Behavioral and physiological changes around estrus events identified using multiple automated monitoring technologies


*Department of Animal and Food Sciences, University of Kentucky, Lexington 40546
†Research Support Office, Royal Veterinary College, University of London, London, NW1 0TU, United Kingdom

ABSTRACT

This study included 2 objectives. The first objective was to describe estrus-related changes in parameters automatically recorded by the CowManager SensOor (Agis Automatisering, Harmelen, the Netherlands), DVM bolus (DVM Systems LLC, Greeley, CO), HR Tag (SCR Engineers Ltd., Netanya, Israel), IceQube (IceRobotics Ltd., Edinburgh, UK), and Track a Cow (Animart Inc., Beaver Dam, WI). This objective was accomplished using 35 cows in 3 groups between January and June 2013 at the University of Kentucky Coldstream Dairy. We used a modified Ovsynch with G7G protocol to partially synchronize ovulation, ending after the last PGF2α injection (d 0) to allow estrus expression. Visual observation for standing estrus was conducted for four 30-min periods at 0330, 1000, 1430, and 2200 h on d 2, 3, 4, and 5. Eighteen of the 35 cows stood to be mounted at least once during the observation period. These cows were used to compare differences between the 6 h before and after the first standing event (estrus) and the 2 wk preceding that period (nonestrus) for all technology parameters. Differences between estrus and nonestrus were observed for CowManager SensOor minutes feeding per hour, minutes of high ear activity per hour, and minutes ruminating per hour; twice daily DVM bolus reticulorumen temperature; HR Tag neck activity per 2 h and minutes ruminating per 2 h; IceQube lying bouts per hour, minutes lying per hour, and number of steps per hour; and Track a Cow leg activity per hour and minutes lying per hour. No difference between estrus and nonestrus was observed for CowManager SensOor ear surface temperature per hour. The second objective of this study was to explore the estrus detection potential of machine-learning techniques using automatically collected data. Three machine-learning techniques (random forest, linear discriminant analysis, and neural network) were applied to automatically collected parameter data from the 18 cows observed in standing estrus. Machine learning accuracy for all technologies ranged from 91.0 to 100.0%. When we compared visual observation with progesterone profiles of all 32 cows, we found 65.6% accuracy. Based on these results, machine-learning techniques have potential to be applied to automatically collected technology data for estrus detection.

Key words: precision dairy farming technology, estrus detection, automated estrus detection, technology, machine learning

INTRODUCTION

Detecting a high percentage of cows in estrus is essential to maintain reproductive performance in dairy herds using AI. The most common form of estrus detection is visual observation, used by 93% of US dairy operations (USDA, 2007). The Dairy Records Management Systems reported mean yearly estrus detection rate on US Holstein herds (including all reproductive management strategies) as 44.9% in 2015 (DRMS, 2015). This low estrus-detection rate may be a result of the extreme decline in Holstein cattle estrus duration (from 18 to less than 8 h) over the last 50 yr (Reames et al., 2011). Increasing age, milk production, and environmental factors (greater ambient temperature, uncomfortable housing, and so on) can also negatively affect length and intensity of estrus expression (Vailes and Britt, 1990; López-Gatius et al., 2005; Palmer et al., 2010).

Automated estrus detection (AED) technologies are an available alternative to supplement or replace visual estrus detection. Parameters with potential for AED include mounting events, activity level, lying time, ruminant events, blood or milk progesterone (P4) levels, feeding time, body temperature, and more (Senger, 1994; Saint-Dizier and Chastant-Maillard, 2012; Fricke et al., 2014). Estrus-related changes in some of
these parameters (mounting events, activity level, lying time, rumination events, and P4) have been quantified repeatedly. However, a lack of consistent data exists surrounding estrus-related changes in feeding time and body temperature. Additionally, not all of these parameters have been measured on the same cows during the same estrus periods.

