video analysis (Supplemental Table S2; <https://doi.org/10.3168/jds.2019-16697>), or (2) the actor was registered at another bin at the exit time of the receiver (meaning that it was impossible that physical contact between the cows could have resulted in this replacement—the actor occupied the bin after the receiver left).

Evaluating the Performance of the Algorithm.

First, we determined the proportion of agonistic interactions in the groups that could possibly be detected by the electronic bins. For this, we calculated the percentage of replacements from displacements at the bins and from all agonistic interactions in the pen. Replacements identified by the human observer were considered the reference to assess replacements identified by algorithm. True positives (**TP**) were replacements identified by both the observer and the algorithm; false positives (**FP**) were replacements found only by the algorithm; and false negatives (**FN**) were replacements found only by the observer. True negatives were not interpretable in our analysis because video-recorded replacement events were used as reference (rather than all feeder visits regardless of whether a replacement occurred). Recall (also called sensitivity) was calculated to measure the quantity of replacements found, and precision was calculated to measure the quality of the revealed replacements:

$$Recall = \frac{TP}{TP + FN},$$

$$Precision = \frac{TP}{TP + FP}.$$

To characterize the tradeoff between recall and precision in the different groups, we calculated both measures for thresholds ranging from 1 to 60 s. To identify the optimal threshold, we used the F-score (i.e., the harmonic mean of recall and precision; Saito and Rehmsmeier, 2015):

$$F\text{-score} = \frac{2 \times Recall \times Precision}{Recall + Precision}.$$

For each group, the threshold associated with the maximum F-score was considered optimal and these thresholds were then used for further analysis.

We investigated the reasons for FN and FP in more detail. First, for TP events, we assessed the association between replacement durations as determined by the algorithm versus the observer and found very good agreement in each group (mean difference <1 s). Therefore, in the detailed analysis, we excluded those FN events where the replacement duration determined by the observer was longer than the threshold of the algorithm. To provide a comparison between the performance of the algorithm and the observer, an independent trained observer reviewed the FN and FP events and decided whether a replacement happened. Then, FN events (i.e., replacements detected only by the first observer) were categorized as “replacement” if the bin did not detect the replacement (e.g., due to technical error) and categorized as “no replacement” if the first observer made an error (i.e., misidentified the cow or coded a replacement instead of displacement). The FP events (i.e., replacements detected only by the electronic bins) were categorized as “no replacement” if there was no aggressive contact between the actor and the receiver, and as “replacement” if a replacement happened but had not been detected by the first observer. This validated data set was then used as a gold standard to recalculate recall and precision and to compare these between the algorithm and the first video observer. Furthermore, we assessed the number of FP replacements that were eliminated by our introduced filtering step (i.e., controlling for the presence of the actor at other bins, see above) and separately assessed replacements for when data from only feed bins or water bins were used.

Dominance Assessment

We investigated whether the algorithm could be used to assess the social hierarchy within the groups, calculating dominance scores (**DS**) for each cow based on the complete observation period, using the normalized David's score (de Vries et al., 2006) as implemented in the R package “steepness” (Leiva and de Vries, 2014). We only used specific data sets for dominance calculations if the ratio of interactions to cows was sufficient (i.e., at least 10 times as many interactions as cows in the group; Sánchez-Tójar et al., 2018). We used 4 different data sets in a stepwise approach to obtain hierarchies: (1) replacements revealed by the algorithm, termed *Bin DS*; (2) replacements at the bins identified by the observer; (3) displacements at the bins identified by the observer; and (4) displacements at the bins and in other zones as identified by the observer, termed *Complete DS*. For each group, we calculated the associations between these dominance hierarchies using Spearman rank correlation (r_s). In a detailed investigation, we compared *Bin DS* and *Complete DS* at the group level based on their steepness (measuring the degree of individual differences in winning dominance encounters; de Vries et al., 2006), and investigated the agreement in individual David's scores between *Bin DS* and *Complete DS*.

