J. Dairy Sci. 99:608–620
http://dx.doi.org/10.3168/jds.2014-8823

Monitoring individual cow udder health in automated milking systems using online somatic cell counts

L. P. Sørensen,*[1] M. Bjerring,† and P. Løvendahl*
*Department of Molecular Biology and Genetics, Center for Quantitative Genetics and Genomics, and
†Department of Animal Science, Research Centre Foulum, Aarhus University, DK-8830 Tjele, Denmark

ABSTRACT

This study presents and validates a detection and monitoring model for mastitis based on automated frequent sampling of online cell count (OCC). Initially, data were filtered and adjusted for sensor drift and skewed distribution using ln-transformation. Acceptable data were passed on to a time-series model using double exponential smoothing to estimate level and trends at cow level. The OCC levels and trends were converted to a continuous (0–1) scale, termed elevated mastitis risk (EMR), where values close to zero indicate healthy cow status and values close to 1 indicate high risk of mastitis. Finally, a feedback loop was included to dynamically request a time to next sample, based on latest EMR values or errors in the raw data stream. The estimated EMR values were used to issue 2 types of alerts, new and (on-going) intramammary infection (IMI) alerts. The new alerts were issued when the EMR values exceeded a threshold, and the IMI alerts were issued for subsequent alerts. New alerts were only issued after the EMR had been below the threshold for at least 8 d. The detection model was evaluated using time-window analysis and commercial herd data (6 herds, 595,927 milkings) at different sampling intensities. Recorded treatments of mastitis were used as gold standard. Significantly higher EMR values were detected in treated than in contemporary untreated cows. The proportion of detected mastitis cases using new alerts was between 28.0 and 43.1% and highest for a fixed sampling scheme aiming at 24 h between measurements. This was higher for IMI alerts, between 54.6 and 89.0%, and highest when all available measurements were used. The lowest false alert rate of 6.5 per 1,000 milkings was observed when all measurements were used. The results showed that a dynamic sampling scheme with a default value of 24 h between measurements gave only a small reduction in proportion of detected mastitis treatments and remained at 88.5%. It was concluded that filtering of raw data combined with a time-series model was effective in detecting and monitoring mastitis status in dairy cows when based on IMI alerts, and by using a dynamically adjusting sampling scheme almost full performance was still obtainable. However, results were less desirable when based on new alerts most likely because of the used gold standard for mastitis, which may not necessarily reflect the onset of and IMI case in contrast to a new alert.
Key words: dairy cattle, mastitis detection, automated milking, online somatic cell count

INTRODUCTION

Mastitis in dairy cattle is a serious disease that causes reduced milk quality and animal welfare, substantial losses due to production loss, increased treatment costs and labor, and higher culling rates (Halasa et al., 2007). Therefore, close monitoring of individual cow udder health is essential for identification of cows in the early stages of an IMI case, as well as timely initiation of treatment and assessment of recovery. Where cows are milked in traditional milking parlors, mastitic cows are identified by the milker, who visually inspects the milk from each quarter for signs of mastitis before milking, sometimes with the aid of sensor technology (i.e., measurements of quarter-based electric conductivity; Hamann and Zecconi, 1998).

In herds using automatic milking systems (**AMS**), no milker is present to visually assess the milk quality of each cow. In Denmark and other milk-producing countries within the European Union, visual control of milk for signs of IMI and color changes is mandatory (EU Directive EC/853/2004). Therefore in AMS, the herd manager must rely on in- or online sensor systems for identification of cows with milk not meeting quality standards (i.e., cows with IMI). Although a range of sensor systems are available, there is a shortage of described and validated detection systems that convert

Received September 5, 2014.
Accepted September 10, 2015.
[1]Corresponding author: LarsPeter.Sorensen@agrsci.dk

multiple sensor-level information into decision support in an effective way, allowing the herd manager to be adequately equipped for the best short- and long-term decisions (Rutten et al., 2013). In the simplest of cases, the herd manager is presented with the raw sensor data only and historic information is not well treated, and because different persons may interpret such data differently this may very well lead to erratic or subjective decisions.

The use of frequently recorded sensor data could allow for a close monitoring of udder health if suitable software interpreting new and historic data were available. The idea of close monitoring of individual cows is 3-fold: (1) to raise alerts if deviations from healthy status occur, (2) to focus on the sick cows for decision making about treatment, and (3) to follow the recovery from IMI as long as it takes. Sensor-based alerts may be detectable long before treatments would usually be initiated, thus allowing for more detailed diagnosis and dedicated treatment. Sensor-based monitoring may also be helpful in detecting recurrent cases where culling would be the ultimate decision; however, the quality and usefulness of any monitoring system depends on its performance. An ideal monitoring system produces a low number of false alerts, that is, high specificity (**SP**), while alerting in a timely manner and with emphasis on the more severe cases (Mollenhorst et al., 2012).

It is generally agreed that IMI cause high SCC levels (Harmon, 1994). The online cell counter (**OCC**; DeLaval International AB, Tumba, Sweden; henceforth DeLaval) was built to use the cell count changes as indicator of IMI and is dedicated for use with AMS for continuous monitoring of cow udder health.

The first objective of our study was to develop an IMI-monitoring system utilizing frequently sampled OCC measurements from AMS-milked dairy cows to provide the dairy manager with daily accurate information about individual cow udder health. The proposed monitoring system aimed to point out cows with IMI and to keep them under surveillance as long as the infection persisted. The second objective was to validate the proposed monitoring system using data from 6 commercial dairy herds.

MATERIALS AND METHODS

This study aimed first at developing an algorithm for detection and monitoring of IMI using OCC data taking a time-series approach on research station data and next to evaluate the monitoring model using OCC data from 6 commercial herds. Technical details of the OCC algorithm and the subsequent optimization procedure are described in the Supplementary Material (http://dx.doi.org/10.3168/jds.2015-8823).

OCC Data

Data for our study was collected via remote access to 1 research herd (**DCRC**; Danish Cattle Research Center, Tjele, Denmark) and 6 commercial dairy herds each using AMS (VMS, DeLaval) fitted with OCC measuring units. The commercial herds had between 103 and 284 Holstein cows and 2 to 5 AMS units. The research herd consisted of 175 cows (2 groups of Holstein and 1 Jersey group) and each group milked in 1 of 3 AMS units. Data from DCRC was used for model development and optimization. It was collected from January 1 to November 30, 2012, and consisted of 150,468 milkings from 387 cow lactations with between 1 and 1,137 milkings. A total of 117,399 milkings were associated with an OCC measurement excluding values equal to zero (i.e., indication of failed measurement). The validation data were collected from the 6 commercial herds from January 1 to December 1, 2012, except 3 herds where OCC units had been out of use for a period and were restarted for this project at April 1, May 16, and May 17, respectively. The edited validation data set consisted of 595,927 milkings from 1,938 cow lactations with between 1 and 1,138 milkings. A total of 519,871 milkings also had OCC measurements. In all cases only milkings between 0 and 305 DIM were used.

Initial Assessment of OCC Versus SCC

A first assessment of the raw OCC data quality was obtained by comparing OCC measurements with test-day SCC data using data for the year prior (2011) to the data collection period (2012) in the 6 commercial herds and the DCRC herd. Milk samples were collected during regular milk recording test days using a basic milk sampler (XMS, DeLaval) attached to the AMS units in each herd. Each milk sample was barcoded and stamped with time of milking. Subsequently, the milk samples were analyzed for SCC at the certified laboratory for DHI analysis (Eurofins, Holstebro, Denmark) using CombiFoss equipment (Foss Electric, Hillerød, Denmark). The ln-transformed OCC and SCC measurements were then compared by linear regression using the REG procedure in SAS (version 9.3, SAS Institute Inc., Cary, NC) to assess the measuring accuracy of each OCC unit.

