Università degli Studi di Padova

Head Office: Università degli Studi di Padova

Department: Medicina Animale, Produzioni e Salute (MAPS)

Ph.D. COURSE IN: Animal and Food Science
SERIES: XXXVI

PRECISION LIVESTOCK FARMING (PLF) APPLICATIONS TO THE DAIRY SECTOR

Coordinator: Prof. Angela Trocino
Supervisor: Prof. Giulio Cozzi

Ph.D. student: Giovanna Ranzato

ABSTRACT

In the modern livestock sector, several applications of Precision Livestock Farming (PLF) have been designed for improving efficiency, productivity, sustainability, and welfare of different farming systems. Precision dairy farming systems, in particular, enable to manage larger herds in a more time-efficient manner, through automated monitoring of individual cow health and welfare. Dairy farmers can be assisted in the identification of unexpected behaviors or in the early detection of pathologies, with the possibility to decide whether and how to act. Furthermore, the analysis of big data from PLF technologies can help dairy farmers in phenotyping the animals for complex traits such as resilience, longevity, or productive life span, bringing to optimized breeding, treatment, and culling decisions. The aim of this thesis was to present different PLF applications to the dairy sector, that can serve as support tools for the optimization of farm management strategies and cows' welfare. A general introductive chapter (**Chapter 1**) focused on an overview about PLF systems, objectives, and limitations, specifically addressed to the dairy sector. Afterwards, a new statistical model within the animal science field (i.e., joint model for longitudinal and time-to-event data) was tested to predict cow's survival using first-parity sensor data as input (**Chapter 2**). The algorithm had good repeatability across farms with modest performances. However, joint models offer such interesting opportunities in terms of applicability and flexibility to justify further research for improving the overall predictive accuracy in the dairy sector. Further research investigated heat wave effects on dairy cows' behaviors registered with sensors (**Chapter 3**). The output revealed that 'heat-sensitive' subjects were more active and spent more time chewing during a heat wave challenge compared to 'heat-tolerant' ones, as an attempt to better dissipate heat load. This suggested that the information provided by high-frequency sensor data can assist farmers in the early identification of cows for which personalized interventions to alleviate heat stress are needed. In a dedicated study, three different mathematical methods to estimate dairy cows' expected production, milk losses, and perturbations of the lactation curve were analyzed and compared (**Chapter 4**). The output of this study can help dairy practitioners in choosing the method that best fits their management strategies, to understand, for example, how the animals cope with challenges, or to optimize their production system. Finally, a pilot study addressed to early detect cows at risk of metabolic disorders was conducted, using milk fatty acids analysis obtained with FTIR spectroscopy (**Chapter 5**). Preliminary reference intervals for de novo, mixed, and preformed fatty acids were calculated for healthy cows' during early and mid-lactation. These reference ranges could help farmers

to screen cows at risk of specific health disorders (e.g., subclinical ketosis) even before clinical signs are visible. In conclusion (**Chapter 6**), this thesis highlighted the potential of PLF in assisting dairy farmers to make better choices about the sustainability and the efficiency of their production system, by providing more objective information about health and productivity of the animals.

RIASSUNTO

Recentemente, il settore zootecnico è stato caratterizzato da numerose applicazioni dell'Allevamento di Precisione (PLF) progettate per migliorare l'efficienza, la produttività, la sostenibilità e il benessere di diversi sistemi di allevamento. In particolare, i sistemi di allevamento di precisione per il settore lattiero-caseario consentono di gestire mandrie più grandi in modo più efficiente nel tempo, attraverso il monitoraggio automatizzato della salute e del benessere delle singole vacche. Gli allevatori possono essere assistiti nell'identificazione di comportamenti inaspettati o nella rilevazione precoce di patologie, con la possibilità di decidere se e come intervenire. Inoltre, l'analisi di 'big data' provenienti dalle tecnologie PLF può aiutare gli allevatori nel fenotipizzare gli animali per tratti complessi come la resilienza o la longevità, portando ad ottimizzare decisioni riguardanti la selezione, il trattamento o la macellazione degli animali. L'obiettivo di questa tesi consiste nel presentare diverse applicazioni PLF al settore lattiero-caseario, in particolare come strumenti di supporto per l'ottimizzazione delle strategie di gestione dell'azienda e del benessere delle vacche. Un capitolo introduttivo generale (**Capitolo 1**) si è concentrato su una panoramica di sistemi, obiettivi, e limitazioni dell'allevamento di precisione, specialmente nel settore lattiero-caseario. Successivamente, è stato testato un nuovo modello statistico nel campo delle scienze animali (i.e., 'joint model') per prevedere la sopravvivenza delle vacche da latte utilizzando i dati dei sensori nei soggetti primipari come input (**Capitolo 2**). L'algoritmo ha avuto una buona ripetibilità tra gli allevamenti considerati, con prestazioni modeste. Tuttavia, i 'joint models' offrono opportunità talmente interessanti in termini di applicabilità e flessibilità da giustificare ulteriori ricerche per migliorarne l'accuratezza predittiva. Ulteriori ricerche hanno indagato gli effetti delle ondate di calore sui comportamenti delle vacche da latte registrati tramite sensori (**Capitolo 3**). L'output ha rivelato che i soggetti 'heat-sensitive' erano più attivi e trascorrevano più tempo a masticare durante le ondate di calore, nel tentativo di dissipare meglio il carico termico. Ciò ha suggerito che le informazioni fornite dai dati ad alta frequenza possono assistere gli allevatori nell'identificazione precoce delle vacche per cui sono necessari interventi personalizzati per alleviare lo stress da caldo. In uno studio dedicato, sono stati analizzati e confrontati tre diversi metodi matematici per stimare la produzione prevista delle vacche, le perdite di latte, e le perturbazioni della curva di lattazione (**Capitolo 4**). L'output di questo studio può aiutare allevatori o veterinari a scegliere il metodo che meglio si adatta alle loro strategie di gestione, per capire, ad esempio, come gli animali affrontano le sfide, o per ottimizzare il sistema produttivo. Infine, è stato condotto uno studio pilota per rilevare le vacche a

rischio di disturbi metabolici, utilizzando l'analisi degli acidi grassi del latte ottenuta con spettroscopia FTIR (**Capitolo 5**). Gli intervalli di riferimento preliminari per acidi grassi de novo, misti, e preformati sono stati calcolati per vacche sane durante le prime fasi della lattazione. Questi intervalli di riferimento potrebbero aiutare gli allevatori ad identificare le vacche a rischio di specifici disturbi di salute (chetosi subclinica) anche prima che siano visibili i segni clinici. In conclusione (**Capitolo 6**), questa tesi ha evidenziato il potenziale dell’allevamento di precisione nell'assistere gli allevatori nelle scelte riguardanti la sostenibilità e l'efficienza del loro sistema produttivo, fornendo informazioni più oggettive sulla salute e sulla produttività degli animali.

INDEX

ABBREVIATIONS

ACT = activity time
AFC = age at first calving
AMS = automatic milking system
AUC = Area Under the Curve
BW = body weight
CIC = correlation information criteria
CV = cross validation
CHEW = chewing time
DIM = days in milk
FA = fatty acids
FTIR = Fourier-transformed mid-infrared
GEE = generalized estimating equations
GP = Grana Padano
HW = heat wave
IW = iterative Wood model
LIE = lying time
ML = milk loss
MY = milk yield
PE = prediction error
PLF = precision livestock farming
PLM = perturbed lactation model
PR = Parmigiano Reggiano
QR = quantile regression
RH = relative humidity
RUM = rumination time
SEAS = season
T = temperature
THI = temperature-humidity index
ULC = unperturbed lactation curve

CHAPTER 1

Artificially generated image (DALL·E 2, OpenAI)

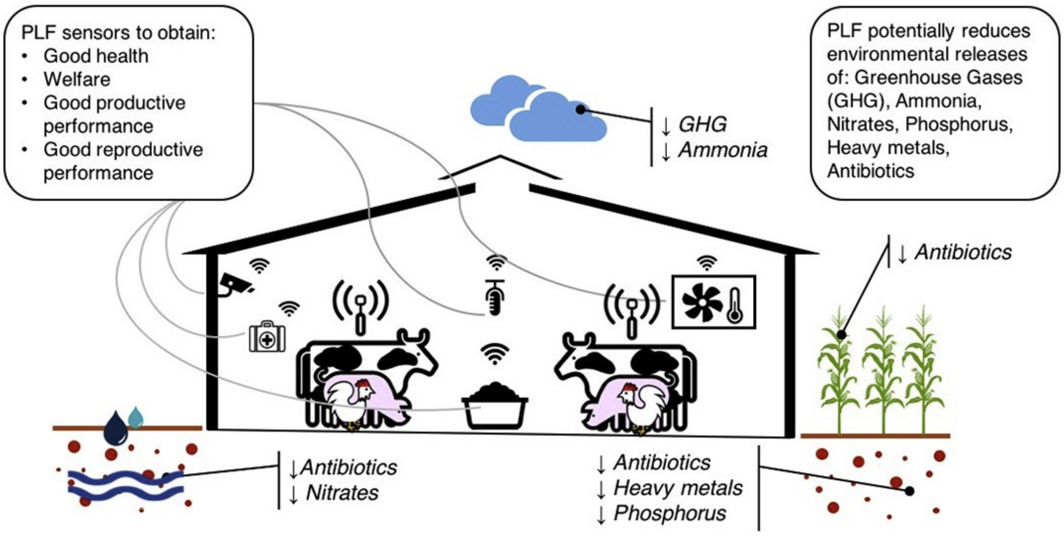
PRECISION LIVESTOCK FARMING

From the beginning of the new century, Precision Livestock Farming (PLF) has represented a paradigm shift in the approach to livestock management (Morrone et al., 2022). It has emerged as a multidisciplinary science dedicated to improving efficiency, productivity, sustainability, and welfare within animal farming systems. The integration of information and communication technologies, automation, bio-sensing, and data analytics have opened the way to innovative solutions to long-standing challenges in the livestock sector (Berckmans, 2014). While conventional livestock management relies on the producer’s experience and, if available, on population-level data to make decisions, PLF emphasizes the continuous, automated monitoring of each animal within a herd, allowing for more personalized strategies that address individual needs in every specific rearing environment (Jiang et al., 2023).

In the context of PLF, various technologies like cameras or sound devices can be used to monitor the animals, and different types of sensor systems (e.g., accelerometers, infrared thermography) allow to register physiological parameters and behavioral data (Stygar et al., 2021). To obtain useful information from the enormous amount of data extracted, data analysis, *machine learning,* and control systems techniques are used (García et al., 2020). Among the numerous applications, animal health monitoring is one of the most crucial. Early detection of diseases allows for timely interventions, minimizing the impact on welfare and productivity (Neethirajan, 2017). Reproduction management represents another area of application: accurate prediction of estrus and parturition timing can contribute to improved reproductive efficiency (Lopes et al., 2016). Some technologies also enable precision and automated feeding, resulting in optimized animal growth, minimized resource wastage, and reduced environmental footprint (Tullo et al., 2019).

As reported by many authors, livestock environmental impact mitigation can be obtained by enhancing productivity levels, reproduction traits, and maintaining good health (Tullo et al., 2019). Rapid intervention, in case of disease, or modification in the management strategy, in case of stress, can actually and effectively improve the animals’ status. Thus, good productive performance that relies on animal health and welfare achieved through PLF, can mitigate the environmental impact of livestock, as shown in Figure 1.

Figure 1. *Precision livestock farming as a strategy to mitigate environmental risks (from Tullo et al., 2019).*



Despite the multiple advantages of PLF, its implementation is not without challenges. The initial high cost of technology, the still limited technical knowledge among farmers, and the ethical concerns about the objectification of animals and further intensification, represent some factors limiting PLF adoption (Banhazi et al., 2012; Schillings et al., 2021). Furthermore, issues related to data management, such as data integration, interpretation, and privacy, pose non-negligible limitations (Schillings et al., 2021). Table 1 reports some advantages and disadvantages emerged among farmers in the adoption of PLF (Lovarelli et al., 2020).

Table 1. *Main positive and negative aspects from the adoption of precision livestock farming.*

Positive effect	Negative effect
Better environmental performances (less GHG[1] emissions, less nitrogen release)	Initial investment costs
Optimization of production	Need of identifying the added value of production from PLF: the outcome of a decision taken with PLF tools minus the decision taken without PLF tools
Focus on single animal health and welfare conditions	Need of experts able to analyze and understand the collected data
Alleviation of food security challenges	Issues related to data privacy
Economic sustainability (economic costs and benefits)	
Improved work conditions (social aspects)	

[1] Greenhouse gases.

PLF IN THE DAIRY SECTOR

Milk and dairy products are an important source of dietary energy, protein, and fat for the global population. As reported by Norton and Berckmans (2017), milk is the EU's first agricultural product, accounting for 15% of agricultural output in terms of value (European Parliament, 2015). The EU dairy sector is supported by 650,000 specialized dairy farmers and 18 million milking cows and has a labor force of about 1.2 million people (European Parliament, 2015). However, since the abolishment of milk quotas in 2015, farmers are facing increased pressures to exploit the economies of scale by increasing the size of their herds. But with larger herds, farmers no longer have the time to care for their animals as they traditionally used to do. Therefore, the implementation of technology on farms is becoming increasingly necessary, both in terms of animal welfare and economic success.

According to Lovarelli et al. (2020), the main PLF applications in the dairy sector regard:

- health and estrus phases: instrumentation to control animals' health, including devices that measure the motor activity and the health state (e.g., accelerometers, cameras, microphones);
- management: instrumentation to control productive variables (e.g., milk quality and quantity);
- environment: instrumentation to control the environmental conditions in the barn such as temperature, humidity, radiation, wind;
- behavior: including cameras to control the behavior of single animals and/or the relations among the animals.