RESULTS

Video Analysis

The numbers of replacements and other displacements are presented in Table 1. In all groups, the majority of displacements occurred at the feed bunk, more than two-thirds of which were replacements. These

replacements represented roughly half of all agonistic interactions in the pen.

Replacement Detection Algorithm

The video validation showed that long bouts were not caused by a technical error but rather by cows truly occupying the bin for a long time (see Supplemental Figure S1 for boxplots of bout durations; <https://doi.org/10.3168/jds.2019-16697>). Cows were detected simultaneously at 2 bins in fewer than 1% of bin visits (Supplemental Table S2); all these events were caused by one cow replacing another without the bin registering the change rather than by a bin staying open after a cow left.

The tradeoff between the recall and precision of the algorithm with increasing threshold is shown in Figure 1A–D. Groups varied in the optimal threshold revealed by the maximum F-score (mean $\pm$ SD: 24.8 ± 2.06 s), and in the threshold where recall and precision were closest to equal (mean $\pm$ SD: 21.3 ± 3.40 s). In all groups, recall and precision values were high (around 0.8) across thresholds, ranging from 15 to 30 s. Few intervals >30 s were replacements and most replacements were 9 to 12 s (Figure 1E–H). Accordingly, thresholds >30 s were associated with only a slight increase in recall and a decline in precision (i.e., the increase in TP was associated with a higher amount of FP). Thresholds corresponding to 9 to 12 s were associated with high precision (i.e., almost all detected replacements were true), but they were worse in recall (i.e., many replacements were missed).

We compared the performance of the algorithm with the performance of the video observer using independent validation of FN and FP replacements. Table 2 shows

Table 1. Number of video observed replacements¹ and other displacements² at the electronic bins, lying stalls, and in the walking alley in 4 groups in 2 facilities³ (FBN 1, 2, UBC 1, 2); the proportion of replacements from agonistic interactions at the bins or in the pen is also shown

Group	Replacements	Other displacements			Proportion of replacements	
		Bins	Lying stalls	Walking alley	Agonistic at bins	All agonistic
FBN 1	232	105	68	159	0.69	0.41
FBN 2	196	83	22	42	0.70	0.57
UBC 1	257	118	16	49	0.69	0.58
UBC 2	357	165	23	27	0.68	0.62

¹Aggressive contact that results in the receiver leaving the bin and the actor is the first cow taking occupancy of the same bin within 60 s.

²Other displacement at the bins: aggressive contact that results in the receiver leaving the bin but the actor is not the first cow taking occupancy of the same bin within 60 s. Displacement in other zones: aggressive contact in which the receiver leaves the lying stall or takes a minimum of 2 steps in the walking alley.

³FBN = Leibniz Institute for Farm Animal Biology (Dummerstorf, Germany); UBC = University of British Columbia Dairy Education and Research Centre (Agassiz, BC, Canada).

Figure 1. Recall, precision, and F-scores of the replacement detection algorithm for different thresholds in 4 groups in 2 facilities (A: FBN 1, B: FBN 2, C: UBC 1, D: UBC 2), where FBN = Leibniz Institute for Farm Animal Biology (Dummerstorf, Germany), and UBC = University of British Columbia Dairy Education and Research Centre (Agassiz, BC, Canada). Gray lines show the threshold with the maximum F-score, dashed gray lines show the point where recall and precision are closest to equal. The distribution for the duration of replacements detected by the electronic bins (E: FBN 1, F: FBN 2, G: UBC 1, H: UBC 2) is also shown.