Model Architecture

The IMI detection algorithm consisted of modules (Figure 1) which are described in details in the Supplementary Material (http://dx.doi.org/10.3168/jds.2015-8823). Briefly, in the first module, the raw data filtering and adjustment module, raw OCC readings were

checked for validity, ln-transformed, and adjusted for sensor device aberrations and drift via single exponential smoothing (Hyndman et al., 2008) at sensor level. The adjusted OCC values were combined with cow data and information from the AMS databases and transferred to the next module, the noise reduction and trend detection module. Here, a time-series approach was taken to reduce variance by means of double exponential smoothing (Hyndman et al., 2008) so that dynamic levels and any trends in data could be estimated (i.e., sudden increases in OCC values indicating IMI). The output from this module, OCC level and trend, were then combined in the EMR module for calculation of elevated mastitis risk (**EMR**) values. The EMR value is continuous (on a 0–1 scale), where values close to zero indicate little or no risk of mastitis and higher values approaching 1 indicate increasing or high risk of mastitis. Alerts were issued based on a fixed EMR threshold value, which was found during an optimization phase (see Supplementary Material; http://dx.doi.org/10.3168/jds.2015-8823). Finally, the time to next sample (**TNS**) module dynamically determined when to request the next OCC measurement based on the calculated EMR values and systematic cow factors.

Definition of Model Alerts. An alert was raised when EMR values exceeded an optimized threshold value (see Supplementary Material for details; http://dx.doi.org/10.3168/jds.2015-8823). Two types of alerts were defined: new alert and IMI alert, where the new alert is the first alert when a new IMI is detected. An IMI alert was defined as any milking where EMR values exceed the threshold value; thus, IMI alerts also included new alerts. The motivation for the 2 levels of model alerts is that a cow can have several IMI alerts during the course of an IMI either because recovery from IMI takes a long time or because the cow suffers from persistent IMI. Herd managers are not interested in seeing alerts for the same cow multiple times, but prefer a single new alert to take action from. When a cow receives a new alert, it will appear on both a new alert and an IMI alert list. After the new alert, it will appear only on the IMI alert list, where it can remain and be monitored until the mastitis case has been cured and udder health status has recovered.

A cow may have more than 1 IMI during any lactation. According to IDF (1987), 2 separate IMI cases can be defined when time between the 2 cases is more than 8 d. A similar rule was applied for the OCC model; if time between 2 alerts (new or IMI) was more than 8 d, the most recent alert was assumed to belong to a new IMI and defined as a new alert.

Gold Standard for Mastitis. It was assumed that mastitis cases were treated with antibiotics to cure the infection and treatments were given by a veterinarian or the herd manager. For the participating herds, any treatment of a cow was assumed reported to the National Danish Cattle Database (SEGES, Aarhus, Denmark). Thus, recorded mastitis treatments were extracted from that database and used as gold standard for IMI. A cow can have multiple treatments during lactation; thus, 2 mastitis records were assumed to belong to 2 separate mastitis cases if time between 2 recordings was more than 8 d (IDF, 1987).

Design of Validation Test

Two approaches were used for validation of OCC model alerts: a direct comparison of treated and untreated cows and a more traditional approach with calculation of model performance parameters. The basis of both approaches was the time-window analysis, which is used to link model alerts with events (Sherlock et al., 2008). In our case, the events were defined by recorded mastitis treatments. The present validation methods

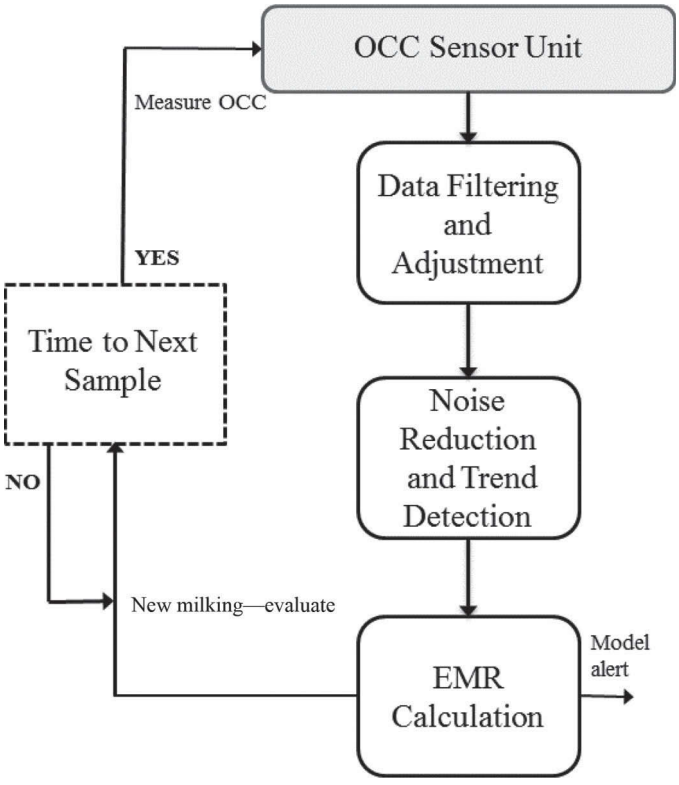

Figure 1. Flowchart of the online cell count (OCC) detection algorithm. Raw OCC data enters the "Data Filtering and Adjustment" module where data are filtered and merged with cow data and historical information. The processed data are then transferred to the second module, where noise is reduced and any trend is detected. Based on the OCC level and trend, elevated mastitis risk (EMR) is calculated, and a possible new mastitis alert is determined in the third module. In the last module, time to next OCC measurement is determined based on EMR values and systemic cow factors.

were adopted from Kamphuis et al. (2013) with some modifications. This setup is illustrated in Figure 2, with a study period of 21 d. The time window, in which alerts were detected, was defined as 4 d and began 2 d before the IMI episode. In our case, a reference mastitis treatment could only be assigned to a date (treatment time was not recorded) so that an IMI episode was defined to last 2 d. A mastitis treatment usually terminates an IMI episode. Therefore, the time window began 2 d before the IMI episode. It was assumed that not all IMI cases were treated (Vaarst et al., 2002; Wolff et al., 2012). Such cases would likely be detected by the detection algorithm but would appear as false positive (**FP**) alerts in the chosen validation setup. However, it was not possible to confirm whether such alerts were true or FP alerts. Thus, a study period of only 21 d was used in the current study, including 15 d before to 5 d after a mastitis treatment. Some cows were treated in early lactation and did not have 15 DIM before the treatment date. For those cows, number of days before treatment could be less than 15 d.

Direct Comparison of Treated and Contemporary Untreated Cows. Using recorded mastitis treatments as gold standard for mastitis are unlikely to capture all IMI cases as detected by the algorithm. The following procedure may therefore give an indication of the suitability of the used gold standard for the present study. A clear discrimination between healthy (untreated) and sick (treated) cows was expected for the calculated EMR values (i.e., based on OCC measurements). To test this, cows treated for mastitis were matched and compared with presumably healthy cows not treated in the same lactation. To do so, all cows within the herd were sorted according to parity into 2 groups, first parity and second parity or greater. Cows within the 2 groups where then sorted and matched by calving date. Each treated cow was matched with 2 untreated cows that calved before and after the treated cow. Any untreated cow calving before a treated cow could only be matched to 1 treated cow, and the same was true for an untreated cow calving after a treated cow. The study period for treated and untreated cows was defined as to cover −15 to 15 d after the date of treatment (31 d) to enable monitoring of the recovery period. As an example, if a cow was treated at 34 DIM, the study period will be 19 to 49 DIM. The same DIM was then extracted from the matching untreated cows calving before and after the treated cow, respectively. The mean EMR values for each group of cows were calculated for each day in the study period (−15 to 15 d). Finally, differences in EMR values (least squares means) for each day in the study period between the 2 groups were tested using *t*-tests as implemented in the HPMIXED procedure in SAS (version 9.3, SAS Institute Inc.), using cow within herd and lactation as random effects. This analysis was performed on data from the 6 commercial farms.

