Invited review: Toward a common language in data-driven mastitis detection research

M. van der Voort,[1]* **D. Jensen,**[2] **C. Kamphuis,**[3] **I. N. Athanasiadis,**[4] **A. De Vries,**[5] **and H. Hogeveen**[1]
[1]Business Economics Group, Wageningen University & Research, 6706 KN Wageningen, the Netherlands
[2]Department of Veterinary and Animal Sciences, University of Copenhagen, 1870 Frederiksberg C, Denmark
[3]Animal Breeding & Genomics, Wageningen University & Research, 6708 PB Wageningen, the Netherlands
[4]Geo-Information Science and Remote Sensing Laboratory, Wageningen University & Research, 6706 KN Wageningen, the Netherlands
[5]Department of Animal Sciences, University of Florida, Gainesville 32611

ABSTRACT

Sensor technologies for mastitis detection have resulted in the collection and availability of a large amount of data. As a result, scientific publications reporting mastitis detection research have become less driven by approaches based on biological assumptions and more by data-driven modeling. Most of these approaches try to predict mastitis events from (combinations of) raw sensor data to which a wide variety of methods are applied originating from machine learning and classical statistical approaches. However, an even wider variety in terminologies is used by researchers for methods that are similar in nature. This makes it difficult for readers from other disciplines to understand the specific methods that are used and how these differ from each other. The aim of this paper was to provide a framework (filtering, transformation, and classification) for describing the different methods applied in sensor data–based clinical mastitis detection research and use this framework to review and categorize the approaches and underlying methods described in the scientific literature on mastitis detection. We identified 40 scientific publications between 1992 and 2020 that applied methods to detect clinical mastitis from sensor data. Based on these publications, we developed and used the framework and categorized these scientific publications into the 2 data processing techniques of filtering and transformation. These data processing techniques make raw data more amendable to be used for the third step in our framework, that of classification, which is used to distinguish between healthy and nonhealthy (mastitis) cows. Most publications (n = 34) used filtering or transformation, or a combination of these 2, for data processing before classification, whereas the remaining publications (n = 6) classified the observations directly from raw data. Concerning classification, applying a simple threshold was the most used method (n = 19 publications). Our work identified that within approaches several different methods and terminologies for similar methods were used. Not all publications provided a clear description of the method used, and therefore it seemed that different methods were used between publications, whereas in fact just a different terminology was used, or the other way around. This paper is intended to serve as a reference for people from various research disciplines who need to collaborate and communicate efficiently about the topic of sensor-based mastitis detection and the methods used in this context. The framework used in this paper can support future research to correctly classify approaches and methods, which can improve the understanding of scientific publication. We encourage future research on sensor-based animal disease detection, including that of mastitis detection, to use a more coherent terminology for methods, and clearly state which technique (e.g., filtering) and approach (e.g., moving average) are used. This paper, therefore, can serve as a starting point and further stimulates the interdisciplinary cooperation in sensor-based mastitis research.

INTRODUCTION

Since the 1980s, research and industry have been putting effort into developing sensors that can automatically detect mastitis by using changes in one or more milk characteristics (Gebre-Egziabher et al., 1979; Fernando et al., 1982). The introduction of automatic milking systems in the 1990s required a system to detect clinical mastitis without visual observations, which boosted the development and applicability of sensors to automatically detect clinical mastitis. The first sensor developed for mastitis detection was electrical

conductivity, which is also one of the most widely studied sensor systems (Rutten et al., 2013). Other sensors that have been developed are direct and indirect SCC measurement (Løvendahl and Sørensen, 2016; Deng et al., 2020), L-lactate dehydrogenase sensors (Jørgensen et al., 2016), color sensors (Song and Tol, 2010), and biosensor-based techniques (Dalen et al., 2019). Despite all these efforts, the detection of mastitis with sensors remains challenging. Currently, none of the published scientific studies of automatic systems to detect clinical mastitis meet the demanded 80% sensitivity and 99% specificity (ISO, 2007; Hogeveen et al., 2010; Dominiak and Kristensen, 2017). With the prospect that advanced methods in machine learning, and the integration of more sources of data can improve clinical mastitis detection (Kamphuis et al., 2010b; Khatun et al., 2018a; Ebrahimi et al., 2019), more studies can be expected on this topic.

The published scientific studies show a wide range of diversity in applied methods, from simple thresholds applied to raw sensor data to complex machine learning methods (Hogeveen et al., 2010; Dominiak and Kristensen, 2017; Slob et al., 2021). One aspect that becomes clear from a review of the existing literature is that there is not only a wide range of applied methods, but researchers also tend to use different terminology for similar methods. This makes it difficult for people from different research disciplines to get a clear understanding of the methods used and how they differ from each other, which may potentially hinder the development of new and better detection models.

The aim of this paper was to provide a framework (filtering, transformation, and classification) for describing the different methods applied in sensor data–based clinical mastitis detection research and use this framework to review and categorize the approaches and underlying methods described in the scientific literature on mastitis detection. We seek to provide a structured way for data-driven mastitis researchers to communicate the approach they have taken and provide more terminological clarity and a starting point for (interdisciplinary) discussions and future research.

FRAMEWORK TO CATEGORIZE APPROACHES

A wide range of data-driven methods are described in the scientific literature to process raw sensor data into useful information for clinical mastitis detection in dairy cows. In this context, we define information as understanding numbers and words with relationships which enriches knowledge (understanding patterns) of the decision maker (e.g., researcher, farmer, advisor) about the disease status of a cow (Zins, 2007). With this knowledge, decisions can be taken. In the context of clinical mastitis detection, raw (sensor) data are typically converted into information in the form of an alert if a given cow is believed to have clinical mastitis, or no alert if it is believed she does not have mastitis. To get from raw data to information, the framework consists of 3 data analysis techniques, namely time-series filtering (henceforth called filtering), transformation, and classification (Figure 1). Various transformation and filtering steps can be used together or separately in sequential order.

The starting point of the framework is raw sensor data, which are directly generated by sensors. In relation to mastitis, raw sensor data include measurements such as electrical conductivity, SCC, L-lactate dehydrogenase, milk color values, milk yield, and milk temperature of the (quarter) milk (Hogeveen et al., 2010). Raw data are difficult to interpret in practice and therefore (pre-) processing of the data is often an important step to make optimal use of the model capacity of classification systems (Eradus and Jansen, 1999). Two pre-
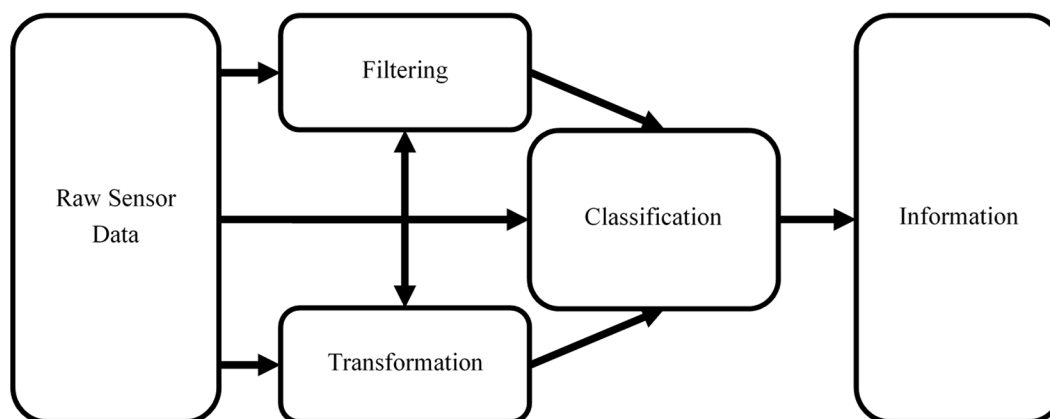
Figure 1. Framework describing the 3 techniques to convert raw sensor data into information in sensor-based mastitis detection research.

processing techniques, filtering and transformation, are typically applied to make raw data more amendable to classification.