Table 2. Evaluation of replacements detected at the electronic bins, comparison of recall and precision of the first video observer and the algorithm in 4 groups in 2 facilities¹ (FBN 1, 2, UBC 1, 2)²

Group	Threshold ³ (s)	True positives ⁴	False positives ⁵		False negatives ⁶		Recall		Precision	
			–	+	–	+	Bins	Observer	Bins	Observer
FBN 1	27	184 (185)	53 (68)	15	10 (2)	3	0.99 (0.99)	0.93	0.79 (0.75)	0.95
FBN 2	22	165	20 (26)	3	2	8	0.96	0.98	0.89 (0.87)	0.99
UBC 1	25	232 (234)	39 (58)	18	7 (5)	7	0.97 (0.98)	0.93	0.87 (0.81)	0.97
UBC 2	25	316	48 (71)	24	0	13	0.96	0.93	0.88 (0.83)	1.00

¹FBN = Leibniz Institute for Farm Animal Biology (Dummerstorf, Germany); UBC = University of British Columbia Dairy Education and Research Centre (Agassiz, BC, Canada).

²The corresponding values without using the additional filtering step to control for the presence of the actor at other bins are shown in parentheses.

³Optimal threshold determined by the maximum F-score.

⁴Replacements found by both the algorithm and the first observer.

⁵Replacements found only by the algorithm using the optimal threshold. + and – represent replacement and no replacement, respectively, according to the independent observer.

⁶Replacements found only by the first observer, and only replacements not longer than the optimal threshold are considered. + and – represent replacement and no replacement, respectively, according to the independent observer.

the number of replacements found by the observer and the algorithm, and the decision (i.e., replacement or no replacement) of the independent observer for replacements originally categorized as FN or FP. Based on this information as a gold standard, we calculated similarly high recall values for the algorithm and for the first observer; the precision of the algorithm was high (around 0.80), but lower than the precision of the observer.

Filtering decreased the number of false detections (cf. Table 2). The number of FP decreased by an average of 27.6%, resulting in an average increase in precision of 0.043. Only a slight decrease in recall was observed in 2 groups because of filtering some replacements where the bin that was previously occupied by the actor was still open when the receiver left the bin.

The contribution of feed and water bins to detecting replacements is shown in Table 3. The water bins recorded only a small proportion of all detected replace-

ments in all groups. In groups where water bin data were used for all days, on average 4.3% of TP, 8.9% of FP, and 26.4% of FN replacements were associated with water bins.

Dominance Assessment

In all groups, there was a strong association between the 4 dominance hierarchies (Table 4; all $r_s > 0.80$). Focusing on individual cows, we investigated whether the algorithm could predict *Complete DS* values. Figure 2 provides evidence that electronic bins performed well in estimating dominance status. However, the algorithm overestimated DS for cows that were lower in the hierarchy and underestimated it for cows at the top of the hierarchy. In other words, the algorithm generally resulted in flatter hierarchies (i.e., less difference in DS between the most and least dominant cows).

Table 3. Number of replacements detected by feed and water bins using the optimal threshold in 4 groups in 2 facilities¹ (FBN 1, 2, UBC 1, 2)²

Group	Threshold (s)	True positives		False positives		False negatives	
		Feeder	Water	Feeder	Water	Feeder	Water
FBN 1	27	176	8	65	3	10	3
FBN 2	22	164	1	23	0	9	1
UBC 1	25	224	8	55	2	10	4
UBC 2	25	291	25	56	16	7	6

¹FBN = Leibniz Institute for Farm Animal Biology (Dummerstorf, Germany); UBC = University of British Columbia Dairy Education and Research Centre (Agassiz, BC, Canada).

²Replacements longer than the optimal threshold are excluded from false negatives (replacements found only by the video observer).

Table 4. Spearman rank correlation coefficients between dominance hierarchies in 4 groups in 2 facilities¹ (FBN 1, 2, UBC 1, 2)²

Item	A	B	C	D
FBN 1				
A	1	0.92	0.93	0.89
B		1	0.98	0.87
C			1	0.90
D				1
FBN 2				
A	1	0.97	0.96	0.96
B		1	0.99	0.99
C			1	1
D				1
UBC 1				
A	1	0.85	0.81	0.81
B		1	0.96	0.95
C			1	0.97
D				1
UBC 2				
A	1	0.97	0.96	0.96
B		1	0.97	0.98
C			1	0.99
D				1

¹FBN = Leibniz Institute for Farm Animal Biology (Dummerstorf, Germany); UBC = University of British Columbia Dairy Education and Research Centre (Agassiz, BC, Canada).