Although new alerts were intended to detect new infections, cows may not have been completely healthy before the treatments used for validation in the current study; IMI may linger (be subclinical) for some time before clinical signs are detected. As a consequence, a new alert may have been issued several days or weeks before treatment, resulting in only IMI alerts being present in the 4-d time window used for model evaluation. Using the above data set with treated and con-

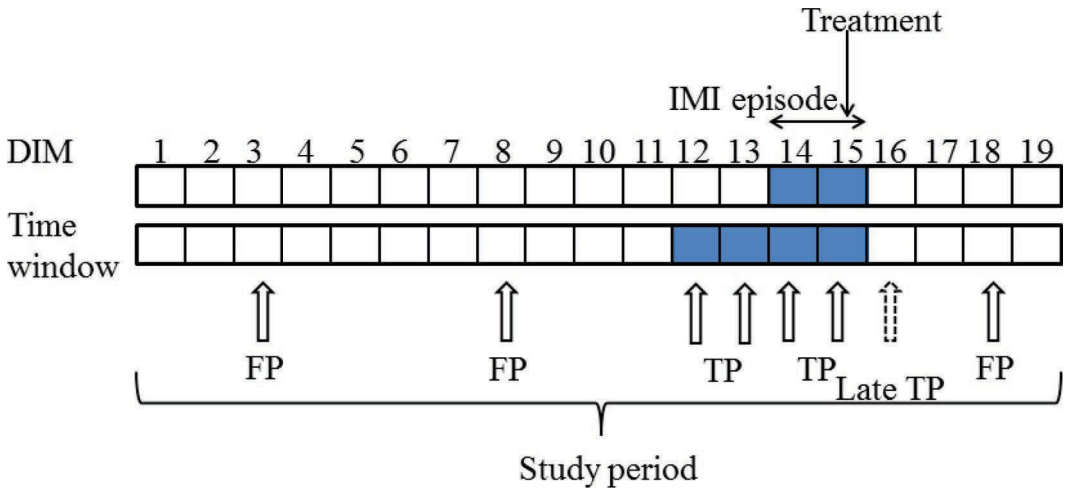

Figure 2. Application of mastitis treatments (downward arrow) to identify episodes of IMI and the use of time windows to link model alerts with IMI episodes. Upward arrows indicate model alerts and model performance is indicated by true positive (TP) alerts and false positive (FP) alerts. False negative alerts (not shown) are present if no TP alerts are present within the time-window. Modified from Kamphuis et al. (2013). Color version available online.

temporary untreated cows, the expected distribution of weekly new and IMI alert rates for both groups were calculated as number of alerts divided by the number of milkings. Eleven weekly intervals were defined for the treated cows, where wk 0 was defined as the week ending with the time of treatment, wk 1 and 2 as the subsequent weeks, and wk −8 to −1 as the weeks before wk 0. Similar intervals were defined for the contemporary untreated cows.

Performance Parameters. True positive (**TP**) model alerts were defined as new (or IMI) alerts issued within the time window. All new alerts issued outside the time window were defined as FP alerts. If no model alerts were issued within the time window, it was considered a false negative (**FN**) alert. Model performance was primarily assessed using new alerts. Traditionally, sensitivity (**SN**) and SP are used for evaluation of detection model performance. With the current validation setup it was not possible to calculate SP directly because untreated cows could not be confirmed as being healthy; however, SN, which refers to the proportion of IMI episodes detected by the OCC model using new alerts, was calculated as (Kamphuis et al., 2013):

$$\text{SN\_new (\%)} = [\text{TP count}/(\text{TP count} + \text{FN count})] \times 100\%.$$

From a practical point of view, farmers may be more interested in the proportion of new alerts that can be associated with an IMI episode. This defines success rate (**SR**) as proposed by Sherlock et al. (2008):

$$\text{SR\_new (\%)} = [\text{TP count}/(\text{TP count} + \text{FP count})] \times 100\%.$$

Finally, a measure for the number of FP alerts per 1,000 milkings (**FAR1000_new**) was defined as (Sherlock et al., 2008):

$$\text{FAR1000\_new} = 1{,}000 \times \text{FP count}/\text{total cow milkings},$$

where the number of cow milkings was calculated for the study period among the treated cows only because we could not confirm or reject model alerts for the untreated cows. According to Hogeveen et al. (2010) FAR1000_new can be converted to SP as:

$$\text{SP\_new (\%)} = 100\% - \text{FAR1000\_new}/10.$$

Finally, the proportion of IMI episodes detected by the algorithm, when IMI alerts were used, was calculated. In contrast to new alerts, several IMI alerts can be present for each IMI case, which complicates the assessment of TP and FP alerts. Thus, FP alerts were ignored and TP alerts were then defined as situations where at least 1 IMI alert was detected within the time window, and FN alerts as cases were no IMI alerts were detected within the time-window. The proportion was calculated as:

$$\text{SN\_IMI (\%)} = [\text{TP\_IMI count}/(\text{TP\_IMI count} + \text{FN\_IMI count})] \times 100\%.$$

Validation Scenarios with Reduced Sampling Intensity. The OCC detection algorithm was validated using the full sampling scenario (**ALL**) and 6 reduced sampling scenarios 3 fixed (**FIX**) and 2 dynamic (**DYN**), to assess the effect of decreasing sampling intensity. Scenario ALL included all possible OCC measurements in the model. In scenarios FIX_24 and FIX_36, sampling intensities were reduced to 1 sample per 24 or 36 h, respectively, and in scenarios DYN_24, DYN_36, DYN_48, and DYN_72, dynamically adjusted sampling schemes were employed using the TNS module with default time between samples of 24, 36, 48, and 72 h, respectively.

RESULTS

Assessment of Raw Data Quality

Results from comparison of OCC measurements with laboratory-based SCC are shown in Figure 3 as average R^2 values in 2011 for each OCC unit. The average R^2 value across herds and OCC units was 0.86 and values from the individual units ranged from 0.71 to 0.93. In herd G (DCRC), AMS 3, which showed a lower R^2 value compared with the AMS 1 and 2, was dedicated to milking Jersey cows. In herd B, AMS 2 was not running before 2012, and in herd C both AMS units were not used in 2011. The results were based on 2 to 12 test dates in the commercial herds and 46 test dates at DCRC.

Model Dynamics

The OCC model reacts to changes in OCC measurements. Following this, a new IMI case was characterized by a rapid increase in OCC values (Figure 4a), which converts to increasing EMR values. In the shown example, a new alert was issued 2 d before the cow was treated, indicating that the algorithm reacted faster with a new alert than the dairy manager. In this case,