To determine the accuracy of a specific AED technology, estrus events identified by the technology algorithm (a set of criteria used to determine “estrus”) are compared with a gold standard such as visual observation, ultrasonography, blood or milk P4 levels, or a combination of these. Correctly identified estrus events are considered true positives (TP), nonalerted estrus events are false negatives (FN), nonalerted nonestrus events are true negatives (TN), and alerted nonestrus events are false positives (FP; Firk et al., 2002). Detecting estrus events is a balance of sensitivity and specificity. Sensitivity, the probability that an event is alerted, is equal to TP/(TP + FN) × 100 (Hogeveen et al., 2010). Specificity, the probability that when an event does not occur no alert is generated, is equal to TN/(TN + FP) × 100. Because neither sensitivity nor specificity account for the prevalence of the event, other comparative measurements are also useful, including accuracy: \[
\frac{(TP + TN)/(TP + TN + FP + FN)}{2} \times 100
\]

The estrus-detection accuracy of a technology depends on 3 factors: (1) how strongly and closely the measured parameters are associated with estrus, (2) how accurately the technology is measuring those parameters, and (3) if the technology manufacturer algorithm is accurately processing the data to create useful estrus alerts. Most technology manufacturer algorithms are proprietary, making it difficult to identify how well each of the 3 factors described above are performing. Machine-learning techniques can replace the manufacturer alert algorithms and evaluate technologies based solely on parameter data collected. Mitchell et al. (1996) and Krieter (2005) previously described the use of machine-learning techniques for estrus detection; however, both studies focused on identifying the day of estrus rather than a more specific period. Additionally, no commercially available AED technologies were evaluated in those analyses.

Our study included 2 objectives. The first objective was to describe estrus-related changes in neck activity, ear activity, leg activity, step count, lying bouts, lying time, rumination, feeding time, reticulorumen temperature, and ear surface temperature as measured using 5 AED technologies on the same cows. The second objective of our study was to explore the estrus-detection potential of machine-learning techniques using parameters collected by AED technologies.

MATERIALS AND METHODS

This study was conducted at the University of Kentucky Coldstream Dairy under Institutional Animal Care and Use Committee protocol number 2013–1069. All lactating cows (n = 82) were housed in 2 groups, separated by a shared, raised feedbunk. Both groups maintained open access to freestalls, one group with sawdust-covered rubber-filled mattresses (PastureMat; Promat, Ontario, Canada) and the other group with sawdust-covered Dual Chamber Cow Waterbeds (Advanced Comfort Technology Inc., Reedsburg, WI). Cows received access to a grass-seeded exercise lot for 1 h/d at 1000 h, weather permitting. All other surfaces accessible to cows (freestall area, feed bunk, holding pen, and alleys) contained grooved concrete. Delivery of a TMR ration containing corn silage, alfalfa silage, whole cottonseed, and grain mix occurred twice daily at 0530 and 1330 h. Milking occurred twice daily at 0430 and 1530 h.

Our study enrolled 32 Holstein cows not bred in their current lactation. Parity, DIM at the beginning of the study protocol, and summit milk production from the current lactation of these cows were (mean ± SD) 2.0 ± 1.2, 77.8 ± 20.5 d, and 39.8 ± 8.8 kg, respectively. Cow ovulations were synchronized in 3 groups of 14, 10, and 8 cows, starting on January 24, March 19, and May 14, respectively. The synchronization protocol (Figure 1) was a modification of the standard Ovsynch (Pursley et al., 1995), preceded by G7G (Bello et al., 2006). In contrast to the standard Ovsynch, the last injection of GnRH (gonadorelin diacetate tetrahydrate, Cystorelin; Merial Limited, Duluth, GA; 100 μg i.m.) was not administered to stimulate estrus expression. Additionally, to stimulate corpus luteum regression, 2 PGF2α injections (dinoprost tromethamine, Lutalysin; Zoetis, Florham Park, NJ; 25 mg i.m.) were given on the last day of the protocol (7 d after the first GnRH injection), 6 h apart (0800 and 1400 h). Day 0 was designated as the last day of the synchronization protocol in each group (Figure 1).

Estrus Confirmation

Visual observation of cows for four 30-min periods at 0330, 1000, 1430, and 2200 h occurred on d 2, 3, 4, and 5 (Figure 1). Two observers were present at each shift, with one assigned to each side of the separated housing area. Study cows were clearly identified using spray paint. Observers recorded the time of each standing estrus event.

Blood samples (10 mL) were collected from cow coccygeal veins on d −2, −1, 0, 1, 2, 7, 9, and 11 (Figure
1. Plasma was separated from centrifuged samples (2,800 rpm for 20 min at room temperature) and stored at −20°C until the concentration of P4 was determined by RIA (Coat-a-Count Progesterone, Siemens Medical Solutions USA Inc., Malvern, PA). Response to the synchronization protocol was confirmed if P4 was greater than 1.0 ng/mL on d −2, −1, and 0, dropped to less than 1.0 ng/mL by d 1, and returned above 1.0 ng/mL by d 9. We used P4 results to determine sensitivity, specificity, and accuracy of visual observation. Only validated standing estrus events were used to describe estrus-related changes in AED parameters and to explore estrus-detection potential of machine-learning techniques.

