Quantifying degree of mastitis from common trends in a panel of indicators for mastitis in dairy cows

S. Højsgaard¹ and N. C. Friggens²

Faculty of Agricultural Sciences, Aarhus University, Tjele, DK-8830, Denmark

ABSTRACT

This paper has 2 objectives. First, it argues that it is beneficial to regard degree of infection with respect to mastitis as a latent quantity varying continuously from 0 (truly healthy) to 1 (full-blown clinical mastitis). This quantity is denoted as degree of infection (DOI). The DOI is based on extracting common characteristics from a panel of indicators measured repeatedly over time. The indicators used in this paper are electrical conductivity (EC), somatic cell count (SCC), and the immune response related enzyme lactate dehydrogenase (LDH). Second, this paper presents a statistical model for such data and a corresponding method for estimating the DOI from a panel of indicators. An empirical proof of concept is provided. Using DOI, there was a significant difference between the DOI of mastitic and healthy control cows beginning 5 d before the mastitic cows were treated for mastitis.

Key words: mastitis, health status, dynamical factor analysis, state space model

INTRODUCTION

Given the available knowledge on mechanisms of infection and development of mastitis, it is generally accepted that a definition of mastitis as a categorical condition with only 2 (or 3) states, healthy and sick (often subdivided into subclinical and clinical), is a crude simplification (Sloth et al., 2003; Green et al., 2004; Detilleux et al., 2006). Moreover, technological on-farm advances in milking parlors mean that a range of mastitis indicators can now be routinely and repeatedly measured with a high level of accuracy (Chagunda et al., 2006a; Kamphuis et al., 2008b). Combining these aspects, it seems logical to develop systems that estimate and predict the degree of infection of the cow with respect to mastitis. Degree of infection (DOI) can

be envisaged as a continuous measure ranging from 0 (truly healthy) to 1 (full-blown clinical mastitis; Friggens et al., 2007).

Electrical conductivity (EC), SCC, and, more recently, immune response related enzymes such as lactate dehydrogenase (LDH), are all accepted as indicators of mastitis (Sloth et al., 2003; Norberg et al., 2004; Chagunda et al., 2006b). In conjunction with mastitis, increases in EC are linked to decreased integrity of the tight junctions between alveolar cells. The increase in immune cells (primarily leucocytes) in response to the bacterial infection is measured by SCC, and LDH is implicated in the inflammatory response. These 3 indicators thus relate to different aspects of mastitis. Because they are all affected by the bacterial challenge in the mammary gland (i.e., the underlying infection process), they are expected to exhibit a common trend (Figure 1). This suggests that they are all manifestations of a latent quantity that is not directly measured, namely DOI.

The aim of this study was to provide proof of concept for the DOI approach by developing a simple method to generate DOI and evaluating it relative to clinical observations of mastitis.

MATERIALS AND METHODS

Source Data

The data used to parameterize and test the model described below were created by a period of intense monitoring and milk sampling on 1 farm, the Danish Cattle Research Centre (Tjele, Denmark). All cows were milked with an automatic milking system (DeLaval VMS, Tumba, Sweden; 3 units, on average 2.3 ± 0.83 milkings/cow per day) in which the milk yield was automatically recorded. The data used were collected in the period between April 2003 and September 2006. The cows were of 3 breeds (Danish Holstein = 123, Danish Red = 130, and Jersey = 79) and were loose housed in 1 barn. The total number of lactations in first, second, and greater parities was 246, 179, and 177, respectively. Electrical conductivity was measured in the milking cluster for each quarter (DeLaval), and the interquartile ratio was calculated (Norberg et al.,

Received June 2, 2009.

Accepted October 27, 2009.

¹Corresponding author: sorenh@agrsci.dk

²Present address: UMR Physiologie de la Nutrition et Alimentation, INRA-AgroParisTech, 16 rue Claude Bernard, 75231 Paris Cedex 05, France.