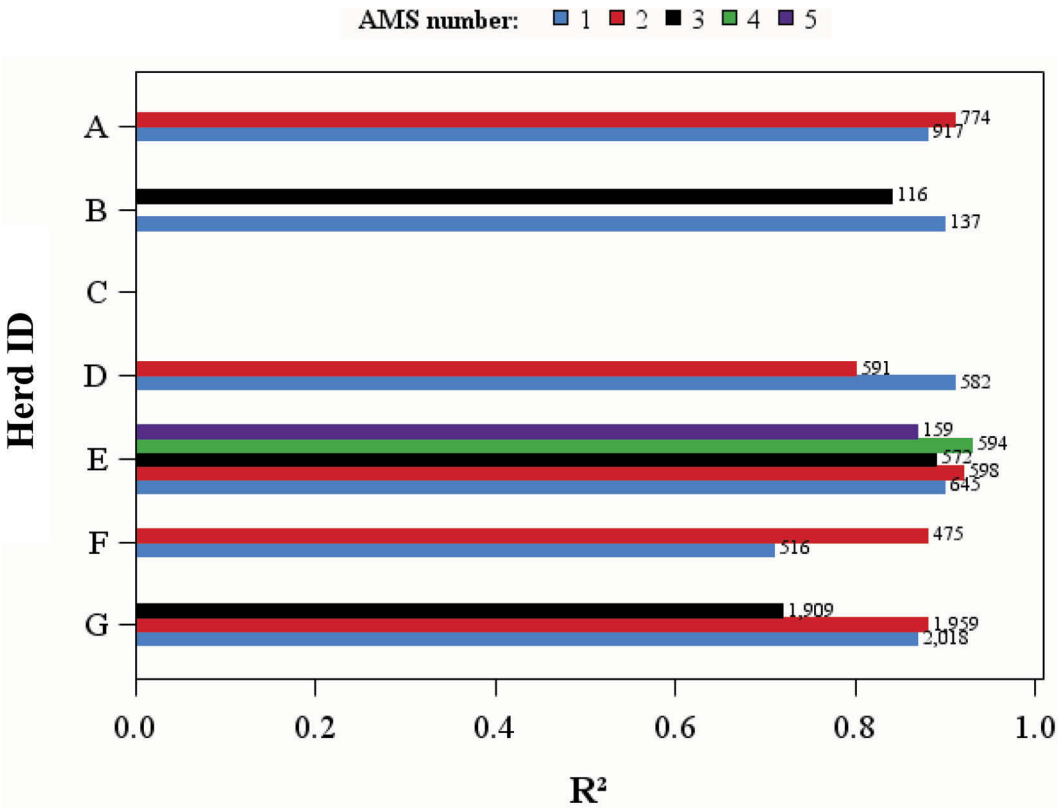

Figure 3. Comparison of online cell count (OCC) and SCC values shown as R^2 values for each automatic milking system (AMS) unit in 7 herds from herd control dates in 2011. Herd B had no OCC data available from AMS unit 2, and in herd C no OCC data were available for 2011. Numbers of samples for each AMS unit within herd are shown next to bars. Color version available online.

the treatment was successful and OCC values returned to preinfection level about 2 wk later. In Figure 4b, an example of a cow likely to suffer from persistent IMI is shown. In such cases, the overall OCC level remains high, sudden increases are less pronounced, and over time a wave-like OCC pattern can be observed. A new alert was issued in the beginning of the lactation and again 19 and 33 d later. This cow was treated twice without success and IMI alerts continued to be issued throughout the rest of data collection period; eventually this cow was put on the cull list.

Model Validation Using Data from Commercial Herds

Comparison of Treated and Untreated Cows. The mean EMR values per day relative to day of treatment are shown in Figure 5. For both first parity and cows in later parities, the differences in mean EMR values per day for treated and untreated cows were significant ($P < 0.05$) for all days in the study period (−15 to 15 d). However, the estimated EMR value increased from d −10 to −3, peaked with EMR of 0.82 (parity 1) or 0.76 (older) at d 0, and then declined. Comparison between EMR values at d −15 and 15 showed a significantly ($P = 0.002$) higher EMR level after treatment for older cows and no difference for parity 1 cows, indicating IMI are more likely to persist in older cows compared with parity 1 cows. The average recovery period after treatment was approximately 7 d for both parity groups. Low, stable mean EMR values were a common characteristic for the untreated cows from both parity groups. For first-parity cows these were around 0.10 and for older cows the mean EMR values were around 0.18.

Figure 6a shows the distribution of new alert rates (number of alerts relative to number of milkings) per week relative to week of treatment. From 2 to 8 wk before the week of treatment the new alert rates for untreated cows were only slightly lower than new alert rates for the treated cows, around 0.30 and 0.52%, respectively. At the time around treatment the difference increased dramatically, with a 4.6 times higher new alert rate for the treated compared with the untreated cows. A similar pattern was observed for the distribution of IMI alerts (Figure 6b) but, in contrast to the distribution of new alerts, the IMI alert rate stayed

high for the treated cows during the week after treatment. Also, the IMI alert rates were much higher than the new alert rates because of differences in the definition of the 2 types of alerts.

Sampling Statistics for Validation Scenarios. The full sampling scenario (ALL) was compared with 6 reduced sampling scenarios (Table 1). In scenario ALL, all milkings were requested for EMR calculations

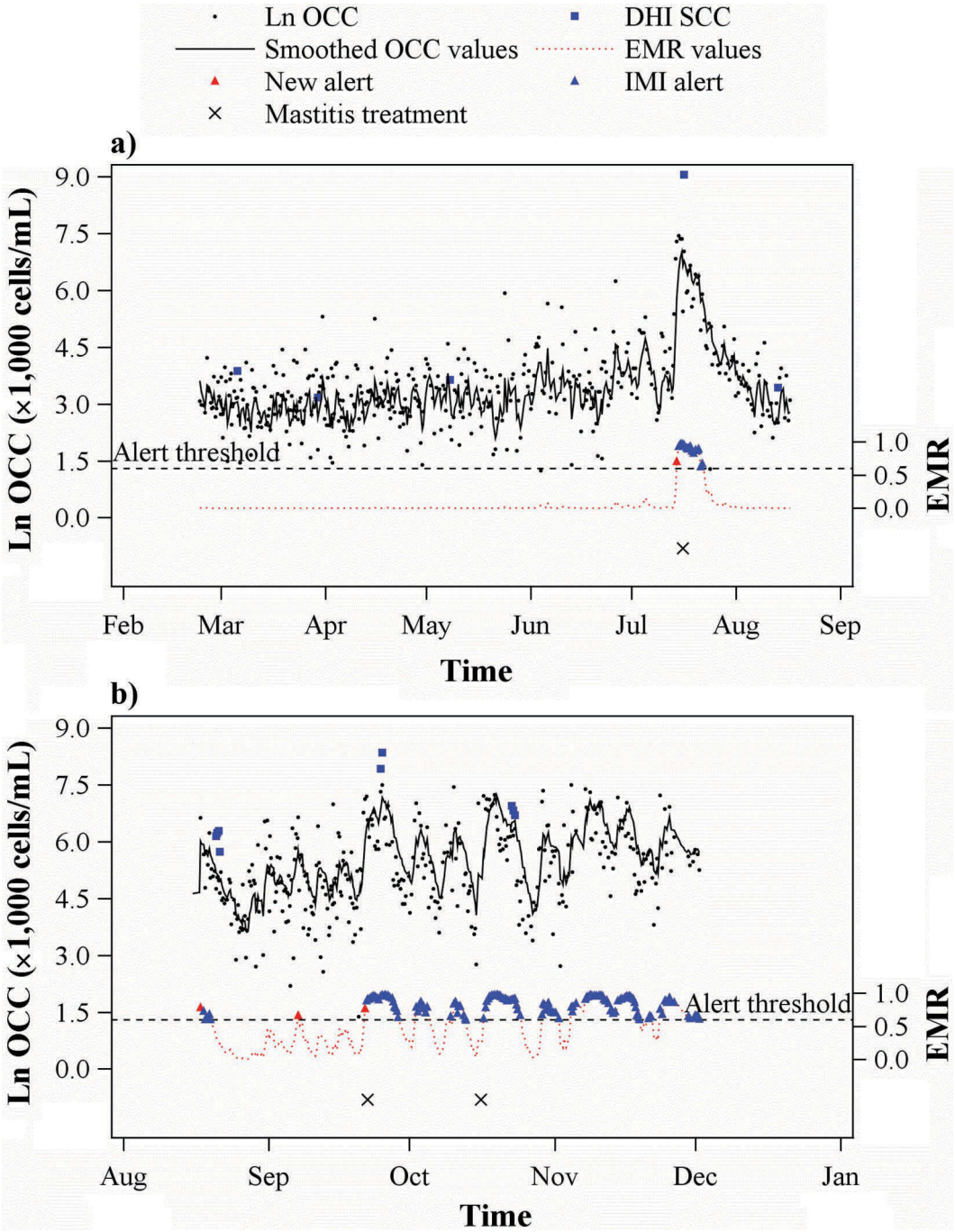

Figure 4. Examples of online cell count (OCC) model dynamics. The cow in (a) is a typical example of an acute IMI case indicated by a rapid increase in OCC values. This cow was treated successfully at July 16 and a new alert was issued on July 14. The cow in (b) is an example of a cow likely to suffer from persistent IMI. The cow was treated without success on September 22 and October 16. The first new alert was issued on August 17, the second on September 7, and the third on September 21. Continuous IMI alerts were seen continuously from September 21. Smoothed OCC values = OCC level and OCC trend from double exponential smoothing algorithm. Milk recording (DHI) SCC showed for comparison. EMR = elevated mastitis risk. Color version available online.

and 86.1% had valid OCC measurements and were used for calculation of EMR values (used milkings). The proportion of requested milkings in the remaining scenarios was lowest in the fixed sampling schemes, FIX_24 and FIX_36, and slightly larger (around 5 percentage points) for the dynamic scenarios. The proportion of used milkings was highest for the 3 fixed sampling schemes and lowest for DYN_72. Large differences were seen among the herds (results not shown), mainly caused by time periods with missing or invalid OCC measurements. The mean time between EMR calculations followed sampling intensity and was shortest for ALL and longest for DYN_72, the most extreme sampling scheme.