### Technologies Evaluated

Each cow was fitted with 5 automated monitoring technologies before beginning synchronization. The CowManager SensOor (Agis Automatisering, Harmelen, the Netherlands), attached to the left ear, used a 3-axis accelerometer to classify each minute into 1 of 6 behaviors (rumination, feeding, resting, low activity, regular activity, or high activity) and reported hourly percentage of time associated with each behavior. Additionally, the CowManager SensOor used a digital surface temperature monitor to evaluate mean hourly ear surface temperature. The behavioral portion of the CowManager SensOor, but not the temperature monitor, was previously validated on dairy cows (Bikker et al., 2014). The DVM bolus (DVM Systems, LLC, Greeley, CO), placed into the reticulorumen using a bolus gun, recorded reticulorumen temperature twice daily using a passive radio-frequency identification transponder. Data download occurred at the time of parlor entrance, where panel readers were located. The HR Tag (SCR Engineers Ltd., Netanya, Israel), held on the left side of the neck using a nylon collar, measured neck activity and rumination time in 2-h blocks using a 3-axis accelerometer and a microphone containing a microprocessor, respectively. The rumination portion of the HR Tag was previously validated on dairy cattle (Schirmann et al., 2009; Burfeind et al., 2011). The IceQube (IceRobotics Ltd., Edinburgh, UK), attached to the left rear leg using a plastic strap, reported number of steps, lying bouts, and lying time every 15 min using a 3-axis accelerometer. The Track a Cow (Animart Inc., Beaver Dam, WI), attached to the front right leg using a nylon strap, used a 3-axis accelerometer to measure hourly activity and lying time.

### Statistical Analysis

All technology parameter data were averaged by 12-h blocks of time. The 12-h block of time used to define estrus depended on the analysis.

#### Estrus versus Nonestrus.

For our analysis, if a cow was observed in standing estrus during visual observation periods (0330, 1000, 1430, or 2200 h), that cow’s estrus was classified as starting 6 h before the first observed standing estrus event and ending 6 h after the first observed standing estrus event. For example, a cow first observed in standing estrus during the 1430 h observation period would have estrus defined as 0830 to 2030 h of that day. The twenty-eight 12-h periods (14 d) before the estrus period were classified as periods of nonestrus. The MIXED procedure of SAS 9.3 (SAS Institute, Inc., Cary, NC) was used to analyze the main effects of estrus status (estrus or nonestrus), parity, DIM at the start of the synchronization protocol, summit milk production, and the interaction of estrus status and selected covariates (parity, DIM at the start of the synchronization protocol, and summit milk production) on all technology parameter data, considering cow as a random effect and time as a repeated measure. All main effects were kept in each model regardless of significance level. Stepwise backward elimination was used to remove nonsignificant interactions ($P \geq 0.05$).

#### Machine Learning.

For this analysis, if a cow was observed in standing estrus during visual observation periods (0330, 1000, 1430, or 2200 h), that cow’s estrus was classified as the 12-h period of time leading up to the first observed standing estrus event. For example, a cow first observed in standing estrus during the 1430 h observation period would have estrus defined as 0230 to 1430 h of that day. This was different from the estrus versus nonestrus analysis because it would not be valuable for machine learning to detect estrus after the
observation of standing estrus. The twenty-eight 12-h periods (14 d) before the estrus period were classified as periods of nonestrus.

Unmodified data, as recorded by 4 of the technologies (CowManager SensOor, HR Tag, IceQube, and Track a Cow), were used for machine learning analysis. The DVM bolus was excluded from this analysis because machine-learning techniques work by finding patterns between parameters and are not meant to be applied to single parameter data sets. The caret package from R version 3.1.1 (R Foundation for Statistical Computing, Vienna, Austria) was used to create a 4-fold cross-validation, including 10 analysis per series, using 70% of all technology parameter data. Three machine-learning techniques were tested: random forest, linear discriminant analysis, and neural network. The goal of the algorithm development was to predict which time block (of the twenty-nine 12-h periods of estrus and nonestrus defined earlier) each data line referenced. After algorithm development, the remaining 30% of all technology parameter data were used to test prediction ability. Sensitivity, specificity, and accuracy of each technology and machine-learning technique combination were calculated relative to observed standing estrus. The exact method was used to calculate 95% confidence intervals for each measurement (Clopper and Pearson, 1934).