Filtering is a collection of pre-processing methods used to remove the random noise and unrealistic values to determine the underlying systematic patterns in the data. Filtering makes data more amendable for further analysis. The representation of the data stays the same, however. Examples of well-known filtering approaches are the moving average (**MA**; e.g., Cavero et al., 2007) and the Kalman filter (e.g., de Mol et al., 1997).

Transformation is a collection of pre-processing methods that makes input data more amendable for analyses by transforming data from one representation to another. It changes the observed value into a transformed value, for example, by applying a logarithmic transformation to make the data more normally distributed or by a standardization transforming the observed value to a standardized value. Examples of well-known data transformations are normalization (e.g., Ebrahimie et al., 2018), standardization (e.g., Kamphuis et al., 2008b), and also the calculation of summary statistics such as mean, median, and the standard deviation (e.g., Steeneveld et al., 2010b).

As a hypothetical example, the raw electrical conductivity values measured by a sensor during milking may be transformed by the sensor system itself, if the sensor system only reports the mean value per quarter recorded during the session (first transformation). The 4 mean quarter-values may then be ordered and used to calculate the interquarter ratio between the highest and the lowest quarter electrical conductivity value (second transformation). Some filtering method such as the Kalman filter could then be applied to the time series of interquarter ratios values throughout the lactation of the cow (first filtering). The forecast errors between observed and expected interquarter ratios values could then be calculated (third transformation) along with the variance of the forecast, which would then enable the transformation into the standardized forecast error (fourth transformation). Filtering and transformation are at the most a collection of pre-processing in-between steps. These pre-processing steps always need to be followed by classification.

Classification is used to convert the data into a categorical value to create an alert for mastitis detection. Alternatively, a continuous value can be presented that can be interpreted by the user, such as "probability." The data input for classification can be either raw or pre-processed by filtering or transformation or both. Various approaches for classification exist such as directly applying a threshold (e.g., Cavero et al., 2007), logistic regression (e.g., Khatun et al., 2018b), decision trees (e.g., Kamphuis et al., 2010a), Bayesian classifiers

(e.g., Jensen et al., 2016), and artificial neural networks (**ANN**; e.g., Mammadova and Keskin, 2015).

PUBLICATIONS ON DATA-DRIVEN CLINICAL MASTITIS DETECTION

We systematically selected peer-reviewed publications and conference proceedings dealing with sensor-based clinical mastitis detection using electronic search engines. The search was conducted using the scientific literature electronic databases CAB (CAB International, Oxon, UK), Medline (National Library of Medicine, Rockville Pike, MD), and ISI Web of Knowledge (Thomson Reuters, New York, NY). Additionally, the snowball method was applied to track down relevant publications in key publications (Wohlin, 2014). The search was first conducted in January 2020 and to ensure the inclusion of the latest publications it was repeated in December 2020. Studies included in this search focused on sensor-based clinical mastitis detection published between 1990 and 2020, and needed to be written in English. Publications from conference proceedings were included if they had not been published in peer-reviewed journals. We excluded publications that (1) did not use sensor data related to clinical mastitis detection, (2) focused on predicting mastitis patterns, or (3) reviewed the literature.

Figure 2 shows the number of peer-reviewed publications on sensor-based clinical mastitis detection in 5-yr intervals between the years 1990 and 2020. In total, 40 publications that met our requirements have been published since 1992, and the number of publications increased considerably after 2005. Seven publications used a classification technique directly, which implied that no data pre-processing with filtering or a transformation technique was applied. In the remaining 33 publications, 6 publications pre-processed the data by applying filtering only, 10 publications only applied transformation, and 17 publications used a combination of filtering and transformation before classification.

Approaches and Use of Methods

Among the 3 techniques comprising the conceptual framework (i.e., filtering, transformation, and classification), many different approaches have been applied in sensor-based clinical mastitis detection research. Table 1 gives an overview of the approaches and underlying methods for filtering, transformation, and classification that were identified in the reviewed literature. Approaches are defined as the overarching term for a group of similar methods, whereas the method is an actual implementation within an approach. In the publications reviewed we found that researchers tend to use
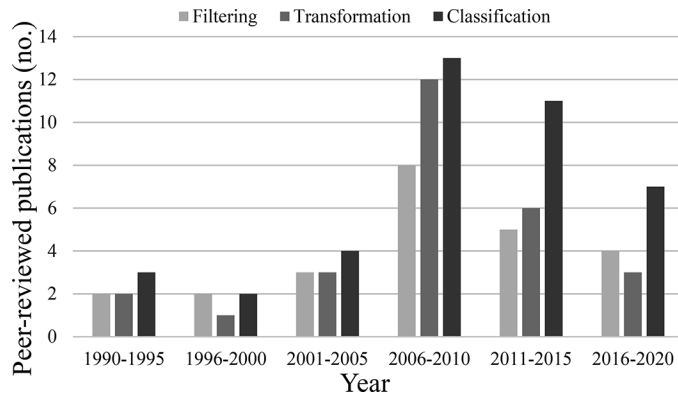
Figure 2. Number of peer-reviewed publications per in-between technique on sensor-based clinical mastitis detection in the period 1990 to 2020.

different terminologies for similar or the same methods. In some publication this is because a derivative of the approach is used, but for others no differences seems to exist between the methods and only a different terminology was used (Table 1).

We found the largest number of different methods for the MA approach. This filtering approach was used in 13 publications, and was mostly referred to as rolling average or MA. The exponentially weighted MA (**EWMA**) was used in 2 publications, but with different terminologies. Regression approaches are found in both filtering (i.e., autoregressive models) and classification (i.e., regression models), but applied different underlying methods, for example, handling categorical or continuous variables.

Under transformation, 11 different approaches were identified. From the 27 publications that used transformation, 12 publications do not specifically point out that they used a transformation method, which challenged the identification of whether transformation took place.

For classification also, different methods were found for similar approaches. In case of ANN, 4 different methods were found in 5 different publications. For the decision tree analysis approach, 5 different methods were found in 3 publications. A frequently used approach within classification was the threshold, which was applied in 19 publications and consistently called by the same name, "threshold." In the next sections, the approaches within each of the 3 techniques are worked out more in depth.

Filtering

In the 24 publications that applied filtering, a variety of filtering approaches has been used, including MA,

Kalman filters, linear regression updates, autoregressive models, wavelet filters, and attribute weighting. A detailed overview of the filtering methods found in our literature review is provided in Table 2.

The MA has been used for mastitis detection as far back as 1992 (Maatje et al., 1992). A variation on the MA approach was found in the EWMA, but also various methods were found within the publications. Maatje et al. (1992), for example, took the MA of 3 electrical conductivity measurements of each quarter, as was also done by Hovinen et al. (2006) (called running average), and compared this to the reference quarter (i.e., quarter with the lowest MA). When the MA deviated by more than 15% from the reference quarter, and when the actual quarter electrical conductivity exceeded 20% above the reference value, the quarter was identified as infected. Khatun et al. (2017) instead used the previous 3 to 9 electrical conductivity observations of each quarter (called rolling average), and Miekley et al. (2013b) used the 5-d MA of electrical conductivity (called MA).

Although dynamic linear models (**DLM**) were described as a better approach to handle time series of variables that are expected to change systematically over time and are cow specific (Van Bebber et al., 1999; de Mol et al., 2001), they have not been frequently used in mastitis detection, as seen both in Table 2 and described by Dominiak and Kristensen (2017). A DLM was first applied to the problem of mastitis detection in the late 1990s (de Mol et al., 1997, 1999, 2001). Furthermore, Jensen et al. (2016) applied a multivariate DLM to various sensor data collected on the individual cow at milking, where the DLM was implemented with the assumption of a time-dependent linear trend. This assumption of a time-dependent trend was not used in earlier studies. In all cases, though, the filtered parameter vectors were used to make one-step forecasts of all modeled variables. In the case of de Mol et al. (1997, 1999, 2001), this classification was based on simple thresholds approaches on the forecast errors. In the case of Jensen et al. (2016), the forecast errors were transformed (i.e., standardized and categorized) based on their direction and magnitude before being used as inputs for a naïve Bayesian network, which was then used for classification.