Precision dairy farming systems enable to manage larger herds in a more time-efficient manner (Rutten et al., 2013), through automated monitoring of individual cow health and welfare. Farmers are assisted in the identification of stress or unexpected behaviors, and in the early detection of pathologies, with the possibility to decide whether and how to act (Rojo-Gimeno et al., 2019). Alternatively, visualizing images as video-recordings helps identifying aggressive behaviors (such as biting, head knock, nose-to-nose cases) and operates to reduce social problems (Oczak et al., 2013).

To have a more general overview, Table 2 summarizes a literature review by Lovarelli et al. (2020) on the use of PLF on dairy cattle farms. Carpentier et al. (2018), for example, investigated whether monitoring of coughing in a calf house had the potential to detect cases of respiratory infection before they became too severe. They developed an algorithm that, using sound data, forecasted the infection with a precision higher than 80%. Shahriar et al. (2016) used an unsupervised learning technique to detect heat events in pasture-based dairy cows. Accelerometer data from the cow collars were used to identify increased activity levels associated with recorded heat events, with an overall accuracy higher than 80%. Zebari et al. (2018) studied whether the number of steps, lying time, lying bouts, dry matter intake, feeding duration, and number of visits to feed were affected by behavioral or silent estrus in lactating dairy cows. Using video cameras and accelerometer data, they found that, during *behavioral* estrus, the number of steps increased, while lying time, lying bouts, dry matter intake, and feeding duration decreased, whereas during silent estrus only feeding duration significantly decreased.

The Italian scenario

Italy has more than 1,600,000 dairy cows across the country; most of the dairy herds are located in the north (78%), with Lombardy region being at first place for the total number of cows (579,042 cows) and the total amount of milk produced per cow (10,938 kg) (CLAL, 2022). The incidence of the dairy cattle sector on the value of Italian animal productions is about 30% (CREA - Research Centre for Agricultural Policies and Bioeconomy, 2022). After Germany, France, and the Netherlands, Italy is the fourth European country for total cow milk production (EUROSTAT, 2020).

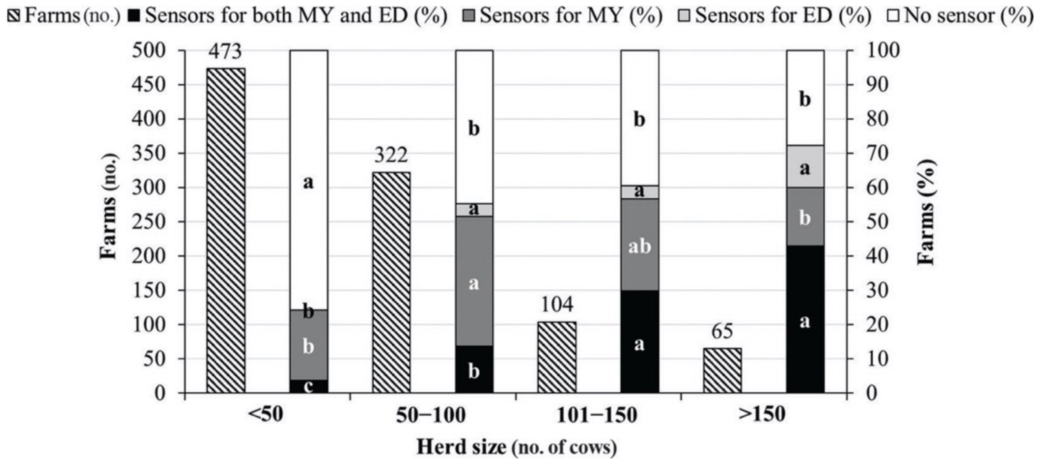
Despite the importance of the dairy cattle sector in the agricultural economy of the country, the diffusion of farm technology in the Italian dairies does not seem to be as extensive as it is in other countries, particularly in Northern Europe (Bianchi et al., 2022). Many Italian researchers have investigated the presence of PLF tools in Italian dairy farms. Lora et al. (2020) conducted a large survey on 964 dairy farms located in Veneto region. Farmers were interviewed by technicians of the regional breeders' association to collect information on the type of sensors installed on farms and the main parameters recorded. Overall, 42% of the surveyed farms had at least one sensor system, and most of them (72%) reared more than 50 cows (Figure 2). Sensors for measuring individual cow milk yield (MY) were the most prevalent type installed (39%), whereas only 15% of farms had systems for estrus detection (Figure 2). More sophisticated parameters, such as rumination, were automatically monitored in less than 5% of the farms. Bianchi et al. (2022) distributed an online questionnaire to dairy farmers from Lombardy. Precision systems that provide information on animal activity (heat detection) and on MY and flow were the most popular and were considered among the most useful. Management of reproduction and milk production were the areas where farmers showed interest for future investments. Younger farmers appeared to have implemented more PLF systems than older ones, and they showed more willingness to invest in more sophisticated precision tools of the latest generation. Similar results to the ones of Lora et al. (2020) and Bianchi et al. (2022) were reported by Abeni et al. (2019) in a survey to assess the propensity to invest in PLF tools carried out in the province of Cremona (Lombardy).

Table 2. *Results of a literature review on the use of precision livestock farming in the dairy industry.*

Authors	Goal	Sensor used	What is analyzed
Arcidiacono et al., 2017	Detect real-time behavior and improve a software	Neck collar with accelerometer	Feeding and standing behavior
Arcidiacono et al., 2018	Detect velocity through software for estrus detection	Neck collar tag and sensor in barn	Visualization of velocity during motion
Benaissa et al., 2018	Test model	Accelerometers	Exposure to the wireless power transfers
Carpenteir et al., 2018	Detect bovine respiratory disease	Microphones for cough sounds	Label cough in calves
Gernand et al., 2019	Heat stress	Temperature and humidity data loggers	THI[3]
Grinter et al., 2018	Validate precision and accuracy of collars	Behavior-monitoring collar	Rumination, heat detection, feeding and resting behavior
Mattachini et al., 2019	Barn and cow monitoring	Hobo sensors, activity sensors, and AMS[2]	Temperature and humidity, leg orientation, lying time, bout frequency and duration, milking data
Mayo et al, 2018	Estrus detection	6 accelerometers compared	Lying time and estrus
Meen et al., 2015	Verify if vocalization is correlated with behavior	Cameras and microphones	Behavior analyzing video and sound recording
Meunier et al., 2018	Identify cow positions/activities in the barn through images	Collar tag, video from performance tags, wireless sensors on the barn	Methodology to evaluate behavior from tools
Potter et al., 2018	Evaluate SCC[1] and milk losses	Algorithm	SCC content and milk
Shahriar et al., 2016	Estrus identification	Accelerometers + algorithm	Motor activity
Tullo et al., 2019	Evaluate lying time and behavior	Accelerometers and temperature-humidity data loggers	Lying time
Van Hertem et al., 2013	Test algorithm	Videocameras for computer vision	Motion
Van Hertem et al., 2013	Develop a model to detect clinical lameness as function of behavior and milk performance	Neck collar tag for neck activity and rumination	Lameness
Vandermeulen et al., 2016	Bovine respiratory disease	Cough monitor with microphone	Continuous cough sound
Viazzi et al., 2014	Test algorithm	Videocameras for computer vision	Back posture
Zebari et al., 2018	Estrus identification	Videocameras plus IceQubes accelerometers for cow activity	Spontaneous behavioral estrus, analysis of progesterone in milk

[1] Somatic cell count; [2] automatic milking system; [3] temperature-humidity index.

Figure 2. *Distribution of the surveyed Italian dairy farms according to herd size, and proportion of farms equipped with sensor systems to monitor at least either individual cow milk yield (MY), estrus (ED), or both. Variables of the same color with different letters (a, b, c) differ (p < 0.05) (from Lora et al., 2020).*



PRECISION DAIRY FARMING TECHNOLOGIES

Automatic milking systems

Automatic milking systems (see an example in Figure 3) were one of the earliest PLF tools developed in the dairy sector. Robotic milking by automatic milking systems (AMS) has revolutionized dairy farming around the world, both in terms of milking process and farm management (John et al., 2016). Milking is controlled by fully automated equipment, and it is no longer performed in defined sessions; rather, the cow can choose when to be milked in AMS, allowing milking to be distributed throughout a 24-h period.

Figure 3. *Robotic arm of an automatic milking system (from delaval.com).*

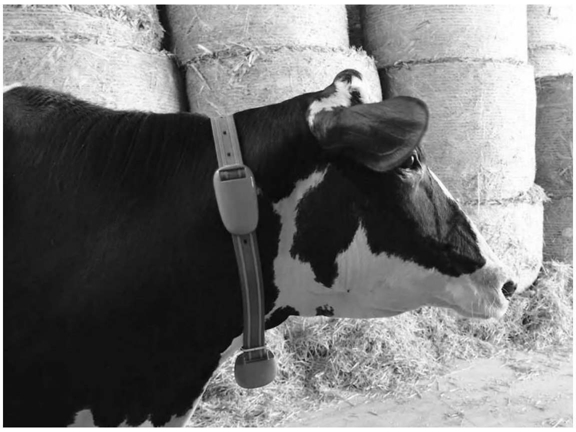
An AMS is equipped with a variety of sensors to control the milking process and detect any abnormalities. Some sensors control the technical functioning of the system, such as cow identification, nipple cleaning, teat-cup attachment, vacuum level, and onset of the milk letdown process. Other sensors control the quality of the milking process by measuring MY and checking on anomalies in the milk and on udder health. Some AMS also have online measurement of milk temperature, cow feed intake, cow body weight, and body condition score. All measurements (e.g., quarter yield, quarter conductivity, color, fat, protein, somatic cell count) are automatically stored in a database, and a dedicated software is used to analyze the collected data and to control the settings and conditions for the cows to be milked. Attention lists and reports are available to the farmer on a computer screen or portable tools. In urgent cases like a breakdown of the equipment, or a severe problem with a cow, the system immediately warns the worker by sending an alarm or a text message (de Koning, 2002).

In Europe, the United States, Australia, and New Zealand, AMS have had a positive effect on the quality of dairy farmers' lives. When operating optimally, in fact, AMS has many benefits: improved cow health, quick and easier health detection, increased milk production, less routine activities, and a more flexible lifestyle for the farmer (Tse et al., 2018).

Accelerometer-based technologies

Wearable accelerometers (see an example in Figure 4) can remotely and efficiently collect data related to measures of cow behavior. The accelerometer measures changes in velocity over time and it is attached to the animal in a specific location so that the orientation of the device can provide detailed information about movement and body position relating to the behavior of interest (Hendriks et al., 2020). Change in velocity can be recorded along 2 (2D accelerometer) or 3 (3D accelerometer) axes, separately, per unit of time; sampling frequencies usually range from 1 to 100 Hz and are predetermined by the product manufacturer. The accelerometers can be offered with different animal attachment solutions (collar, leg, ear, halter), depending on the provider and the behavior of interest (Stygar et al., 2021). Measures of behavior may concern activity (heat detection), lying bouts, lying time, eating time, rumination time, chewing time, and visits to the feed bunk.

Figure 4. *Accelerometer-based collar (from tdm.it).*



Load cells

In the systematic review by Stygar et al. (2021), load cells, together with accelerometer-based systems, were the most spread technologies on dairy farms. Load cells can be used for many purposes, such as tracking the feeding program, measuring milk quantity, and weighing of the animals. Dickinson et al. (2013), for instance, described the functioning of an automatic walk-over weighing system with load cells placed under the weighing platform within load bars (see an example in Figure 5). As a cow traverses the platform, the weight of the cow bends the load cell and produces a change in electrical resistance between 2 points, which is converted to a weight. Liveweight, cow ID, and time of weighing are then automatically recorded.

Figure 5. Automatic *walk-over weighing system through load cells (from odonovaneng.ie).*

Other technologies

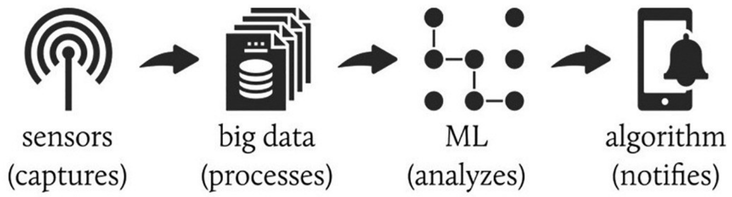
As herd size increases, there is a growing demand for tools to quickly trace individual cows that need particular attention (for insemination or medical treatments). For localization, GPS sensors and RFID tags are commonly used on farms (Achour et al., 2022). Video recordings are sometimes utilized to monitor the herd: specific cameras can be dedicated to body thermal monitoring (thermal cameras), body condition scoring, feeding monitoring, and general behavior monitoring (Stygar et al., 2021b). Sound analysis can be performed to monitor the animal welfare when microphones are installed on the barn or on wearable sensors (Meen et al., 2015). Bolus sensors can be inserted into the rumen to measure body temperature, pH, and rumen activity, as well as for animal identification (Mottram et al., 2008; Stygar et al., 2021b).

Portable infrared spectrometers (near infrared – NIR, or mid infrared – MIR) can be a versatile technology that rapidly gives information on the chemical-physical composition of raw materials, total mixed ration, milk, and feces (Evangelista et al., 2021). In addition, thermal images obtained by infrared thermography can be used to detect lesions or pathologies (Machado et al., 2021).

BIG DATA IN PRECISION DAIRY FARMING

Big data plays a key role in applying advanced technologies to dairy farming practices. Statistical methods, machine learning algorithms, and artificial intelligence can make use of this extensive data to analyze, predict and notify farmers in case there is something abnormal (Neethirajan, 2020). A graphical representation of this process is shown in Figure 6.

