



## Estimating the nonlinear association of online somatic cell count, lactate dehydrogenase, and electrical conductivity with milk yield

John Bonestroo,<sup>1,2,3\*</sup> Mariska van der Voort,<sup>2</sup> Nils Fall,<sup>2</sup> Ulf Emanuelson,<sup>2</sup> Ilka Christine Klaas,<sup>1</sup> and Henk Hogeveen<sup>3</sup>

<sup>1</sup>DeLaval International AB, Gustaf De Laval's väg 15, 147 21 Tumba, Sweden

<sup>2</sup>Swedish University of Agricultural Sciences, Department of Clinical Sciences, PO Box 7054, SE-750 07 Uppsala, Sweden

<sup>3</sup>Wageningen University and Research, Business Economics Group, Hollandseweg 1, 6706 KN Wageningen, the Netherlands

### ABSTRACT

Reduction of milk yield is one of the principal components in the cost of mastitis. However, past research into the association between milk yield and mastitis indicators is limited. Past research has not been based on online or in-line daily measurements and has not fully explored nonlinearity and the thresholds at which milk yield starts to decrease. In dairy herds with automated milking systems equipped with sensors, mastitis indicators of individual cows are measured on an intraday frequency, which provides unprecedented avenues to explore such effects in detail. The aim of this observational study was primarily to investigate the nonlinear associations of lactate dehydrogenase (LDH), electrical conductivity (EC), and somatic cell count (SCC) with milk yield at various stages of lactation, parity, and mastitis chronicity status (i.e., whether the cow had  $SCC \geq 200,000$  SCC/mL for the last 28 d). We also investigated thresholds at which mastitis indicators (LDH, EC, and SCC) started to be negatively associated with milk yield. We used data from 21 automated milking system herds measuring EC and online SCC. Of these herds, 7 of the 21 additionally measured online LDH. We operationalized milk yield as milk synthesis rate in kilograms per hour. Applying a generalized additive model, we estimated the milk synthesis rate as a function of the 3 mastitis indicators for 3 different subgroups based on parity, stage of lactation, and mastitis chronicity. Partial dependence plots of the mastitis indicators were used to evaluate the milk synthesis rate to study if the milk synthesis rate was associated with mastitis indicators at a specific level. Results showed that milk synthesis rate decreased with increasing SCC, LDH, and EC, but in a nonlinear fashion. The thresholds at which milk synthesis rate started to decrease

were 2.5 LnSCC (12,000 SCC/mL) to 3.75 LnSCC (43,000 SCC/mL), 0 to 1 LnLDH (1–2.7 U/L), and 5.0 to 6.0 mS/cm for EC. Additionally, another substantial decrease of milk synthesis rate was observed at thresholds of 5.625 LnSCC (277,000 SCC/mL) and 3 LnLDH (20 LDH U/L) but not for EC. Having chronic mastitis decreased milk synthesis rate in all models. The identified nonlinearities between mastitis indicators and milk synthesis rate should be incorporated in statistical models for more accurate estimations of milk loss due to mastitis.

**Key words:** mastitis, correlation, generalized additive model, milk loss

### INTRODUCTION

Mastitis is one of the most important diseases on commercial dairy farms, and the costliest consequence of it is loss of milk production (Hogeveen et al., 2019), both in subclinical (72% of the subclinical mastitis cost) as well as in clinical mastitis cases (48% of the clinical mastitis cost; Aghamohammadi et al., 2018). To support farmer decision-making regarding udder health, insight into the milk production losses due to mastitis is important. Milk production losses due to subclinical mastitis can be estimated and linked to the level of a specific mastitis indicator (e.g., milk loss of 1 kg/d with 200,000 SCC/mL).

Many milking systems have electrical conductivity (EC) sensors and could be equipped with a range of sensors that measure mastitis indicators of individual cows on an intraday frequency, such as SCC and lactate dehydrogenase (LDH). The higher the mastitis indicator, the more severe the inflammation is in the udder, resulting in a larger milk yield loss (Hortet et al., 1999; Hagnestam-Nielsen et al., 2009).

Somatic cell count is the most widely studied mastitis indicator for estimating losses of milk yield. Measuring SCC in DHI tests or in an experimental setting showed that milk yield is negatively associated with increasing SCC with greater losses at higher SCC (Jones et

Received September 29, 2021.

Accepted December 4, 2021.

\*Corresponding author: [John.bonestroo@delaval.com](mailto:John.bonestroo@delaval.com)

al., 1984; Hortet et al., 1999; Hagnestam-Nielsen et al., 2009). Previous studies found or suggested a negative effect on milk yield at thresholds of 148,000 SCC/mL (Tyler et al., 1989), 100,000 SCC/mL (Halasa et al., 2009), 50,000 SCC/mL (Hortet and Seegers, 1998), 7,400 SCC/mL (Dürr et al., 2008), and 12,400 SCC/mL (Gonçalves et al., 2018). A significantly reduced milk yield due to subclinical chronic mastitis was found by Hadrich et al. (2018).

Electrical conductivity can also be used as an indicator of inflammation of the mammary gland (International Dairy Federation, 2011). A few studies found a negative association between EC and milk production or milk loss (Oshima et al., 1990; Nielsen et al., 1993). Similar to EC, LDH is also less well studied in relation to milk yield. Nyman et al. (2014) found that milk yield was negatively associated with LDH, analyzed using milk samples in a laboratory, together with other cow factors (such as DIM and parity).

Although the studies mentioned used a multitude of different approaches and data collection protocols, they did not use data from online sensors of LDH and SCC that are used in commercial dairy farms. Moreover, all these studies assume a form of relationship between milk yield and the mastitis indicator beforehand using linear models (Hagnestam-Nielsen and Østergaard, 2009), logarithmic transformations (Green et al., 2006; Dürr et al., 2008; Hagnestam-Nielsen and Østergaard, 2009), or a combination of linear segments (Dürr et al., 2008; Gonçalves et al., 2018). However, the exact form, which could be a combination of effects (e.g., linear, exponential, and cyclical), is less studied. By choosing a functional form beforehand, the researcher might be limited by the inflexibility of the chosen parametric functional form to model these effects. In addition to the functional form, the threshold of SCC, LDH, and EC, at which a negative association with milk yield can be seen, has been studied to a limited extent. In the case of SCC, various thresholds have been suggested (Tyler et al., 1989; Hortet et al., 1999; Dürr et al., 2008; Halasa et al., 2009; Gonçalves et al., 2018), whereas EC and LDH thresholds based on milk loss have not been studied.

The primary aim of this observational study was to investigate the nonlinear associations of LDH, EC, and SCC with milk yield at differing stages of lactation, parity, and chronicity status during cow lactation. Chronicity status in this study is defined as a mean increase of SCC above 200,000 SCC/mL for at least 28 d. A second aim was to investigate thresholds at which mastitis indicators start to be negatively associated with milk yield. To achieve these aims, we applied a generalized additive model on milk yield, LDH, SCC, EC, and DIM to estimate milk synthesis rate over cow

lactation cycles, and to assess the association between these mastitis indicators and milk synthesis rate.

## MATERIALS AND METHODS

### *Herd Selection and Data*

Data were available in a central DeLaval database from 21 dairy herds from Canada, the Netherlands, Finland, and Sweden with automated milking systems (AMS; VMS series, DeLaval International AB) that measured EC and were equipped with an online cell counter (OCC; DeLaval International AB) to measure SCC. Farms were selected from the database on the presence of an OCC. Of these 21 herds with OCC, 7 also were equipped with a DeLaval Herd Navigator (DeLaval International AB) to measure LDH. Data were retrieved from a database of DeLaval International AB for the period from January 2017 to April 2020. The majority of the herds kept Holstein cows. The number of lactating cows ranged from 66 to 603 cows, with a mean of 194 cows. The 305-d herd average milk yield ranged from 5,712 to 11,979 kg, with an average of 9,870 kg. These data were reported per milking and included the herd identification, cow identification, DIM, time of milking, milk yield, SCC, LDH, and the EC. For SCC and LDH, a sampling algorithm based on the risk of mastitis determined when a sample was taken (Mazeris, 2010). We could not check the setting of the algorithm in the farms. However, we did observe fewer samples of SCC and LDH than number of milkings in the analyzed data set.