²Correlation coefficients were calculated based on A, replacements revealed by the algorithm using the optimal threshold; B, observer replacements at the bins; C, observer displacements at the bins; D, observer displacements at the bins and in other zones.

DISCUSSION

We evaluated an algorithm based on combined electronic feed and water bin data to detect replacements at the feed bunk and estimate the dominance hierarchy in groups.

Video Analysis

We used 2 to 3 d of continuous video to assess agonistic behavior within groups of lactating dairy cows. Previous work showed that this period provides a reasonable snapshot of the agonistic behavioral structure within a group (Foris et al., 2019). Milking times were determined based on video observation in this study, but the electronic feeder system (combined with tri-axial accelerometers) also allows for automated detection of times when cows are away (Thompson et al., 2017). Regarding the spatial distribution of agonistic interactions, our results indicate that if cows are not overstocked for lying stalls, most social competition happens at the feed bunk, supporting previous findings (Val-Laillet et al., 2008a). In our study, we only considered aggressive interactions with clear outcomes and did not include subtle interactions or indirect dominance effects without aggressive contact. To our knowledge, this is the first study to assess the proportion of replacements from these agonistic interactions

in dairy cows. In all groups, replacements represented the majority of agonistic interactions at the feed bunk and a high percentage of all agonistic interactions in the pen. This finding supports the use of electronic bins for the automatic recording of social competition.

Replacement Detection Algorithm

When using automated data, it is important to control for systematic errors. We included quality checks in the data preparation procedure. Assessing outliers at the bins may detect if a bin does not close when a cow leaves and could serve as an indicator for mechanical or electronic problems. The simultaneous detection of a cow at 2 bins can be the result of the above-mentioned issue, or it can occur if a cow replaces another in a way that the bin does not detect (termed “stealing” by Ruuska et al., 2014). The latter issue was marginal in our data, but such data checks may be of more value in long-term assessments.

To assess the use of electronic bins for monitoring social competition, we investigated how the algorithm threshold influenced recall and precision, and how these vary in different facilities and in groups with varying stocking densities. We assessed replacements by video only when an agonistic contact from the actor resulted in the occupation of that bin within 60 s. Our results confirmed that this limit was appropriate because the vast majority of replacements took place within 30 s (Figure 1). The optimal thresholds were similar in all 4 groups, but higher stocking densities were associated with somewhat higher optimal thresholds. Given that most replacements were detected at the feed bins, this trend might be attributed to stocking density at the feeder. Based on the tradeoff between recall and precision, we suggest using thresholds between 20 and 30 s when applying the algorithm in groups under similar housing conditions. In this range, F-scores were high, but lower thresholds were often associated with much lower recall, and higher thresholds corresponded to slightly lower F-scores. Previous studies examined data from feed bins (Huzzey et al., 2014) or water bins (McDonald et al., 2019), but in the current study we used these in combination. We observed too few replacements at the water bins to assess the data from the 2 sources separately. The optimal thresholds for feed (26 s) and water (29 s) bins identified in these earlier studies correspond well with the range identified here. Despite the small proportion of replacements at water bins observed in our study, we encourage combining feed and water bin data whenever possible; data from drinkers may be especially helpful during periods of heat stress. In addition, the data from the water bins can contribute to filtering FP replacements. Our

filtering step increased the precision of the algorithm in all observed groups. The plausibility check of replacements by controlling if the actor was detected at another location when the receiver left the bin appears to be a promising approach for further improvements. Combining this algorithm with data from other sensors,

including accelerometers (Ledgerwood et al., 2010) and location tracking (Pastell et al., 2018), could improve performance.