Figure 1. Average values of log SCC (cells/mL), lactate dehydrogenase (LDH; $\mu\text{mol}/\text{min}$ per liter), and electrical conductivity (EC; mS) plotted against time from calving for the control cows.

2004). Proportional samples of composite milk (all quarters combined) were collected automatically from each cow at each milking. The automatic milk sampling system was emptied in the morning and in the afternoon. Milk samples were kept at 4°C until laboratory analysis, which was done within 24 h of sampling.

Milk samples were analyzed for LDH and SCC. Samples were analyzed for LDH activity ($\mu\text{mol}/\text{min}$ per liter) using a fluorometer (Fluostar, BMG Labtech, Offenburg, Germany) as described by Larsen (2005). For this experiment, milk samples were taken for bacteriology each time the veterinarian diagnosed a mastitis; the resulting proportions of the different pathogen types were typical of Danish dairy farms (36% *Staphylococcus aureus*, 25% CNS, 10% *Streptococcus uberis*, 5% *Escherichia coli*, 5% *Streptococcus dysgalactiae*, and 13% no culture). The SCC was measured using Fossomatic 5000 automatic equipment (Foss Electric, Hillerød, Denmark). The average frequencies of LDH and SCC measurements in the data set were 0.9 and 2/d, respectively.

The data used in the analysis were as follows. The SCC variable was the logarithm of the SCC (cells/mL). The EC variable was the logarithm of the logarithm of the interquartile ratio of EC (mS) as calculated by Norberg et al. (2004). This double logarithmic transformation was used to make the distribution of EC

approximately symmetrical. Finally, the LDH variable was the activity of LDH in milk ($\mu\text{mol}/\text{min}$ per liter).

Modeling Framework

The general framework was a panel of indicators that were recorded at typically nonequidistant times t_j on a single individual (a cow). Let $y(t_j) = [y^1(t_j), \dots, y^3(t_j)]$ denote the panel of indicators at time t_j . Typically, only some of the indicators were recorded at a given time point, so some components of $y(t_j)$ would often be missing. For a healthy cow, each indicator $y^k(t_j)$ was assumed to be composed of slowly varying, cow-specific level $\beta^k(t_j)$, a more rapidly fluctuating, short-term component $r^k(t_j)$, and an error term $v^k(t_j)$, such that

$$y^k(t_j) = \beta^k(t_j) + r^k(t_j) + v^k(t_j). \quad [1]$$

Onset of a disease (mastitis in this study) tended to generate a rapid change in $r^k(t_j)$. Each of the indicators (LDH, SCC, and EC) can be regarded as manifestations of the DOI. To make this assumption operational, it was assumed that each of these indicators was linearly related to a latent variable (i.e., that there was a common pattern in these components) that, for reasons explained later, was denoted by $\text{DOI}_{\text{free}}(t_j)$. This leads to the model assumption

$$y^k(t_j) = \beta^k(t_j) + \lambda^k \text{DOIfree}(t_j) + v^k(t_j), \quad [2]$$

where λ^k is proportionality constant defining the amount of change in indicator $y^k(t_j)$ when $\text{DOIfree}(t_j)$ changed 1 unit, and $v^k(t_j)$ denotes residual variation from this linear relationship. Hence, the slowly varying components $r^k(t_j)$ in equation [1] are in equation [2] assumed to have a common form across all indicators, namely $\lambda^k \text{DOIfree}(t_j)$. A method for estimating $\text{DOIfree}(t_j)$ is presented below.