Other filtering approaches (i.e., autoregressive models, wavelet filtering, and differencing) were used in 5 publications. Wavelet filtering aims to transform signals into coefficients in which it reconstructs the sensor data in such a way that it removes the noise from the sensor data. The wavelet filter was found to be a relative new filtering approach and the first such study in relation to sensor-based mastitis detection was found in Miekley et al. (2012).

Table 1. Techniques with the approaches and underlying methods used in the 40 publications between 1992 and 2020

Technique	Approach (no. of publications)	Methods
Filtering	Moving average (MA) (13)	Simple MA
		MA model
		Running average
		Rolling average
		Running mean
		Gaussian transfer function
		Fractional Brownian motion
		Steady-state model
		Autoregressive MA
		Locally weighted (polynomial) regression
	Exponentially weighted MA (2)	Exponentially weighted MA
		Exponential smoothing
	Autoregressive model (3)	Autoregressive model
		Vector autoregressive model
		Linear regression iteratively updated
	Differencing (1)	—
	Dynamic linear model (6)	Kalman filter
	Wavelet filtering (2)	—
	Attribute weighting method (1)	—
Transformation	Standardization (4)	—
	Normalization (6)	—
	Summary statistics (6)	—
	Linear interpolation (2)	—
	Mathematical transformation (3)	—
	Categorization (1)	—
	Fuzzyfication (4)	—
	Interquarter ratio (7)	—
	Cumulative sum (CUSUM; 2)	—
	Principal component analysis (3)	—
	Subset selection (1)	—
Classification	Threshold (19)	Simple threshold
		CUSUM control charts
		Shewhart control charts
	Artificial neural network (ANN) (6)	Neural network
		ANN: backpropagation
		ANN: multilayer feed forward
		ANN: self-organizing map
		Multilayer perceptron
	Regression model (7)	Linear mixed model
		Regression analysis
		Partial least squares regression
		General linear model
		Logistic regression
		Generalized linear model
	Discriminant function analysis (1)	—
	Fuzzy inference (6)	Fuzzy logic
		Adaptive neuro-fuzzy inference system
	Decision tree (3)	Stump decision tree
		Parallel decision tree
		Random forest
		J48 (C4.5) algorithm
		Gradient boosting
	Naïve Bayesian classifier (2)	Naïve Bayesian networks
	Support vector machines (2)	—
	Latent class analysis (1)	Latent class model
		Latent variable method
		Factor analysis

Transformation

Various approaches for transforming either the raw or filtered data were identified in the publications included in this review. Table 3 provides an overview of these approaches and methods.

From the 40 publications included in this review, 27 publications used at least one transformation approach. Interquarter ratios were the most commonly used method, appearing in 7 publications. The advantage of the interquarter ratio is that it simplifies 4 values (the 4 per-quarter electrical conductivity observations) to

Table 2. Names and descriptions of the approaches used for filtering in relation to sensor-based mastitis detection research, as published in 40 publications on sensor-based mastitis detection between 1992 and 2020

Approach	Description	Publications[1]
Moving average (MA)	At each time step, the filtered value, z_t, is defined as the simple mean of the n most recently observed values, with n being a predefined integer value called the window length. Forecasts for a new observation at a given time t is given as the filtered value, z_{t-1}, at time $t-1$. The forecast variance to time t is estimated as $\sigma_{z_t}^2 \approx \dfrac{\sigma^2}{n}$, where σ^2 is the variance of the observed values in the rolling window.	Maatje et al. (1992) Nielen et al. (1995a) Cavero et al. (2006) Hovinen et al. (2006) Cavero et al. (2007) Cavero et al. (2008) Kamphuis et al. (2008b) Steeneveld et al. (2010c) Sun et al. (2010) Miekley et al. (2013b) Huybrechts et al. (2014) Khatun et al. (2017) Dalen et al. (2019)
Exponentially weighted MA	At each time step, the filtered value is defined according to the following equation: $z_t = \lambda \cdot k_t + (1-\lambda) \cdot z_{t-1}$, where λ is a weight factor that can take values between 0 and 1, and k_t is the observed value at time t. Forecasts for a new observation at a given time t are given as the filtered value, z_{t-1}, at time $t-1$. The forecast variance is estimated as $\sigma_{z_t}^2 \approx \sigma^2 \cdot \left(\dfrac{\lambda}{2-\lambda}\right)$, where σ^2 is the variance of all the observed values up to time $t-1$.	Cavero et al. (2007) Ankinakatte et al. (2013)
Autoregressive model	At each time step, the filtered value is given by the following equation: $z_t = \beta_0 + \beta_1 \cdot z_{t-1} \ldots + \beta_n \cdot z_{t-n} + \varepsilon_t$, where β_0 is the intercept, β_1 to β_n are the model coefficients applied to the 1 to n previous observations of the time series, and ε_t is the normally distributed error term at time t.	de Mol and Ouweltjes (2001) Miekley et al. (2013a) de Mol and Woldt (2001)
Differencing	A method used to remove trends or seasonality from time-series data. Trends are removed using lag -1 differencing, which takes the difference between 2 neighboring values: $Y_t - Y_{t-1}$. Cyclic patterns are removed using lag $-M$ differencing: $Y_t - Y_{t-M}$, where M is the number of time steps per cycle. Differencing is typically used as an initial step before applying the above listed filtering methods.	Miekley et al. (2013a)
Dynamic linear model	At each time step, the filtered value of one or more observable variables, as well as any unobservable trend components that might apply to these variables, are defined by the system equation: $\boldsymbol{\theta}_t = \mathbf{G}_t \cdot \boldsymbol{\theta}_{t-1} + w_t$, where $\boldsymbol{\theta}_t$ is the parameter vector, $\mathbf{G}_t$ is the system matrix, and the error term is defined as $w_t \approx N(\underline{0}, \mathbf{W})$, with $\mathbf{W}$ being the systematic co-variance matrix. Forecasts of the next observation are made according to the observation equation: $Y_t = \mathbf{F}_t' \cdot \boldsymbol{\theta}_t + v_t$, where $\mathbf{F}_t'$ is the (transposed) design matrix, and the error term is defined as $v_t \approx N(\underline{0}, \mathbf{V})$, with $\mathbf{V}$ being the observational co-variance matrix. Given the forecast errors for each observation time, the parameter vector values are updated using the Kalman filter (West and Harrison, 1997). The forecast variance is estimated during Kalman filtering. Notice that (unlike other filtering methods in this table) the forecasted values might deviate from the filtered values at time $t-1$, depending on the definition of the design matrix.	de Mol et al. (1997) de Mol et al. (1999) de Mol et al. (2001) de Mol and Woldt (2001) Friggens et al. (2007) Jensen et al. (2016)
Wavelet filtering	The entire time series up to a given observation time, t, is filtered using several wavelets (i.e., short finite wave patterns of various shapes), which always sum to 0. These wavelets can be stretched or squeezed to maximize the fit of the wavelet to the data. Forecasts for each observation time, t, are defined as the filtered value at that time.	Miekley et al. (2012) Miekley et al. (2013a)
Attribute weighting method	Specifying a set of attributes that describe the value-relevant properties of outcomes. Therefore, single-attribute value functions are assessed over the levels of each attribute, and attribute weights that govern the rate of substitution of value across attributes are determined.	Ebrahimie et al. (2018)