### *Data Cleaning and Preparation*

As a first step in the analyses, several data cleaning and data preparation activities were undertaken. All cow lactations had missing data as not all mastitis indicators were measured at every milking. The raw data set consisted of 5,990,883 milkings from 6,372 cows. Milkings outside the range of 5 to 305 DIM were discarded (removed 735,977 milkings). Furthermore, quarter EC values outside of the range from 3 to 10 mS/cm were set to “not available” (in 53,015 milkings) as done by Anglart et al. (2020). To ensure that the decreases in milk yield were not due to teat blinding (not milking 1 or more udder quarters during lactation), milkings with quarter milk yields equal to 0 were removed (removed 640,958 milkings). We used milk synthesis rate (see below), assuming an approximately linear relationship between time interval and milk yield. This was not the case for short- and long-time intervals (Hogeveen et al., 2001); therefore, we removed them. As such, records where the time interval between the current milking

and the previous milking was outside of the range of 4 to 24 h were discarded (removed 66,044 milkings). Records on days where the number of milkings for a cow was equal or greater than 5 were discarded (removed 11,577 milkings). This resulted in a data set of 4,536,637 milkings from 7,352 lactations from 5,805 cows from 21 herds.

A set of variables was created to facilitate statistical analysis. We used milk synthesis rate (kg/h) as the dependent variable. Each interval between milkings is different in AMS farms, leading to a large variation in time intervals between milkings (Hogeveen et al., 2001). Therefore, to obtain a comparable milk yield-based measure, we divided the milk yield (in kg/milking) by the time interval (in h) between milkings to obtain milk synthesis rate. We used online SCC and LDH as independent variables. These 2 variables were transformed using the natural logarithm (LnSCC and LnLDH). Furthermore, we used the mean EC of the 4 quarters as the third independent variable (mean EC). Mean EC was chosen to compare the milk loss results for LDH and SCC, as it was a cow-level indicator, similar to SCC and LDH. In addition, the subgroup variable “chronicity status” was created to represent whether the cow was chronic or not. A milking day observation was labeled as chronic if a cow had weekly SCC geometric averages equal or higher than 200,000 cells/mL for a period of 4 consecutive weeks or more before the current milking day (Bonestroo et al., 2021) based on available SCC samples. Last, we also created a cow lactation variable (CowLactation) that combined the unique animal identification number with the parity to identify unique cow lactations.

We aggregated the multiple individual milkings on a given day by using the maximum daily values of LnSCC, LnLDH, mean EC, and averaged the milk synthesis rate (reducing the data from 4,536,637 milkings to 1,687,508 milking days). The daily maximum value was used to capture the severity of the increase. When some values were missing for specific milkings but not for all milkings on specific days, these values were ignored in determining the maximum. When there was no observation of the mastitis indicator at all during a day, no daily maximum value of that day was given. Lactations with less than 100 SCC day observations or on average 1 SCC sample per 3 d (to allow determination of the chronicity status based on SCC throughout the lactation) were discarded. This requirement allowed us to define chronic observations anywhere in the lactation. We chose 100 d as a threshold because a lower value (e.g., 29 d) would only let us define observations coming from chronic cows very sparingly, whereas using a requirement of higher number of samples limited the data set to such a substantial amount that we had little

data left. Lactations with less than 100 d of observations of a specific mastitis indicator were also discarded in the data set for that indicator-specific model. For each of the 3 mastitis indicators, a separate data set was created. Because not all mastitis indicators were always reported, these 3 data sets differed in number of observations. The selection steps reduced the data further from 1,687,508 milking days to 788,572 milking days of the SCC data set (4,516 lactations and 3,352 cows and 21 herds), 179,335 milking days of the LDH data set (1,394 lactations and 1,116 cows and 7 herds), and 1,146,320 milking days of the mean EC data set (4,515 lactations and 3,350 cows and 21 herds).

To analyze the association between milk synthesis rate and mastitis indicators for different levels of parity, DIM, and chronicity, 3 subgroups were created and analyzed separately. The first subgroup was formed according to 3 DIM levels (5–28, 29–60, and 61–305 DIM) as multiple authors have found differences in milk loss between stages of lactation (Hagnestam-Nielsen et al., 2009; Gonçalves et al., 2018). These cut-offs were determined by selecting the median DIM where the day-to-day change in milk synthesis rate was maximal (28 DIM) and where the milk synthesis rate peaked (60 DIM) in our data set. The second subgroup was based on parity (first lactation, second lactation, and third or more lactation) as it can be expected that multiparous cows, which give more milk, will have a higher milk loss when subclinical mastitis occurs (Dürr et al., 2008; Hagnestam-Nielsen et al., 2009; Gonçalves et al., 2018). The last subgroup was formed according to chronicity (nonchronic and chronic mastitis) as cows with chronic mastitis tend to have higher milk losses (Hadrlich et al., 2018). The differences in milk synthesis loss in the various levels of parity, stage of lactation, and chronicity were studied separately using separate regression models, as discussed in the next section.

### Statistical Analysis

We applied a generalized additive model using the R package *mgcv* (Wood, 2012) in R 3.6.1 (<https://www.r-project.org/>) to model milk synthesis rate per hour. Milk synthesis rate was estimated as a function of the mastitis indicator and DIM, for each subgroup, respectively (Eq. 1, 2, and 3). The DIM and CowLactation were treated as confounders. Depending on the subgroup that was analyzed, the subgroup value in these equations can take the form of the parity, stage of lactation, or chronicity status. A generalized additive model is an extension of a general linear model where the dependent variable can depend linearly on unknown smoothing functions in combination with normal regression coefficients and random effects (as

used in the general linear mixed model). The smooth functions can be fitted with data and can have any form (e.g., linear, quadratic, plateauing, or a combination of them). The function form does not have to be predefined, allowing a very flexible estimation of the association between a mastitis indicator and milk synthesis rate. Last, we included a random effect of each individual cow lactation (random cow lactation effect) using the CowLactation variable to control for nonindependence of observations. In generalized additive models, different link functions can be used to model the relation between the dependent variable and the independent variables, as it is an extended general linear model. Milk synthesis rate was assumed to have a scaled-t distribution rather than a normal Gaussian distribution because it was expected that milk synthesis rate would have more extreme observations than a normal distribution. We used the following models:

$$\begin{aligned} \text{Milk synthesis rate} = & \text{intercept} + \text{subgroup intercept} \\ & + f_{LnSCC}(\text{LnSCC}) \times \text{subgroup} + f_{DIM}(\text{DIM}) \\ & \times \text{subgroup} + \text{random cow lactation effect}, \quad [1] \end{aligned}$$

$$\begin{aligned} \text{Milk synthesis rate} = & \text{intercept} + \text{subgroup intercept} \\ & + f_{LnLDH}(\text{LnLDH}) \times \text{subgroup} + f_{DIM}(\text{DIM}) \\ & \times \text{subgroup} + \text{random cow lactation effect}, \quad [2] \end{aligned}$$

$$\begin{aligned} \text{Milk synthesis rate} = & \text{intercept} + \text{subgroup intercept} \\ & + f_{Mean\ EC}(\text{mean EC}) \times \text{subgroup} + f_{DIM}(\text{DIM}) \\ & \times \text{subgroup} + \text{random cow lactation effect}, \quad [3] \end{aligned}$$

where  $f_{DIM}$  is a nonlinear smoothing function modeling the milk synthesis rate over the lactation cycle with a cubic spline basis that was estimated separately for every subgroup. The  $f_{DIM}$  was not plotted in the results for brevity, but it takes the form similar to a Wood lactation curve as found in literature (Wood, 1967), and where  $f_{LnSCC}$ ,  $f_{LnLDH}$ , and  $f_{Mean\ EC}$  are a nonlinear smoothing function modeling the association between LnSCC, LnLDH, and mean EC with milk synthesis rate. To enable the analysis, we used a baseline where the mastitis indicators were not associated with decreases in milk synthesis rate. As such, this study assumed, before the analysis, that a level of 1,000 SCC/mL, 1 U/L of LDH, and 4 mS/cm mean EC would have no effect on milk synthesis rate. These levels were close to the minimally observable levels, and were chosen due to the low thresholds found and proposed for SCC (Dürr et al., 2008; Gonçalves et al., 2018). These assumptions were needed to estimate  $f_{LnSCC}$ ,  $f_{LnLDH}$ , and  $f_{Mean\ EC}$ .