Figure 6. *Process for obtaining information within PLF (from Neethirajan, 2020).*



To get insights into the type of big data recorded by sensor systems, these are the main categories (Lokhorst et al., 2019):

- *time-series data:* sequences of values or events obtained over repeated measurements of time;
- *streaming* data: they are constantly arriving, for instance, from remote sensors or surveillance systems, and should be processed in an online fashion;
- *sequence* data: sequences of ordered elements or events that are recorded with or without a concrete notion of time;
- *graph* data: where problems are modeled as graphs, like in biological networks;
- *spatial* data: place-related data or remote-sensing data;
- *multimedia* data: images, videos, audio, text mark-ups.

Based on the type of data, specific models and algorithms are used to automatically learn patterns and make inference. The main categories of machine learning algorithms are (Lokhorst et al., 2019):

- *supervised learning*: the outcome of interest is known for each record used for model development, which can be 'regression' or 'classification'. For regression, the outcome variable has a numerical value; possible techniques involve linear regression, polynomial regression, and multivariate adaptive regression splines. For classification, the outcome variable is categorical; possible techniques include logistic/multinomial regression, neural networks, decision trees, naive Bayes model, and support vector machines;
- *unsupervised learning*: the outcome of interest is unknown for each record used for model development, which can be 'clustering' *or 'dimensionality reduction'*. Clustering techniques include, for example, K-means and hierarchical clustering.

Dimensionality-reduction techniques include, for example, principal component analysis;

- *reinforcement learning*: a mapping function is learnt to maximize a reward function. This technique is used, for example, in Markov decision processes.

Lokhorst et al. (2019) expect that the full potential of big data, within the precision dairy farming area, will be reached when multiple big data characteristics (*Volume*, *Velocity*, *Variety*, and *Other V's*; De Mauro et al., 2016) and sources (animal, groups, farms, and chain parts) will be used simultaneously, adding value to operational and strategic decisions.

PRECISION PHENOTYPING OF DAIRY COWS

Besides health and welfare monitoring, many researchers within the dairy sector have recently investigated the possibility of using big data extracted from sensor systems for precision phenotyping of complex traits such as resilience, longevity, or productive life span. Cows with a long productive life span typically exhibit good reproductive performance, few health problems, and efficient milk production (Adriaens et al., 2020). Early culling and short longevity thus clearly have a negative influence on the economic efficiency of a herd. Correct and timely identification of resilient animals, namely the ones that avoid early culling by coping well with the farm management conditions, would allow for optimization of breeding, treatment, and culling decisions (Adriaens et al., 2020). Today, many breeding decisions are still made based on habit; to rely on the cows' actual performance on farm, sensor data can be combined into mathematical models to be used for both decision support and precision phenotyping.

Adriaens et al. (2020), for example, investigated whether resilience and productive life span of dairy cows could be predicted using proxies of first-parity sensor data. The authors analyzed high-frequency milk and activity data retrieved from the AMS of 27 different dairy farms. Using a multivariate linear regression, they concluded that proxies of first-lactation sensor data had the potential to predict cows' resilience rankings within farms. They did not find a common model structure across all farms due to too much variability in culling, reproduction, and health management strategies. In the wake of Adriaens et al. (2020), Ouweltjes et al. (2021) developed a data-driven random forest algorithm with daily aggregated sensor data as input that provided at least as good resilience predictions as the models with sensor-derived proxies as input. In the context of MY dynamics modelling, Ben Abdelkrim et al. (2021) developed a model that allows dairy cows' precision phenotyping by estimating their individual milk production potential (as if there was no perturbation), the characteristics of each perturbation occurring during the lactation, and the overall consequence of perturbations on milk losses. Similarly, Adriaens et al. (2021) proposed an iterative procedure fitting a model on daily MY data to determine the potential production, and a criterion to identify perturbations of the lactation curve.

NEAR FUTURE OF PLF

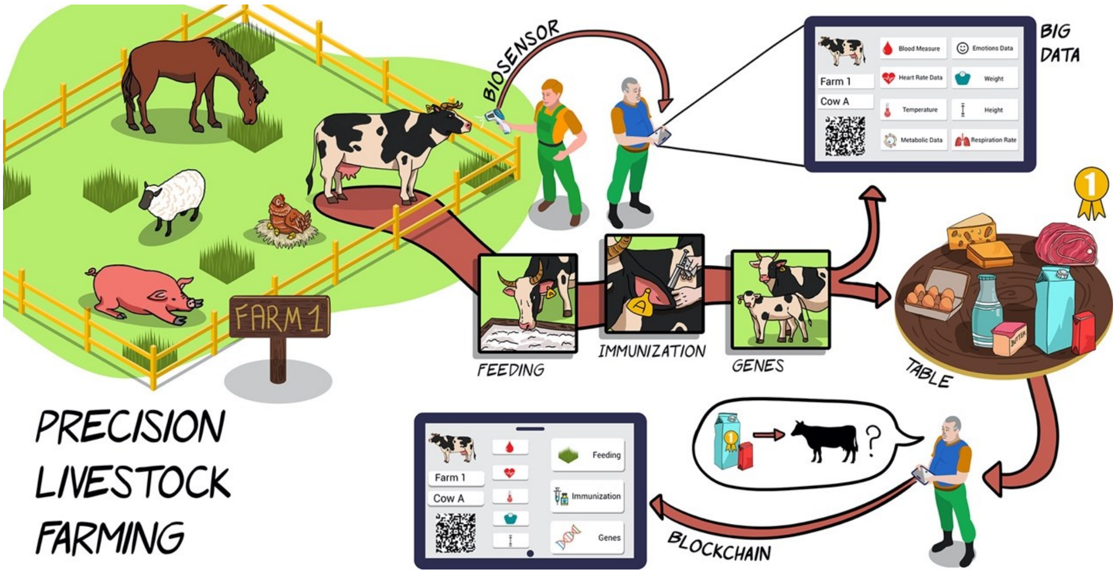
To meet the growing demand for animal products while addressing concerns about environmental sustainability, public health, and animal welfare, the livestock sector along with animal scientists will increasingly rely on PLF technologies. Moreover, as climate change intensifies, the risk of diseases, heat stress, and other health issues among animals may increase. Therefore, there is greater urgency to identify health problems and disease outbreaks early on, understand disease transmission, and take preventative measures to avoid large-scale economic losses (Neethirajan & Kemp, 2021).

Blockchain may represent a new type of PLF technology (Figure 7). It could provide important benefits to livestock farming, including decentralized transactions throughout the life of an animal from farm to table, and transactions that could contribute to more efficient auditing systems for certification organizations (Neethirajan & Kemp, 2021; Picchi et al., 2020). Blockchain could be extremely useful in detecting and tracking disease breakouts, such as H1N1 swine flu, Mad Cow diseases, Avian influenza, and Salmonella (Neethirajan & Kemp, 2021). In addition, it could help trace harmful foods back to the source, increasing traceability and accountability for problematic practices within livestock farming (Lin et al., 2018). Bioengineers and data scientists may play a significant role in formulating appropriate criteria for deciding which type of blockchain solution is the most beneficial for specific livestock farming sectors (Neethirajan & Kemp, 2021).

A branch of PLF that will see new developments is automated video-based animal detection. Computer vision and image analysis are still nonfunctional within farming (Wurtz et al., 2019). These technologies, for example, cannot yet track individual animals, at least not for a sufficiently long period to obtain meaningful information about the behaviors of interest. Furthermore, many of the studies testing computer vision techniques in livestock have been performed on pigs. More work is necessary to assess their applicability to other species (Neethirajan & Kemp, 2021). However, computer science is constantly growing; it is just a matter of time before new deep learning algorithms with greater predictive ability are implemented in agriculture and livestock areas.

To widespread PLF technologies on farms, the information, communication, and telecommunications industry must address accessibility issues, as well as push to create easy-to-use software and data visualization. At the same time, as technology advances, PLF systems will become increasingly accessible to farmers around the world (Alonso et al., 2020).

Figure 7. *Blockchain within precision livestock farming (from Neethirajan & Kemp, 2021).*



Background of the thesis

The general purpose of this Ph.D. thesis was to analyze different precision dairy farming applications concerning, partially or entirely, the concepts of decision support, precision phenotyping, and resilience. Some of the studies included in the thesis were, in fact, carried out within the European project 'GenTORE' (Genomic management Tools to Optimise Resilience and Efficiency; see https://www.gentore.eu/, Figure 8). This was a 5-year (2017 – 2022) EU funded project involving 21 academic and non-academic partners (e.g., breeding associations, farm technology suppliers). The main aim of the project was to give metrics for quantifying resilience and efficiency in dairy and beef cattle, which are very difficult to be measured under commercial conditions. The balance of resilience and efficiency determines the animal's ability to adapt to changes, which is crucial especially in view of future challenges. Livestock, in fact, will be exposed to increasing challenges under different production systems and grazing environments, making the need for resilient systems particularly urgent (GenTORE, 2017). The involvement of PLF technologies enable to develop data-based metrics and tools, and to perform predictive modelling to optimize resilience and efficiency both at animal level and system level. Within GenTORE, improved methods for genomic analysis and new indicators for on-farm phenotyping were developed, such as on-farm management indices and new ways to measure local production environments. Last but not least, most of the works presented in the thesis were realized in collaboration with 'Livestock Technology' research group of KU Leuven (Geel, Belgium). The cooperation with this group was fundamental for the analysis of dairy cows' lactation dynamics and their modelization starting from high-frequency milk sensor data.

Figure 8. *Logo of the European GenTORE project.*

CHAPTER 2

Artificially generated image (DALL·E 2, OpenAI)

Joint Models to predict dairy cow survival from sensor data recorded during the first lactation

G. Ranzato[1,2], I. Adriaens[2,3], I. Lora[1], B. Aernouts[1], J.M.E. Statham[4,5], D. Azzolina[6], D. Meuwissen[2], I. Prosepe[7], A. Zidi[1], G. Cozzi[1]

[1] *Department of Animal Medicine, Production and Health (MAPS), University of Padova, Viale dell'Università 16, 35020 Legnaro (PD), Italy*

[2] *Division of Animal and Human Health Engineering, Department of Biosystems, KU Leuven, Kleinhoefstraat 4, 2440 Geel, Belgium*

[3] *Animal Breeding and Genomics, Wageningen University and Research, P.O. Box 338, 6700 AH Wageningen, The Netherlands*

[4] *RAFT Solutions Ltd., Sunley Raynes Farm, Galphay Road, Ripon HG4 3AJ, UK*

[5] *Harper & Keele Veterinary School, 2 Church Bank, Keele, Newcastle ST5 5NS, UK*

[6] *Department of Environmental and Preventive Sciences, University of Ferrara, Corso Ercole I d'Este 32, 44121 Ferrara, Italy*

[7] *Unit of Biostatistics, Epidemiology and Preventive Sciences, Department of Cardiac, Thoracic, Vascular Sciences and Public Health, University of Padova, Via L. Loredan 18, 35131 Padova, Italy*

Animals (2022), 12, 3494

Summary

Early predictions of cows' probability of survival to different lactations would help farmers in making successful management and breeding decisions. For this purpose, this research explored the adoption of joint models for longitudinal and survival data in the dairy field. An algorithm jointly modelled daily first-lactation sensor data (milk yield, body weight, rumination time) and survival data (i.e., time to culling) from 6 Holstein dairy farms. The algorithm was set to predict survival to the beginning of the second and third lactations (i.e., second and third calving) from sensor observations of the first 60, 150, and 240 days in milk of cows' first lactation. Using 3-time-repeated 3-fold cross-validation, the performance was evaluated in terms of Area Under the Curve and expected error of prediction. Across the different scenarios and farms, the former varied between 45% and 76%, while the latter was between 3.5% and 26%. Significant results were obtained in terms of expected error of prediction, meaning that the method provided survival probabilities in line with the observed events in the datasets (i.e., culling). Furthermore, the performances were stable among farms. These features may justify further research on the use of joint models to predict the survival of dairy cattle.

INTRODUCTION

Cow survival is a complex trait that depends on multiple factors, such as milk production, fertility, health, and farm management conditions (van der Heide et al., 2020). If survival is computed from the day of the first calving, it coincides with the productive life of the animal, which represents a very important trait in the dairy practice (van Pelt et al., 2015). Typically, cows with longer productive lives are more resilient, exhibiting good productive and reproductive performances and having few health problems that they overcome rapidly (Adriaens et al., 2020). Nowadays, the average cow productive life ranges from 2.5 to 3.5 lactations (Dallago et al., 2021; Schuster et al., 2020), while a dairy cow is biologically capable of a life span up to 20 years (Hoffman & Valencak, 2020). Additionally, the research by Bach (2011) reported a decline in survival rates of first-parity cows. When dairy cows do not manage to survive beyond the first lactation, the rearing costs are not paid back; cows start making profit for the farmer only during the second lactation, reaching the full production potential during the third lactation (Cabrera, 2018). Moreover, Grandl et al. (2019) showed that cows that do not complete the first lactation perform particularly unfavorably with regard to their greenhouse gas emissions per unit of produced milk. Moreover, from an ethical perspective, short longevity is typically an indicator of poor animal welfare, being a sign of impaired biological functions and health conditions (Bruijnis et al., 2013).