[1]The publications where the various types of filtering were applied. Some studies use more than one approach to filter the data and thus appear multiple times in this table.

a single value. Moreover, it enables the comparison of mastitis indicators within the cow. Data transformation by various simple summary statistics and normalization were the second most applied approach; both appeared in 6 publications. Claycomb et al. (2009) normalized maximum quarter electrical conductivity of the individ-

Table 3. Names and descriptions of the approaches used for transformation in relation to sensor-based mastitis detection research, as published in 40 publications on sensor-based mastitis detection between 1992 and 2020

Approach	Description	Publications[1]
Standardization	Standardization of a numerical variable with an arbitrary normal distribution: $x \sim N\left(\mu,\sigma^2\right)$; adjusts the values to make them follow a standard normal distribution, $x \sim N\left(0,1^2\right)$, by adjusting for the (known or estimated) mean and SD of the original distribution: $x' = \dfrac{\left(x - \mu\right)}{\sigma}$, where x is the observation, μ is the mean of a given distribution, σ^2 is the standard deviation of a given distribution, and x' is the standardized variable.	de Mol and Woldt (2001) Miekley et al. (2012) Huybrechts et al. (2014) Jensen et al. (2016)
Normalization	Normalization is generally used to adjust numerical values to be within a specific range. Most commonly used (from our reading of the literature) is the Min-Max feature scaling, which scales the values to the range [0;1]: $x' = \dfrac{x - \min\left(x\right)}{\max\left(x\right) - \min\left(x\right)}$, where x' is the normalized variable, x is the observation, min is the minimum of the given observations, and max is the maximum of the given observations.	Cavero et al. (2007) Kamphuis et al. (2008b) Claycomb et al. (2009) Sun et al. (2010) Khatun et al. (2017) Ebrahimie et al. (2018)
Summary statistics	Condensed numerical descriptions of a collection of data, such as mean, median, SD, range, and so on.	Nielen et al. (1995a) Norberg et al. (2004) Cavero et al. (2006) Cavero et al. (2008) Kamphuis et al. (2010a) Steeneveld et al. (2010c)
Linear interpolation	The unobserved values between 2 observed values are estimated using an assumed (linear) relationship over time between the 2 observed values.	de Mol and Ouweltjes (2001) de Mol and Woldt (2001)
Mathematical transformation	The data are transformed using some transformation function (e.g., a sigmoidal function or a logarithmic function).	Friggens et al. (2007) Mollenhorst et al. (2010) Kamphuis et al. (2010b)
Categorization	A numerical value is categorized by assigning to it a hard membership to a category, based on which of a set of predefined intervals, the value is found within.	Jensen et al. (2016)
Fuzzyfication	A soft categorization, in which a numerical value is assigned a relative membership [0;1] to multiple categories with overlapping intervals. Fuzzyfication is an essential transformation step to use the fuzzy logic method of classification.	de Mol and Woldt (2001) Cavero et al. (2006) Mikail and Keskin (2015) Mammadova and Keskin (2015)
Interquarter ratio (IQR)	Can be calculated when values are measured at each teat of the udder separately. Different specific definitions of IQR are used in different studies, including the following: relative deviation of quarter electrical conductivity (EC) from the lowest EC quarter, the EC value of a given quarter divided by the average value of the 2 lowest quarters, the maximum quarter EC value divided by the minimum quarter EC value, and EC of the quarter minus the mean EC of all quarters.	de Mol et al. (1999) de Mol and Ouweltjes (2001) de Mol and Woldt (2001) Norberg et al. (2004) Hovinen et al. (2006) Mollenhorst et al. (2010) Steeneveld et al. (2010c)
Cumulative sum (CUSUM)	A CUSUM is a sequence of partial sums of a given sequence of numerical values. For example, the simplest form of the cumulative sums of the sequence {a,b,c...} is a,a + b,a + b + c ... The CUSUM is typically applied to the forecast errors produced by various time-series filtering methods. Lower limits on the absolute deviation from 0 may be applied to define the magnitude of the forecast error required to update the CUSUM value. When the CUSUM is only allowed to reach values in either the positive or negative direction, it is referred to as a one-sided CUSUM. As an example, for a one-sided CUSUM which only allowed to reach positive values, the CUSUM value to time t will be updated as follows: $\text{CUSUM}_t = \text{CUSUM}_{t-1} + \max\left(0, e_t - k\right)$, where e_t is the forecast error at time t and k is the reference value which depends on the assumed statistical distribution and change of most interest. A multivariate form of the CUSUM (MCUSUM) also exists. A threshold which depends on the false alarm rate is usually applied to the CUSUM value to classify the observations as being either in or out of control.	Miekley et al. (2013a) Huybrechts et al. (2014)
Principal component analysis (PCA)	PCA is a procedure for converting a set of multivariate observations into a set of values of linearly uncorrelated variables called principal components. By definition, the first principal component accounts for the most variability in the data, and each succeeding component in turn accounts for a decreasing amount of variation. In the case of mastitis detection (and similar research), PCA can be used to transform multivariate observations into (usually) the 2 or 3 most informative principal components. These are subsequently used as inputs for the later classification step.	Sun et al. (2010) Soyeurt et al. (2012) Miekley et al. (2013c)
Subset selection	A subset of the available data is selected and used in further steps. For mastitis detection, this can be when all EC measurements during a milking session are recorded, but only the first, middle, and last minute of data are used.	Nielen et al. (1995b)

[1]The publications where the various transformations were applied. Some studies use more than one approach to transform the data and thus appear multiple times in this table.

ual cow against the running mean of the herd over the 5 most recent milkings. In the publication of Kamphuis et al. (2008b) SCC data were $\log_{10}$ transformed. Subset selection and principal component analysis, applied in 2 and 3 publications, respectively, were used to select observations or variables that best represent the data. Overall, most publications used data transformation approaches to prepare data for classification and made necessary adjustments to make the values of different variables with different magnitudes and variances more similar to each other.

Classification

Classification was applied in all publications and is used a binary system to distinguish between healthy and nonhealthy (mastitis) cows. Table 4 provides an overview of the approaches used to classify. As per definition, all of the reviewed publications used a classification method. In 19 of the 40 publications, thresholds were used, making it by far the most widely used classification method in sensor-based mastitis detection research. Thresholds were combined with filtering (5 out of the 19 publications), transformation (4 out of the 19 publications), or both (10 out of the 19 publications). Setting a decision threshold was already applied by Maatje et al. (1992) and was still used by Dalen et al. (2019). The ways in which thresholds were applied varied between the publications. For example, Maatje et al. (1992) compared the MA of each electrical conductivity quarter of the individual cow to the reference quarter (i.e., quarter with the lowest MA). When the deviation between the MA of a given quarter and the reference quarter was >15% or >20% of the MA of the reference quarter during 2 successive milkings, a mastitis alert was raised. de Mol and Ouweltjes (2001) and de Mol et al. (2001) set thresholds on the difference between actual and predicted values (i.e., the forecast errors) of their Kalman filter. When the forecast error fell outside the 95%, 99%, or 99.9% confidence interval, a mastitis alert was raised. Other publications used thresholds based on general knowledge on SCC, electrical conductivity, and online cell counts to raise a mastitis alert (Friggens et al., 2007; Mollenhorst et al., 2010; Khatun et al., 2017; Dalen et al., 2019).

Alternatively to the use of thresholds, the classification can be done either by (a) regression, that is, first apply a regression model and then a threshold to convert the predicted value into a binary alert, or a classification method that produces a label, typically "healthy" or "sick." All published papers in this review labeled clinical mastitis as a binary trait (healthy or mastitis). This approach might be triggered by the performance requirements for automatic clinical mastitis

detection (specificity and sensitivity), as well as the practical situation where farmers use these models to find clinical mastitis cases to start a timely treatment (Mollenhorst et al., 2012a). However, this does not mean that this binary approach is necessarily correct and there are publications that speak about the desire for researchers and farmers to have a prediction on a continuous scale, or multiple classes (e.g., healthy, subclinical, clinical with no immediate treatment, and clinical with immediate treatment; Steeneveld et al., 2009; Højsgaard and Friggens, 2010; Kamphuis et al., 2011). Acknowledging different severity levels of mastitis is expected to improve mastitis detection (Dominiak and Kristensen, 2017; Dalen et al., 2019; Hyde et al., 2020; Bonestroo et al., 2021).