Transformation to the Unit Interval—DOI

The quantity $\text{DOIfree}(t_j)$ given in equation [2] varies freely and, when $\lambda^k > 0$, high values of the indicators [relative to the baseline level $\beta^k(t_j)$] correspond to high values of $\text{DOIfree}(t_j)$ and vice versa. Therefore, $\text{DOIfree}(t_j)$ is in itself a measure of DOI that motivates the name for this quantity. It is more convenient to represent the DOI as a continuum on the interval [0;1], where 0 corresponds to no infection and 1 to full-blown clinical mastitis. The approach taken here is by using a sigmoid transformation of $\text{DOIfree}(t_j)$ to obtain the DOI:

$$\text{DOI}(t_j) = 1 / \{1 + \exp[-a - b \text{DOIfree}(t_j)]\}. \quad [3]$$

The parameters a and b control the horizontal displacement and the steepness of the transition from 0 to 1 and can be chosen arbitrarily. However, in practice they must be chosen so as to reflect the range of variation in DOIfree . The parameters were estimated using logistic regression on a subset of data from cows that had a recorded mastitis treatment using data in the time window ± 30 d from treatment and assuming the cows were not healthy in the period ± 10 d.

Factor Analytical State Space Model

Collecting all indicators and related quantities in equation [2] into vectors gives

$$y(t_j) = \beta(t_j) + \Lambda \text{DOIfree}(t_j) + v(t_j), \quad [4]$$

where Λ is a 3×1 matrix of regression coefficients and $v(t_j)$ is the vector of residual terms. It was assumed that $v(t_j) \sim N(0, \Psi)$, where Ψ is a diagonal matrix so that the residuals $v(t_j)$ were uncorrelated. The interpretation of uncorrelated residuals is that the indicators are independent conditional on (after adjusting for) $\text{DOIfree}(t_j)$. Notice that for known $\beta(t_j)$, equation [2] has the form of a factor analytical model (see, for example, Seber, 1984, page 213). However, in classical factor analysis,

the dependency of measurements over time is ignored and this would be inappropriate in this setting. An alternative approach to analyzing the multivariate time series in equation [4] is by dynamic factor analysis (see, for example, Zuur et al., 2003), which has inspired the approach taken here.

The slowly varying levels $\beta^k(t_j)$ are modeled as

$$\beta^k(t_j) = \beta^k(t_{j-1}) + w(t_j), \quad [5]$$

where $w(t_j) \sim N(0, W)$, where W is the variance.

Equation [5] indicates that the change in level $\beta^k(t_j) - \beta^k(t_{j-1})$ from t_{j-1} to t_j is a random quantity with zero mean, and that this quantity is the same for all indicators. Alternative ways of modeling the slow-moving trend can be considered (see the Discussion section).

Time dependency was also put into the model for $\text{DOIfree}(t_j)$. Specifically, it was assumed that

$$\begin{aligned} \text{DOIfree}(t_j) = & [\text{DOIfree}(t_{j-1}) - \text{DOIfree}(t_{j-2})] \\ & + \text{DOIfree}(t_{j-1}) + u(t_j), \end{aligned} \quad [6]$$

where $u(t_j) \sim N(0, U)$, where U is the variance.

The term $[\text{DOIfree}(t_{j-1}) - \text{DOIfree}(t_{j-2})]$ on the right side of equation [6] can be interpreted as the trend in the development of DOIfree at time t_{j-1} , and the subsequent term $\text{DOIfree}(t_{j-1})$ gives the level of DOIfree at time t_{j-1} . Therefore, equation [6] postulates a local linear development in DOIfree . The random perturbation $u(t_j)$ will allow the local linear development to vary over time.

Equations [4], [5], and [6] define a linear state-space model (Durbin and Koopman, 2001). The latent variables DOIfree can be estimated using, for example, a Kalman filter. Without going into details, it is noted that $\text{DOIfree}(t_j)$ becomes a weighted sum of the indicators $y(t_j)$ and $\text{DOIfree}(t_{j-1})$. For the present study, a bootstrap filter (Doucet et al., 2001, page 11) was used because it was easy to implement and modify such a filter in R (R Development Core Team, 2008).