Dairy farmers would benefit from a tool able to provide information about the future prospect of the first-parity cows in their herds. Based on survival predictions at farm level, they could select the ones that better cope with the existing housing and management conditions, optimizing culling decisions and breeding schemes. To date, no decision support tools have been implemented to help farmers in selecting the cows that are more likely to thrive in their own farm environment. Nowadays, some possibilities can arise from the great amount of information provided by the increasing number of sensor systems operating on many dairy farms (Lora et al., 2020; Steeneveld & Hogeveen, 2015). These new technologies provide a constant flow of high-frequency repeated measures of parameters, such as MY and quality (e.g., somatic cell count) or a cow's activity (e.g., locomotion and rumination), which can reflect changes in the physiological and health status of the animal (King & DeVries, 2018; Rutten et al., 2013). These measurements can be used to predict cow survivability using new statistical methods. These methods are based on the joint modelling of longitudinal and time-to-event data (Rizopoulos, 2012). Joint models are used in the field of biomedicine to predict patients' survival probabilities based on temporal trajectories of disease-specific biomarkers and to discriminate between patients with a low or high risk of mortality. These models are versatile, being easily adapted to different recording periods of longitudinal data, time points of survival prediction, and variables to be used in the models. Furthermore, joint models avoid deriving biologically meaningful proxies from time-series data, since they directly estimate the information provided by the raw (nearly unprocessed) longitudinal data.

The aim of the present study was to explore the adoption of a joint model that used first-lactation longitudinal sensor data of MY, body weight (BW), and rumination time (RUM) to predict cows' survival to subsequent lactations.

MATHERIALS AND METHODS

Data

Data were retrieved from 6 Holstein dairy farms (3 British, 2 Belgian, and 1 Italian) equipped with AMS of Lely Industries (Lely Industries N.V., Maasluis, The Netherlands). Farms were selected based on data availability and on farmers' willingness to participate in the study. Daily records of individual cow MY, BW, and RUM were collected from the AMS database, to be used as potential indicators of lactating cows' health status (Ouweltjes et al., 2021; Steensels et al., 2016) and, therefore, as information possibly related to their survival. Dates of cows' birth, calving, and culling were also retrieved from the farm databases. The time period covered by all the datasets varied between 2013 and 2020. Descriptive information for each farm is reported in Table 1.

Table 1. *Overview of the available datasets.*

Farm	Time Period	Cows (n)	t_1 [1]	Culled before t_1	t_2 [2]	Culled before t_2
Italian	2014–2020	98	414	12%	828	41%
Belgian 1	2014–2020	169	422	18%	843	34%
Belgian 2	2013–2019	182	397	9%	793	21%
British 1	2013–2019	266	384	9%	768	24%
British 2	2013–2019	101	402	11%	805	26%
British 3	2013–2019	226	400	6%	799	17%

[1] Average number of days between the first and second calving; [2] average number of days between the first and third calving.

Data processing

Data processing and analysis were performed with RStudio software (R version 4.1.2; RStudio PBC, Boston, MA) and equally conducted for each dataset (i.e., farm).

The survival time (T) of each cow was computed as the number of days between the first calving and the culling, coinciding with the productive life of the animal. Culling dates were derived from the last date on which milk production was registered. If no culling date was available, the cow was considered still alive at the final date of the dataset (i.e., censored), and the survival time was computed as the difference in days between the final date and the date of the first calving; the cow was removed if she had not yet completed the first lactation at the end date of the dataset. The age at first calving (AFC) of each cow was expressed as a 3-category variable: ‘low’ if it was below the first quartile of herd AFC, ‘medium’ if it was within the interquartile range, and ‘high’ if it was above the third quartile. The season of the data recording period (SEAS) was transformed into a binary variable: ‘warm’ if between April and October, ‘cold’ if otherwise.

Individual cow raw sensor data of MY, BW, and RUM recorded during first lactations were used in the study. Farm databases provided daily MY and BW in kilograms, while RUM data consisted of 2-hourly measures that were summed into single daily records expressed in minutes. According to Adriaens et al. (2020), values of each sensor variable that fell outside of 3 SD from the respective herd means were treated as outliers and removed from the dataset, except when they were present more than 30 times for the same cow. The rationale was to clean the dataset of errors in the data recording while keeping the information related to possible real disturbances (such as diseases). This was assuming a cow had an actual ‘abnormal behavior’ when outliers characterized a total of at least 30 days of the whole lactation time. Table 2 reports means and SD of daily MY, BW, and RUM for every farm.

Table 2. *Means and SD of the recorded sensor data.*

Farm	MY[1]		BW[2]		RUM[3]	
	Mean	SD	Mean	SD	Mean	SD
Italian	32.8	6.60	595	62.1	458	94.2
Belgian 1	27.5	6.33	530	65.9	470	98.0
Belgian 2	32.4	6.57	548	109	487	126
British 1	31.9	8.01	634	65.1	491	102
British 2	24.2	6.11	567	60.0	500	120
British 3	33.9	6.84	578	59.0	484	125

[1] Milk yield (kg/d); [2] body weight (kg/d); [3] rumination time (min/d).

All the cows culled before 50 days in milk of the first lactation (i.e., $T < 50$) were deleted from the dataset to examine only animals with a reasonable amount of sensor observations. Moreover, we considered first-lactation sensor measurements in the interval 5 – 305 DIM; the starting point was set at 5 DIM to avoid missing data associated with the very first days after calving, while the maximum observed time was set at 305 DIM, as it is the standard lactation length used for genetic evaluations in cattle (Græsbøll et al., 2016). After the data-filtering and cleaning procedures, a cow was removed from the dataset if she remained with less than 90% daily observations with respect to the first-lactation length (maximum 305 DIM).

Algorithm development

An algorithm based on multivariate joint modelling of longitudinal and time-to-event data was built to predict cow survival from raw daily data of MY, BW, and RUM recorded during the first lactation; 'multivariate' refers to the presence of 3 longitudinal variables to be modelled simultaneously.

The joint modelling technique has been studied by Rizopoulos (2012). It consists of two steps: (i) description of the evolution of the longitudinal variable over time using a (generalized) linear mixed model (G. Fitzmaurice et al., 2008) and (ii) estimation of the survival probabilities using the estimated evolution within a survival Cox model (Kalbfleisch & Prentice, 2002). Assuming $i = 1, \dots, n$ is the statistical unit (e.g., patient) and $k = 1, \dots, K$ identifies the different longitudinal outcomes, the evolution over time t of each outcome y_{ik} can be described by the following linear mixed model:

$$\begin{cases} y_{ik}(t) = x_i^{\mathrm{T}}(t)\beta_k + z_i^{\mathrm{T}}(t)b_{ik} + \varepsilon_{ik}(t) \\ b_{ik} \sim \mathcal{N}(0, D_k), \qquad \varepsilon_{ik}(t) \sim \mathcal{N}\left(0, \sigma_k^2\right) \end{cases},$$

where x_i are the predictors associated with the fixed effects β_k, z_i are the predictors associated with the random effects b_{ik}, and ε_{ik} is the error term. Both the vector of the random effects and the vector of the errors have a normal distribution. The correlation between the different longitudinal variables y_{ik} is then captured by setting a multivariate normal distribution for the random effects $b_i = (b_{i1}, \dots, b_{iK})^{\mathrm{T}} \sim \mathcal{N}(0, D)$. Assuming $m_{ik}(t) = x_i^{\mathrm{T}}(t)\beta_k + z_i^{\mathrm{T}}(t)b_{ik}$ is the 'true' value of each outcome at time t, we can define the following multivariate joint model (i.e., Cox hazard model containing the evolution processes of the longitudinal outcomes):

$$h_i\left(t \middle| \mathcal{M}_{i1}(t), \dots, \mathcal{M}_{iK}(t)\right) = h_0(t) \exp\left(\gamma^{\mathrm{T}}\omega_i + \sum_{k=1}^{K} \alpha_k m_{ik}(t)\right).$$

The equation $\mathcal{M}_{ik}(t) = \{m_{ik}(s), 0 \leq s \leq t\}$ represents the longitudinal history of m_{ik} until t, where $h_0(t)$ is the baseline hazard function at time t, α_k measures the association between m_{ik} and the risk of an event, and ω_i are baseline variables. The joint estimation process is carried out with a Markov Chain Monte Carlo algorithm (van Ravenzwaaij et al., 2018).

According to this theoretical approach, in the present study, the K longitudinal variables were represented by first-lactation daily sensor data: $\left(MY_i(t), BW_i(t), RUM_i(t)\right) = \left(y_{i1}(t), y_{i2}(t), y_{i3}(t)\right)$. The evolution of each $y_{ik}, k = 1, 2, 3$ over t was described by the following linear mixed model:

$$y_{ik}(t) = \beta_{0k} + \beta_{1k}\mathrm{ns}(t) + \beta_{2k}AFC_i + \beta_{3k}SEAS_i(t) + b_{i0k} + b_{i1k}\mathrm{ns}(t) + \varepsilon_{ik}, \qquad i = 1, \dots n,$$

where n was the number of cows in the dataset. The fixed effects $\beta_k = (\beta_{0k}, \beta_{1k}, \beta_{2k}, \beta_{3k})^{\mathrm{T}}$ were respectively associated with the intercept of the model, the time

t expressed as DIM (5 ≤ DIM ≤ 305), the cow's AFC, and SEAS at t. More specifically, the time was modelled with a natural cubic spline (ns). The spline was set to have one knot at the median DIM of the dataset (resulting in 2 different cubic sub-polynomials) when RUM was the longitudinal outcome. For MY and BW, the splines were set to have 3 knots at the 3 quartiles of DIM of the dataset (resulting in 4 different cubic sub-polynomials) to capture the well-defined shapes of the trend of these two traits over an entire lactation (VandeHaar & St-Pierre, 2006). The random effects $b_{ik} = (b_{i0k}, b_{i1k})^{\mathrm{T}}$ were respectively associated with the cow-specific intercept and the cow-specific time slope. The random intercept was necessary to capture the variation of the parameters of the i^{th} animal from those in the dataset, while the random slope allowed the evolution in time described by $\mathrm{ns}(t)$ to be different from one cow to another. The correlation between y_{i1}, y_{i2}, and y_{i3} was captured using $b_i = (b_{i01}, b_{i11}, b_{i02}, b_{i12}, b_{i03}, b_{i13})^{\mathrm{T}} \sim \mathcal{N}(0, D)$ with unstructured covariance matrix D. Assuming $m_{ik}(t) = \beta_{0k} + \beta_{1k}\mathrm{ns}(t) + \beta_{2k}AFC_i + \beta_{3k}SEAS_i(t) + b_{i0k} + b_{i1k}\mathrm{ns}(t)$ (i.e., the sensor value without error), we defined the following multivariate joint model:

$$h_i\big(t \big| \mathcal{M}_{i1}(t), \mathcal{M}_{i2}(t), \mathcal{M}_{i3}(t)\big) \qquad \text{(Eq.1)}$$
$$= h_0(t)\exp\big(\gamma_1 AFC_i + \alpha_1 m_{i1}(t) + \alpha_2 m_{i2}(t) + \alpha_3 m_{i3}(t)\big),$$

where the event was represented by 'the cow was culled by the last date of the dataset'. The risk of being culled at t could then be associated with the first-lactation levels of MY, BW, and RUM at t, adjusted by the animal's AFC (baseline variable).

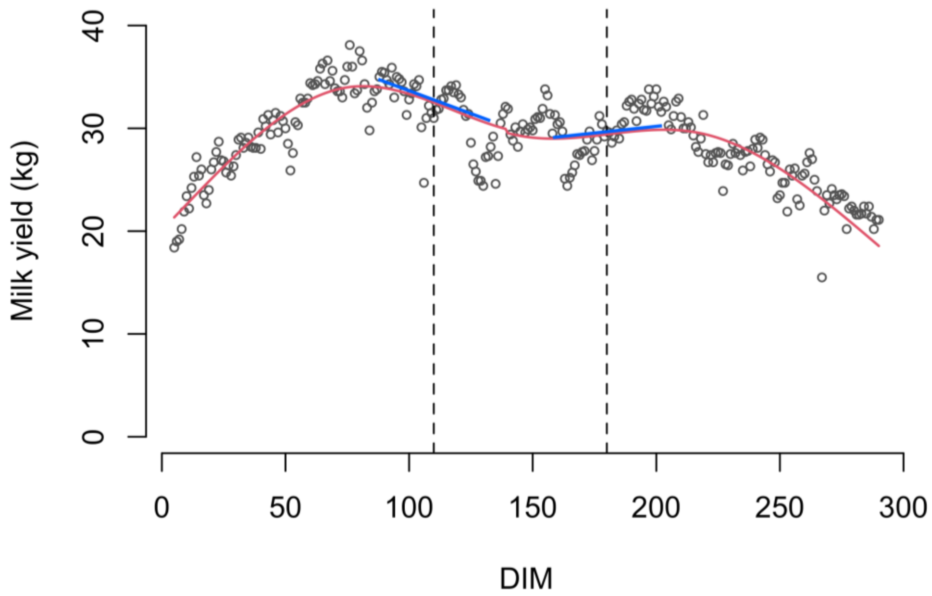
We supposed it was more likely that the risk of being culled at t could be associated with the slopes of the trajectories of the sensor variables at t, and not with their current values as in the previous model specification (Eq.1). In this way, the joint estimation process could identify fluctuations in the sensor measurements resulting from possible disturbances (such as diseases) and examine their relationship with the cow at risk of being culled. An illustrative example is reported in Figure 1 for the MY variable related to one cow; the lactation curve deviates from the typical lactation curve of dairy cattle, and this deviation is captured by the slope. The final model used in the study was then expressed by the following equation:

$$h_i\big(t \big| \mathcal{M}_{i1}(t), \mathcal{M}_{i2}(t), \mathcal{M}_{i3}(t)\big) = h_0(t)\exp\big(\gamma_1 AFC_i + \alpha_1 m'_{i1}(t) + \alpha_2 m'_{i2}(t) + \alpha_3 m'_{i3}(t)\big),$$

where $m'_{ik}(t) = \frac{d}{dt}\{\beta_{0k} + \beta_{1k}\mathrm{ns}(t) + \beta_{2k}AFC_i + \beta_{3k}SEAS_i(t) + b_{i0k} + b_{i1k}\mathrm{ns}(t)\}$ was the time-dependent slope of the sensor variable $k, k = 1,2,3,$ for cow i (i.e., the first derivative of $m_{ik}(t)$).