In total, 7 publications applied a form of regression approach. Nielen et al. (1995a) applied a logistic regression model to relate udder SCC with a mastitis indicator for infected and noninfected cows. A generalized linear model was applied by Norberg et al. (2004) to distinguish between healthy and nonhealthy cows based on electrical conductivity. We did not find the use of nonlinear regression methods in any of the publications reviewed.

In total, 6 publications on mastitis detection used ANN in the past 20 years. The first publications were published in 1995 and conducted by (Nielen et al., 1995a,b) and the latest publication is from Mammadova and Keskin (2015). The data source for the ANN differed, including raw electrical conductivity data (Cavero et al., 2008; Mammadova and Keskin, 2015), summary statistics on the MA of electrical conductivity (Nielen et al., 1995a), and MA of normalized electrical conductivity data (Sun et al., 2010) and EWMA (Ankinakatte et al., 2013).

Fuzzy logic was applied in 5 of the publications and was first used in 2001 (de Mol and Woldt, 2001). Fuzzy logic is perceived as a more sophisticated approach compared with regression analysis (Kamphuis et al., 2008b). The input data in the fuzzy logic models varied from raw data to filtered data and transformed data (de Mol and Woldt, 2001; Cavero et al., 2006; Mikail and Keskin, 2015).

Applications of decision trees were found in 2 publications (Kamphuis et al., 2010a; Ebrahimie et al., 2018). The advantage of decision trees in comparison with the previously described categorization approaches is that they provide an easy-to-interpret set of rules for how exactly an observation is classified. Decision trees are better than most other methods at handling unbalanced data sets. Single tree analysis were often combined with boosting and bagging methods to get a clear view on improvements. A related method, random forest, was only used in 1 publication (Ebrahimie et al., 2018),

Table 4. Names and descriptions of the approaches used for classification in relation to sensor-based mastitis detection research, as published in 40 publications on sensor-based mastitis detection between 1992 and 2020

Approach	Description	Publications[1]
Threshold	A threshold value is defined for and applied to either the raw, filtered, or transformed data to classify the cow as healthy or sick. An example of a threshold method is the Shewhart control chart. These charts contain a baseline value, which the values are expected to be normally distributed around, as well as an upper and lower control limit. The basic rule is that an alarm is generated when observations fall outside the limits. Other threshold rules may be applied [e.g., that the values must be outside of the limits for a certain number of consecutive observations, or all be on the same side of baseline for a (higher) number of consecutive observations].	Maatje et al. (1992) de Mol et al. (1997) de Mol et al. (1999) de Mol and Ouweltjes (2001) de Mol et al. (2001) Norberg et al. (2004) Cavero et al. (2007) Friggens et al. (2007) Hovinen et al. (2006) Claycomb et al. (2009) Mollenhorst et al. (2012b) Miekley et al. (2012) Miekley et al. (2013a,c) Huybrechts et al. (2014) Addis et al. (2016) Jensen et al. (2016) Khatun et al. (2017) Dalen et al. (2019)
Artificial neural network (ANN)	The ANN is a data processing method, which is loosely inspired by biological brains. In their most basic form, input data are transformed to output values (e.g., a numerical value in the [0;1] range in the base of binary classification) through a series of matrix multiplications, each of which is followed by a nonlinear transformation.	Nielen et al. (1995a,b) Cavero et al. (2008) Sun et al. (2010) Ankinakatte et al. (2013) Mammadova and Keskin (2015)
Regression model	A logistic regression model describes the logarithm of the odds (logit) of a binary variable, for example, mastitis, being true (1), given a linear combination of one or more independent variables: $ln\left(\dfrac{p}{1-p}\right) = \beta_0 + \beta_1 \cdot x_1 + \ldots + \beta_n \cdot x_n$, where p is the probability of the binary variable being true, β_0 is the intercept, and β_1 to β_n are the model coefficients applied to the values of the 1 to n independent variables, (i.e., x_1 to x_n).	Nielen et al. (1995a) Kaşikçi et al. (2012) Ankinakatte et al. (2013) Mammadova and Keskin (2013) Sathiyabarathi et al. (2016) Khatun et al. (2018b) Soyeurt et al. (2012)
Discriminant function analysis (DFA)	DFA is a method for reducing the dimensions of multivariate observations with known classes in such a way that the distance between the means of the 2 (or more) classes is maximized and the variance within the classes is minimized. Once such a transformation has been defined based on training data, new observations can be transformed. The transformed new observation is classified as belonging to the class where it has the smallest distance to the class mean.	Norberg et al. (2004)
Fuzzy inference	Fuzzy inference contains 3 phases: (1) fuzzification, which turns numeric input variables into linguistic ones, (2) fuzzy inference, where linguistic variables are applied to simple if–then rules, creating a new linguistic (likelihood) variable, and (3) defuzzification, where the cumulative result of the linguistic likelihood variable is transformed back into single value, which is then compared with a threshold to discriminate between health and sick.	de Mol and Woldt (2001) Kamphuis et al. (2008b) Cavero et al. (2006) Mikail and Keskin (2015) Mammadova and Keskin (2015)
Adaptive neuro-fuzzy inference	Adaptive neuro-fuzzy inference is a type of ANN that integrates fuzzy logic principles. Its inference system corresponds to a set of fuzzy if–then rules that have learning capability to approximate nonlinear functions.	Mammadova and Keskin (2015)
Decision tree	A decision tree classifies an observation based on a series of yes or no questions about variables describing that observation.	Kamphuis et al. (2010a,b) Ebrahimie et al. (2018)
Random forest	A random forest consists of a multitude of decision trees, each of which is made with a bootstrapped version of the training data. The output of a random forest is the aggregate of the outputs of the trees that make up the random forest.	Ebrahimie et al. (2018)
Naïve Bayesian classifier (NBC)	The NBC is a simple form of a Bayesian network, in which all input variables are assumed to be mutually independent.	Steeneveld et al. (2010a,b) Jensen et al. (2016)
Support vector machines (SVM)	SVM represent the data as points in a higher-dimensional space, then defines the hyperplane which best separates the 2 classes of the data. The position of the hyperplane is the midpoint between 2 data points, one from each class. These data points are called support vectors. Which data points are used for support vectors is determined by cross-validation on the training set. New observations are classified based on what side of the hyperplane they are located on.	Mammadova and Keskin (2013) Miekley et al. (2013b)
Latent class analysis (LCA)	LCA is a measurement model in which individuals can be classified into mutually exclusive and exhaustive types, or latent classes, based on individual patterns of behavior on a set of categorical indicator variables. Each individual is probabilistically assigned to a class.	Addis et al. (2016)

[1]The publications where the various classification were applied. Some studies use more than one approach to classify the data and thus appear multiple times in this table.

even though this method generally outperforms a single decision tree in direct comparisons (Ali et al., 2012). Both decision trees are, however, difficult to interpret, making them less useful to understand biological aspects of mastitis, thus making the random forest less desirable for studies where understanding the biology of mastitis is prioritized over predicting mastitis.

The naïve Bayesian classifier approach was found in 3 publications and was found to be a relatively novel method in mastitis detection. It appears to be a powerful tool in presenting probability distributions and for probabilistic inference (Steeneveld et al., 2010b; Jensen et al., 2016). The naïve Bayesian classifier is, like decision trees, good in handling missing data, captures the dependencies between target variables (e.g., mastitis yes or no) and each of the input variables in isolation, and describes prior and posterior probabilities of variables.