Direct or video observation is commonly used as the gold standard to evaluate behavioral detection methods. However, human observers are also prone to error.

Figure 2. Dominance scores (DS value: normalized David's score) of cows based on all observed agonistic interactions in the pen (Complete DS) and based on replacements detected by the algorithm using the optimal threshold (Bin DS) in 4 groups in 2 facilities (A: FBN 1, B: FBN 2, C: UBC 1, D: UBC 2), where FBN = Leibniz Institute for Farm Animal Biology (Dummerstorf, Germany), and UBC = University of British Columbia Dairy Education and Research Centre (Agassiz, BC, Canada). Cows are presented in descending Complete DS order. For both dominance hierarchies the equations describing the fitted line are provided in parentheses. Number of interactions (n) used for calculating DS values and Spearman rank correlation coefficients (r_s) between Complete DS and Bin DS values of cows are also presented.

We analyzed discrepancies between the observer and the algorithm. Recall was slightly lower for the observer compared with the algorithm in 3 groups (Table 2), indicating that the observer might miss more replacements than the algorithm. Precision was higher for the observer in all groups, because the algorithm detected a replacement even if a cow occupied a bin without any agonistic contact with the previous cow. However, the precision of the algorithm was generally high, confirming that, in most cases, short time intervals between cows at the same bin were associated with agonistic interactions. Our results indicate that the algorithm is comparable to human video observation in detecting replacements at the feed bunk. Multiple studies have used this replacement detection algorithm with feed bin data (e.g., Sepúlveda-Varas et al., 2016; Crossley et al., 2017; Jensen and Proudfoot, 2017). The use of the algorithm facilitates long-term assessments for large groups that would be impractical using human observers, and more generally facilitates the inclusion of agonistic behaviors in future work. Groups in this study were small compared with those in usual commercial conditions. However, stocking density at the feed bunk may have more influence on the optimal algorithm settings than group size per se. Therefore, we expect good algorithm performance in larger groups equipped with the same system where stocking density is comparable to that in our study. In addition, the presented approach can be used as a framework for evaluating different systems that are more suitable for commercial farms. Replacement detection might also be possible with other methods, such as location tracking or monitoring attendance at specific feeding areas.

Dominance Assessment

Studies of dominance in dairy cattle have a long history (e.g., Schein and Fohrman, 1955) and have questioned whether patterns of social competition correspond to classical properties of dominance (Wierenga, 1990; Val-Laillet et al., 2008b, 2009). On commercial dairy farms, group composition is frequently changed according to the physiological status and nutritional needs of cows, affecting social behavior and production (e.g., von Keyserlingk et al., 2008; Schirrmann et al., 2011). Long-term data collection on agonistic interactions may improve our understanding of dominance in dynamic groups, but it is not feasible using traditional observation methods. Our second objective was to assess whether the algorithm can be used to assess the dominance structure. The algorithm-based dominance showed very good agreement with dominance based on all observed agonistic interactions at the feed bunk, indicating that automated replacements can be used to