These functions are also nonlinear smoothing functions with a cubic spline basis. We used the BAM function, which is a generalized additive model with discretization of covariate values that makes it more time and memory efficient when having large data sets (Wood, 2012, 2017). Each of the 3 models (Eq. 1, 2, and 3) were estimated separately for each subgroup (parity, stage of lactation, and chronicity), thus leading to the fitting of 9 models in total (3 mastitis indicators times 3 subgroup variables).

To visualize the associations, we plotted  $f_{LnSCC}$ ,  $f_{LnLDH}$ , and  $f_{Mean\ EC}$  for each mastitis indicator and each of the subgroups. The value of the mastitis indicator, at which it started to be negatively associated with milk synthesis rate, was identified as a threshold. This point was found by determining the maximum positive milk synthesis rate difference in the partial plot (the highest point in the plot) and was highlighted in the partial effect plots. Further points of potential substantial decreases in milk synthesis rate start, after this initial threshold, will be described by how the rate of the decrease changes abruptly (e.g., whether the line starts to decrease considerably more).

The residuals of all models were checked for normality, homoscedasticity, and autocorrelation using qq-plots, fitted values-residual plots, and autocorrelation plots. During the analysis, we detected substantial autocorrelation for all models. The autocorrelation problem was solved by adapting the model. The BAM function used in the mgcv library (Wood, 2012) does not allow to estimate a first order autoregressive (AR1) residual structure, but it does allow for a predefined AR1 parameter. Consequently, we allowed an AR1 residual structure by first estimating a model without an AR1 parameter, then estimating the residual autocorrelation at the first lag, and inserting that value as the AR1 parameter when fitting the final model using autocorrelation functions of the R package itsadug (Van Rij et al., 2017). The inclusion of the AR1 structure reduced the autocorrelation problem to an insubstantial level.

## RESULTS

### Descriptive Statistics

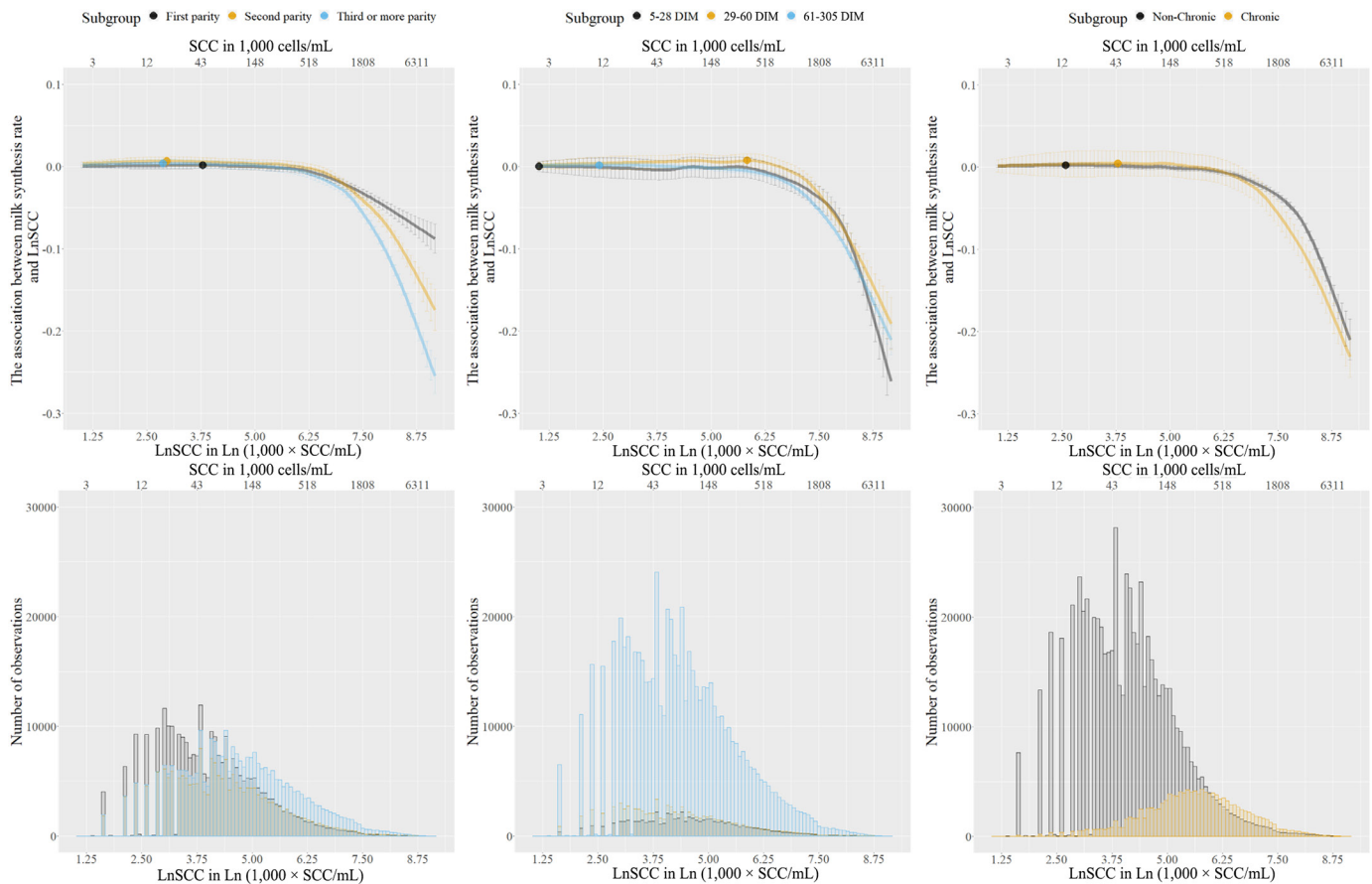
The number of AMS per herd ranged from 1 to 9 with a mean of 2.7. After the data selection process, the number of cows analyzed in each herd varied with a mean of 159, a minimum of 22, and a maximum of 512. Two herds had an especially small number of lactations because we required 100 SCC-day observations per lactation for all lactations, SCC were not sampled

every day, and only data on completed lactations from mid-2018 were available. The mean parity across herds was 2.42 with a standard deviation of 0.36, a herd parity mean minimum of 1.63, and a herd parity mean maximum of 3.22. The mean milk synthesis rate was 1.47 kg/h across herds with a standard deviation of 0.17, with a herd mean minimum of 0.99 and a herd mean maximum of 1.74.

In the SCC data set, the mean LnSCC across herds was 4.39 (80,640 SCC/mL) with a standard deviation of 0.45, a herd mean minimum of 3.58 (35,873 SCC/mL), and a herd mean maximum of 5.30 (200,337 SCC/mL). In the LDH data set, the mean LnLDH across herd means was 3.07 (21.54 U/L) with a standard deviation of 0.24, a herd mean minimum of 2.83 (16.95 U/L), and a herd mean maximum of 3.54 (34.47 U/L). In the EC data set, the mean EC across herds was 4.66 mS/cm with a standard deviation of 0.22, a herd mean minimum of 4.28 mS/cm, and a herd mean maximum of 5.09 mS/cm.