RESULTS AND DISCUSSION

Progesterone analysis indicated that 29 of the 32 cows (90.6%) ovulated after completing the synchronization protocol. Eighteen cows (62.1%) were observed standing to be mounted during the visual observation periods. Failure to detect the remaining 11 cows may have resulted from unexpressed estrus or short estrus lengths that were unobserved because of noncontinuous observation.

A researcher error resulted in some data not properly saved from the computer. Consequently, 4 cows observed in estrus were missing lying time data as measured by Track a Cow and were removed from the statistical analysis. Additionally, a technology malfunction resulted in no data measured by the IceQube for 1 other cow, which was also removed from the statistical analysis. Remaining technology parameter statistical analyses included all 18 cows observed in standing estrus.

**Estrus versus Nonestrus**

**Activity.** All activity measures increased during estrus compared with nonestrus (Table 1). The percent activity change between nonestrus and estrus for high ear activity as measured by CowManager SensOor, neck activity as measured by HR Tag, number of steps as measured by IceQube, and leg activity as measured by Track a Cow were 309.4, 118.5, 280.4, and 237.4%, respectively (Table 1). The range of increase in activity may have resulted from differing accelerometer attachment locations. Overall, similar estrus-associated increases in numbers of steps (2 to 4 times) have been reported previously (Kiddy, 1977; Redden et al., 1993; Roelofs et al., 2005a).

The interaction of DIM at the start of synchronization and estrus status significantly influenced all measures of activity (Table 2). Cows that started the synchronization protocol at a later DIM displayed greater estrus-related activity levels than cows starting the synchronization protocol at earlier DIM. Addition-

Table 1. Comparison of automated monitoring technology\(^1\) parameters (adjusted means ± SE) during estrus (6 h before and after first observed standing event\(^2\)) and nonestrus (the 14 d before estrus)

<table>
<thead>
<tr>
<th>Category</th>
<th>Parameter</th>
<th>n(^3)</th>
<th>Estrus</th>
<th>Nonestrus</th>
<th>P-value(^4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
<td>HR Tag neck activity (units/2 h)</td>
<td>18</td>
<td>61.62 ± 2.04</td>
<td>28.20 ± 0.78</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>IceQube number of steps (per h)</td>
<td>17</td>
<td>300.82 ± 10.92</td>
<td>79.07 ± 4.13</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>CowManager SensOor high ear activity (min/h)</td>
<td>18</td>
<td>17.40 ± 0.66</td>
<td>4.25 ± 0.39</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Track a Cow leg activity (units/h)</td>
<td>18</td>
<td>321.14 ± 11.87</td>
<td>95.17 ± 7.16</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Lying time and lying bouts</td>
<td>IceQube lying bouts (per h)</td>
<td>17</td>
<td>0.35 ± 0.09</td>
<td>0.72 ± 0.07</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>IceQube lying time (min/h)</td>
<td>17</td>
<td>10.19 ± 1.91</td>
<td>24.82 ± 0.95</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Track a Cow lying time (min/h)</td>
<td>14</td>
<td>6.56 ± 2.55</td>
<td>18.18 ± 1.81</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Rumination and feeding time</td>
<td>HR Tag rumination (min/2 h)</td>
<td>18</td>
<td>20.47 ± 2.68</td>
<td>32.96 ± 0.54</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>CowManager SensOor rumination (min/h)</td>
<td>18</td>
<td>12.90 ± 1.07</td>
<td>22.96 ± 0.57</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>CowManager SensOor feeding time (min/h)</td>
<td>18</td>
<td>16.93 ± 0.99</td>
<td>8.93 ± 0.65</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Temperature</td>
<td>DVM bolus reticulorumen temperature (°C)</td>
<td>18</td>
<td>39.29 ± 0.21</td>
<td>38.86 ± 0.18</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>CowManager SensOor ear surface temperature (°C)</td>
<td>18</td>
<td>24.17 ± 1.20</td>
<td>22.97 ± 0.83</td>
<td>0.20</td>
</tr>
</tbody>
</table>

\(^1\)CowManager SensOor (Agis Automatisering, Harmelen, the Netherlands); DVM bolus (DVM Systems, LLC, Greeley, CO); HR Tag (SCR Engineers Ltd., Netanya, Israel); IceQube (IceRobotics Ltd., Edinburgh, Scotland); and Track a Cow (Animart Inc., Beaver Dam, WI).

\(^2\)Observations for standing estrus occurred for 30-min periods at 0330, 1000, 1430, and 2200 h daily.