Estimation of Static Parameters

The parameters Λ and Ψ were assumed to be population specific. These were estimated on the basis of a batch of control cows, that is, cows that were not treated for mastitis at any time in the current lactation and that had not had any systematic deviations in logSCC greater than 2.25 (according to Friggens et al., 2007). This was done by first subtracting averages from the indicators and then fitting the factor analysis model [4] to the residuals. The variance parameters W

Table 1. Distribution across breed and parity for true mastitis cases and true healthy control cases

Parity	Mastitis cases (n)			Healthy cases (n)		
	Danish Red	Holstein	Jersey	Danish Red	Holstein	Jersey
1	5	9	6	2	12	9
2	8	9	2	10	11	2
>3	8	4	7	12	4	9

and U in equations [5] and [6] were chosen by visual inspection of data to be $W = 0.0001$ and $U = 1.0$, so the standard deviation of W was 100 times smaller than the standard deviation of U . An alternative approach would have been to estimate these parameters as part of the model fitting, but this becomes cumbersome.

Test Data

Two subsets of cows were identified. These were cows that could, with reasonable certainty, be classified as either having clinical mastitis or being completely free of clinical mastitis (hereafter called healthy controls) within specific time windows. The basis for the health definitions was veterinary treatment records; in Denmark, only veterinarians are permitted to initiate treatments and they are legally required to record them. Cows were selected for veterinary examination by the farm staff on the basis of the cows showing physical signs of mastitis and from weekly visual inspection of milk. A mastitic cow was one that had a recorded veterinary treatment for mastitis with the day of treatment being designated days from mastitis = 0. Additionally, to account for the possibility of mastitis cases being overlooked or misdiagnosed, a refinement of the basic health definitions was made. When a positive systematic deviation in a cow's time series of logSCC greater than 2.25 occurred, the cow was assumed to have had mastitis and to not be healthy in the period -10 to $+10$ d relative to the SCC peak (as described by Friggens et al., 2007). The threshold of 2.25 was chosen because the average increase in logSCC for cows identified as having had mastitis on the basis of this threshold was the same as the average increase in logSCC for cows identified by the veterinarian as having had mastitis. If 2 records of veterinary treatment of mastitis occurred within 8 d, they were treated as 1 case (IDF, 1997). For each mastitis case, days from treatment was calculated. Generally speaking, number of days from treatment is positive and increases with time after the mastitis case. It is negative and more negative with time before the mastitis case. Between 2 mastitis cases, the change from positive days from treatment (after the previous case)

to negative days from treatment (before the next case) occurs midway between the 2 cases. In this study, a mastitis case window was defined as being from -30 to $+30$ d from treatment. For each mastitis case, matching control periods were chosen from the pool of cows of the same breed and lactation number class (1, 2, 3+) that had no veterinary treatment for mastitis in that lactation until at least 30 d after the DIM of the mastitis case being matched (Friggens et al., 2007). The chosen control period spanned the same DIM of the case being matched on. Mastitis cases occurring after 350 DIM were excluded. This resulted in a data set containing 58 true mastitis cases and 71 true healthy control cases. The distribution across breed and parity is shown in Table 1.

Statistical Analysis

At which time points the DOIfree profile of the true mastitic cows were significantly different from those of the true healthy cows was analyzed using a mixed model with breed, parity, days from treatment, and the interaction between days from treatment and case-type (healthy vs. mastitic) as fixed effects. The model used was

$$\text{DOIfree}_{ipd} = \text{breed}_i + \text{parity}_p + \text{day}_d + (\text{case} \times \text{day})_{kd} + A_{ip} + \varepsilon_{ipd}, \quad [7]$$

where i refers to a cow, p to the parity ($p = 1, 2, 3+$), d to day from treatment ($d = -30, \dots, 30$), and k to case type (healthy, mastitic). Further, A_{ip} and ε_{ipd} are random effects, where the former accounts for repeated measurements on the same cow. It was assumed that $A_{ip} \sim N(0, \sigma_A^2)$ and $\varepsilon_{ipd} \sim N(0, \sigma_\varepsilon^2)$, and all were independent. The model was fitted using PROC MIXED of SAS (version 9.2, SAS Institute, Cary, NC). The nominator degrees of freedom was estimated using Satterthwaite's method as implemented in PROC MIXED. Not including case type as a main effect allowed direct estimation, for each day, of the difference between healthy and mastitic cows and the significance of this difference.