assess social hierarchy at the feed bunk. These results agree with previous findings showing that the success index of focal cows is highly correlated between automatically detected replacements and video-recorded displacements at the feed bunk (Huzzey et al., 2014). However, different resources in the pen may be associated with different hierarchies according to the individual motivation of cows to gain access. Previous work found a low positive correlation between the competitive success of cows at the feed bunk and at the lying stalls (Val-Laillet et al., 2008a). In the current study, cows were overstocked for feed and water bins but understocked for lying stalls in all groups, and the number of agonistic interactions was low in areas other than the feed bunk. Indeed, we recorded too few displacements at the lying stalls to meet the minimum for dominance calculation (cf. Sánchez-Tójar et al., 2018), and so could not determine whether the competition at the feed bunk and at the lying stalls followed different patterns. A comparison of *Bin DS* and *Complete DS* suggests that differences in competitiveness of cows are less pronounced at the feed bunk (i.e., a subordinate cow may replace a dominant cow, perhaps according to individual differences in satiety). Other work has shown that many bidirectional displacement relationships are observed at the feeder (Val-Laillet et al., 2008b). High stocking density at the lying stalls and frequent regrouping may influence the number of agonistic interactions in different pen areas and therefore also the reliability of algorithm-based dominance estimates. In this study, the group with the lowest proportion of replacements had a higher number of displacements at the lying stalls and in the walking alley than the other groups. The relatively high stocking density at the bins in this group could be one reason for higher social stress and more aggressive encounters independently from resources. However, this did not seem to affect our findings; correlations between the different dominance hierarchies were high, comparable with those observed in other groups. Despite the potential variation in hierarchies at different locations, high correlations between *Bin DS* and *Complete DS* show that the algorithm can reasonably estimate the dominance status of cows, at least under housing conditions similar to those in this study. The use of electronic bins for automatic dominance assessment may help us understand associations between the social environment of cows and other factors such as health or productivity.

CONCLUSIONS

Combined electronic feed and water bin data were used to detect agonistic replacements at the feed bunk in 4 groups of cows housed at different densities and in

different facilities. We found that a 20- to 30-s interval between visits by 2 different cows at the same bin was optimal to identify replacements. The algorithm performed well in terms of recall and precision, showing similar reliability to human observations of video. The algorithm can also be used to assess dominance structure, facilitating work on social relationships in dairy cows.

ACKNOWLEDGMENTS

We are grateful to J. Lomb for her help in extracting data from the UBC herd. We also thank the faculty, staff, and students at The University of British Columbia's Dairy Education and Research Centre and the Animal Welfare Program for their help in collecting this data. We also thank the staff members at the experimental cattle facility of the FBN; the technicians for supporting our work: D. Sehland, and K. Wendland; S. Düpjan for providing cameras; and I. Müntzel, P. Löbner, and J. Dickert for video analysis. This study was part of a project funded by The Natural Sciences and Engineering Research Council of Canada's Discovery research grants awarded to M. A. G. von Keyserlingk and D. M. Weary. This study was also part of the joint project "PHÄNOMICS" with financial support of the German Federal Ministry of Education and Research (BMBF Grant No. 0315536G).