## Generalized Additive Model Analyses

**SCC Results.** Figure 1 provides the nonlinear association between LnSCC and milk synthesis ( $f_{LnSCC}$ ) and the frequency of LnSCC observations for different parity, stage of lactation, and chronicity classes (Figure 1–C). Table 1 summarizes the results of the regression models. The milk synthesis rate was negatively associated with LnSCC over a specific threshold. The large dot in Figure 1 marks the point on the line where milk synthesis rate started to decrease, and thereby milk losses occurred. For most cases, this threshold was approximately between 2.5 LnSCC (12,000 SCC/mL) and 3.75 LnSCC (43,000 SCC/mL), whereas occasional lower and higher thresholds were found in the analysis depending on subgroup. Moreover, the milk synthesis rate started to decrease even further a second time when LnSCC increased, and at an increasing speed and nonlinearly. This occurred approximately after 5.625 LnSCC/mL (~277,000 SCC/mL) for all subgroups.



**Figure 1.** Estimated associations between milk synthesis rate and LnSCC and number of observations for parity, stage of lactation, and chronicity subgroups. The dots indicate that the start of milk synthesis rate decreases, and thereby milk losses increase from that point.

Table 1. Regression summary of somatic cell count analysis<sup>1</sup>

Item	Parity groups model	Coefficient (SE)	Chronicity model	Coefficient (SE)	Stage of lactation model	Coefficient (SE)
Parametric terms	Intercept	1.27 (0.01)***	Intercept	1.48 (0.005)***	Intercept	0.99 (5.84)
	Intercept second parity	0.27 (0.01)***	Intercept chronicity	-0.04 (0.01)***	Intercept of DIM 29-60	0.68 (5.86)
	Intercept ≥third parity	0.35 (0.01)***		EDF (Ref. df)	Intercept of DIM 61-305	0.52 (5.84)
Smoothing terms		EDF (Ref. df)				EDF (Ref. df)
	DIM of first parity	32.66 (36.87) <sup>2</sup> ***	DIM of nonchronic	36.35 (38.60)***	DIM of 5-28	3.94 (3.99)***
	DIM of second parity	33.11 (37.16)***	DIM of chronic	17.14 (21.67)***	DIM of 29-60	3.95 (4.58)***
	DIM of ≥third parity	34.68 (37.97)***			DIM of 61-305	9.59 (12.26)***
	LnSCC of first parity	7.34 (8.61)***	LnSCC of nonchronic	12.97 (14.58)***	LnSCC of 5-28 DIM	9.71 (11.36)***
	LnSCC of second parity	10.02 (11.56)***	LnSCC of chronic	11.86 (13.82)***	LnSCC of 29-60 DIM	9.71 (11.36)***
LnSCC of ≥third parity	12.26 (13.79)***			LnSCC of 61-305 DIM	12.26 (14.01)***	
Adjusted R <sup>2</sup>		0.81				0.78

<sup>1</sup>EDF = estimated degrees of freedom; Ref. df = reference degrees of freedom.

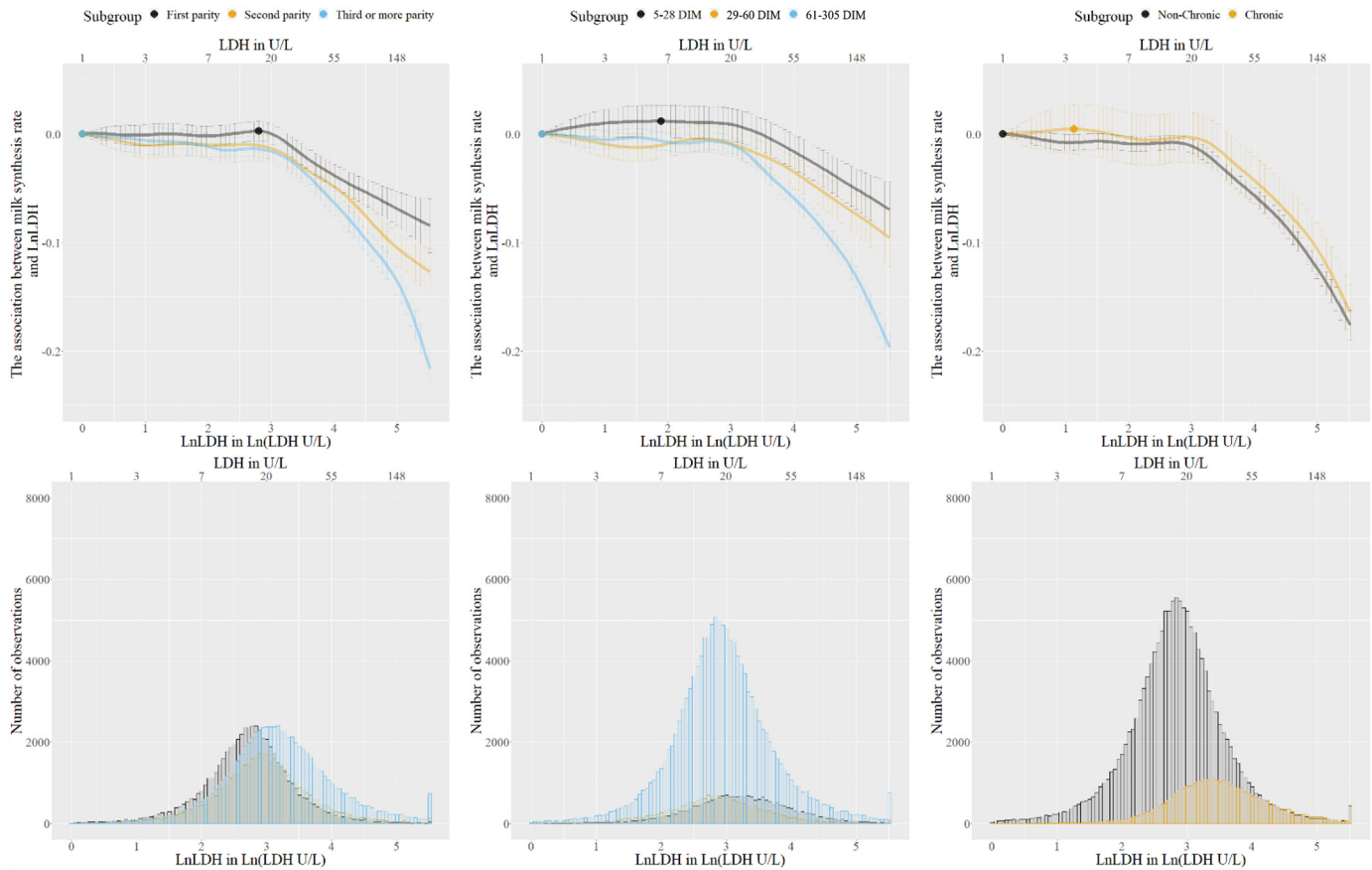
<sup>2</sup>Significance of smoothing terms can be determined by a method explained in detail in Wood (2013).

\*\*\* $P < 0.0001$ .