RESULTS

Trends in Indicator Data

Figure 1 shows the average values of each indicator (SCC, LDH, and EC) plotted against time from calving for the control cows. A more detailed view, provided in Figure 2, shows profiles of SCC, LDH, and EC for each of 12 control cows that did not have and had never had mastitis (in that lactation). Data were centered around averages and standardized to unit variance and subsequently displaced to enhance visual interpretability.

Figures 1 and 2 support the hypothesis stipulated in connection with equation [5], namely that the indicators evolved roughly in parallel.

Figure 3 shows similar plots for cows that had all been treated for mastitis. The dashed vertical lines indicate the time of treatment for mastitis. Figure 3 also supports the assumption stipulated in connection with equation [2], namely that there was a common pattern in the behavior of the indicators, and in particular that short-term changes in the indicators tended to follow each other.

Figure 2. Indicators plotted against time from calving for 12 cows that do not have and have never had mastitis (in that lactation). The vertical dashed lines indicate the time of the control point on which cases and controls are matched. Indicators are log SCC (cells/mL), lactate dehydrogenase (LDH; $\mu\text{mol}/\text{min}$ per liter), and electrical conductivity (EC; mS). Data are centered around averages and standardized to unit variance and subsequently displaced to enhance visual interpretability.

Estimates of Static Parameters

The parameters Λ and Ψ were estimated from a batch of 20 cows that had not been treated for mastitis using the factor analysis model [2]. The estimated regression coefficients (Λ) for SCC, LDH, and EC were 0.891, 0.152, and 0.127, respectively, and the corresponding conditional variances (the diagonal entries of Ψ) were 0.004, 0.087, and 0.368, respectively. Stipulation of the parameters a and b in equation [3] was done as follows. A batch of 20 cows that had been treated for mastitis and for which

there were values of DOIfree in a time window of 30 d before treatment up to 30 d after treatment were used. Each cow was classified as being not healthy in a time window of -10 to $+10$ d around the day of treatment and healthy outside that window. Using these classifications as responses and the values DOIfree as explanatory variables, a logistic regression was performed because this had the same form as equation [3] and related this DOI function to the probability of a given level of DOIfree belonging to a nonhealthy cow. The resulting estimates were $a = -3.5$ and $b = 2.3$.

Figure 3. Indicators plotted against time from calving for 12 cow-lactations for which a mastitis treatment has taken place. Indicators are log SCC (cells/mL), lactate dehydrogenase (LDH; $\mu\text{mol}/\text{min}$ per liter), and electrical conductivity (EC; mS). Data are centered around averages and standardized to unit variance and subsequently displaced to enhance visual interpretability. The vertical dashed lines indicate the times of treatment.

DOI for Mastitic and Healthy Cows

Estimated values DOI obtained from equation [3] were subsequently smoothed. Figure 4 shows the DOI for 12 control cows (i.e., cows that had not been treated for mastitis), the same cows as presented in Figure 2. Figure 5 shows the DOI for 12 cows that had a mastitis treatment, the same cows as presented in Figure 3.

Evaluation by Comparison of Cases and Controls

Following the evaluation scheme established by Friggens et al. (2007), we estimated DOI for a group of case and control cows. Means of these (together with standard errors of mean) are shown in Figure 6. Data are

aligned according to day from treatment. Using model [7], the values of DOI_{free} were compared at each time point. The *P*-values are shown in Figure 7; the figure shows a significant difference in DOI_{free} in a window of 5 d around the treatment and control point.