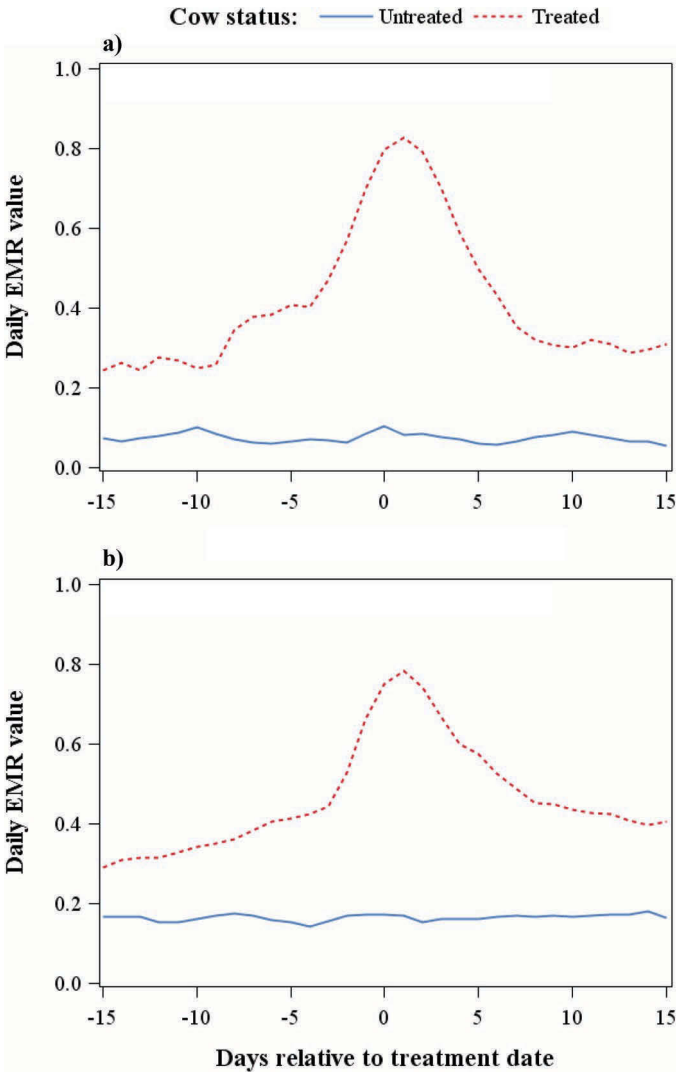

Figure 5. Least squares means of daily elevated mastitis risk (EMR) values relative to treatment day (d 0) for treated and contemporary untreated cows. Results are shown for first parity (a) and older cows (b). Color version available online.

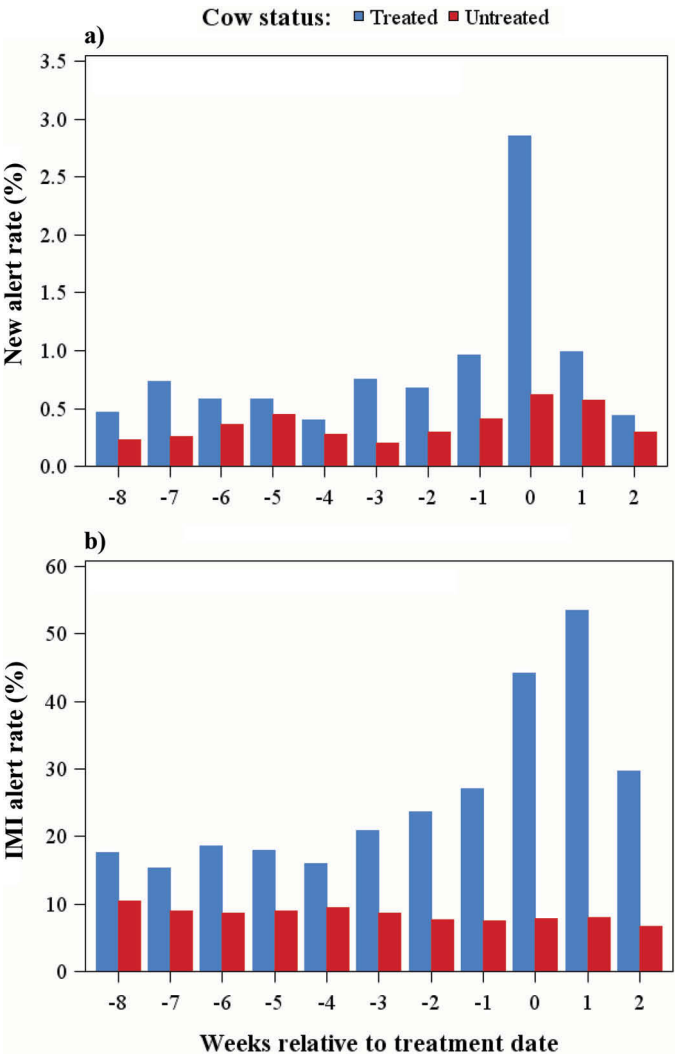

Figure 6. Distributions of new (a) and IMI (b) alert rates (%) per week for cows treated for mastitis and contemporary untreated cows. Weeks are shown relative to day of treatment, which is the last day of wk 0. Color version available online.

Performance of OCC Model. Performance parameters for the 7 validation scenarios are presented in Table 2. The sensitivity, SN_IMI, was calculated based on IMI alerts which indicate the model's ability to correctly identify IMI episodes given by mastitis treatments. In the default scenario ALL, 194 out of 218 (89.0%) IMI episodes were confirmed within the validation window. For scenarios DYN_24 and DYN_36, SN_IMI was only marginally lower. Differences (results not shown) among the herds were large for SN_IMI. In the worst case (DYN_72) 31.0% of IMI episodes in 1 herd were detected, whereas in the best case (DYN_24) 94.1% of IMI episodes in another herd were detected. Differences between herds may reflect management differences regarding treatment or differences in the

Table 1. Realized sampling intensity and acquired useful data from a base scenario (ALL) and 6 reduced sampling scenarios applied to 6 commercial herds; range of values from the individual herds is shown in brackets

	Validation scenario[2]						
Item	ALL	DYN_24	DYN_36	DYN_48	DYN_72	FIX_24	FIX_36
Requested milkings (%)	100	45.3 [41.6–48.4]	36.3 [32.5–40.2]	31.1 [26.3–35.4]	25.1 [19.5–28.8]	30.9 [30.2–33.1]	21.7 [21.2–23.4]
Used milkings[1] (%)	86.1 [81.2–91.9]	80.0 [68.0–89.4]	78.0 [64.5–88.4]	76.6 [61.5–87.9]	75.4 [60.2–88.4]	85.6 [80.9–91.6]	85.0 [80.6–91.2]
Mean time (h) between EMR calculations	10.2 [9.4–10.7]	24.4 [23.3–26.0]	31.1 [29.6–33.9]	37.0 [34.9–42.4]	46.5 [43.0–57.6]	28.9 [28.7–29.3]	41.1 [40.9–41.5]

[1]Proportion of requested milkings with valid online cell count measurement used for calculation of elevated mastitis risk (EMR).