Some differences in thresholds and the steepness of the decrease in milk synthesis rate between subgroups were seen. These differences in thresholds were caused by minor differences in the LnSCC and milk synthesis rate association on an overall approximately flat line on the lower levels of LnSCC. Therefore, the differences between thresholds should be interpreted carefully. In Figure 1A, the decrease in milk synthesis rate was steeper at higher levels of parity (indicating an increasing decrease in milk synthesis rate in older cows) of LnSCC for the multiparous subgroups. Moreover, the parity subgroup differences in intercept between first-parity cows and second-parity cows and first-parity cows and ≥third-parity cows were 0.27, with a standard error of 0.01, and 0.35, with a standard error of 0.01, respectively ( $P < 0.01$ ). Regardless of LnSCC, cows with a higher parity tended to produce more milk. In Figure 1B, no clear difference in milk synthesis rate in the stage of lactation subgroups could be seen, and the difference in intercept between 5 to 28, 29 to 60, and 61 to 305 DIM subgroups was also not significant ( $P > 0.05$ ). In Figure 1C, the chronic subgroups had approximately the same form; the chronic group was steeper in its decrease and was lower than its nonchronic counterpart. The difference in intercepts between the chronic and nonchronic subgroups, indicating long-term effects on milk synthesis rate, was -0.04 with a standard error of 0.01 ( $P < 0.01$ ).

**LDH Results.** Figure 2 provides the nonlinear association between LnLDH and milk synthesis rate ( $f_{LnLDH}$ ) and the frequency of LnLDH observations for different parity, stage of lactation, and chronicity classes (Figure 2A-C). Table 2 summarizes the results of the regression models. The dot in Figure 2 marks the point where milk synthesis rate started to decrease, and thereby milk losses increased when LnLDH increased. It can be seen from the results that LnLDH was negatively associated with milk synthesis rate after the threshold for all parity, stage of lactation, and chronicity subgroups. The thresholds ranged from approximately 0 to 3 LnLDH (1-20 U/L) for all subgroups. Despite the similarity in the general form and level of the smoothing function, the differences in thresholds were large. The differences in thresholds seemed to be caused by minor differences in the shape of the association between LnLDH and milk synthesis rate between the subgroups. In other words, the difference in milk loss between the thresholds was limited. Nevertheless, the milk synthesis rate decreased noticeably more after approximately 3 LnLDH (20 U/L) in all subgroups.

Several dissimilarities in thresholds and the steepness of the decrease in milk synthesis rate between subgroups were seen. In Figure 2A, multiparous cows showed a larger decrease in milk synthesis rate as



**Figure 2.** Estimated associations between milk synthesis rate and LnLDH and number of observations for parity, stage of lactation, and chronicity subgroups. The dots indicate that the start of milk synthesis rate decreases, and thereby milk losses increase from that point. Occasionally, the orange dot is covered by the blue dot in the parity and the stage of lactation subgroups. LDH = lactate dehydrogenase.

sociated with higher LnLDH than primiparous cows. Even more, the parity subgroup differences in intercept between first-parity cows and second-parity cows and first-parity cows and  $\geq$ third-parity cows were 0.27 with a standard error of 0.01 and 0.36 with a standard error of 0.02 ( $P < 0.01$ ), respectively. Regardless of LnLDH, cows with a higher parity produced significantly more milk. In Figure 2B, 61 to 305 DIM observations showed a larger decrease in milk synthesis rate than the 5 to 28 and 29 to 60 DIM observations. The stage of lactation subgroup difference in intercept between 5 to 28, 29 to 60, and 61 to 305 DIM subgroups were not significantly different ( $P > 0.1$ ). In Figure 2C, the line of nonchronic cows was slightly lower than the line of chronic cows, but the chronic cow intercept in the model was  $-0.08$  with a standard error of 0.01 ( $P < 0.01$ ). Regardless of current LDH, the milk synthesis rate for chronic cows was lower.

**Mean EC Results.** Figure 3 provides the nonlinear association between mean EC and milk synthesis ( $f_{Mean\ EC}$ ) and the frequency of mean EC observations for different parity, stage of lactation, and chronicity

classes (Figure 3A–C). Table 3 summarizes the results of the regression models. Figure 3 indicates that the association between mean EC and milk synthesis rate was highly nonlinear. In addition, the threshold of milk synthesis rate decrease was within the range of 5.0 to 6.0 mS/cm mean EC for all subgroups. This threshold was found at a high percentile of the mean EC distribution compared with LnSCC and LnLDH (see bottom panels in Figure 2 and 3). Mean EC remained negatively associated with milk synthesis rate after the threshold for all parity, stage of lactation, and chronicity subgroups.

Between the subgroups, several differences and similarities could be seen in Figure 3. The differences in thresholds between subgroups were limited as they all fell between 5.0 and 6.0 mS/cm. The differences in the function forms between subgroups should be interpreted with care as a large section of the decrease in milk synthesis rate was based on a small area of the mean EC distribution. The limited number of observations explained the increase in milk synthesis rate at 7.5 mS/cm for the second-parity subgroup in Figure 3A. In Figure 3A, the milk synthesis rate of the multiparous

Table 2. Regression summary of lactate dehydrogenase (LDH) analysis<sup>1</sup>

Item	Parity groups model	Coefficient (SE)	Chronicity model	Coefficient (SE)	DIM model	Coefficient (SE)
Parametric terms	Intercept	1.26 (0.01)***	Intercept	1.49 (0.01)***	Intercept	2.03 (3.58)
	Intercept second parity	0.27 (0.02)***	Intercept chronicity	-0.08 (0.01)***	Intercept of DIM 29-60	-0.45 (3.62)
	Intercept ≥third parity	0.36 (0.02)***		EDF (Ref. df)	Intercept of DIM 61-305	-0.53 (3.57)
Smoothing terms		EDF (Ref. df)		EDF (Ref. df)		EDF (Ref. df)
	DIM of first parity	22.08 (27.81) <sup>2</sup> ***	DIM of nonchronic	28.74 (34.31)***	DIM of 5-28	3.56 (3.85)***
	DIM of second parity	24.93 (30.77)***	DIM of chronic	13.35 (17.23)***	DIM of 29-60	3.42 (4.09)***
	DIM of ≥third parity	26.66 (32.39)***			DIM of 61-305	7.11 (9.19)***
	LnLDH of first parity	9.07 (11.13)***	LnLDH of nonchronic	11.20 (13.42)***	LnLDH of 5-28 DIM	5.63 (7.17)***
	LnLDH of second parity	9.53 (11.66)***	LnLDH of chronic	8.93 (10.93)***	LnLDH of 29-60 DIM	6.79 (8.54)***
Adjusted R <sup>2</sup>	12.51 (14.73)***			LnLDH of 61-305 DIM	12.00 (14.23)***	
	0.77				0.73	

<sup>1</sup>EDF = estimated degrees of freedom; Ref. df = reference degrees of freedom.

<sup>2</sup>Significance of smoothing terms can be determined by a method explained in detail in Wood (2013).

\*\*\* $P < 0.0001$ .