DISCUSSION

The aim of this study was to provide proof that several time series measurements of udder health status indicators could be combined, using a simple method, to generate a DOI measure. The results of evaluating this DOI measure relative to clinical observations of mastitis indicate that the proposed method performs well in prediction of impending mastitis (Figure 6).

Figure 4. Degree of infection (DOI) and DOI_{free} plotted against time from calving for a group of control cows. The vertical dashed lines indicate the time of the control point on which cases and controls are matched.

This in itself is not completely unexpected because the method combines known indicators of mastitis; however, this DOI approach offers several benefits.

The DOI approach encompasses 2 conceptual issues that we believe are of importance for further research into quantifying animal health status. Infection has traditionally been viewed as a dichotomous quantity (healthy vs. sick). This is convenient from the point of view of clinical treatment and of measuring the efficacy of treatments with respect to clinical symptoms (i.e., specificity or sensitivity). However, this does not reflect the underlying etiology and development of the infection (Green et al., 2004; Detilleux et al., 2006). In this context, it has been proposed that infection be

regarded not as a dichotomous quantity but as a continuously varying quantity (Sloth et al., 2003; Friggens et al., 2007) that can be quantified by a measure such as DOI. The conversion from DOI_{free} to DOI using the sigmoid transformation in equation [3] is analogous to a logit function that could be applied in modeling the probability of a cow being sick as a function of DOI_{free}. In other words, if classical binary data for a population of cows were used to derive a function for the probability of a cow having clinical mastitis based on an indicator measure, it would have the same form as the DOI function.

The other conceptual issue implicit in the DOI approach concerns the nature of the true disease status

Figure 5. Degree of infection (DOI) and DOI_{free} plotted against time from calving for cows that had a mastitis treatment (the same cows as in Figure 3). The vertical dashed lines indicate the time of treatment.

Figure 6. Mean value of DOI free calculated for each day in a 60-d window around time of treatment for cases and time of control point for control cows. The bars indicate standard errors of means. DOI = degree of infection.

of an animal. This is a latent entity (i.e., we cannot directly measure it). We assume that a given indicator of mastitis is correlated in some way to the true disease status of the animal but, by definition, not perfectly. This issue lies behind the continuing debate on the relative strengths and weaknesses of different measures for detecting mastitis (Sloth et al., 2003; Cavero et al., 2008). Using the DOI approach, we model DOI as a latent entity and use the correlations between individual indicators to estimate DOI. Because both the approach to modeling DOI and the conceptual issues it addresses are general, it is expected that this approach is of value not only to mastitis but also to a wide range of other diseases.

The choices of the specific form of the slowly moving trends in equations [5] and [6] are open to criticism. The choices made here are intended to make relatively few assumptions, but clearly they can be refined. For example, the value of the indicators tends to decrease rapidly in a period immediately after calving and slowly increase throughout the rest of the lactation. This

natural behavior of the indicators can be embedded in equation [5]. The linearity assumption behind equation [6] is purely operational. However, any choice of model for these slowly moving components will tend to give a lag problem in the sense that the model will always lag behind the development in data, both when data goes up and down. This is the phenomenon causing the cases to end at a lower level than where they started in Figure 6.

The strong conceptual basis for the DOI approach also has practical benefits. One of the limitations within mastitis research has been the difficulty of obtaining reliable reference measures against which to test new mastitis management tools (Kamphuis et al., 2008a); this is particularly the case for subclinical mastitis. One way of dealing with this is to identify clinically sick animals and then examine them in a pre- and postinfection time window and compare them with healthy controls (e.g., Figure 6). This is a valuable approach for proof of concept but usually involves discarding the majority of the data set to find the relevant cases. The

Figure 7. Plot of *P*-values for the hypothesis of no difference between cases and controls against time under model [7] (see text). Values above 0.25 do not appear in the plot.