Other reported approaches in classification were support vector machine (Mammadova and Keskin, 2013; Miekley et al., 2013b) and latent class analysis (Addis et al., 2016). The 2 publications on support vector machine were both published in 2013 and studied the usefulness of this approach in mastitis detection. Support vector machine was described as performing well compared with other classical statistical or machine learning approaches in relation to biological problems because it can handle nonvector inputs as variable length sequences or graphs, which are common in biology.

DISCUSSION AND CONCLUSIONS

This study reviews the scientific publications in clinical mastitis detection based on sensor-based data in the period 1990 and 2020. In total, 40 publication were found with the highest number of publications found in the period 2006 to 2010. After this period we found a decreasing tendency in number of publication. A reason for finding fewer publications after 2010 might be that mastitis detection sensor systems are well established in dairy farming and so fewer people are interested in performing research on this topic. Despite this, it became clear from the literature that the performances of mastitis detection sensor systems did still not reach the demanded performance levels as defined by Rasmussen (2005) and that of International Organization for Standardization (ISO, 2007). Another explanation could be found in the peak in publication in 2010, 6 in total, of which 4 originated from the same research group, whereas the publications found in later years are from a larger number of research groups.

Categorization of approaches and methods used in the publications studied in this review was challenging.

First, a large number of publications did not explain how raw data are handled. One needs to carefully read many of the publications and read between the lines to find out what methods the authors really used. Most publications did not explicitly mention the technique, approach or method used. For example the paper of Norberg et al. (2004) does not specifically state which type of model was used (only stated they used the GLM approach), but from the context in the paper it is most likely they used a logistic regression. Expertise in the methods underlying the approaches was therefore needed to categorize the publications. However, before being able categorize publications, even within our group of experienced co-authors there were intense discussions regarding the categorization of approaches and methods. A reason for the difficulty in defining the different techniques, approaches, and methods were the differences in our backgrounds. Terminology differed between the research disciplines of our interdisciplinary team of authors (from computer science, statistics, dairy science, and agricultural economics) as well as in the teams of authors in the reviewed publications. Overall this challenge showed that a good understanding of methods is needed to be able to identify if something is a new method or if the method resulted in improved mastitis detection. In addition to the ever-changing definition of the problem, and lack of open data, a lack of common terminology does not allow comparisons across systems. This paper can serve as a reference for people from various research backgrounds who need to be able to collaborate and communicate efficiently about the topic of mastitis detection and the methods used in this context. The framework we used in this paper could also be used by other researchers to get a more structured overview of scientific publications, which subsequently could improve the understanding of these papers.

A wide range of methods is described in the literature. Often they seem to be novel when new terminology is used, but frequently they are just another term for a method or approach that was already used in earlier research. Despite this, improvements in algorithms can be made in better filtering, better transformations, and better classification. New ways to improve mastitis detection performances are emerging (Hogeveen et al., 2010; Dominiak and Kristensen, 2017; Slob et al., 2021). The utilization of multiple data sources is indicated as a way to improve the detection of mastitis (Kamphuis et al., 2008a; Jensen et al., 2016; Sathiyabarathi et al., 2016). However, the fact is that they can still be considered as "small data," whereas a lot of advancement offered by popular approaches such as in deep learning comes from the availability of huge data sets. Other advanced ways to improve mastitis detection

include the use of image analysis of the udder with a thermal imaging system, or taking pictures of the milk to detect flakes (Zhang et al., 2020). Therefore, deep convolutional neural network approaches are known to be useful for handling image data for image classification (Xie et al., 2019), face recognition (Hansen et al., 2018), or even some regression purposes (e.g., Jensen et al., 2019). In the future we also expect more complex machine learning pipelines from large interdisciplinary projects such as the Virtual Dairy Farm Brain project of the University of Wisconsin–Madison (Liang et al., 2018).

Based on these observations we strongly advise authors of future sensor-based mastitis and other animal health detection research to clearly provide the name of the method used, refer to publications using the same method (name) to get more coherence between publications, make explicit mention if the method used is new in this field of research, and use a uniform name. The framework we used in this paper (filtering, transformation, and classification) and the underlying categorization of approaches and methods is a starting point in relation to definitions that are used in sensor-based animal disease detection research. We encourage future research on sensor-based animal disease detection to use a more coherent terminology

for methods that support readers to better distinguish methods used between studies. Therefore, this paper serves as a starting point to provide a common language toward further interdisciplinary data-driven mastitis research.

ACKNOWLEDGMENTS

REFERENCES

Addis, M. F., V. Tedde, G. M. G. Puggioni, S. Pisanu, A. Casula, C. Locatelli, N. Rota, V. Bronzo, P. Moroni, and S. Uzzau. 2016. Evaluation of milk cathelicidin for detection of bovine mastitis. J. Dairy Sci. 99:8250–8258. https://doi.org/10.3168/jds.2016-11407.

Ali, J., R. Khan, N. Ahmad, and I. Maqsood. 2012. Random forests and decision trees. Int. J. Comput. Sci. 9:272–278.

Ankinakatte, S., E. Norberg, P. Løvendahl, D. Edwards, and S. Højsgaard. 2013. Predicting mastitis in dairy cows using neural networks and generalized additive models: A comparison. Comput. Electron. Agric. 99:1–6. https://doi.org/10.1016/j.compag.2013.08.024.

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.

Cavero, D., K. H. Tölle, C. Buxadé, and J. Krieter. 2006. Mastitis detection in dairy cows by application of fuzzy logic. Livest. Sci. 105:207–213. https://doi.org/10.1016/j.livsci.2006.06.006.

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. https://doi.org/10.1016/j.livsci.2007.05.012.

Cavero, D., K. H. Tölle, G. Rave, C. Buxadé, and J. Krieter. 2007. Analysing serial data for mastitis detection by means of local regression. Livest. Sci. 110:101–110. https://doi.org/10.1016/j.livsci.2006.10.006.

Claycomb, R. W., P. T. Johnstone, G. A. Mein, and R. A. Sherlock. 2009. An automated in-line clinical mastitis detection system using measurement of conductivity from foremilk of individual udder quarters. N. Z. Vet. J. 57:208–214. https://doi.org/10.1080/00480169.2009.36903.

Dalen, G., A. Rachah, H. Nørstebø, Y. H. Schukken, and O. Reksen. 2019. The detection of intramammary infections using online somatic cell counts. J. Dairy Sci. 102:5419–5429. https://doi.org/10.3168/jds.2018-15295.

de Mol, R. M., A. Keen, G. H. Kroeze, and J. M. F. H. Achten. 1999. Description of a detection model for oestrus and diseases in dairy cattle based on time series analysis combined with a Kalman filter. Comput. Electron. Agric. 22:171–185. https://doi.org/10.1016/S0168-1699(99)00016-2.

de Mol, R. M., G. H. Kroeze, J. M. F. H. Achten, K. Maatje, and W. Rossing. 1997. Results of a multivariate approach to automated oestrus and mastitis detection. Livest. Prod. Sci. 48:219–227. https://doi.org/10.1016/S0301-6226(97)00028-6.

de Mol, R. M., and W. Ouweltjes. 2001. Detection model for mastitis in cows milked in an automatic milking system. Prev. Vet. Med. 49:71–82. https://doi.org/10.1016/S0167-5877(01)00176-3.

de Mol, R. M., W. Ouweltjes, G. H. Kroeze, and M. M. W. B. Hendriks. 2001. Detection of estrus and mastitis: Field performance of a model. Appl. Eng. Agric. 17:399–407. https://doi.org/10.13031/2013.6201.

de Mol, R. M., and W. E. Woldt. 2001. Application of fuzzy logic in automated cow status monitoring. J. Dairy Sci. 84:400–410. https://doi.org/10.3168/jds.S0022-0302(01)74490-6.