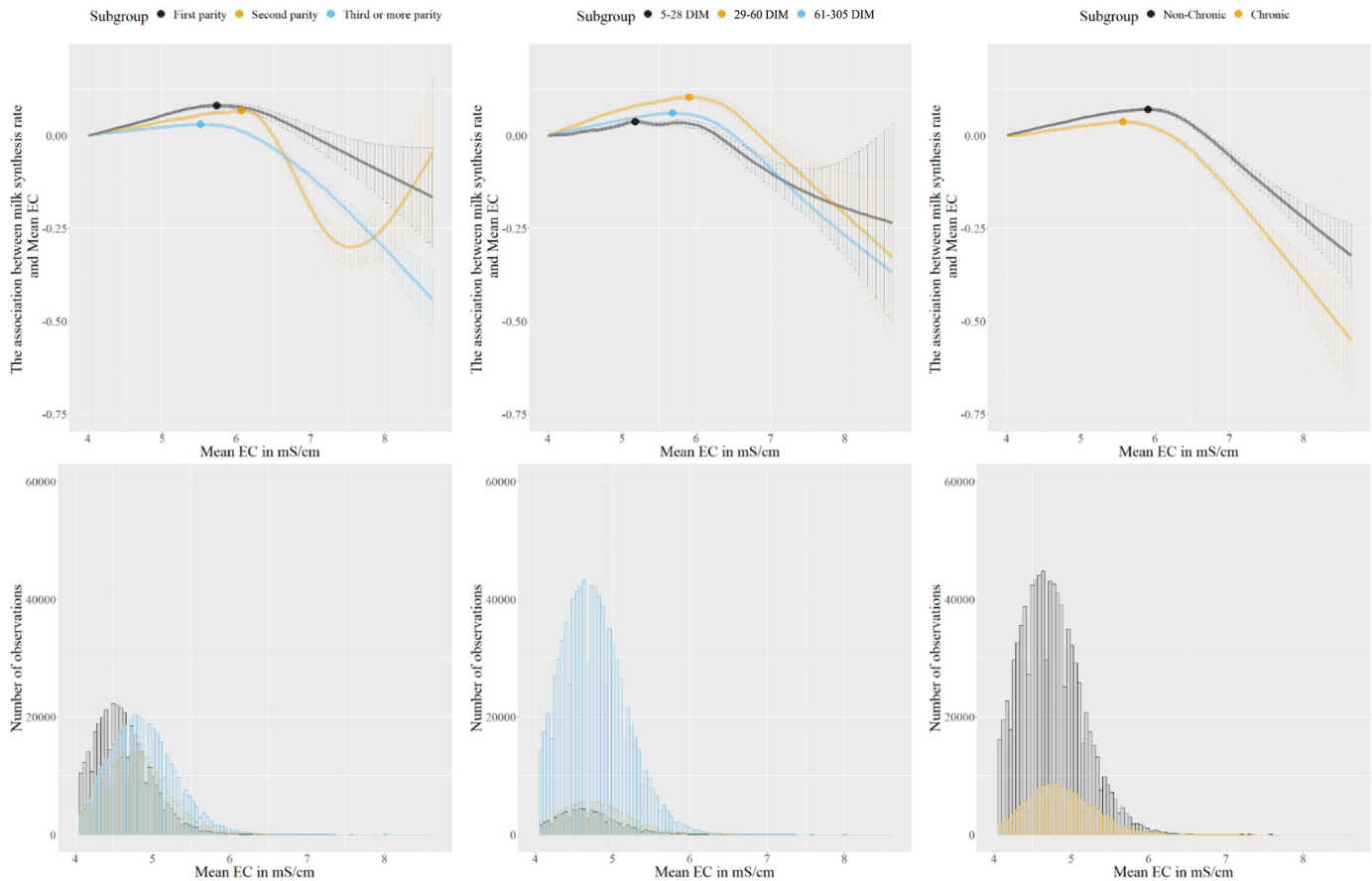
subgroups decreased more when mean EC increased than in the first-parity subgroups. Furthermore, the parity subgroup differences in intercept between first-parity cows and second-parity cows and first-parity cows and ≥third-parity cows were 0.28 with a standard error of 0.01 and 0.36 with a standard error of 0.01, respectively ( $P < 0.01$ ). Regardless of mean EC, cows with a higher parity produced significantly more milk. In Figure 3B, the milk synthesis rate of the 29 to 60 and 61 to 305 DIM subgroups decreased more than the milk synthesis rate of 5 to 28 DIM subgroup when mean EC increased, whereas the difference in intercepts between the stage of lactation subgroups was not significant ( $P > 0.1$ ). In Figure 3C, the milk synthesis rate of the chronic subgroups decreased more than for the nonchronic subgroup when mean EC increased. In addition, the chronic mastitis subgroup differences in intercept between chronic cows and nonchronic cows was -0.04 with a standard error of 0.003 ( $P < 0.01$ ). Milk synthesis rate for chronic cows was lower when controlling for the current level of mean EC.

## DISCUSSION

Our aim was to investigate the association between milk synthesis rate and online LDH, EC, and SCC at differing stages of lactation, parity, and chronicity statuses. As a second aim, we wanted to investigate the differences in thresholds at which online LDH, mean EC, and SCC levels start to be associated negatively with milk synthesis rate. We found strong nonlinearities after a linear phase in the association between milk synthesis rate and different subgroups. Estimating the nonlinearity correctly would lead to a more accurate estimation of milk loss. In past research, the resulting nonlinear functions of the associations between the studied mastitis indicators with milk yield have not been found in this detail (Dürr et al., 2008; Gonçalves et al., 2018). Each of these mastitis indicators is tied to a connected, but dissimilar, mechanism in the immune response of a cow (Viguiet et al., 2009), which may be the explanation for the differences we see. An immune response leads to increases in SCC in the milk when polymorphonuclear neutrophils, white blood cells, are released into the milk to engulf the pathogen. Next, apoptosis occurs, and somatic cells can be found in the milk. Differences in EC occur by tissue damage and breaching of the blood-milk barrier (Viguiet et al., 2009). The LDH, on the other hand, is released in the milk when a pathogen is engulfed and killed by a polymorphonuclear neutrophil (Viguiet et al., 2009).

This study looked at differences in associations between parity, stage of lactation, and chronicity subgroups. We found nonlinear as well as linear charac-





**Figure 3.** Estimated associations between milk synthesis rate and mean electrical conductivity (EC) and number of observations for parity, stage of lactation, and chronicity subgroups. The dots indicate that the start of milk synthesis rate decreases, and thereby milk losses increase from that point.

teristics in the association between milk synthesis rate and LDH, SCC, and mean EC. Typically, a linear phase with no clear association or a proportionally small decreasing association would be followed by a nonlinear phase where a quadratic decrease (i.e., decrease at an increasing speed) could be seen. This justified the use of nonlinear models for LDH, SCC, and EC measures, and supported the use of the threshold model in Dürr et al. (2008) and Gonçalves et al. (2018) wherein SCC has no effect on milk yield up to a certain threshold, and a negative effect on milk yield after the threshold. After reaching this threshold, milk synthesis rate would decrease at an increasing rate, and this is also reported by others (Jones et al., 1984; Hagnestam-Nielsen et al., 2009; Hand et al., 2012). The milk synthesis difference was higher for the multiparous subgroups than their primiparous counterparts for all mastitis indicators. This is also seen in Hagnestam-Nielsen et al. (2009) and Hand et al. (2012) in the case of SCC. Multiparous cows often have a higher milk production; therefore, the losses can be greater. For the stage of lactation

subgroups, no clear differences were seen in the SCC models; however, the milk synthesis rate in the 5 to 28 and 29 to 60 DIM subgroups decreased less than 61 to 305 DIM in the LnLDH models, and the 29 to 60 and 61 to 305 DIM subgroup decreased more than 5 to 28 DIM for mean EC. In contrast to our SCC results, Hagnestam-Nielsen et al. (2009), Dürr et al. (2008), and Gonçalves et al. (2018) reported increasing milk losses related to SCC with increasing DIM, although this relationship could also be parabolic. These conflicting results could be caused by the limited number of DIM subgroups in our study. The results of the chronic subgroups are difficult to compare with results of other researchers due to differences in the definition and operationalization of mastitis chronicity. Nevertheless, Hadrich et al. (2018) also found that milk losses increased when the number of past consecutive observations with a higher SCC increased.