DOI method provides the means to generate a continuous reference measure that is applicable to all of the data. As modern on-farm technology is becoming more widespread, farmers are increasingly having access to continuous time series measurements of more than 1 mastitis indicator (Friggens et al., 2007; Kamphuis et al., 2008b). The DOI method allows these to be combined into 1 robust mastitis index. This method allows the different indicators to be measured with differing frequencies and deals with missing values (i.e., gaps in the time series).

The approach proposed here for using a panel of indicators for deriving DOI has the advantage of being extended to other indicators. The steps to be considered are 1) a model for normal behavior of each indicator and 2) a verification of a common trend in the deviations of the indicators from their normal levels. For example, it is known that onset of mastitis causes milk yield to decrease (Green et al., 2006). Therefore, including milk yield as an indicator can be envisioned.

This was not done in the present study because the use of milk yield raises several issues relating to feature extraction that we could reasonably ignore with respect to the other parameters used here. There is substantial variation among cows in milk yield, variation that is not related to degree of mastitis. The same applies to variation in yield through lactation. The mastitis-related variation in milk yield could be modeled as deviations from the normal milk yield curve, but this implies the use of an appropriate form of yield curve (rather than a simple linear function) for deriving the slowly varying trend of normal milk yield. There is also the issue of which feature of the deviations in milk yield is most indicative of mastitis (e.g., amplitude, rate of change, area under the curve, and so on). These are important issues for maximizing the information obtained from an indicator and minimizing inappropriate noise (Kamphuis et al., 2008a) that merit further study (both biologically and statistically).

The model presented here exploits the biological fact that the indicators are related to each other in the sense that when one indicator increases so do the others. The model fuses information from individual indicators to give a stronger common expression. The use of several indicators should provide robustness because of, for example, different pathogen profiles. This, however, remains to be tested. Furthermore, the indicators can be held against each other so that outliers or faulty measurement devices or both can be detected. It would be interesting to investigate the added value of including any additional indicator. For example, in a setting where LDH was measured at every milking, the additional value of including EC might not be overwhelming. However, measurements of EC are

essentially free after an initial overhead, whereas there is currently a cost associated with every single measurement of LDH.

It would be appealing to take a completely model-based approach to the problem discussed in this paper. One possible approach to analyzing the multivariate time series with a view to extracting common underlying characteristics is by dynamic factor analysis (see, for example, Zuur et al., 2003 for a thorough description). A dynamical factor analysis model will often be implemented by a Kalman filter (see, for example, Durbin and Koopman, 2001) or a related type of filter such as particle filters (see, for example, Doucet et al., 2001). One benefit of applying such a model would be the ability to associate measures of uncertainty with the DOI values and to make longer term predictions. However, parameter estimation in dynamic factor analysis models is in general not straightforward to implement. A more modest approach would have been to impose a proper time series structure on the latent variables $\text{DOIfree}(t_j)$. Whether such an approach is beneficial in relation to estimating the degree of infection is also a topic to be investigated.

CONCLUSIONS

Several time series measurements of udder health status indicators could be combined, using a simple method, to generate a DOI measure. The DOI measure detected significant differences between mastitic cows and healthy controls up to 5 d before clinical signs and treatment for mastitis. This method has general applicability to other diseases.

ACKNOWLEDGMENTS

We gratefully acknowledge the valuable contribution of the farm staff at the Danish Cattle Research Centre (Tjele, DK-8830, Denmark), Jens Clausen, Carsten Berthelsen, Peter Løvendahl, Jes Nielsen, and Connie Middelhed (all at the Faculty of Agricultural Sciences, Aarhus University, Tjele, DK-8830, Denmark) for their efforts in securing this massive milk data set. This study was part of the ILSORM project funded by the Danish Ministry of Food, Agriculture and Fisheries and the Danish Cattle Industry via Finance Committee Cattle.