REFERENCES

- Anonymous. 2016. Guidelines for the treatment of animals in behavioural research and teaching. *Anim. Behav.* 111:I–IX. [https://doi.org/10.1016/S0003-3472\(15\)00461-3](https://doi.org/10.1016/S0003-3472(15)00461-3).
- Boyland, N. K., R. James, D. T. Mlynski, J. R. Madden, and D. P. Croft. 2013. Spatial proximity loggers for recording animal social networks: Consequences of inter-logger variation in performance. *Behav. Ecol. Sociobiol.* 67:1877–1890. <https://doi.org/10.1007/s00265-013-1622-6>.
- Boyland, N. K., D. T. Mlynski, R. James, L. J. N. Brent, and D. P. Croft. 2016. The social network structure of a dynamic group of dairy cows: From individual to group level patterns. *Appl. Anim. Behav. Sci.* 174:1–10. <https://doi.org/10.1016/j.applanim.2015.11.016>.
- CCAC (Canadian Council on Animal Care). 2009. CCAC Guidelines on the Care and Use of Farm Animals in Research, Teaching and Testing. Canadian Council on Animal Care, Ottawa, ON, Canada. https://www.ccac.ca/Documents/Standards/Guidelines/Farm_Animals.pdf.
- Chapinal, N., D. M. Veira, D. M. Weary, and M. A. G. von Keyserlingk. 2007. Technical note: Validation of a system for monitoring individual feeding and drinking behavior and intake in group-housed cattle. *J. Dairy Sci.* 90:5732–5736. <https://doi.org/10.3168/jds.2007-0331>.
- Crossley, R. E., A. Harlander-Matauschek, and T. J. DeVries. 2017. Variability in behavior and production among dairy cows fed under differing levels of competition. *J. Dairy Sci.* 100:3825–3838. <https://doi.org/10.3168/jds.2016-12108>.
- de Vries, H., J. M. G. Stevens, and H. Vervaecke. 2006. Measuring and testing the steepness of dominance hierarchies. *Anim. Behav.* 71:585–592. <https://doi.org/10.1016/j.anbehav.2005.05.015>.
- Foris, B., M. Zebunke, J. Langbein, and N. Melzer. 2019. Comprehensive analysis of affiliative and agonistic social networks in lactating dairy cattle groups. *Appl. Anim. Behav. Sci.* 210:60–67. <https://doi.org/10.1016/j.applanim.2018.10.016>.
- Guzhva, O., H. Ardö, A. Herlin, M. Nilsson, K. Åström, and C. Bergsten. 2016. Feasibility study for the implementation of an automatic system for the detection of social interactions in the waiting area of automatic milking stations by using a video surveillance system. *Comput. Electron. Agric.* 127:506–509. <https://doi.org/10.1016/j.compag.2016.07.010>.
- Guzhva, O., H. Ardö, M. Nilsson, A. Herlin, and L. Tufvesson. 2018. Now you see me: Convolutional neural network based tracker for dairy cows. *Front. Robot. AI* 5:1–9. <https://doi.org/10.3389/frobt.2018.00107>.
- Gygax, L., G. Neisen, and H. Bollhalder. 2007. Accuracy and validation of a radar-based automatic local position measurement system for tracking dairy cows in free-stall barns. *Comput. Electron. Agric.* 56:23–33. <https://doi.org/10.1016/j.compag.2006.12.004>.
- Huzzey, J. M., D. M. Weary, B. Y. F. Tiau, and M. A. G. von Keyserlingk. 2014. Short communication: Automatic detection of social competition using an electronic feeding system. *J. Dairy Sci.* 97:2953–2958. <https://doi.org/10.3168/jds.2013-7434>.
- Jensen, M. B., and K. L. Proudfoot. 2017. Effect of group size and health status on behavior and feed intake of multiparous dairy cows in early lactation. *J. Dairy Sci.* 100:9759–9768. <https://doi.org/10.3168/jds.2017-13035>.
- Koene, P., and B. Ipema. 2014. Social networks and welfare in future animal management. *Animals (Basel)* 4:93–118. <https://doi.org/10.3390/ani4010093>.
- Ledgerwood, D. N., C. Winckler, and C. B. Tucker. 2010. Evaluation of data loggers, sampling intervals, and editing techniques for measuring the lying behavior of dairy cattle. *J. Dairy Sci.* 93:5129–5139. <https://doi.org/10.3168/jds.2009-2945>.
- Leiva, D., and H. de Vries. 2014. steepness: Testing steepness of dominance hierarchies. R package version 0.2–2. <https://CRAN.R-project.org/package=steepness>.
- McDonald, P. V., M. A. G. von Keyserlingk, and D. M. Weary. 2019. Technical note: Using an electronic drinker to monitor competition in dairy cows. *J. Dairy Sci.* 102:3495–3500. <https://doi.org/10.3168/jds.2018-15585>.
- Neave, H. W., J. Lomb, M. A. G. von Keyserlingk, A. Behnam-Shabang, and D. M. Weary. 2017. Parity differences in the behavior of transition dairy cows. *J. Dairy Sci.* 100:548–561. <https://doi.org/10.3168/jds.2016-10987>.
- Pastell, M., L. Frondelius, M. Järvinen, and J. Backman. 2018. Filtering methods to improve the accuracy of indoor positioning data for dairy cows. *Biosyst. Eng.* 169:22–31. <https://doi.org/10.1016/j.biosystemseng.2018.01.008>.
- R Core Team. 2017. R: A language and environment for statistical computing. Version 3.4.2 R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Ruuska, S., W. Hämäläinen, A. Sairanen, E. Juutinen, L. Tuomisto, M. Järvinen, and J. Mononen. 2014. Can stealing cows distort the results of feeding trials? An experiment for quantification and prevention of stealing feed by dairy cows from roughage intake control feeders. *Appl. Anim. Behav. Sci.* 159:1–8. <https://doi.org/10.1016/j.applanim.2014.08.001>.
- Saito, T., and M. Rehmsmeier. 2015. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One* 10:e0118432. <https://doi.org/10.1371/journal.pone.0118432>.
- Sánchez-Tójar, A., J. Schroeder, and D. R. Farine. 2018. A practical guide for inferring reliable dominance hierarchies and estimating their uncertainty. *J. Anim. Ecol.* 87:594–608. <https://doi.org/10.1111/1365-2656.12776>.
- Schein, M. W., and M. H. Fohrman. 1955. Social dominance relationships in a herd of dairy cattle. *Br. J. Anim. Behav.* 3:45–55. [https://doi.org/10.1016/S0950-5601\(55\)80012-3](https://doi.org/10.1016/S0950-5601(55)80012-3).
- Schirrmann, K., N. Chapinal, D. M. Weary, W. Heuwieser, and M. A. G. von Keyserlingk. 2011. Short-term effects of regrouping on be-