\(^3\)Number of cows included in statistical analysis.

\(^4\)The reported P-value represents the main effect of estrus status (estrus or nonestrus) alone, independent of covariate effects.
ally, the interaction of parity and estrus status significantly influenced activity as measured by the IceQube and Track a Cow (Table 2). In both cases, as parity increased, estrus-related activity decreased. In agreement, López-Gatius et al. (2005) found that with each additional parity, walking activity decreased 21.4%. Other studies have identified a similar relationship (Rolfós et al., 2005a; Yániz et al., 2006). In our study, parity only influenced estrus-related activity levels when monitored using leg-mounted technologies, indicating that later-parity cows increase head and neck movements during estrus but do not walk around as much as younger cows. Activity, as measured by Track a Cow, was also significantly influenced by the interaction of summit milk production and estrus status (Table 2). As summit milk production increased, estrus-related activity increases were suppressed. The relationship between greater milk production and decreased estrus-related activity has previously been established (López-Gatius et al., 2005; Yániz et al., 2006; Reith et al., 2014). Why all activity measurement devices did not observe this effect is unclear.

**Lying Time and Lying Bouts.** All lying measures decreased during estrus compared with nonestrus (Table 1). The percent change between nonestrus and estrus for lying bouts as measured by IceQube, lying time as measured by IceQube, and lying time as measured by Track a Cow were similar at −51.4, −58.9, and −63.9%, respectively. Time spent lying decreases around estrus because of increased activity and restlessness (Esslemont and Bryant, 1976; Livshin et al., 2005; Jonsson et al., 2011).

The interaction of DIM at the start of synchronization and estrus status significantly influenced lying bouts as measured by IceQube and lying time as measured by Track a Cow (Table 2). Cows that started the synchronization protocol at a later DIM expressed shorter lying time as measured by Track a Cow and fewer lying bouts as measured by IceQube during estrus than cows starting the synchronization protocol at earlier DIM. Why lying time as measured by IceQube was not affected in the same way is unclear. No measures of lying activity were significantly influenced by the interactions of parity or summit milk production with estrus status.

**Rumination and Feeding Time.** Both measures of rumination time decreased during estrus compared with nonestrus (Table 1). The percent change in rumination time between nonestrus and estrus for the CowManager SensOor and the HR Tag were −43.8 and −37.9%, respectively. Reith and Hoy (2012) evaluated 265 estrus events, finding that rumination on the day of estrus decreased 17% (74 min), but with large variation between herds (14 to 24%). In a follow-up study that

### Table 2. Effect of estrus status, parity, DIM at the start of synchronization, summit milk production, and selected interactions on automated monitoring technology parameters

<table>
<thead>
<tr>
<th>Category</th>
<th>Parameter</th>
<th>Estrous Parity</th>
<th>DIM</th>
<th>Summit Milk Production</th>
<th>Estrus × Parity</th>
<th>Estrus × DIM</th>
<th>Estrus × Summit</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity IceQube</td>
<td>number of steps (per h)</td>
<td>0.42</td>
<td>0.01</td>
<td>0.01</td>
<td>0.44</td>
<td>&lt;0.01</td>
<td>0.01</td>
<td>&gt;0.01</td>
</tr>
<tr>
<td>Track a Cow leg</td>
<td>activity (units/h)</td>
<td>&lt;0.01</td>
<td>0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>IceQube lying</td>
<td>time (min/h)</td>
<td>&lt;0.01</td>
<td>0.04</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Track a Cow lying time (min/h)</td>
<td></td>
<td>0.29</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>CowManager SensOor ear surface temperature (°C)</td>
<td></td>
<td>0.38</td>
<td>0.85</td>
<td>0.03</td>
<td>0.13</td>
<td>&lt;0.01</td>
<td>0.03</td>
<td>0.02</td>
</tr>
</tbody>
</table>

1 Observations for standing estrus occurred for 30-min periods at 0330, 1000, 1430, and 2200 h daily. Estrus was defined as the 6 h before and after the first observed standing event.

2 CowManager SensOor (Agis Automatisering, Harmelen, the Netherlands); DVM bolus (DVM Systems, LLC, Greeley, CO); HR Tag (SCR Engineers Ltd., Netanya, Israel); IceQube (IceRobotics Ltd., Edinburgh, Scotland); and Track a Cow (Animart Inc., Beaver Dam, WI).
looked at 453 estrous cycles, rumination time decreased 19.6% (83 min) on the day of estrus (Reith et al., 2014). Pahl et al. (2015) also found a decrease in rumination on the day of (19.3%) and the day before (19.8%) inseminations leading to pregnancy. The comparatively large decreases in rumination around estrus found in the current study could be the result of a narrower estrus window (12 h) as compared with the previous studies (1 d).