The modelling was carried out with R package 'JMbayes2' (Rizopoulos et al., 2022).

Figure 1. *Tangent (blue) lines to the estimated milk yield trajectory (red curve) at time points t = 110 DIM and t = 180 DIM for one cow (5 ≤ DIM ≤ 290) of one farm randomly chosen. The joint model examines the association between the slope of the tangent line at t and the risk of being culled at t.*



Algorithm Evaluation

To evaluate the performance of the algorithm, avoiding data underfitting or overfitting, repeated 3-fold cross-validation (CV) was used in every farm dataset. All the cows of the dataset were randomly partitioned into 3 groups of similar sizes; then 2 of these groups were used to train the model, and the third group was used to test it. This operation was repeated 3 times, rotating the groups (Stone, 1974). The same procedure was again repeated 3 times in total, and the mean performance across all folds from all runs was reported (i.e., mean of 9 single results per farm).

During the training, 67% of the animals in the dataset were used to fit the joint model. The model was trained on sensor data recorded during 5 – 305 DIM of the first lactation and on the cows' observed survival times, and the effect of the trajectory of each sensor variable on the risk of being culled was estimated. The testing used 33% of the cows to evaluate the prediction performance. The model accuracy in predicting cow survival was tested under 6 different scenarios: 2 different time points of prediction (i.e., second and third calving) from sensor data recorded during 3 different observation periods of the cow's first lactation (i.e., 60, 150, and 240 DIM). Survival was therefore predicted at t_1 = 'second calving' and t_2 = 'third calving', respectively, and estimated as once and twice the average calving interval (in days) after the date of the first calving for all the cows of the farm. A summary of the values of t_1 and t_2, along with the number of cows that were culled before them, is reported for each dataset in Table 1.

Given that $\mathcal{Y}_i(v) = \{y_{ik}(s), 5 \leq s \leq v, v = 60, 150, 240, k = 1,2,3\}$ represented the available first-lactation sensor measurements for a 'new' cow i of the testing set that had provided MY, BW, and RUM values up to v, individualized predictions of the survival probabilities up to $t_j, j = 1{,}2$, for cow i was obtained by estimating

$$\pi_i(u|v) = \Pr\{T_i \geq u | T_i > v,\ \mathcal{Y}_i(v),\ \mathcal{R}\}$$

where $v < u \leq t_j$, and $\mathcal{R}$ denoted the sample on which the model was fitted (i.e., the training set). Providing measurements up to time v implied that the cow was still alive at v (i.e., $T_i > v$); in every testing set, only the animals that had survived at least up to 240 DIM (i.e., the maximum period of days considered) were then examined. Assuming a specific threshold value $c \in (0,1)$ (here $c = 0.5$), cow i was finally predicted 'culled at t_j', $j = 1{,}2$, if $\pi_i(t_j|v) \leq c$. Two measures of prediction accuracy were accordingly computed based on the value of $\pi_i(t_j|v)$: the Area Under the Curve (AUC) (Heagerty & Zheng, 2005) and the expected prediction error (PE) (Gerds & Schumacher, 2006). The AUC measured the ability of the model to distinguish between the classes 'culled at t_j' and 'still on farm at t_j', representing a measure of its discrimination capability (0 ≤ AUC ≤ 1). The PE measured the accuracy of the obtained survival predictions by computing the average squared distance between the survival status (i.e., culled or alive) and the predicted survival probability, making it a measure of the calibration capability of the model (0 ≤ PE ≤ 1). The higher the AUC, the better the model performed at predicting the cows that were culled within t_j as actually 'culled at t_j' and the cows that were still on the farm at t_j as 'still on farm at t_j'; the lower the PE, the more the survival predictions were aligned with the observed events (i.e., culling) within t_j.

RESULTS

To clearly illustrate the algorithm training phase, Table 3 shows the output of the fitting obtained in one training set (148 cows; 40,995 observations) of the repeated CV procedure for one of the available farms. In this case, the longitudinal modelling process highlighted the presence of between-cow variability, expressed by the estimated SD of the random effects for the three sensor outcomes (MY, BW, and RUM). Focusing on the survival process, the slope of RUM (α_3) was negatively associated with the risk of being culled, keeping all other variables constant. This implied that a lower value of the slope was associated with poorer survival probability.

Table 3. *Output from the fitted multivariate joint model in one dataset.*

Survival outcome		
Parameters	coeff.[1]	*p*
$\gamma_{1,1}$ (AFC[2] medium)	0.997	$*$
$\gamma_{1,2}$ (AFC high)	1.35	$*$
α_1 (slope MY[3])	0.005	n.s.
α_2 (slope BW[4])	0.178	n.s.
α_3 (slope RUM[5])	-0.633	$**$

Longitudinal outcomes						
Parameters	**MY** ($k = 1$)		**BW** ($k = 2$)		**RUM** ($k = 3$)	
Fixed	coeff.	*p*	coeff.	*p*	coeff.	*p*
β_{0k} (intercept)	28.0	$***$	469	$***$	472	$***$
$\beta_{1,1k}$ $(\mathrm{ns(DIM)}$ 1)[6]	6.42	$***$	95.2	$***$	−18.6	$**$
$\beta_{1,2k}$ $(\mathrm{ns(DIM)}$ 2)	4.34	$***$	68.9	$***$	34.6	$***$
$\beta_{1,3k}$ $(\mathrm{ns(DIM)}$ 3)	13.8	$***$	86.3	$***$	-	-
$\beta_{1,4k}$ $(\mathrm{ns(DIM)}$ 4)	−13.5	$***$	99.7	$***$	-	-
$\beta_{2,1k}$ (AFC medium)	−0.268	n.s.	−3.64	n.s.	−21.3	$***$
$\beta_{2,2k}$ (AFC high)	1.24	$*$	50.9	$***$	−24.8	$***$
β_{3k} (SEAS[7] warm)	−0.021	n.s.	0.918	n.s.	−9.03	$*$
Random	SD		SD		SD	
b_{i0k} (intercept)	4.61		48.3		129	
$b_{i1,1k}$ $(\mathrm{ns(DIM)}$ 1)	7.56		43.3		191	
$b_{i1,2k}$ $(\mathrm{ns(DIM)}$ 2)	6.64		44.8		106	
$b_{i1,3k}$ $(\mathrm{ns(DIM)}$ 3)	11.4		74.5		-	
$b_{i1,4k}$ $(\mathrm{ns(DIM)}$ 4)	10.1		45.4		-	

$***$ $p < 0.001$, $**$ $p < 0.01$, $*$ $p < 0.05$, n.s. $p \geq 0.05$

[1] Mean estimate of the effect; [2] age at first calving; [3] milk yield (kg/d); [4] body weight (kg/d); [5] rumination time (min/d); [6] natural spline of days in milk (1, 2, 3, 4: sub-polynomials); [7] season of the recording period.

The mean AUC and mean PE over the 9 CV runs (3 × 3 folds) are reported in Table 4. For some farms (‘Belgian 2′ and ‘British 2′), there were no culling events registered within t_1 (i.e., second calving) in any testing set of the CV procedure; therefore, the performance metrics at t_1 could not be estimated. To determine the significance of the performance metrics over 0.50 for AUC and below 0.25 for PE (i.e., algorithm performing random guessing between ‘culled’ and ‘alive’ (Goldstein-Greenwood, 2021), we constructed a 95% confidence interval using the mean and the standard deviation obtained from the 9 CV repetitions for each farm in each scenario. The PE values were always significantly lower than 0.25 at t_1 (i.e., second calving) and, in most cases, at t_2 (i.e., third calving) (Table 4); PE was generally low at t_1, suggesting that the model accurately predicted the events within the second calving. The AUC was significantly higher than 0.50 only in a few cases, both for the predictions at t_1 and at t_2, remaining generally close to 0.50 (Table 4). Only one farm reported an average AUC of 0.76 at t_1, with 240 DIM of first lactation sensor observations to obtain predictions. It is worth noting that this was the dataset that, across training sets, had the highest number of significant associations between the sensor variables and survival, meaning that the sensor information was, in this case, particularly useful for predicting the animals' survival. These results revealed that the algorithm had a good calibration capability (PE), but the same did not apply for its discrimination capability (AUC). However, the average model performance metrics tended to improve with more days of longitudinal information (i.e., 240 vs. 150 vs. 60 DIM) and when predicting survival at closer endpoints (i.e., t_1 vs. t_2). Furthermore, the results from Levene's tests (Levene, 1960) conducted in each scenario to verify the homogeneity of variances of the AUC and PE among farms revealed that the performance metrics of the algorithm were stable. Only AUC values estimated at the third calving with 60 or 150 DIM information had different variances among farms (respectively, p = 0.01 and p = 0.02).

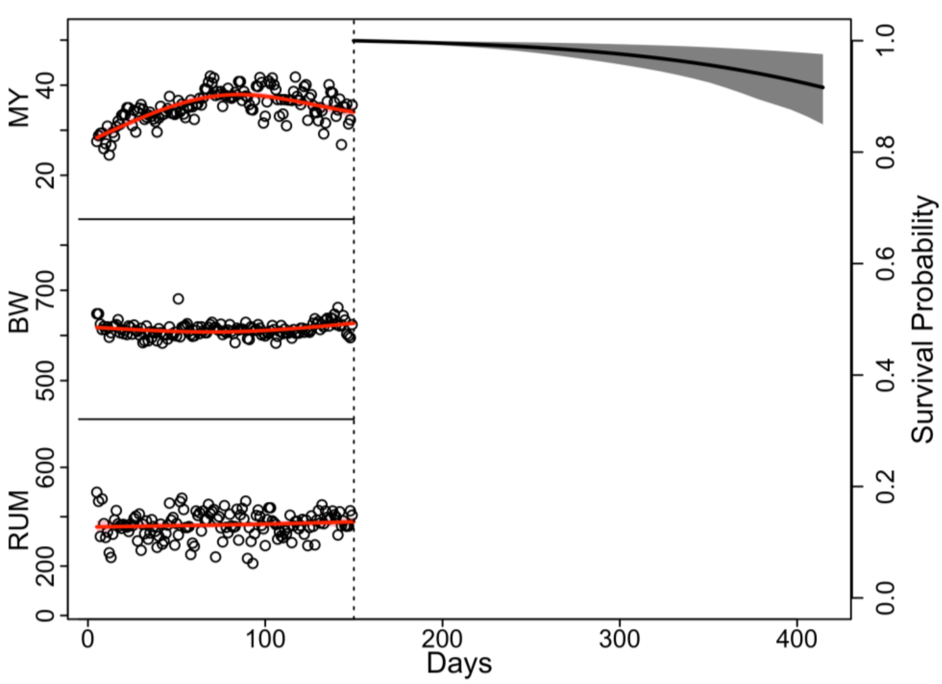
Table 4. *Predictive accuracy measures (Area Under the Curve – AUC; prediction error – PE) of the algorithm.*

DIM	Farm	AUC		PE	
		t_1[1]	t_2[2]	t_1	t_2
60	Italian	0.558	0.505	0.098 †	0.263
	Belgian 1	0.580 *	0.497	0.091 †	0.228 †
	Belgian 2	-	0.451	-	0.146 †
	British 1	0.526	0.519	0.061 †	0.202 †
	British 2	-	0.498	-	0.196 †
	British 3	0.476	0.508	0.037 †	0.164 †
150	Italian	0.605	0.556	0.096 †	0.256
	Belgian 1	0.578	0.513	0.085 †	0.225 †
	Belgian 2	-	0.475	-	0.143 †
	British 1	0.562	0.526 *	0.060 †	0.202 †
	British 2	-	0.520	-	0.194 †
	British 3	0.535	0.514	0.036 †	0.164 †
240	Italian	0.616	0.566 *	0.096 †	0.259
	Belgian 1	0.597 *	0.533	0.083 †	0.229
	Belgian 2	-	0.507	-	0.143 †
	British 1	0.577	0.539	0.060 †	0.200 †
	British 2	-	0.593 *	-	0.189 †
	British 3	0.763 *	0.559 *	0.035 †	0.158 †

* Significantly higher than 0.5; † significantly lower than 0.25.
[1] Average second calving time; [2] average third calving time.

Figure 2 represents a possible output of the algorithm, obtained by a farmer for a 'new' cow of his/her herd. The farmer may decide to keep this cow for breeding purposes, given that at 150 DIM of the first lactation, she has a predicted probability of surviving to the second calving equal to 90%.

Figure 2. *Predicted survival function from 150 DIM of the first lactation (dotted line) to the second calving (414 days post-first calving) for one cow of one farm, randomly chosen. The multivariate joint model estimates the evolutions of the sensor data (MY = milk yield (kg/d), BW = body weight (kg/d), RUM = rumination time (min/d)) until 150 DIM of the first lactation (red curves) and, based on those, predicts the survival function until the second calving (black curve).*



DISCUSSION

This study explored the possibility of using joint models to predict dairy cow survival at different lactations, starting from raw daily sensor data recorded on-farm during different (early) stages of the first lactation. The algorithm implemented in this work could represent the basis for a prognostic model-based tool able to inform farmers of the future prospect of each first-parity cow in their herds. This may be very useful in the early adjustment of herd breeding and management decisions, improving farm efficiency and sustainability; farmers could, for instance, optimize the use of dairy sexed and beef semen or decide whether to give another chance to those cows that are not pregnant after two or three inseminations.