[2]Dynamic (DYN) sampling scheme via the time next to sample module with default time between samples of x h; fixed (FIX) sampling scheme with 1 sample per x h.

presence of causative pathogens. Across all scenarios FAR1000_new was around 7 or less.

The proportion of correctly identified IMI episodes within the time-window when based on new alerts (SN_new) was highest for FIX_24, closely followed by the default scenario (ALL), and lowest for DYN_72. Both fixed sampling schemes were better than their dynamic counterparts. The 4 dynamic sampling schemes produced the lowest SN_new performance values.

The success rate, seen as the proportion of new alerts issued within the study period (SR_new) that could be matched with a mastitis treatment was highest for FIX_36 and closely followed by ALL and the other scenarios. For all performance parameters differences among the herds were large (Table 2).

DISCUSSION

The aims of our study were to develop and validate a model for detection and monitoring of IMI. Model output was based on OCC measurements from AMS incorporating a dynamic sampling scheme providing optimal balance between running cost and performance. These goals were achieved using an algorithm with a modular architecture-integrating data filtering, time-series dynamics, risk estimation, and dynamic sampling intensity control; however, performance parameters based on new alerts, compared with IMI alerts, were less satisfactory. The model responded to both slow and rapid increases in OCC level by issuing new and sustained alerts. The model enables the herd manager to react to mastitis cases by going through a short list of new alerts at any time and initiate necessary treatments and closely monitor the recovery phase.

Data Quality

Data used for the OCC model must first pass an effective filter so as to avoid data of poor quality. Missing or faulty measurements were likely caused by clogged

Table 2. Validation results for the online cell count detection model using different thinning schemes in 6 commercial herds[1]

	Validation scenario[3]						
Item[2]	ALL	DYN_24	DYN_36	DYN_48	DYN_72	FIX_24	FIX_36
FAR1000_new	6.5 [4.7–8.5]	6.6 [4.7–9.3]	7.0 [4.5–10.1]	7.7 [4.7–10.4]	7.1 [3.1–11.6]	6.9 [5.1–9.2]	6.5 [2.3–8.4]
SP_new (%)	99.4 [99.1–99.5]	99.3 [99.1–99.5]	99.3 [99.0–99.6]	99.2 [99.0–99.5]	99.3 [98.8–99.7]	99.3 [99.1–99.5]	99.4 [99.2–99.8]
SN_IMI (%)	89.0 [77.3–93.1]	88.5 [76.5–94.1]	88.1 [81.8–92.0]	64.2 [44.8–72.3]	54.6 [31.0–76.0]	86.7 [77.3–92.1]	80.7 [70.6–88.1]
SN_new (%)	42.7 [32.0–55.2]	39.5 [25.0–51.7]	39.0 [29.2–41.2]	32.1 [17.2–40.9]	28.0 [13.8–36.4]	43.1 [32.0–48.3]	42.2 [33.3–52.9]
SR_new (%)	56.7 [44.4–64.0]	54.4 [37.5–62.5]	52.8 [36.8–66.7]	45.8 [33.3–60.0]	44.2 [23.1–66.7]	55.6 [42.1–63.6]	56.8 [44.4–78.6]

[1]Values in brackets indicate the range of values from the individual herds. The total number of mastitis cases to be detected was 218 [17–101].

[2]FAR1000_new = false alert rate per 1,000 milkings based on new alerts; SP_new = specificity based on new alerts; SN_IMI = sensitivity based on IMI alerts (i.e., proportion of treated mastitis cases that appear on the IMI list); SN_new = sensitivity based on new alerts; SR_new = success rate based on new alerts.

[3]Dynamic (DYN) sampling scheme via the time next to sample module with default time between samples of x h; fixed (FIX) sampling scheme with 1 sample per x h.

tubes (i.e., clots from mastitis milk), camera failure, out of fluid, or other maintenance-requiring problems. This indicates that proper maintenance by both herd manager and service technician is important. Common for the nonmissing OCC measurements (nonzero values) not accepted by the filtering step were very low OCC values compared with the forecasted OCC values. Failure to remove such values could result in FP alerts. Whenever a faulty OCC measurement was detected, it would trigger the TNS module to request a new OCC sample at the first upcoming milking unless the dynamic sampling function was off. Measuring accuracy when compared with DHI SCC was generally high (mean $R^2 = 0.86$), which is comparable to laboratory-based result present by Ruegg et al. (2005), who found a correlation of 0.92 ($R^2 = 0.85$) and a nonsignificant different between $\log_{10}$(SCC) and $\log_{10}$(DCC). The latter is based on DeLaval Cell Counter (DCC, DeLaval); the OCC measuring unit uses similar technology.

Carryover has been found to vary widely between AMS installations of any make (up to 20%; Løvendahl and Bjerring, 2006; Løvendahl et al., 2010) and is complicated to adjust for because it is has a random nature. Carryover between cows was not accounted for in this version of the detection model and may theoretically cause some FP alerts. However, in a modified version of the OCC model, carryover was accounted for by simply requiring 2 consecutive EMR values above the EMR threshold before a new alert was issued (results not shown); but this approach did not improve model performance, so it was decided to ignore carryover issues in this version of the model. Because carryover also affects DHI herd testing samples, it should preferably be dealt with at the source (i.e., within the AMS system rather than by model adjustment).

Model Performance

Performance of the OCC model was investigated using 2 different approaches; first a comparison of EMR values between treated and untreated cows, and next an evaluation of SN, SP, and other performance parameters. The latter approach was extended with a data dilution study to investigate effects of reduced sampling intensity.

For the direct comparison of treated and untreated cows, treated cows were matched by parity and DIM and study periods were defined for the untreated cows. Although average EMR values were higher for treated cows for all days in the study period, only the treated cows showed a substantial peak before and around the time of treatment. Difference in OCC level between treated and untreated cows were expected and the results clearly reflect this. However, not all IMI cases are necessarily treated and recorded, as discussed by Vaarst et al. (2002) and Wolff et al. (2012). Thus, categorization of cows using recorded IMI treatments could make the differences between healthy cows and cows with mastitis less distinct. This was not the case in the present study, which showed that recorded mastitis treatments can be as useful as the gold standard for IMI when no better option is available.

An important performance characteristic of a mastitis-detection system is a low false alert rate (Mollenhorst et al., 2012). It was found that the FAR1000_new was well below the recommended maximum of 10 FP alerts per 1,000 milkings (e.g., Kamphuis et al., 2013). When FAR1000_new was converted to SP_new (Hogeveen et al., 2010), SP_new between 99.2 and 99.4% were achieved, which are above the limit of 99% as suggested by ISO (2007). Chagunda et al. (2006) also used historical data for model validation; using a similar approach they obtained an estimated SP of 99%, SN of 82%, and FAR1000 of 4.9. However, their approach has been criticized by Hogeveen et al. (2010) for excluding cows that had only somewhat elevated cell counts and thereby an unknown udder health status. This criticism can also be held against the current study and other studies with strict classifications of healthy and mastitic cows (e.g., Nielen et al., 1995; Norberg et al., 2004).

The OCC-detection model was able to confirm a high proportion of the recorded mastitis treatments used as gold standard in the present study (SN = 89.0% when all available data were used). This fulfills the recommendations to mastitis-detection systems and is higher than results from other studies based on data from commercial dairy farms (Kamphuis et al., 2010a,b; Mollenhorst et al., 2010). However, from a practical point of view, IMI alerts are not very useful because of multiple alerts over time (similar to using raw OCC data).