Differences between technology-measured rumination times (2.66 min/h during estrus and 6.48 min/h during nonestrus) could be the result of differing recording methods. The CowManager SensOor used an accelerometer to identify ear movement associated with rumination. The HR Tag used a microphone system that rested on the cow’s neck to identify the regurgitation and rechewing of cud. Both systems have been validated with high correlations to visual observation (CowManager SensOor: r = 0.93; HR Tag: r = 0.93; Bikker et al., 2014; Schirmann et al., 2009). However, the CowManager SensOor validation was conducted on a per-minute basis whereas the HR Tag validation was conducted on a 2-h basis, meaning results are not directly comparable.

One explanation for decreased rumination around estrus is decreased feed intake (Maltz et al., 1997; Diskin and Sreenan, 2000). Conversely, feeding time as measured by the CowManager SensOor in our study increased by 8.00 min/h during estrus compared with nonestrus (Table 1). Other researchers agree that feeding behavior may not always decrease around estrus. De Silva et al. (1981) found no change in feed intake during the 3-d period surrounding estrus, and Lukas et al. (2008) found DMI increased 0.61 kg/d during estrus. The method by which the CowManager SensOor measured feeding time in the current study depended on the ability of an accelerometer to distinguish ear movements related to feeding; this is not a true measure of intake. Therefore, the reported increase in feeding time may not represent an actual increase in DMI, but rather an increase in head movements similar to those occurring when a cow is eating.

Feeding time was not significantly influenced by the interaction of DIM at the start of synchronization, parity, or summit milk production with estrus status. The interaction of DIM at the start of synchronization and estrus status significantly influenced both measures of rumination (Table 2). Cows that started the synchronization protocol at a later DIM expressed a larger decrease in rumination during estrus than cows starting the synchronization protocol at earlier DIM. This result is consistent with the other observations of estrus expression in our study (activity and lying time), as DIM at the start of synchronization increased. Neither measure of rumination was significantly influenced by the interactions of parity or summit milk production with estrus status.

**Temperature.** Reticulorumen temperature as measured by the DVM bolus increased 0.43°C during estrus ($P < 0.01$; Table 1). Ear surface temperature as measured by the CowManager SensOor increased 1.20°C during estrus ($P = 0.20$; Table 1). Although the numeric increase in ear surface temperature during estrus was greater than that of the reticulorumen temperature, it also displayed a larger variation as evident in the greater standard error (Table 1). Ear surface temperature is influenced both by core body temperature and ambient temperature (Mader and Kreikemeier, 2006); therefore, ear surface temperature was expected to be less than and fluctuate more than reticulorumen temperature (a measure of core body temperature alone). CowManager SensOor temperature measurements are not marketed for estrus detection use, likely because of this variation.

The temperature increases observed in this study (0.51–1.27°C) are similar to previously reported estrus-related temperature changes. Both Maatje and Rossing (1976) and McArthur et al. (1992) found that milk temperature increased 0.3°C around estrus. Other researchers have found that vaginal temperature increased 0.10 to 1.02°C around estrus (Lewis and Newman, 1984; Kyle et al., 1998). Piccione et al. (2003) found that rectal temperatures, though not measured automatically, displayed an even greater increases during estrus (1.3°C). These estrus-related temperature increases have reportedly lasted for 6.8 ± 4.6 h in dairy cows and 6.5 ± 2.7 h in beef cows (Redden et al., 1993; Kyle et al., 1998).

Differences in temperature measurements may have resulted from the difference in frequency of measurement between the 2 technologies. The CowManager SensOor sampled temperature each minute and reported a mean hourly ear surface temperature, whereas the DVM bolus recorded reticulorumen temperature only twice daily at the time the cow entered the parlor for milking. Reticulorumen temperature readings at those times likely did not accurately represent the entire 12-h period between milkings and, therefore, would not be comparable to ear surface temperature as measured by the CowManager SensOor. Newer versions of the DVM bolus can continuously monitor temperature, which could reduce variation between the 2 technologies.