The performances of the algorithm were compared with the results of the few similar studies dealing with dairy cow survival predictions and/or longitudinal sensor data extracted from AMS. Van der Heide et al. (2019) predicted survival to the second lactation using breeding and phenotypic variables from different moments in the heifer's life. The authors compared three different machine-learning methods for many performance metrics, including AUC. Average AUC was 0.67 when using the information available at 6 weeks post-first calving (i.e., 40–50 DIM) and 0.68 when using the information at 200 DIM. The performance of these models was then higher compared to our average results (AUC = 0.54 ± 0.05 at second calving using 60 DIM, and AUC = 0.64

± 0.09 at second calving using 240 DIM; mean ± SD), but their ability to correctly identify non-surviving animals was very low (average positive predictive value of 0.17). The same authors tried to improve these performances by using ensemble-learning approaches (van der Heide et al., 2020), which were expected to have better performances and more robustness, but the results remained quite poor (average positive predictive value of 0.20). Adriaens et al. (2020) studied the possibility of predicting lifetime resilience and the productive life of dairy cows starting from sensor-derived proxies of first-parity daily sensor data, obtaining a mean classification performance ('low' vs. 'medium' vs. 'high' lifetime resilience rank) of 47 ± 8% (± SD), when using milk yield features alone, and of 56 ± 12% when using lactation and activity features together. Ouweltjes et al. (2021) assessed the performance of different models that included milk yield, body weight, rumination, and activity sensor data of cows in first lactation to predict lifetime resilience. Model performances, expressed in the percentage of correctly classified cows ('low' vs. 'medium' vs. 'high' lifetime resilience rank), ranged between 45 ± 8% (mean ± SD) and 51 ± 6%.

The results of this research, in line with the results of the other works, confirm that cow survival is a complex trait, difficult to accurately predict (van der Heide et al., 2020). It indeed combines several different factors, such as fertility, health, milk production, farm management, and environmental conditions (Olechnowicz et al., 2016). With the only information at our disposal (i.e., AFC, SEAS, MY, BW, RUM), we could capture a small portion of these aspects; for instance, having information on disease occurrence would have likely improved the predictive performance of the algorithm. Furthermore, to build an algorithm applicable to all the farms with MY, BW, and RUM data from AMS, we had to ignore all the local and evidence-based farm management rules, which are particularly relevant when developing decision support tools for dairy farms (Adriaens et al., 2020).

We identified two main strengths of the methodology presented in this study. First, in contrast to other works with similar research goals (Adriaens et al., 2020; Poppe et al., 2020), the present joint modelling approach has the practical advantage of not requiring the translation of sensor time-series data into biologically meaningful sensor features. Using raw sensor data to obtain longevity predictions avoids proper feature definition and a lot of pre-processing (thus reducing the chance of errors) and provides at least the same performance as models with pre-processed data, as demonstrated by Ouweltjes et al. (2021). Second, joint models have the advantage of being very flexible; they allow for the dynamic update of predicted survival probabilities as additional longitudinal data are recorded, as well as for the easy change of the final time point of prediction based on the target the user wants to test. These features may justify future research to improve the current performance within a farm. The model, for instance, could be tested by including additional variables from automated technologies (e.g., cow activity, somatic cell count) or cows' additional information from other sources (e.g., test days, health records).

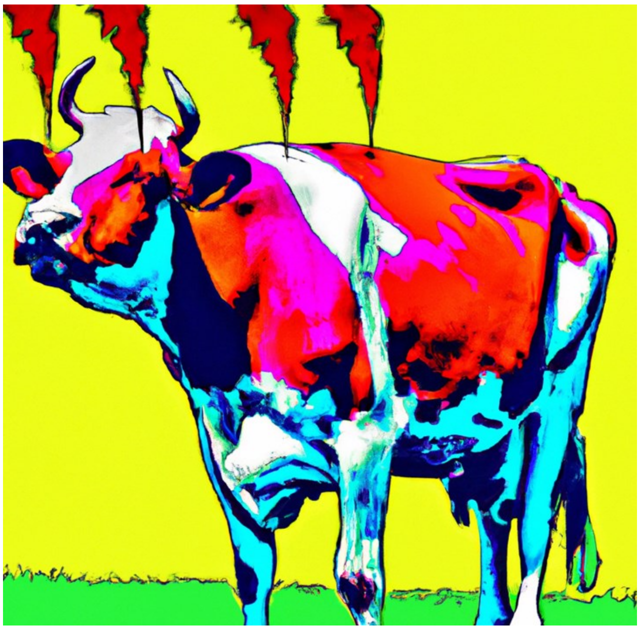
CONCLUSIONS

This study explored the potential of using joint models for longitudinal and time-to-event data to predict dairy cow survival at different lactations from raw sensor data recorded during different stages of the cow's first lactation. The algorithm tested in this study had a modest performance in terms of discrimination accuracy (Area Under the Curve) but good results in terms of calibration accuracy (expected error of prediction), as well as good repeatability across different farms. The interesting opportunities that joint models offer in applicability and flexibility should justify further research in the attempt to improve the overall predictive accuracy in the dairy field.

ACKNOWLEDGMENTS

This research was funded by the European Union's Horizon 2020, GenTORE project grant number 727213.

CHAPTER 3



Artificially generated image (DALL·E 2, OpenAI)

Sensor-based behavioral patterns can identify heat-sensitive lactating dairy cows

G. Ranzato[1,2], I. Lora[1], B. Aernouts[2], I. Adriaens[2], F. Gottardo[1], G. Cozzi[1]

[1] *Department of Animal Medicine, Production and Health (MAPS), University of Padova, Viale dell'Università 16, 35020 Legnaro (PD), Italy*

[2] *Division of Animal and Human Health Engineering, Department of Biosystems, KU Leuven, Kleinhoefstraat 4, 2440 Geel, Belgium*

Accepted for publication in *International Journal of Biometeorology* (2023)

Summary

Heat stress impairs the health and performance of dairy cows, yet only a few studies have investigated the diversity of cattle behavioral responses to heat waves. This research was conducted on an Italian Holstein dairy farm equipped with precision livestock farming sensors to assess potential different behavioral patterns of the animals. Three heat waves, defined as at least five consecutive days with mean daily temperature-humidity index higher than 72, were recorded in the farm area during the summer of 2021. Individual daily milk yield data of 102 cows were used to identify 'heat-sensitive' animals, meaning the cows that, under a given heat wave, experienced a milk yield drop that was not linked with other health events (e.g., mastitis). Milk yield drops were detected as perturbations of the lactation curve estimated by iteratively using Wood's equation. Individual daily minutes of lying, chewing, and activity were retrieved from ear-tag-based accelerometer sensors. Semi-parametric generalized estimating equations models were used to assess behavioral deviations of heat-sensitive cows from the herd means under heat stress conditions. Heat waves were associated with an overall increase in the herd's chewing and activity times, along with an overall decrease of lying time. Heat-sensitive cows spent approximately 15 min/d more chewing and performing activities ($p < 0.05$, respectively). The findings of this research suggest that the information provided by high-frequency sensor data could assist farmers in identifying cows for which personalized interventions to alleviate heat stress are needed.

INTRODUCTION

The effect of environmental heat on livestock species is a topic of growing concern, especially in light of the current climate change (Vitali et al., 2015). Future scenarios for the temperature-humidity index (THI) are not promising: Segnalini et al. (2013)

forecasted an increase of THI in the Mediterranean area which will cause growing thermal discomfort to the animals, with negative consequences on their welfare, performance, health, and survival (Vitali et al., 2015). In addition to the overall increase in temperatures, heat waves (HW) are becoming more and more frequent, intense, and extended (Beniston et al., 2007). In the next decades, the Mediterranean basin is expected to be dominated by increased droughts and HW (Gao & Giorgi, 2008) to the point that it has been categorized as a global warming hotspot (Segnalini et al., 2013).

Lactating dairy cows produce a large quantity of metabolic heat, that under heat stress periods is coupled with a compromised cooling capability because of environmental conditions (West, 2003). This brings heat load in the cows to raise, causing an increase of body temperature and the inability to maintain thermal energy balance (Becker et al. 2020). As adaptive response, the animals show physiological and behavioral changes (Islam et al., 2021): increased respiration rate (de Andrade Ferrazza et al. 2017), decreased lying time (Allen et al. 2015), increased shade utilization (Brown-Brandl et al. 2003), increased water intake (Coimbra et al. 2012), and reduced feed intake resulting in reduced milk yield (Bohmanova et al. 2007). Each of these adaptations aims at mitigating metabolic heat production and promoting the dissipation of body temperature (Islam et al., 2021). Furthermore, thermal stress may partially suppress the innate immune functions in lactating cows, leading to a higher risk of clinical diseases such as mastitis and metritis (Becker et al., 2020).

Monitoring dairy cows is crucial to identify and manage heat stress to limit its negative impact on welfare, health, and production (Hut et al., 2022). Nowadays, many sensor systems are commercially available to replace visual observation of the animals, which can be impractical in large commercial herds (Barriuso et al., 2018). Accelerometer-based systems are the most widely available and validated technology for continuous, real-time, and autonomous monitoring of core behaviors like eating, rumination, lying, and walking (Allen et al., 2015; Islam et al., 2021; Stygar et al., 2021b). The information provided by sensor-based behavioral data can be used to assist farmers in the early identification of climate-related distress (Abeni & Galli, 2017), minimizing the negative economic and welfare implications of heat stress.

Several studies have already investigated the main overall effects of heat stress on dairy cows' behavior, as described above. However, to our knowledge, only Islam et al. (2021) analyzed changes in cows' behavior based on the animals' different responses to thermal stress. To further explore this topic, we retrospectively identified the 'heat-sensitive' cows of a dairy herd based on their drop in milk yield associated with summer heat waves. Individual high-frequency sensor data were analyzed to detect different behavioral patterns of heat-sensitive animals, and to explore the potential of sensor systems to early identify cows for which personalized interventions to alleviate heat stress are needed.

MATERIALS AND METHODS

Dataset

The study used data (Ranzato, 2023) of 369 Holstein-Friesian cows (43% primiparous, 57% multiparous) from a dairy farm located in the Po Valley, Italy (45.263791, 9.021119; temperate climate). The animals were housed in two barns equipped with high-volume low-speed horizontal ceiling fans and with cooling showers along the feeding alleys ('new ventilation system') in the period between 25 DIM and 280 DIM. During the post-partum days and in the late lactation phases, the cows were moved to two barns equipped with only old vertical fans ('old ventilation system'). For the period from 28th March to 30th September 2021 covered by the study, cows were fed total mixed rations based on maize silage, grass silage, maize and soybean meals (average dry matter content of 53%, and crude protein, NDF, starch, ether extracts and ash mean contents of 15.8%, 31%, 28.8%, 4.3% and 6.2% of dry matter, respectively).

Behavioral data regarding lying (LIE), chewing (CHEW), and activity (ACT) times were collected by an ear-tag-based accelerometer (Smartbow GmbH, Weibern, Austria) that registered the time budgets of each cow by measuring head and ear movements (Krieger et al., 2019). The ear tag captured and sent acceleration data once per second (1 Hz); daily minutes of LIE (lying + standing = 1440 min/d), CHEW (chewing + 'not chewing' = 1440 min/d), and ACT (activity + inactivity = 1440 min/d) were used in the study. Individual cow health events and individual daily milk yield (MY) data, automatically recorded in the milking parlor, were also retrieved from the farm databases.

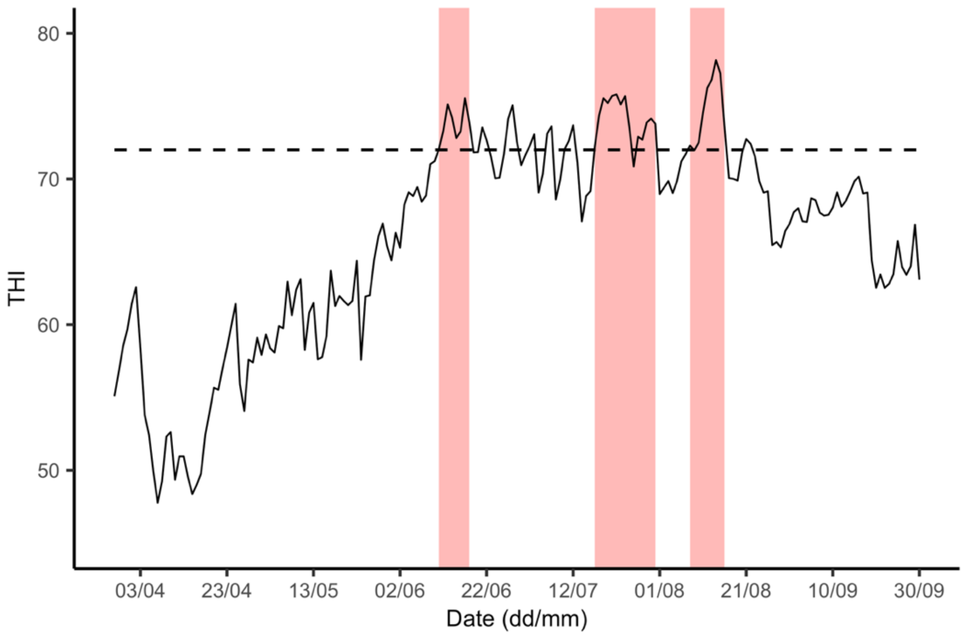
Climate data were restored from the online archive of the local environmental protection agency (Agenzia Regionale Protezione Ambiente, 2022), by referring to the nearest weather station to the farm (7 km of distance). The average daily temperature (T, in °C) and average daily relative humidity (RH, in %) were used to compute average daily THI according to the equation by Kelly and Bond (1971):

$$THI = (1.8 \cdot T + 32) - (0.55 - 0.55 \cdot RH/100) \cdot \{(1.8 \cdot T + 32) - 58\}.$$

A HW is generally described as a prolonged period of excessively hot weather, but no official definition is available (Maggiolino et al., 2022). Commonly, a THI of 72 is considered the threshold after which milk production starts to decrease in Holstein cows because of thermal discomfort (Heinicke et al., 2018; Segnalini et al., 2013). In this study, a HW was then defined as a period of at least 5 consecutive days (Frich et al., 2002) with a mean daily THI ≥ 72. If successive HW were less than three days apart from each other, they were considered as one HW (therefore, one HW could contain days with THI < 72). Three HW were identified during the study period (Fig.1): from 11/06 (dd/mm) to 18/06; from 17/07 to 31/07; from 08/08 to 16/08. Table 1 gives an overview of the meteorological characteristics of each HW.