When results were based on new alerts only, SN_new were at 43.1 and 45.9%, much lower than the requested 80% (ISO, 2007; Mein and Rasmussen, 2008). However, in our case, 47.5% of all new alerts were observed from 60 to 5 d before the validation window. Including these alerts, an SN_new of 90.8% could be achieved. The reason for the widespread distribution of new alerts is purely speculative, but it is likely that persistent or untreated IMI were present before the defined time window or some IMI cases were subclinical some time before they became clinical and were eventually treated. Persistent infections may be untreatable and flare up at irregular intervals (Sears et al., 1990), causing multiple new alerts over time. Another possibility might be that the dairy manager was simply not aware of any IMI given the available IMI detection at the time of

treatment. Raw OCC data were available, but it is not confirmed that this data were used for detecting cows to be treated for mastitis in the commercial herds. In theory, new alerts are useful from a practical aspect, but it is clear from the present study that validation performance based on new alerts is poor when based on historical data. From a practical point of view, new alerts are ideal as an indicator for when to check a cow. However, from the present study it is unclear if the poor model performance when based on new alerts is because of poor model performance or because of the chosen validation setup and gold standard. This shows the need for further model validation, preferably a cross-sectional study in one or more herds.

Effects of Reduced Sampling Intensity

Another objective, but less important, in the development of the OCC detection model was reduction of running costs (i.e., fewer measurements) without compromising model performance. A feedback mechanism (TNS module) was implemented to determine when a new measurement was needed. The module could either use a fixed or a dynamic schedule. The fixed schemes had a minimum sampling interval of 24 or 36 h and the dynamic schemes had a default minimum interval of 24, 36, 48, or 72 h. The fixed schedules reduced sampling intensity to 30.9 and 21.7% for FIX_24 and FIX_36, respectively, of all milkings. Less reduction was obtained using the dynamic schedules, with sampling intensity going down to 47.7 and 40.0% for DYN_24 and DY_36, respectively. Extending default sampling intervals to 48 and 72 h still required 31.1 and 25.1% of the milkings, respectively for OCC measurement. The higher proportion of requested milkings in the 4 dynamic sampling schemes was caused by increased sampling intensity in early lactation and when EMR values above the EMR threshold were observed. It should also be noted that missing OCC measurements and invalid values automatically increased the number of requested milkings in the dynamic but not the fixed scenarios.

The scenarios FIX_24 and FIX_36 had the lowest sampling frequency and also the lowest performance compared with their dynamic counterparts. The dynamic sampling schemes are preferable because of their higher SN_IMI, especially if DYN_24 is used. Few other studies have investigated the effect of reduced sampling intensity of a mastitis indicator because running costs of some sensors (i.e., electric conductivity) used in AMS are insignificant. Running costs were also considered by Chagunda et al. (2006), who used a dynamic sampling scheme similar to the one presented here. They used 72% of the requested milkings, which is more than was used in DYN_24 in the present study. However, they only simulated 1 cow for approximately 40 d with 2 mastitis episodes which makes comparison difficult.

Our results showed that reduction of sampling frequency is almost possible without compromising model performance. Using all available samples, 89.0% of all IMI episodes were detected by the model. Using DYN_24 and DYN_36 the proportion was reduced to 88.5 and 88.1%, respectively; similarly, FAR1000_new was almost not affected with an increase of 0.1 and 0.5 for DYN_24 and DYN_36, respectively, compared with ALL. Extending the default sampling interval to 48 or 72 h had a pronounced negative effect on SN_IMI and the other performance indicators. Only FAR1000_new was less affected by the increasing interval. The 2 fixed scenarios, especially FIX_36 gave an unacceptable reduction in SN_IMI. For those reasons, a dynamic sampling scheme was preferred when reductions were requested. Of the dynamic sampling schemes, DYN_24 showed the least reduction in performance compared with scenario ALL. Thus, based on the results in the current study, we recommend a default sampling scheme of 1 sample per cow per day. Going beyond 36 h cannot be recommended because too much performance is lost. However, given the poor model performance when based on new alerts, further studies may be necessary to find the optimal setting for dynamic sampling scheme.

Choice of Gold Standard and Validation Method

A wide range of reference data or gold standards and another wide range of validation methods have been proposed in the literature (Hogeveen et al., 2010; Rutten et al., 2013). For the present study we were restricted by the available data (longitudinal study), as many were in previous studies (e.g., Chagunda et al. 2006). Therefore, our results are best compared with those using approximately similar methods and gold standards; focus on those parameters has been assumed to be the most relevant to herd owners and decision makers.

The gold standard used in the present study was records of treated mastitis cases, as was previously used in other studies (e.g., Chagunda et al. 2006). This put the focus on mastitis cases deemed serious enough to be treated (assumption). Some mastitis cases are not treated because of subjective decisions by the herd manager (Vaarst et al., 2002; Wolff et al., 2012). Despite this, a clear difference in EMR values was still apparent between treated and untreated cases. Some risk of bias exists because the herd manager had access to the OCC measurements and may have used that for the decision to treat a cow, so that the optimal blindfolded approach was not necessarily obtained. Alternatively, a cross-sectional study seems better, as all cows in the

herd (or group) can be checked and udder health status established. However, to get trustworthy results using such an approach the herd needs to be tested at several points in time, which is laborious and costly.

Online cell count or SCC from DHI sampling schemes may be considered a gold standard for mastitis on its own and is used worldwide as an indicator for mastitis (Laevens et al., 1997), and especially for subclinical mastitis (IDF, 2011). Schukken et al. (2003) suggested a threshold of 200,000 cells/mL to distinguish between infected and uninfected quarters. This cut-off value had an SN of approximately 75% and an SP of approximately 90%. However, in the present study SCC was found to be very similar to OCC, and both are measured in composite milk samples; thus, the possibility of comparing between quarters within cow was not available.

Another possible gold standard for defining sick cows is the use of quantitative PCR (Taponen et al., 2009) for identifying mastitis pathogens in milk samples. This method can be used to screen a dairy herd for mastitis (i.e., sample all cows) and is currently used for selecting cows for dry-cow therapy in Denmark (e.g., Cederlöf et al., 2012). However, the procedure is costly and time consuming because sampling needs to be done manually to eliminate carryover effects, and each herd is likely to be sampled more than once. Low mastitis prevalence may also cause difficulties in matching model alert and gold standard because of a possible low number of daily new alerts issued by the OCC model. Finally, the PCR test detects both viable and nonviable pathogens, is not straightforward to interpret the results (i.e., what threshold to use), and it is not clear whether a positive PCR results reflects an IMI.

The importance of time windows and study periods was discussed in other studies (Hogeveen et al., 2010). For the present study, the approaches of 2 present reports (Sherlock et al., 2008; Kamphuis et al., 2013) were adopted with some modifications. This gave a solid base for validation. However, it was discovered that many new and IMI alerts were raised several days before the mastitis cases were treated. When these cases were studied in detail (data not shown) they were characterized by higher EMR and fluctuating high OCC values, which were seen as signs of persistent infections. Therefore, a clear improvement of the model would be a way to define such cases and keep them separated from the group of noninfected cows.

Aspects Related to Herd Udder Health

The IMI-detection model presented in our study is, to our knowledge, the first utilizing frequently measured OCC measurements as mastitis indicator. The validation results based on historical data were satisfying but only when based on IMI alerts, which are of little practical relevance. Further information about the onset and recovery of a mastitis case could give a clearer match between model alerts and realized events. This would require and independent validation study where udder health status of all cows in a herd is closely monitored by humans at every milking during several days. More importantly, the effects of persistent infections need to be considered. To take the alerts and EMR to a simpler and more easily implementable level, another module needs to be built to assign cows to a few health classes so that standard operating procedures can be assigned to each health class to make it easier for herd managers to monitor overall udder health status changes in time. Rutten et al. (2013) introduced a 4-level framework for the use of sensor information in dairy farm management: level I, technique (sensor and any algorithm for producing sensor data); level II, data interpretation (detection algorithm); level III, integration of information (decision support and monitoring); and level IV, decision making (farmer autonomous). For the OCC-detection model, requirements for levels I and II are fulfilled with the current study. Using the OCC model, the dairy manager can rely on an automatic translation of the OCC measurements into a simple alert system to get a daily list of cows that needs to be checked. However, the alert system can be expanded (e.g., by assigning cows to different udder health classes) so that the OCC detection model will fulfill the level III requirements (see above) and will enable longitudinal monitoring of herd udder health; work is already in progress regarding this.