We found substantial variation in thresholds for the onset of milk loss between the different subgroups in all SCC models as well as the LDH models, which was

**Table 3.** Regression summary of mean electrical conductivity (EC)<sup>1</sup>

Item	Parity groups model	Coefficient (SE)	Chronicity model	Coefficient (SE)	DIM model	Coefficient (SE)
Parametric terms	Intercept	1.25 (0.01)***	Intercept	1.45 (0.004)***	Intercept	1.25 (5.55)
	Intercept second parity	0.28 (0.01)***	Intercept chronicity	-0.04 (0.003)***	Intercept of DIM 29-60	0.33 (5.57)
	Intercept ≥third parity	0.36 (0.01)***		EDF (Ref. df)	Intercept of DIM 61-305	0.23 (5.55)
Smoothing terms		EDF (Ref. df)				
	DIM of first parity	32.81 (36.96) <sup>2</sup> ***	DIM of nonchronic	36.16 (38.54)***	DIM of 5-28	3.93 (3.98)***
	DIM of second parity	32.98 (37.06)***	DIM of chronic	16.37 (20.65)***	DIM of 29-60	4.05 (4.70)***
	DIM of ≥third parity	34.17 (37.71)***			DIM of 61-305	9.28 (11.76)***
	Mean EC of first parity	10.37 (12.21)***	Mean EC of nonchronic	13.42 (15.12)***	Mean EC of 5-28 DIM	10.67 (12.30)***
Mean EC of second parity	15.63 (16.97)***	Mean EC of chronic	11.64 (13.62)***	Mean EC of 29-60 DIM	10.38 (12.05)***	
Mean EC of ≥third parity	11.64 (13.53)***			Mean EC of 61-305 DIM	12.96 (14.43)***	
Adjusted R <sup>2</sup>		0.80				0.76

<sup>1</sup>EDF = estimated degrees of freedom; Ref. df = reference degrees of freedom.<sup>2</sup>Significance of smoothing terms can be determined by a method explained in detail in Wood (2013).\*\*\**P* < 0.0001.

driven by minor differences in overall similar functional forms. Nevertheless, similar low thresholds (mostly 2.5 LnSCC or 12,000 SCC/mL) for SCC were found compared with milk loss thresholds reported in the past (7,400–12,400 SCC/mL; Dürr et al., 2008; Gonçalves et al., 2018). We found some occasional deviation from 2.5 LnSCC between parity groups and stage of lactations, which is also found by Gonçalves et al. (2018). Other research pointed toward a dilution effect of SCC on milk yield (Green et al., 2006), and the dilution effect may cause an overestimation of the milk yield loss with increasing SCC. Due to the nonparametric nature of generalized additive models, the threshold was not a single parameter in the model, and we could not estimate the uncertainty of the threshold, whereas other authors report occasional substantial uncertainty in thresholds (Dürr et al., 2008; Gonçalves et al., 2018). Moreover, parametric piecewise models may have difficulty converging when finding complex (quadratic) parametric functions (Gonçalves et al., 2018). Nevertheless, we would argue that, if one were interested in only the threshold, one should fit a parametric piecewise model in which the threshold is a parameter. It would give a more precise description of the threshold together with an estimation of the uncertainty of the threshold.

This study had several limitations that may constrain the conclusions drawn. The disadvantage of observational secondary data, as used in this study, is that one cannot control all factors (e.g., farmer decision-making, availability of data, and LDH or SCC sampling algorithm) that can influence the association between inflammation indicators and milk production, potentially reducing the internal and external validity. Additionally, we did not have LDH data for all farms, reducing the generalizability of the LDH results relative to the SCC or mean EC results. In terms of farmer decision-making, farmers may have decided to retain high-producing cows with chronic mastitis more than lower-producing cows with chronic mastitis. This may cause a bias in the data where chronic cows are less affected by higher values of the mastitis indicators, reducing generalizability. The (automated) sampling strategy for LDH or SCC was partly based on the risk of having mastitis. Hence, cows that were sampled were more likely to have udder problems, even when they had low levels of mastitis indicators, potentially reducing the estimated milk loss (i.e., the difference between a cow with a low SCC and the cow with a high SCC) and generalizability. Furthermore, the time intervals between milkings differed substantially. We assumed that the effect of time interval on milk yield was linear by calculating the milk synthesis rate as milk yield divided by the time interval. However, milk yield

is nonlinearly associated with time interval (Hogeveen et al., 2001), especially during the first hours of the time interval between milkings. We have attempted to solve this issue by excluding observations with extreme time intervals between milkings of less than 4 h. Also, we took the daily average of milk synthesis rate, reducing the influence of short or long milking intervals. Therefore, milk synthesis rate can be regarded as an adequate measurement of a cow's milk production capacity. The accuracy of SCC, LDH, and EC to measure udder inflammation may differ between indicators and between the same indicator measured by equipment from different brands, but it should be inconsequential with regard to the estimates of milk yield losses associated with the indicators because we are not claiming to assess milk yield losses due to (subclinical) mastitis. Nevertheless, for the OCC, Nørstebø et al. (2019) showed that the correlation between laboratory SCC and OCC was 0.82. The research results regarding the ability of LDH to detect mastitis of Chagunda et al. (2006) and Friggens et al. (2007) were integrated in the development of the Herd Navigator system and hence can be used by farmers. The EC is a commonly used mastitis indicator (Nielen et al., 1992), but to our knowledge, the diagnostic properties of EC specific to a DeLaval system are not reported. In variable creation, we decided to create the cow-level mean EC variable by taking the mean of the quarter conductivity values rather than using the quarter conductivity values. This decision was made to make the estimated cow-level milk yield losses comparable to the losses associated with SCC and LDH. Nevertheless, the availability of quarter level EC, in combination with quarter level milk yield, would make it possible to estimate the milk losses due to increased EC at the quarter level, which would improve the accuracy of the estimates. That possibility, however, was beyond the scope of this article.

The overall statistical properties of the models were acceptable. The statistical fit of the SCC, LDH, and mean EC models was good in that the adjusted  $R^2$  of the models ranged from 0.78 to 0.81, 0.73 to 0.77, and 0.76 to 0.80, respectively (see Tables 1, 2, and 3). Without an adjustment for autocorrelation, we saw substantial autocorrelation in all models, and this affected the estimated associations between the mastitis indicators and milk synthesis rate. In our models, we have attempted to correct for autocorrelation by allowing an AR1-correlation structure. This resulted in a decrease in autocorrelation to less than approximately 0.2 on all lags, although the autocorrelation did not completely disappear. Nevertheless, the autocorrelation was minimal; therefore, we think that our estimates were not substantially influenced. No large deviations from the

assumptions of homoscedasticity and normality could be seen using qq-plots and fitted versus residuals plots.

## CONCLUSIONS

We found that the negative associations between SCC, EC, and LDH with milk yield were nonlinear and had a similar function form between parity groups, at different lactation stages, and at different chronicity statuses of the cow but occasionally differed in their level of decrease. Nevertheless, multiparous cows incurred larger milk losses than primiparous cows, whereas the effect of stage of lactation differed between indicators. Chronicity had a negative association with milk synthesis rate. In contrast to mean EC, milk synthesis rate started to decrease substantially more for SCC and LDH at higher SCC (277,000 SCC/mL) and LDH (20 U/L) levels. The study highlighted the nonlinearities that exist in the associations between different mastitis indicators that can be useful to more accurately predict mastitis-related milk loss.

## ACKNOWLEDGMENTS

This research was funded by an industry PhD program of the Swedish Government (Stockholm, Sweden; reference number N2017/036895/SK) and DeLaval International AB (Tumba, Sweden). John Bonestroo and Ilka C. Klaas are employed by DeLaval International AB. The authors have not stated any other conflicts of interest.

## REFERENCES

- Aghamohammadi, M., D. Haine, D. F. Kelton, H. W. Barkema, H. Hogeveen, G. P. Keefe, and S. Dufour. 2018. Herd-level mastitis-associated costs on Canadian dairy farms. *Front. Vet. Sci.* 5:100. <https://doi.org/10.3389/fvets.2018.00100>.
- Anglart, D., C. Hallén-Sandgren, U. Emanuelson, and L. Rönnegård. 2020. Comparison of methods for predicting cow composite somatic cell counts. *J. Dairy Sci.* 103:8433–8442. <https://doi.org/10.3168/jds.2020-18320>.
- Bonestroo, J., M. van der Voort, N. Fall, H. Hogeveen, U. Emanuelson, and I. C. Klaas. 2021. Progression of different udder inflammation indicators and their episode length after onset of inflammation using automatic milking system sensor data. *J. Dairy Sci.* 104:3458–3473. <https://doi.org/10.3168/jds.2019-18054>.
- 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.
- Dürr, J. W., R. I. Cue, H. G. Monardes, J. Moro-Méndez, and K. M. Wade. 2008. Milk losses associated with somatic cell counts per breed, parity and stage of lactation in Canadian dairy cattle. *Livest. Sci.* 117:225–232. <https://doi.org/10.1016/j.livsci.2007.12.004>.
- 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.