- havior of prepartum dairy cows. *J. Dairy Sci.* 94:2312–2319. <https://doi.org/10.3168/jds.2010-3639>.
- Sepúlveda-Varas, P., K. L. Proudfoot, D. M. Weary, and M. A. G. von Keyserlingk. 2016. Changes in behaviour of dairy cows with clinical mastitis. *Appl. Anim. Behav. Sci.* 175:8–13. <https://doi.org/10.1016/j.applanim.2014.09.022>.
- Thompson, A. J., D. M. Weary, and M. A. G. von Keyserlingk. 2017. Technical note: Mining data from on-farm electronic equipment to identify the time dairy cows spend away from the pen. *J. Dairy Sci.* 100:3975–3982. <https://doi.org/10.3168/jds.2016-11713>.
- Tresoldi, G., D. M. Weary, L. C. P. M. Filho, and M. A. G. von Keyserlingk. 2015. Social licking in pregnant dairy heifers. *Animals (Basel)* 5:1169–1179. <https://doi.org/10.3390/ani5040404>.
- Val-Laillet, D., A. M. de Passillé, J. Rushen, and M. A. G. von Keyserlingk. 2008b. The concept of social dominance and the social distribution of feeding-related displacements between cows. *Appl. Anim. Behav. Sci.* 111:158–172. <https://doi.org/10.1016/j.applanim.2007.06.001>.
- Val-Laillet, D., V. Guesdon, M. A. G. von Keyserlingk, A. M. de Passillé, and J. Rushen. 2009. Allogrooming in cattle: Relationships between social preferences, feeding displacements and social dominance. *Appl. Anim. Behav. Sci.* 116:141–149. <https://doi.org/10.1016/j.applanim.2008.08.005>.
- Val-Laillet, D., D. M. Veira, and M. A. G. von Keyserlingk. 2008a. Short communication: Dominance in free-stall-housed dairy cattle is dependent upon resource. *J. Dairy Sci.* 91:3922–3926. <https://doi.org/10.3168/jds.2008-1332>.
- von Keyserlingk, M. A. G., D. Olenick, and D. M. Weary. 2008. Acute behavioral effects of regrouping dairy cows. *J. Dairy Sci.* 91:1011–1016. <https://doi.org/10.3168/jds.2007-0532>.
- Wierenga, H. K. 1990. Social dominance in dairy cattle and the influences of housing and management. *Appl. Anim. Behav. Sci.* 27:201–229. [https://doi.org/10.1016/0168-1591\(90\)90057-K](https://doi.org/10.1016/0168-1591(90)90057-K).