CONCLUSIONS

A simple continuous interpretation of OCC values via EMR values was used to issue new and IMI alerts. The new alerts are intended for notification of the herd manager of new mastitis cases and have an acceptable error rate (high SP). However, SN was low, indicating a mismatch between new alerts and treatments during the cause of an IMI case. The IMI alerts, on the other hand, are excellent for monitoring both infections and recovery phase and gave a much higher SN, above the recommended level, compared with new alerts. Finally, the results showed that it is possible to apply a dynamic sampling scheme without compromising model performance.

ACKNOWLEDGMENTS

This project was supported by the Danish Council for Independent Research, Technology and Production

Sciences (grant no. 10-080857; Copenhagen, Denmark), The Danish Milk Levy Fund (Aarhus, Denmark), and DeLaval (Tumba, Sweden). We gratefully acknowledge the Danish Cattle Federation (SEGES, Aarhus, Denmark) for providing us with data from DCRC, and the 6 commercial dairy farms for giving us access to OCC and herd data.

REFERENCES

Cederlöf, S. E., N. Toft, B. Aalbaek, and I. C. Klaas. 2012. Latent class analysis of the diagnostic characteristics of PCR and conventional bacteriological culture in diagnosing intramammary infections caused by Staphylococcus aureus in dairy cows at dry off. Acta Vet. Scand. 54:65.

Chagunda, M. G. G., N. C. Friggens, M. D. Rasmussen, and T. Larsen. 2006. A model for detection of individual cow mastitis based on an indicator measured in milk. J. Dairy Sci. 89:2980–2998.

Halasa, T., K. Huijps, O. Østerås, and H. Hogeveen. 2007. Economic effects of bovine mastitis and mastitis management: A review. Vet. Q. 29:18–31.

Hamann, J., and A. Zecconi. 1998. Evaluation of the electrical conductivity of milk as a mastitis indicator. Bull 334. Int. Dairy Fed., Brussels, Belgium.

Harmon, R. J. 1994. Physiology of mastitis and factors affecting somatic cell counts. J. Dairy Sci. 77:2103–2112.

Hogeveen, H., C. Kamphuis, W. Steeneveld, and H. Mollenhorst. 2010. Sensors and clinical mastitis—The quest for the perfect alert. Sensors (Basel) 10:7991–8009.

Hyndman, R. J., A. B. Koehler, J. K. Ord, and R. D. Snyder. 2008. Forecasting with Exponential Smoothing. Springer-Verlag, Berlin, Germany.

International Dairy Federation (IDF). 1987. Bovine mastitis. Definition and guidelines for diagnosis. Bull. Int. Dairy Fed. Int. Dairy Fed., Brussels, Belgium.

International Dairy Federation (IDF). 2011. Suggested interpretation of mastitis terminology. Bull. Int. Dairy Fed. Int. Dairy Fed., Brussels, Belgium.

ISO (International Organization for Standardization). 2007. Automatic milking systems—Requirements and testing. Annex C: Example of methods of evaluating detection systems for milk demand as abnormal due to blood or changes in homogeneity. ISO 20966:2007. ISO, Geneva, Switzerland.

Kamphuis, C., B. Dela Rue, G. Mein, and J. Jago. 2013. Development of protocols to evaluate inline mastitis detection systems. J. Dairy Sci. 96:4047–4058.

Kamphuis, C., H. Mollenhorst, A. Feelders, D. Pietersm, and H. Hogeveen. 2010a. Decision-tree induction to detect clinical mastitis with automatic milking. Comput. Electron. Agric. 70:60–68.

Kamphuis, C., H. Mollenhorst, J. A. P. Heesterbeek, and H. Hogeveen. 2010b. Detection of clinical mastitis with sensor data from automatic milking systems is improved using decision-tree induction. J. Dairy Sci. 93:3616–3627.

Laevens, H., H. Delyker, Y. H. Schukken, L. Meulemeester, R. Vandermeersch, E. Muelenaere, and A. Kruif. 1997. Influence of parity and stage of lactation on somatic cell count in bacteriologically negative cows. J. Dairy Sci. 80:3219–3226.

Løvendahl, P., and M. Bjerring. 2006. Detection of carryover in automated milk sampling equipment. J. Dairy Sci. 89:3645–3652.

Løvendahl, P., M. Bjerring, and T. Larsen. 2010. Determination of carryover in automated milking, recording and sampling systems using fluorescent tracers. Pages 147–152 in Proc. ICAR 37th Annual Meeting, Riga, Latvia. ICAR, Rome, Italy.

Mein, G. A., and M. D. Rasmussen. 2008. Performance evaluation of systems for automated monitoring of udder health: would the real gold standard please stand up? Pages 259–266 in Mastitis Control: From Science to Practice. T. J. G. M. Lam ed. Wageningen Academic Publishers, Wageningen, the Netherlands.

Mollenhorst, H., L. J. Rijkaart, and H. Hogeveen. 2012. Mastitis alert preferences of farmers milking with automatic milking systems. J. Dairy Sci. 95:2523–2530.

Mollenhorst, H., P. P. J. van der Tol, and H. Hogeveen. 2010. Somatic cell count assessment at the quarter or cow milking level. J. Dairy Sci. 93:3358–3364.

Nielen, M., Y. H. Schukken, A. Brand, and S. Haring. 1995. Comparison of analysis techniques for on-line detection of clinical mastitis. J. Dairy Sci. 78:1050–1061.

Norberg, E., H. Hogeveen, I. R. Korsgaard, N. C. Friggens, K. H. M. N. Sloth, and P. Løvendahl. 2004. Electrical conductivity of milk: Ability to predict mastitis status. J. Dairy Sci. 87:1099–1107.

Ruegg, P. L., C. Hulland, and B. Reith. 2005. Performance of the direct cell counter used on milk samples obtained from fresh cows. Pages 291–292 in 44th Natl. Mastitis Counc. Annual Meet. Proc., Orlando, FL. Natl. Mastitis Counc. Inc., Madison, WI.

Rutten, C. J., A. G. J. Velthuis, W. Steenevel, and H. Hogeveen. 2013. Sensors to support health management on dairy farms. J. Dairy Sci. 96:1928–1952.

Schukken, Y. H., D. J. Wilson, F. Welcome, L. Garrison-Tikofsky, and R. N. Gonzalez. 2003. Monitoring udder health and milk quality using somatic cell counts. Vet. Res. 34:579–596.

Sears, P. M., B. S. Smith, P. B. English, P. S. Herer, and R. N. Gonzalez. 1990. Shedding pattern of *Staphylococcus aureus* from bovine intramammary infections. J. Dairy Sci. 73:2785–2789.

Sherlock, R., H. Hogeveen, G. Mein, and M. D. Rasmussen. 2008. Performance evaluation of systems for automated monitoring of udder health: Analytical issues and guidelines. Pages 275–281 in Mastitis Control: From Science to Practice. T. J. G. M. Lam ed. Wageningen Academic Publishers, Wageningen, the Netherlands.

Taponen, S., L. Salmikivi, H. Simojoki, M. T. Koskinen, and S. Pyörälä. 2009. Real-time polymerase chain reaction-based identification of bacteria in milk samples from bovine clinical mastitis with no growth in conventional culturing. J. Dairy Sci. 92:2610–2617.

Vaarst, M., B. Paarup-Laursen, H. Houe, C. Fossing, and H. J. Andersen. 2002. Farmer's choice of medical treatment of mastitis in Danish dairy herds based on qualitative research interviews. J. Dairy Sci. 85:992–1001.

Wolff, C., M. Espetvedt, A.-K. Lind, S. Rintakoski, A. Egenvall, A. Lindberg, and U. Emanuelson. 2012. Completeness of the disease recording systems for dairy cows in Denmark, Finland, Norway and Sweden with special reference to clinical mastitis. BMC Vet. Res. 8:131.