Ear surface temperature as measured by CowManager SensOor was not significantly influenced by the interactions of DIM at the start of synchronization, parity, or summit milk production with estrus status. Reticulorumen temperature as measured by DVM bolus was significantly influenced by the interactions of both DIM at the start of synchronization and parity.
with estrus status (Table 2). Cows that started the synchronization protocol at a later DIM expressed a larger increase in reticulorumen temperature during estrus than cows starting the synchronization protocol at earlier DIM. Additionally, as parity increased, a smaller estrus-related increase in reticulorumen temperature was observed. Both of these results contribute to the overall conclusion that, as DIM at the beginning of the synchronization protocol decreased and parity increased, weaker estrus expression was observed.

**Machine Learning**

Because of the low number of observed estrus events in our study (n = 18), when 70% of the data were used for the machine learning training sets, data from only 5 cows were left for the machine learning testing sets. Consequently, results should be interpreted carefully, keeping in mind the small sample size. Table 3 shows the sensitivity, specificity, and accuracy accomplished using different combinations of each of the 5 technologies and 3 machine-learning techniques (random forest, linear discriminant analysis, or neural network). Confidence intervals are reported for each measure of performance to emphasize the difficulty in drawing conclusions from the small data set.

Using the random forest machine-learning technique, the CowManager SensOor and IceQube produced the greatest accuracy (98.6%; Table 3). The CowManager SensOor also produced the greatest accuracy (100%) when using linear discriminant analysis, whereas the IceQube produced the greatest accuracy (100%) when using neural networks (Table 3). The number and variety of parameters measured by both the CowManager SensOor (4 parameters measured) and IceQube (3 parameters measured) likely gave them an advantage in these analyses over the other technologies which measured only 2 parameters each (HR Tag and Track a Cow). Similarly, Peralta et al. (2005) showed that although visual observation, activity monitoring, and mounting detection alone produced low estrus-detection sensitivities (49.3, 37.2, and 48.0%, respectively), combining all 3 produced an acceptable sensitivity of 80.2%. Redden et al. (1993) also found that by combining 2 parameters (activity and vaginal temperature), which alone each produced an 80% estrus detection rate, a 90% estrus-detection rate was possible.

Of the remaining technologies, all machine learning results were similar. Accuracy of the HR Tag and Track a Cow ranged from 96.6 to 97.9% and from 91.0 to 97.2%, respectively. Compared with other studies that have tested similar machine-learning techniques for estrus detection, these results are high. Krieter (2005) applied the neural network technique, combining activity and time since last estrus, to a testing set of 74 estrus events. That method accomplished a sensitivity, specificity, and error rate of 77.5, 99.6, and 9.1%, respectively. Mitchell et al. (1996) applied machine-learning techniques to milk yield, milking order, and time since last estrus to identify 69% of estrus events in a 44-cow testing set, but experienced a large number of FP (74%). Both of those analyses predicted the day of estrus, whereas the current study focused on predicting a 12-h period before estrus. Narrowing the

<table>
<thead>
<tr>
<th>Technique</th>
<th>Technology</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random forest</td>
<td>CowManager SensOor</td>
<td>100.00 (47.82–100.00)</td>
<td>98.57 (84.93–99.83)</td>
<td>98.62 (95.11–99.83)</td>
</tr>
<tr>
<td></td>
<td>HR Tag</td>
<td>60.00 (14.66–94.73)</td>
<td>99.29 (96.08–99.98)</td>
<td>97.93 (94.07–99.57)</td>
</tr>
<tr>
<td></td>
<td>IceQube</td>
<td>80.00 (28.36–99.49)</td>
<td>99.29 (96.08–99.98)</td>
<td>98.62 (95.11–99.83)</td>
</tr>
<tr>
<td></td>
<td>Track a Cow</td>
<td>100.00 (47.82–100.00)</td>
<td>97.14 (92.85–99.22)</td>
<td>97.24 (93.09–99.24)</td>
</tr>
<tr>
<td>Linear discriminant analysis</td>
<td>CowManager SensOor</td>
<td>100.00 (47.82–100.00)</td>
<td>100.00 (97.40–100.00)</td>
<td>100.00 (47.82–100.00)</td>
</tr>
<tr>
<td></td>
<td>HR Tag</td>
<td>100.00 (47.82–100.00)</td>
<td>97.86 (93.87–99.56)</td>
<td>97.93 (94.07–99.57)</td>
</tr>
<tr>
<td></td>
<td>IceQube</td>
<td>100.00 (47.82–100.00)</td>
<td>97.86 (93.87–99.56)</td>
<td>97.93 (94.07–99.57)</td>
</tr>
<tr>
<td></td>
<td>Track a Cow</td>
<td>100.00 (47.82–100.00)</td>
<td>96.43 (91.86–98.83)</td>
<td>96.55 (92.14–98.87)</td>
</tr>
<tr>
<td>Neural network</td>
<td>CowManager SensOor</td>
<td>100.00 (47.82–100.00)</td>
<td>98.57 (94.93–99.83)</td>
<td>98.62 (95.11–99.83)</td>
</tr>
<tr>
<td></td>
<td>HR Tag</td>
<td>100.00 (47.82–100.00)</td>
<td>96.43 (91.86–98.83)</td>
<td>96.55 (92.14–98.87)</td>
</tr>
<tr>
<td></td>
<td>IceQube</td>
<td>100.00 (47.82–100.00)</td>
<td>100.00 (97.40–100.00)</td>
<td>100.00 (97.49–100.00)</td>
</tr>
<tr>
<td></td>
<td>Track a Cow</td>
<td>100.00 (47.82–100.00)</td>
<td>90.71 (84.64–94.96)</td>
<td>91.03 (85.16–95.14)</td>
</tr>
</tbody>
</table>