Deng, Z., H. Hogeveen, T. J. G. M. Lam, R. van der Tol, and G. Koop. 2020. Performance of online somatic cell count estimation in automatic milking systems. Front. Vet. Sci. 7:221. https://doi.org/10.3389/fvets.2020.00221.

Dominiak, K. N., and R. Kristensen. 2017. Prioritizing alarms from sensor-based detection models in livestock production - A review on model performance and alarm reducing methods. Comput. Electron. Agric. 133:46–67. https://doi.org/10.1016/j.compag.2016.12.008.

Ebrahimi, M., M. Mohammadi-Dehcheshmeh, E. Ebrahimie, and K. R. Petrovski. 2019. Comprehensive analysis of machine learning models for prediction of sub-clinical mastitis: Deep Learning and Gradient-Boosted Trees outperform other models. Comput. Biol. Med. 114:103456. https://doi.org/10.1016/j.compbiomed.2019.103456.

Ebrahimie, E., F. Ebrahimi, M. Ebrahimi, S. Tomlinson, and K. R. Petrovski. 2018. Hierarchical pattern recognition in milking parameters predicts mastitis prevalence. Comput. Electron. Agric. 147:6–11. https://doi.org/10.1016/j.compag.2018.02.003.

Eradus, W. J., and M. B. Jansen. 1999. Animal identification and monitoring. Comput. Electron. Agric. 24:91–98. https://doi.org/10.1016/S0168-1699(99)00039-3.

Fernando, R. S., R. B. Rindsig, and S. L. Spahr. 1982. Electrical conductivity of milk for detection of mastitis. J. Dairy Sci. 65:659–664. https://doi.org/10.3168/jds.S0022-0302(82)82245-5.

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. https://doi.org/10.3168/jds.2007-0148.

Gebre-Egziabher, A., H. C. Wood, J. D. Robar, and G. Blankenagel. 1979. Evaluation of automatic mastitis detection equipment. J. Dairy Sci. 62:1108–1114. https://doi.org/10.3168/jds.S0022-0302(79)83383-4.

Hansen, M. F., M. L. Smith, L. N. Smith, M. G. Salter, E. M. Baxter, M. Farish, and B. Grieve. 2018. Towards on-farm pig face recognition using convolutional neural networks. Comput. Ind. 98:145–152. https://doi.org/10.1016/j.compind.2018.02.016.

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. https://doi.org/10.3390/s100907991.

Højsgaard, S., and N. C. Friggens. 2010. Quantifying degree of mastitis from common trends in a panel of indicators for mastitis in dairy cows. J. Dairy Sci. 93:582–592. https://doi.org/10.3168/jds.2009-2445.

Hovinen, M., A. M. Aisla, and S. Pyörälä. 2006. Accuracy and reliability of mastitis detection with electrical conductivity and milk colour measurement in automatic milking. Acta Agric. Scand. A Anim. Sci. 56:121–127. https://doi.org/10.1080/09064700701216888.

Huybrechts, T., K. Mertens, J. De Baerdemaeker, B. De Ketelaere, and W. Saeys. 2014. Early warnings from automatic milk yield monitoring with online synergistic control. J. Dairy Sci. 97:3371–3381. https://doi.org/10.3168/jds.2013-6913.

Hyde, R. M., P. M. Down, A. J. Bradley, J. E. Breen, C. Hudson, K. A. Leach, and M. J. Green. 2020. Automated prediction of mastitis infection patterns in dairy herds using machine learning. Sci. Rep. 10:4289. https://doi.org/10.1038/s41598-020-61126-8.

ISO. 2007. Automatic milking installations — Requirements and testing (ISO No. 20966). Accessed Feb. 2, 2021. https://www.iso.org/standard/35593.html.

Jensen, D. B., H. Hogeveen, and A. De Vries. 2016. Bayesian integration of sensor information and a multivariate dynamic linear model for prediction of dairy cow mastitis. J. Dairy Sci. 99:7344–7361. https://doi.org/10.3168/jds.2015-10060.

Jensen, D., M. van der Voort, C. Kamphuis, I. Athanasiadis, A. De Vries, and H. Hogeveen. 2019. Comparison of data driven mastitis detection methods. Pages 626–632 in Precision Livestock Farming '19. Teagasc.

Jørgensen, C. H., A. R. Kristensen, S. Østergaard, and T. W. Bennedsgaard. 2016. Use of inline measures of L-lactate dehydrogenase for classification of posttreatment mammary *Staphylococcus aureus* infection status in dairy cows. J. Dairy Sci. 99:8375–8383. https://doi.org/10.3168/jds.2016-10858.

Kamphuis, C., H. Mollenhorst, A. Feelders, D. Pietersma, and H. Hogeveen. 2010a. Decision-tree induction to detect clinical mastitis with automatic milking. Comput. Electron. Agric. 70:60–68. https://doi.org/10.1016/j.compag.2009.08.012.

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 by using decision-tree induction. J. Dairy Sci. 93:3616–3627. https://doi.org/10.3168/jds.2010-3228.

Kamphuis, C., H. Mollenhorst, and H. Hogeveen. 2011. Sensor measurements revealed: Predicting the Gram-status of clinical mastitis causal pathogens. Comput. Electron. Agric. 77:86–94. https://doi.org/10.1016/j.compag.2011.03.012.

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 clinical mastitis and abnormal milk. Comput. Electron. Agric. 62:169–181. https://doi.org/10.1016/j.compag.2007.12.009.

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. https://doi.org/10.3168/jds.2008-1160.

Kaşıkçı, G., Ö. Çetin, E. B. Bingöl, and M. C. Gündüz. 2012. Relations between electrical conductivity, somatic cell count, California mastitis test and some quality parameters in the diagnosis of subclinical mastitis in dairy cows. Turk. J. Vet. Anim. Sci. 36:49–55. https://doi.org/10.3906/vet-1103-4.

Khatun, M., C. E. F. Clark, N. A. Lyons, P. C. Thomson, K. L. Kerrisk, and S. C. Garciá. 2017. Early detection of clinical mastitis from electrical conductivity data in an automatic milking system. Anim. Prod. Sci. 57:1226–1232. https://doi.org/10.1071/AN16707.

Khatun, M., P. C. Thomson, K. L. Kerrisk, N. A. Lyons, C. E. F. Clark, J. Molfino, and S. C. García. 2018a. Development of a new clinical mastitis detection method for automatic milking systems. J. Dairy Sci. 101:9385–9395. https://doi.org/10.3168/jds.2017-14310.

Khatun, M., P. C. Thomson, K. L. Kerrisk, N. A. Lyons, C. E. F. Clark, J. Molfino, and S. C. García. 2018b. Development of a new clinical mastitis detection method for automatic milking systems. J. Dairy Sci. 101:9385–9395. https://doi.org/10.3168/jds.2017-14310.

Liang, D., H. Delgado, and V. Cabrera. 2018. A Virtual Dairy Farm Brain. In 13th European IFSA Symposium. International Farming Systems Association (IFSA) Europe.

Løvendahl, P., and L. P. Sørensen. 2016. Frequently recorded sensor data may correctly provide health status of cows if data are handled carefully and errors are filtered away. Biotechnol. Agron. Soc. Environ. 20:3–12.

Maatje, K., P. J. M. Huijsmans, W. Rossing, and P. H. Hogewerf. 1992. The efficacy of in-line measurement of quarter milk electrical conductivity, milk yield and milk temperature for the detection of clinical and subclinical mastitis. Livest. Prod. Sci. 30:239–249. https://doi.org/10.1016/S0301-6226(06)80013-8.

Mammadova, N., and I. Keskin. 2013. Application of the support vector machine to predict subclinical mastitis in dairy cattle. ScientificWorldJournal 2013:1–9. https://doi.org/10.1155/2013/603897.