Data mining, processing steps, and statistical analyses were carried out with RStudio software (R version 4.1.2; RStudio PBC, Boston, MA).

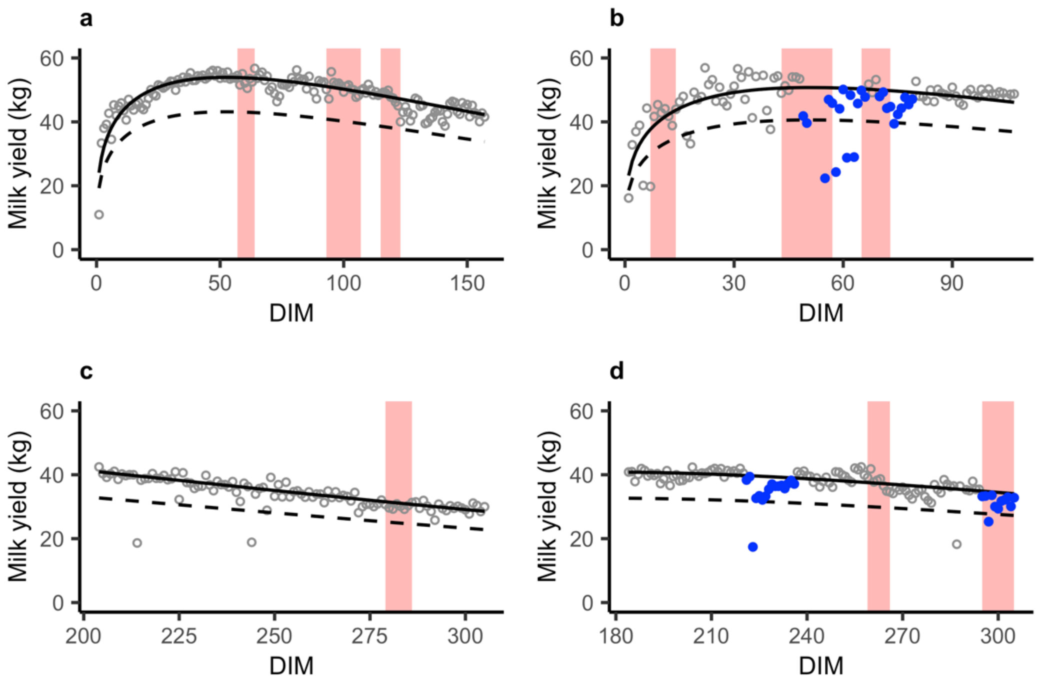
Figure 1. *Average daily temperature-humidity index (THI) during the period of the study (28th March – 30th September 2021) and identified heat waves in pink-colored areas (the dashed line represents the threshold of THI = 72).*



Heat-sensitive cows

Daily MY data were used to identify the cows that were more sensitive to thermal heat. Given that at any production level dairy cattle show an inverse relationship between milk yield and heat stress (Becker et al., 2020; Ravagnolo et al., 2000; West, 2003), we assumed 'heat-sensitive' cows to be the ones that started at least one consistent drop in milk production during a HW. Following the work by Adriaens et al. (2021), drops in milk production were identified as perturbations in the lactation curve compared with the theoretical production for that lactation (i.e., potential milk production when no disturbances are present). Unlike in Adriaens et al. (2021), we didn't have access to complete lactations data due to the restricted time period of the study. Therefore, lactations were selected based on the following criteria: (i) MY data were available from before DIM 30 for at least 100 days, or (ii) MY data were available from beyond DIM 150 for at least 50 days. These filters were necessary to grasp a proper image of the lactation curves in the observation period. We removed records beyond DIM 305 for standardization purposes, because the last part of the lactation curve can be influenced by the gestation stage and feed changes towards dry-off (Adriaens et al., 2021; Ben Abdelkrim et al., 2021). After data editing to remove recording errors (e.g., MY = 0 kg/d), we kept only lactations with no more than 2 gaps of at most 5 days each. To determine the theoretical shape of the lactation curves of the 108 cows left in the dataset, a Wood model (Wood, 1967) was iteratively fitted on the MY data of each animal (for more details see Adriaens et al., 2021). Next, the periods classified as perturbations of the milk production curve were identified as at least 5 days of successively negative residuals with at least one day of MY lower than 80% of the theoretical curve. To illustrate this methodology, MY data of 4 cows are plotted in Figure 2, and MY perturbations, if present, are highlighted in blue.

Figure 2. *Estimated theoretical milk production curve (solid line; respective 80% in dashed line) and identified perturbations (blue dots) for different lactations during the observation period (**a**: 1 ≤ DIM ≤ 157; **b**: 5 ≤ DIM ≤ 180; **c**: 204 ≤ DIM ≤ 305; **d**: 221 ≤ DIM ≤ 305). The pink-colored areas identify the heat waves during the observation period.*



Cows with registered pathologies or health events that influenced their lactation curve (e.g., mastitis) were removed from the dataset to retain only animals with MY perturbations potentially due to heat stress. Nonetheless, for some lactations, perturbations of MY were detected also outside HW. These were handled with the following criteria: (i) when they started before a HW and overlapped with the HW itself, the respective records during HW days were removed to consider only perturbations originated under thermal stress; (ii) when they started during a HW and continued under the next HW, they were kept as prolonged heat stress effects; (iii) when they were completely outside HW, they were kept as perturbations caused by unknown reasons. Two examples of cows found to be heat-sensitive to one or more heat waves are shown in Figure 2b and Figure 2d.

Data processing and statistical analysis

A categorical variable identifying the stage of the lactation was created according to the observed time period of each animal: 'early lactation' when DIM ≤ 100, 'mid lactation' when 100 < DIM ≤ 200, and 'late lactation' when DIM > 200 (Niozas et al., 2019). A binary variable was created to indicate the type of barn where the cows were housed (i.e., 'old ventilation system' vs. 'new ventilation system'). On specific dates, the sum of daily minutes of each behavior (i.e., LIE, CHEW, ACT) with its complementary (i.e., standing, not chewing, inactivity) was lower than 1440 min/d, probably due to recording errors: when the sum was below 1320 min/d (i.e., more than 2 hours of the day were missing), that record was deleted from the dataset; when the sum was between 1320 and 1439 min/d, the respective values were reproportioned to sum to 1440 min/d. The final dataset contained 102 cows, each one with a number of records ranging between 48 and 159 days of data, for a total amount of 12,949 records.

To detect different behavioral trends of heat-sensitive cows, we first used linear mixed-effects models, which are traditionally applied for analyzing longitudinal data. They produced non-normal residuals, even when using transformations of the response variables (i.e., LIE, CHEW, ACT). Accordingly, we used a semi-parametric technique that handles repeated measures, referred to as generalized estimating equations (GEE; R package 'geepack', Halekoh et al., 2006). Generalized estimating equations are used to estimate the parameters of a (generalized) linear model specifying a working correlation structure that accounts for within-subject correlation of the response variable (Hardin & Hilbe, 2013). Different correlation structures can be specified, including independence of observations, exchangeable correlation, first-order autoregressive structure, and unstructured correlation; the most appropriate working correlation structure is the one that produces the smallest correlation information criterion (CIC; Hin and Wang 2009). Generalized estimating equations models are population average models, meaning that the estimated effects are interpreted as for (generalized) linear models but at 'population' level (Hubbard et al., 2010).

The GEE models for the different behaviors were specified as follows:

$$\begin{aligned}\mu_{ij} = \beta_0 &+ \beta_1 multiparous_i + \beta_2 mid\ lactation_{ij} + \beta_3 late\ lactation_{ij} \\ &+ \beta_4 new\ ventilation\ system_{ij} + \beta_5 HW_{ij} + \beta_6 perturbation_{ij} \\ &+ \beta_7 HW_{ij} \cdot perturbation_{ij},\end{aligned} \quad \text{(Eq.1)}$$

where μ_{ij} represents the mean of the response variable y_{ij} ($j = 1, \ldots, 159$ daily measurement of cow $i = 1, \ldots, 102$) corresponding to LIE, CHEW, or ACT times. The variance of y_{ij} is a function of a known variance function v of the mean and a known scale parameter ϕ $\left(V(y_{ij}) = v(\mu_{ij})\phi\right)$, accounting for within-subject correlation of the observations. The information on the HW was expressed by a binary time-dependent variable ('no HW' as reference class, and 'HW'), as well as for the presence of drops in the milk curve ('no perturbation' as reference class, and 'perturbation'). The interaction term referred to heat-sensitive cows, i.e., cows experiencing one or more MY perturbations ($perturbation = 1$) started during a HW ($HW = 1$). The reference categories of the variables used to adjust the comparison between cows were: 'primiparous' for the parity information, 'early lactation' for the variable indicating the stage of the lactation, and 'old ventilation system' for the variable distinguishing the barn based on the type of ventilation system.

The effect of each variable in Eq.1 was quantified by the estimation of the related regression parameter. The effect of HW and $perturbation$ had to be averaged over the levels of the other variable involved in the interaction term (e.g., HW effect $= \hat{\mu}_{HW} - \hat{\mu}_{no\ HW} = \hat{\beta}_5 + \hat{\beta}_7 perturbation = \{\hat{\beta}_5 + (\hat{\beta}_5 + \hat{\beta}_7)\}/2$). The behavioral variation of heat-sensitive cows with respect to the farm means during HW (i.e., HW = 1, $perturbation$ = 1 vs. HW = 1, $perturbation$ = 0) was estimated by $\hat{\beta}_6 + \hat{\beta}_7$.

RESULTS AND DISCUSSION

Three HW were identified during the summer of 2021 (June – September) in the area of the Po Valley where the farm was located. The percentage of heat-sensitive cows in the herd, i.e., the percentage of cows that experienced one or more MY perturbations during a HW, increased from the first to the last HW (Table 1). This result was somehow expected, as the last HW (from 8th to 16th August) was the most severe with a maximum daily THI of 78.2. A THI ≥ 75, in fact, generates alarming conditions for both the welfare and performance of dairy cows (Segnalini et al., 2013). However, we cannot exclude that a heat stress carry-over effect may also have played a role in the increase of the percentage of heat-sensitive cows as a function of the number of HW (Herbut et al., 2018).

Table 1. *Characteristics of the three heat waves (HW) identified during the study period (28th March – 30th September 2021): mean daily temperature (T), mean daily relative humidity (RH), mean daily temperature-humidity index (THI), minimum daily THI, maximum daily THI, number of cows (overall n = 102), number of records (overall n = 12,949), percentage of heat-sensitive cows (i.e., cows that started one or more perturbations of the lactation curve during the heat wave).*

	T (°C)	**RH** (%)	**THI**			**Cows** (n)	**Records** (n)	**Heat-sensitive cows** (%)
			mean	min.	max.			
HW 1(11/06[1] – 18/06)	25.6	61.6	73.8	72.1	75.6	85	670	12
HW 2 (17/07 – 31/07)	24.8	75.3	74.1	70.9	75.8	66	945	18
HW 3 (08/08 – 16/08)	25.5	72.2	74.8	72.0	78.2	57	478	37

[1] dd/mm.

Sensor technologies provide opportunities to constantly monitor dairy cattle behavior, and they may assist farmers in the early identification of thermal stress symptoms (Abeni & Galli, 2017; Hut et al., 2022). To our knowledge, the ear-tag-based accelerometer sensor used in this study has been validated for chewing monitoring (Borchers et al., 2016; Reiter et al., 2018), but no references could be found for lying and activity times in dairy cows. Only Roland et al. (2018) reported that the ear tag reached a satisfying accuracy in detecting posture (i.e., lying vs. standing) and a substantial agreement for some activities in dairy calves. The average sensor-based LIE time of the herd, recorded in the period 28th March – 30th September 2021, was 693 min/d, while mean CHEW and ACT times were respectively 596 min/d and 1097 min/d (Table 2). Data available in the literature show that lactating dairy cows spend 660 to 840 min/d lying down under thermoneutral conditions (Becker et al., 2020). Chewing time may vary across dairy herds as chewing activity can be affected by physical properties of the diet (Beauchemin et al., 2003) and selective feed intakes of the animals (Maulfair et al., 2010), but also by farmers' decisions on cows' grouping (Grant & Albright, 2001). Even activity is a rather variable measure for which it is difficult to define a reference range, as it can vary due to barn design, herd management, and especially type of sensor system used for its recording.

Table 2. *Minimum, first quartile (Q_1), median, mean, third quartile (Q_3), and maximum of the behavioral data retrieved from ear-tag-based accelerometer sensors (lying, LIE; chewing, CHEW; activity, ACT).*

	LIE (min/d)	**CHEW** (min/d)	**ACT** (min/d)
min.	156	125	442
Q_1	592	537	1023
median	700	592	1089
mean	693	596	1087
Q_3	801	650	1153
max.	1212	1085	1426

Results from the three GEE models fitted on LIE, CHEW, and ACT times are reported in Table 3. No model simplification strategy was applied. The working correlation structure that produced the smallest CIC was, for all the models, the exchangeable structure, meaning that all pairs of observations within a subject could be considered equally correlated.