REFERENCES

- Cavero, D., K. H. Tölle, C. Henze, C. Buxadé, and J. Krieter. 2008. Mastitis detection in dairy cows by application of neural networks. *Livest. Sci.* 114:280–286.
- Chagunda, M. G. G., N. C. Friggens, T. Larsen, and M. D. Rasmussen. 2006a. A biological model for detecting individual cow mastitis

- risk based on lactate dehydrogenase: Model description. *J. Dairy Sci.* 89:2980–2998.
- Chagunda, M. G. G., T. Larsen, M. Bjerring, and K. L. Ingvarsten. 2006b. L-Lactate dehydrogenase and N-acetyl- β -D-glucosaminidase activities in bovine milk as indicators of non-specific mastitis. *J. Dairy Res.* 73:431–440.
- Detilleux, J., F. Vangroenweghe, and C. Burvenich. 2006. Mathematical model of the acute inflammatory response to *Escherichia coli* in intramammary challenge. *J. Dairy Sci.* 89:3455–3465.
- Doucet, A., N. Freitas, and N. Gordon. 2001. *Sequential Monte Carlo Methods in Practice*. Springer, New York, NY.
- Durbin, J., and S. J. Koopman. 2001. *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford, UK.
- Friggens, N. C., M. G. G. Chagunda, M. Bjerring, C. Ridder, S. Højsgaard, and T. Larsen. 2007. Estimating degree of mastitis from time-series measurements in milk: A test of a model based on lactate dehydrogenase measurements. *J. Dairy Sci.* 90:5415–5427.
- Green, L. E., Y. H. Schukken, and M. J. Green. 2006. On distinguishing cause and consequence: Do high somatic cell counts lead to lower milk yield or does high milk yield lead to lower somatic cell count? *Prev. Vet. Med.* 76:74–89.
- Green, M. J., P. R. Burton, L. E. Green, Y. H. Schukken, A. J. Bradley, E. J. Peeler, and G. F. Medley. 2004. The use of Markov chain Monte Carlo for analysis of correlated binary data: Patterns of somatic cells in milk and the risk of clinical mastitis in dairy cows. *Prev. Vet. Med.* 64:157–174.
- IDF. 1997. Recommendations for presentation of mastitis-related data. Pages 16–21 in Bulletin 321. International Dairy Federation, Brussels, Belgium.
- Kamphuis, C., D. Pietersma, R. van der Tol, M. Wiedemann, and H. Hogeveen. 2008a. Using sensor data patterns from an automatic milking system to develop predictive variables for classifying mastitis and abnormal milk. *Comput. Electron. Agric.* 62:169–181.
- Kamphuis, C., R. Sherlock, J. Jago, G. Mein, and H. Hogeveen. 2008b. Automatic detection of clinical mastitis is improved by in-line monitoring of somatic cell count. *J. Dairy Sci.* 91:4560–4570.
- Larsen, T. 2005. Determination of lactate dehydrogenase (LDH) activity in milk by a fluorometric assay. *J. Dairy Res.* 72:209–216.
- Norberg, E., H. Hogeveen, I. R. Korsgaard, N. C. Friggens, K. H. M. N. Sloth, and P. L. Løvendahl. 2004. Electrical conductivity in milk: Ability to predict mastitis status. *J. Dairy Sci.* 87:1099–1107.
- R Development Core Team. 2008. R version 2.8.1. www.R-project.org.
- Seber, G. A. F. 1984. *Multivariate Observations*. John Wiley & Sons Inc., New York, NY.
- Sloth, K. H. M. N., N. C. Friggens, P. L. Løvendahl, P. H. Andersen, J. Jensen, and K. L. Ingvarsten. 2003. Potential for improving description of bovine udder health status by combined analysis of milk parameters. *J. Dairy Sci.* 86:1221–1232.
- Zuur, A. F., I. T. Jolliffe, R. Dekker, and J. J. Beukema. 2003. Estimating common trends in multivariate time series using dynamic factor analysis. *Environmetrics* 14:665–685.