- Gonçalves, J. L., R. I. Cue, B. G. Botaro, J. A. Horst, A. A. Valloto, and M. V. Santos. 2018. Milk losses associated with somatic cell counts by parity and stage of lactation. *J. Dairy Sci.* 101:4357–4366. <https://doi.org/10.3168/jds.2017-13286>.
- 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. <https://doi.org/10.1016/j.prevetmed.2006.04.012>.
- Hadrich, J. C., C. A. Wolf, J. Lombard, and T. M. Dolak. 2018. Estimating milk yield and value losses from increased somatic cell count on US dairy farms. *J. Dairy Sci.* 101:3588–3596. <https://doi.org/10.3168/jds.2017-13840>.
- Hagnestam-Nielsen, C., U. Emanuelson, B. Berglund, and E. Strandberg. 2009. Relationship between somatic cell count and milk yield in different stages of lactation. *J. Dairy Sci.* 92:3124–3133. <https://doi.org/10.3168/jds.2008-1719>.
- Hagnestam-Nielsen, C., and S. Østergaard. 2009. Economic impact of clinical mastitis in a dairy herd assessed by stochastic simulation using different methods to model yield losses. *Animal* 3:315–328. <https://doi.org/10.1017/S1751731108003352>.
- Halasa, T., M. Nielsen, A. P. W. De Roos, R. Van Hoorne, G. de Jong, T. Lam, T. Van Werven, and H. Hogeveen. 2009. Production loss due to new subclinical mastitis in Dutch dairy cows estimated with a test-day model. *J. Dairy Sci.* 92:599–606. <https://doi.org/10.3168/jds.2008-1564>.
- Hand, K. J., A. Godkin, and D. F. Kelton. 2012. Milk production and somatic cell counts: A cow-level analysis. *J. Dairy Sci.* 95:1358–1362. <https://doi.org/10.3168/jds.2011-4927>.
- Hogeveen, H., W. Ouweltjes, C. De Koning, and K. Stelwagen. 2001. Milking interval, milk production and milk flow-rate in an automatic milking system. *Livest. Prod. Sci.* 72:157–167. [https://doi.org/10.1016/S0301-6226\(01\)00276-7](https://doi.org/10.1016/S0301-6226(01)00276-7).
- Hogeveen, H., W. Steeneveld, and C. A. Wolf. 2019. Production diseases reduce the efficiency of dairy production: A review of the results, methods, and approaches regarding the economics of mastitis. *Annu. Rev. Resour. Econ.* 11:289–312. <https://doi.org/10.1146/annurev-resource-100518-093954>.
- Hortet, P., F. Beaudeau, H. Seegers, and C. Fourichon. 1999. Reduction in milk yield associated with somatic cell counts up to 600 000 cells/ml in French Holstein cows without clinical mastitis. *Livest. Prod. Sci.* 61:33–42. [https://doi.org/10.1016/S0301-6226\(99\)00051-2](https://doi.org/10.1016/S0301-6226(99)00051-2).
- Hortet, P., and H. Seegers. 1998. Loss in milk yield and related composition changes resulting from clinical mastitis in dairy cows. *Prev. Vet. Med.* 37:1–20.
- International Dairy Federation. 2011. Suggested Interpretation of Mastitis Terminology (Revision of Bulletin of IDF N° 338/1999). Bulletin of the International Dairy Federation. International Dairy Federation.
- Jones, G. M., R. E. Pearson, G. A. Clabaugh, and C. W. Heald. 1984. Relationships between somatic cell counts and milk production. *J. Dairy Sci.* 67:1823–1831. [https://doi.org/10.3168/jds.S0022-0302\(84\)81510-6](https://doi.org/10.3168/jds.S0022-0302(84)81510-6).
- Mazeris, F. 2010. DeLaval herd navigator: proactive herd management. Pages 26–27 in Proceedings of First North American Conference on Precision Dairy Management.
- Nielen, M., H. Deluyker, Y. H. Schukken, and A. Brand. 1992. Electrical conductivity of milk: Measurement, modifiers, and meta-analysis of mastitis detection performance. *J. Dairy Sci.* 75:606–614.
- Nielen, M., Y. H. Schukken, J. Van De Broek, A. Brand, H. A. Deluyker, and K. Maatje. 1993. Relations between on-line electrical conductivity and daily milk production on a low somatic cell count farm. *J. Dairy Sci.* 76:2589–2596. [https://doi.org/10.3168/jds.S0022-0302\(93\)77593-1](https://doi.org/10.3168/jds.S0022-0302(93)77593-1).
- Nørstebo, H., G. Dalen, A. Rachah, B. Heringstad, A. C. Whist, A. Nødtvedt, and O. Reksen. 2019. Factors associated with milking-to-milking variability in somatic cell counts from healthy cows in an automatic milking system. *Prev. Vet. Med.* 172:104786. <https://doi.org/10.1016/j.prevetmed.2019.104786>.
- Nyman, A.-K., K. Persson Waller, T. W. Bennedsgaard, T. Larsen, and U. Emanuelson. 2014. Associations of udder-health indicators with cow factors and with intramammary infection in dairy cows. *J. Dairy Sci.* 97:5459–5473. <https://doi.org/10.3168/jds.2013-7885>.
- Oshima, M., T. Yoshida, K. Koyama, and T. Moriyama. 1990. A method for assessing loss in milk yield due to subclinical mastitis. *J. Dairy Res.* 57:319–324. <https://doi.org/10.1017/S0022029900026960>.
- Tyler, J. W., M. C. Thurmond, and L. Lasslo. 1989. Relationship between test-day measures of somatic cell count and milk production in California dairy cows. *Can. J. Vet. Res.* 53:182–187.
- Van Rij, J., M. Wieling, R.H. Baayen, and H. Van Rijn. 2017. itsadug: Interpreting time series and autocorrelated data using GAMMs. R Package, version 2.
- Viguer, C., S. Arora, N. Gilmartin, K. Welbeck, and R. O’Kennedy. 2009. Mastitis detection: Current trends and future perspectives. *Trends Biotechnol.* 27:486–493. <https://doi.org/10.1016/j.tibtech.2009.05.004>.
- Wood, P. D. P. 1967. Algebraic model of the lactation curve in cattle. *Nature* 216:164–165. <https://doi.org/10.1038/216164a0>.
- Wood, S. 2012. mgcv: Mixed GAM Computation Vehicle with GCV/AIC/REML smoothness estimation. Accessed Sep. 29, 2021. <https://cran.r-project.org/web/packages/mgcv/index.html>.
- Wood, S. N. 2013. On p-values for smooth components of an extended generalized additive model. *Biometrika* 100:221–228. <https://doi.org/10.1093/biomet/ass048>.
- Wood, S. N., Z. Li, G. Shaddick, and N. H. Augustin. 2017. Generalized additive models for gigadata: Modeling the UK black smoke network daily data. *J. Am. Stat. Assoc.* 112:1199–1210. <https://doi.org/10.1080/01621459.2016.1195744>.

## ORCID

- John Bonestroo  <https://orcid.org/0000-0001-8815-6357>  
 Mariska van der Voort  <https://orcid.org/0000-0003-0503-259X>  
 Nils Fall  <https://orcid.org/0000-0001-5597-2358>  
 Ulf Emanuelson  <https://orcid.org/0000-0001-7889-417X>  
 Ilka Christine Klaas  <https://orcid.org/0000-0002-1397-8505>  
 Henk Hogeveen  <https://orcid.org/0000-0002-9443-1412>