1Sensitivity = TP/(TP + FN); specificity = TN/(TN + FP); accuracy = (TP + TN)/(TP + TN + FP + FN); TP = true positive, TN = true negative, FP = false positive, and FN = false negative.
2CowManager SensOor (Agis Automatisering, Harmelen, the Netherlands); HR Tag (SCR Engineers Ltd., Netanya, Israel); IceQube (IceRobotics Ltd., Edinburgh, Scotland); and Track a Cow (Animart Inc., Beaver Dam, WI).
3Machine learning models attempted to identify the 12-h period before the first observed standing estrus event from the twenty-eight 12-h periods leading up to observed standing estrus. The analysis included 18 cows observed in standing estrus, with 70% used for training and 30% used for testing. Observations for standing estrus occurred for 30-min periods at 0330, 1000, 1430, and 2200 h daily. Data from only 14 cows were used for Track a Cow lying time and data from only 17 cows were used for all IceQube parameters.
The estrus period may be more accurate given that multiple researchers have found mean estrus duration to be less than 24 h (Kerbrat and Disenhaus, 2004; Roelofs et al., 2005c; Sveberg et al., 2011). Another explanation for the improved results in our study is the low number of observations in the testing set. Only 5 cows were included in the testing set, resulting in a small number of potential TP (n = 5), a large number of potential TN (n = 140), and wide confidence intervals.

Estrus-detection ability of machine-learning techniques was superior to visual observation. When visual observation was compared with P4 results of all 32 cows, a 62.1% sensitivity, 100% specificity, and 65.6% accuracy of estrus detection were achieved. Noncontinuous monitoring likely limited the ability of visual observation to detect short periods of estrus. Additionally, using secondary signs of estrus to define estrus rather than standing events alone likely would have increased estrus-detection rate (Roelofs et al., 2005c).

The ability to continuously monitor cows using automated monitoring technologies, allowing detection of short estrus periods and estrus periods not including mounting, likely contributed to improved performance over visual observation. However, results should be interpreted carefully given that only 18 cows, all of which exhibited standing estrus, were included in the machine learning analysis, whereas 32 cows, some exhibiting standing estrus and some not, were included in the visual observation results. Cows not displaying standing estrus could not be included in the machine learning analysis because the study design did not allow for identification of exact ovulation time.

CONCLUSIONS

Neck activity, ear activity, leg activity, step count, lying bouts, lying time, rumination, feeding time, and reticulorumen temperature may be useful as predictors of estrus. Ear surface temperature, as monitored in this study, holds less potential for detecting differences between periods of estrus and nonestrus. Additionally, applying machine-learning techniques to automatically collected technology data shows potential for estrus detection.

ACKNOWLEDGMENTS

The authors thank Agis Automatisering (Harmelen, the Netherlands), DVM Systems LLC (Greeley, CO), SCR Engineers Ltd. (Netanya, Israel), IceRobotics Ltd. (Edinburgh, Scotland), and Animart Inc. (Beaver Dam, WI) for supplying the technologies used in this study along with technical support. Additionally, we thank Joey Clark, Susan Hayes, and the rest of the staff at the University of Kentucky Coldstream Dairy for their assistance with this study. This study would not have been possible without visual estrus detection assistance from multiple graduate and undergraduate students.

REFERENCES

Clopper, C., and E. S. Pearson. 1934. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26:404–413.