Mammadova, N. M., and I. Keskin. 2015. Application of neural network and adaptive neuro-fuzzy inference system to predict subclinical mastitis in dairy cattle. Indian J. Anim. Res. 49(OF):671–679. https://doi.org/10.18805/ijar.5581.

Miekley, B., E. Stamer, I. Traulsen, and J. Krieter. 2013a. Implementation of multivariate cumulative sum control charts in mastitis and lameness monitoring. J. Dairy Sci. 96:5723–5733. https://doi.org/10.3168/jds.2012-6460.

Miekley, B., I. Traulsen, and J. Krieter. 2012. Detection of mastitis and lameness in dairy cows using wavelet analysis. Livest. Sci. 148:227–236. https://doi.org/10.1016/j.livsci.2012.06.010.

Miekley, B., I. Traulsen, and J. Krieter. 2013b. Mastitis detection in dairy cows: The application of support vector machines. J. Agric. Sci. 151:889–897. https://doi.org/10.1017/S0021859613000178.

Miekley, B., I. Traulsen, and J. Krieter. 2013c. Principal component analysis for the early detection of mastitis and lameness in dairy cows. J. Dairy Res. 80:335–343. https://doi.org/10.1017/S0022029913000290.

Mikail, N., and I. Keskin. 2015. Subclinical mastitis prediction in dairy cattle by application of fuzzy logic. Pak. J. Agric. Sci. 52:1101–1107.

Mollenhorst, H., L. J. Rijkaart, and H. Hogeveen. 2012a. Mastitis alert preferences of farmers milking with automatic milking systems. J. Dairy Sci. 95:2523–2530. https://doi.org/10.3168/jds.2011-4993.

Mollenhorst, H., L. J. Rijkaart, and H. Hogeveen. 2012b. Mastitis alert preferences of farmers milking with automatic milking systems. J. Dairy Sci. 95:2523–2530. https://doi.org/10.3168/jds.2011-4993.

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. https://doi.org/10.3168/jds.2009-2842.

Nielen, M., Y. H. Schukken, A. Brand, S. Haring, and R. T. Ferwerda-Van Zonneveld. 1995a. Comparison of analysis techniques for on-line detection of clinical mastitis. J. Dairy Sci. 78:1050–1061. https://doi.org/10.3168/jds.S0022-0302(95)76721-2.

Nielen, M., M. H. Spigt, Y. H. Schukken, H. A. Deluyker, K. Maatje, and A. Brand. 1995b. Application of a neural network to analyse on-line milking parlour data for the detection of clinical mastitis in dairy cows. Prev. Vet. Med. 22:15–28. https://doi.org/10.1016/0167-5877(94)00405-8.

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. https://doi.org/10.3168/jds.S0022-0302(04)73256-7.

Rasmussen, M. D. 2005. Visual scoring of clots in foremilk. J. Dairy Res. 72:406–414. https://doi.org/10.1017/S0022029905000993.

Rutten, C. J., A. G. J. Velthuis, W. Steeneveld, and H. Hogeveen. 2013. Invited review: Sensors to support health management on dairy farms. J. Dairy Sci. 96:1928–1952. https://doi.org/10.3168/jds.2012-6107.

Sathiyabarathi, M., S. Jeyakumar, A. Manimaran, H. A. Pushpadass, M. Sivaram, K. P. Ramesha, D. N. Das, M. A. Kataktalware, G. Jayaprakash, and T. K. Patbandha. 2016. Investigation of body and udder skin surface temperature differentials as an early indicator of mastitis in Holstein Friesian crossbred cows using digital infrared thermography technique. Vet. World 9:1386–1391. https://doi.org/10.14202/vetworld.2016.1386-1391.

Slob, N., C. Catal, and A. Kassahun. 2021. Application of machine learning to improve dairy farm management: A systematic literature review. Prev. Vet. Med. 187:105237. https://doi.org/10.1016/j.prevetmed.2020.105237.

Song, X., and R. Tol. 2010. Automatic detection of clinical mastitis in astronaut A3™ milking robot. Pages 154–155 in Proc. First North Am. Conf. Precis. Dairy Manag. Precision Dairy Management.

Soyeurt, H., C. Bastin, F. G. Colinet, V. M. R. Arnould, D. P. Berry, E. Wall, F. Dehareng, H. N. Nguyen, P. Dardenne, J. Schefers, J. Vandenplas, K. Weigel, M. Coffey, L. Thé Ron, J. Detilleux, E. Reding, N. Gengler, and S. McParland. 2012. Mid-infrared prediction of lactoferrin content in bovine milk: Potential indicator of mastitis. Animal 6:1830–1838. https://doi.org/10.1017/S1751731112000791.

Steeneveld, W., L. C. van der Gaag, H. W. Barkema, and H. Hogeveen. 2009. Providing probability distributions for the causal pathogen of clinical mastitis using naive Bayesian networks. J. Dairy Sci. 92:2598–2609. https://doi.org/10.3168/jds.2008-1694.

Steeneveld, W., L. C. van der Gaag, H. W. Barkema, and H. Hogeveen. 2010a. Simplify the interpretation of alert lists for clinical mastitis in automatic milking systems. Comput. Electron. Agric. 71:50–56. https://doi.org/10.1016/j.compag.2009.12.011.

Steeneveld, W., L. C. van der Gaag, W. Ouweltjes, H. Mollenhorst, and H. Hogeveen. 2010b. Discriminating between true-positive and false-positive clinical mastitis alerts from automatic milking systems. J. Dairy Sci. 93:2559–2568. https://doi.org/10.3168/jds.2009-3020.

Steeneveld, W., L. C. van der Gaag, W. Ouweltjes, H. Mollenhorst, and H. Hogeveen. 2010c. Discriminating between true-positive and false-positive clinical mastitis alerts from automatic milking systems. J. Dairy Sci. 93:2559–2568. https://doi.org/10.3168/jds.2009-3020.

Sun, Z., S. Samarasinghe, and J. Jago. 2010. Detection of mastitis and its stage of progression by automatic milking systems using artificial neural networks. J. Dairy Res. 77:168–175. https://doi.org/10.1017/S0022029909990550.

Van Bebber, J., N. Reinsch, W. Junge, and E. Kalm. 1999. Monitoring daily milk yields with a recursive test day repeatability model (Kalman filter). J. Dairy Sci. 82:2421–2429. https://doi.org/10.3168/jds.S0022-0302(99)75493-7.

West, M., and J. Harrison. 1997. Bayesian Forecasting and Dynamic Models. 2nd ed. Springer.

Wohlin, C. 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. Pages 1–10 in Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering. Association for Computing Machinery.

Xie, B., H. K. Zhang, and J. Xue. 2019. Deep convolutional neural network for mapping smallholder agriculture using high spatial resolution satellite image. Sensors (Basel) 19:2398. https://doi.org/10.3390/s19102398.

Zhang, X., X. Kang, N. Feng, and G. Liu. 2020. Automatic recognition of dairy cow mastitis from thermal images by a deep learning detector. Comput. Electron. Agric. 178:105754. https://doi.org/10.1016/j.compag.2020.105754.

Zins, C. 2007. Conceptual approaches for defining data, information, and knowledge. J. Am. Soc. Inf. Sci. Technol. 58:479–493. https://doi.org/10.1002/asi.20508.

ORCIDS

M. van der Voort ⓘ https://orcid.org/0000-0003-0503-259X
D. Jensen ⓘ https://orcid.org/0000-0002-9437-0605
C. Kamphuis ⓘ https://orcid.org/0000-0001-5552-036X
I. N. Athanasiadis ⓘ https://orcid.org/0000-0003-2764-0078
A. De Vries ⓘ https://orcid.org/0000-0003-4511-0388
H. Hogeveen ⓘ https://orcid.org/0000-0002-9443-1412