Cow's parity affected all the recorded behaviors (Table 3). Multiparous cows had longer LIE and CHEW times being less active than primiparous ones. Behavioral differences across parities have been described earlier, although mostly focused on the animals' transition period (Azizi et al., 2010; Neave et al., 2017). As an effect of hierarchical differences between primiparous and multiparous cows, younger animals entering the milking herd for the first time spend less time lying down and increase their daily activity. Cows in mid lactation tended to be more active compared to cows in early lactation ($0.01 < p < 0.05$), while no differences across lactation stages were found for the other behaviors. Also the type of ventilation system did not appear to affect, on average, the behavior of the animals.

Table 3. *Results from the Generalized Estimating Equations models on cows' daily minutes of lying (LIE), chewing (CHEW), and activity (ACT): estimated parameters (coeff.) associated to the variables of the model, standard errors in brackets, and corresponding observed levels of significance.*

Variables	**LIE** (min/d)		**CHEW** (min/d)		**ACT** (min/d)	
	coeff. (SE)	*p*	coeff. (SE)	*p*	coeff. (SE)	*p*
intercept	679 (18.7)	***	557 (9.19)	***	1096 (12.7)	***
multiparous	86.7 (22.1)	***	40.0 (13.1)	**	-54.8 (13.5)	***
mid lactation	-3.70 (9.78)	n.s.	5.95 (5.89)	n.s.	15.2 (6.45)	*
late lactation	-18.7 (11.2)	n.s.	4.75 (7.30)	n.s.	6.15 (8.45)	n.s.
new ventilation system	-14.4 (8.36)	n.s.	7.32 (5.84)	n.s.	1.05 (8.12)	n.s.
HW[1]	-37.0 (3.30)	***	38.8 (3.92)	***	43.8 (3.11)	***
perturbation	12.0 (7.53)	n.s.	-10.3 (4.58)	*	-23.9 (6.30)	***
HW × perturbation	-9.87 (12.5)	n.s.	25.5 (13.1)	*	39.4 (21.3)	**

[1] Heat wave.

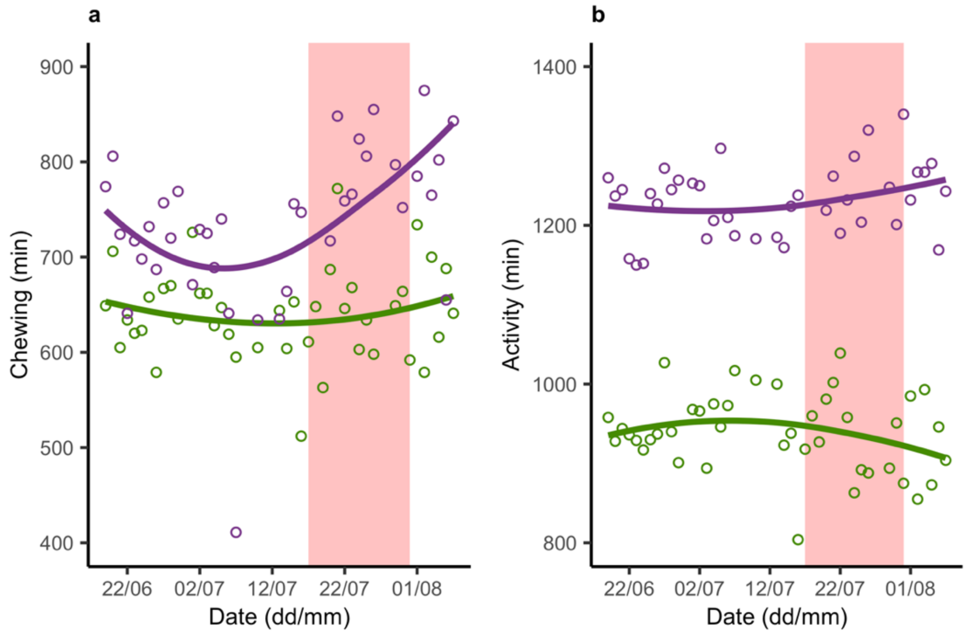
Overall, the average LIE time of the herd decreased during HW periods by 42 min/d. By standing, in fact, cattle expose a greater body surface to the air which helps heat loss due to the convection phenomenon (Allen et al., 2015). Chewing time increased by 52 min/d under prolonged periods of environmental heat. This result may seem in contradiction with some works affirming a tendency of chewing activity to decrease under heat stress conditions (Karimi et al., 2015; Maia et al., 2020). Considering chewing time as the summation of rumination and eating times (Perdomo et al., 2020), the increase of CHEW found in this study could be determined by an increase of eating time due to the presence of cooling systems along the feeding alleys. However, the accuracy of accelerometer ear tags in monitoring chewing activity can be variable depending on the conditions of use, and it can be biased by other movements of the animal's head

(Beauchemin, 2018). The average ACT time increased by 64 min/d during HW. The ear tag recorded a cow being active when the animal was actively moving either standing or lying down. Therefore, ACT could include walking, exploring, drinking, urination, defecation, grooming, head swing, and estrus expression (Becker et al., 2020; Zambelis et al., 2019). Consistent with this finding, Abeni and Galli (2017) reported higher daily activity times associated with higher THI exposure in dairy cows: the animals tend to have more frequent movements of the head, recorded by the sensor, when the environmental temperature increases (Cook et al., 2007). Brzozowska et al. (2014) stated that also the number of steps per day increases during heat stress periods.

Focusing on the heat-sensitive cows ($\hat{\beta}_7$, $HW * perturbation$ effect in Table 3), their behavioral patterns were similar to those of all the cows exposed to HW, but more severe variations were detected for CHEW and ACT times. Heat-sensitive cows chewed 15 min/d more and were 16 min/d more active with respect to the herd means during HW. To our knowledge, only Islam et al. (2021) compared the behavior of 'heat-susceptible' and 'heat-tolerant' heifers by making a distinction based on their panting scores. These authors compared sensor-based eating, rumination, and lying times of the two groups of cows during 3 different HW events, but without involving thermoneutral conditions in the study. They found that heat-susceptible animals spent more time eating (+17 min/d; $p < 0.001$) and less time lying down (-61 min/d; $p = 0.039$) compared to heat-tolerant ones, suggesting that heat-tolerance can be at the expense of reduced production either by inborn genetic merit or by adaptive reduced feed intake. Similarly, we could assume that heat-sensitive cows increased their daily eating time resulting in longer chewing times.

Our results indicate that cows belonging to the same herd, therefore under the same environmental and management conditions, can have different behavioral adaptations to heat stress. Figure 3 is an example of different CHEW (Figure 3a) and ACT (Figure 3b) patterns of two cows of the herd, one of which resulted to be sensitive to the HW from 17th to 31st July. The identification of the more sensitive animals to thermal distress through the monitoring of their behavioral responses could allow targeted interventions by the farmers to alleviate heat stress symptoms. Farmers could decide to create specific groups of heat-sensitive cows to be housed in areas where the cooling is more effective, or to adjust their feeding schedule as limiting feed availability during the hottest hours of the day can reduce heat stress (Davis et al., 2003). In parallel, they could select 'heat-tolerant' cows (i.e., the ones that didn't experience any MY perturbations during HW) for breeding purposes, and optimize breeding schemes and culling decisions (Ranzato et al., 2022). Recent studies have in fact shown that selection for heat-tolerant cows' genotypes is feasible and leads to improvements in milk production and feed intake during and after heat stress events (Liu et al., 2017).

Figure 3. *Chewing (**a**) and activity (**b**) patterns of two cows during one period of summer 2021. One cow (green line) was not sensitive to the heat wave from 17th to 31st July (pink-colored area), the other (purple line) was sensitive to the same heat wave.*



As sensors systems are in continuous progress in the dairy farming sector, it is possible that heat stress sensitivity will be even more accurately assessed combining behavioral information with additional parameters automatically detected. Chen et al. (2018) found that cows more sensitive to thermal stress have higher rectal temperature than less sensitive ones, along with upregulated metabolisms and downregulated neurodegenerative disease pathways. Liu et al. (2017) obtained differences in lipid concentration of milk between heat-sensitive and heat-tolerant cows during heat stress. Finally, Herbut et al. (2018) affirmed that the time of the year and the breed of the cows may have a big impact on when the animals become sensitive to increasing heat loads.

Under heat stress conditions, a consistent decrease in daily milk production is usually registered 48 hours after the thermal stress onset (Spiers et al., 2004). The advantage of referring to high-frequency behavioral data for heat stress detection is that the sensor system could immediately give an alarm when, for example, chewing and activity times overpass specific thresholds, thus preceding the actual milk yield drop. Further research involving more farms and covering more years of recorded behavioral data could be useful to lay the foundations for a decision-support tool for dairy farmers.

CONCLUSIONS

The identification of more sensitive animals to thermal distress through automatic monitoring of deviations in their behavior could allow targeted interventions by the farmers to alleviate heat stress symptoms. This study aimed at exploring differences in

Holstein-Friesian dairy cows' lying, chewing, and activity times, recorded by ear-tag-based accelerometer sensors, when exposed to heat waves. 'Heat-sensitive' cows were identified by the presence of one or more drops in milk yield that started during a given heat wave. The percentage of animals that resulted to be sensitive to heat stress increased progressively from the first to the last heat wave detected during the observation period. Heat-sensitive cows revealed significant deviations from the herd mean behaviors during heat waves. In particular, they spent more daily time chewing and being active, probably in the attempt to better cope with the environmental heat. The analysis of individual high-frequency data from sensor systems gives therefore the chance to early identify those cows for which personalized interventions to alleviate thermal stress are needed. Further research involving more years of data and dairy farms could strengthen the promising outcomes of this study.

ACKNOWLEDGMENTS

This work is part of GenTORE project, which has received funding from the European Union's Horizon 2020 research and innovation program under grant No. 727213.

CHAPTER 4

Artificially generated image (DALL·E 2, OpenAI)

Comparison of three mathematical methods to estimate dairy cows' lactation performances

G. Ranzato[1,2], B. Aernouts[2], I. Lora[1], I. Adriaens[2,3,4], A. Ben Abdelkrim[5], M.J. Gote[2], G. Cozzi[1]

[1] *Department of Animal Medicine, Production and Health (MAPS), University of Padova, Viale dell'Università 16, 35020 Legnaro (PD), Italy*

[2] *Division of Animal and Human Health Engineering, Department of Biosystems, KU Leuven, Kleinhoefstraat 4, 2440 Geel, Belgium*

[3] *BioVism, Department of Data Analysis and Mathematical Modelling, Ghent University, 9000 Ghent, Belgium*

[4] *Animal Breeding and Genomics, Wageningen University and Research, P.O. Box 338, 6700 AH Wageningen, The Netherlands*

[5] *Lactanet, Sainte-Anne-de-Bellevue, QC, H9X 3R4 Canada*

Submitted to *Journal of Dairy Science*

Summary

Milk yield dynamics and production performances reflect how dairy cows cope with their environment. To optimize farm management, time-series of individual cow milk yield have been studied in the context of precision livestock farming, and many mathematical models have been proposed to translate raw data into useful information for the stakeholders of the dairy chain. To gain better insights on the topic, this study aimed at comparing three recent methods that allow to estimate individual cow milk lactation performances, using daily data recorded by the automatic milking systems of 14 dairy farms (7 Holstein, 7 Italian Simmental) from Belgium, the Netherlands, and Italy. Iterative Wood model (IW), perturbed lactation model (PLM), and quantile regression (QR) were compared in terms of estimated total unperturbed (i.e., expected) milk production and, with respect to it, estimated total milk loss. The IW and PLM can also be used to identify perturbations of the lactation curve and were thus compared in this regard. The outcome of this study may help a given end-user in choosing the most appropriate method according to his specific requirements. If there is a specific interest in the post-peak lactation phase, IW fits best. If one wants to accurately characterize the perturbations of the lactation curve, PLM is the most suitable method. If there is need for a fast and 'rough' approach on a very large dataset, QR is the best choice. Finally, as an example of application, PLM was used to analyze the effect of cows' parity, calving season, and breed on their estimated lactation performances.

INTRODUCTION

Dairy cows' lactation curve represents MY as a reproducible pattern that can be expressed mathematically as a function of time: an ascending phase leading up to the peak production, and a following descending phase (Li et al., 2022). Many factors can affect the lactation dynamics and, accordingly, milk production throughout the lactation. Farm housing and management strategies are just as crucial as genetic and environmental factors (Adriaens et al., 2021). Aspects like the type of flooring or stocking density can affect dairy herd performance through the occurrence of lameness, and feed-bunk space may influence MY due to competition and stress at the feeding rack (Bach et al., 2008). Recently, precision livestock farming has contributed to improving production performances through the continuous monitoring of cows' health, reproduction, and welfare, optimized milking routines, and suitable feed and nutrition strategies (Balaine et al., 2020).

A substantial part of the economic challenge of a dairy farm is linked to losses of milk production that manifest in altered MY dynamics, seen as perturbations of the lactation curve (Hertl et al., 2014). Perturbations are mainly the result of health problems or diseases, impaired feed quality, and extreme weather conditions. Understanding how cows cope with those challenges could help to gain insights into their resilience and robustness (Adriaens et al., 2020; Ben Abdelkrim et al., 2021). As a result, the correct identification of robust and resilient dairy cows would allow the optimization of breeding, treatment, and culling decisions (Adriaens et al., 2020; Ranzato et al., 2022).

Perturbations cause drops in MY that pull the fitted lactation curve downwards. Many phenotyping tools have been proposed to estimate dairy cows' expected milk production in the absence of perturbations. Thanks to high-frequency milk meter data and advanced computation, the most recent tools enable to study lactation dynamics in great detail. Adriaens et al. (2021) proposed to iteratively use Wood's model (Wood, 1967) to determine the expected production and characterize perturbations of the lactation curve, Ben Abdelkrim et al. (2021) implemented a lactation model with explicit representation of perturbations, and Poppe et al. (2020) suggested to use a 4th order quantile regression model to make the resulting lactation curve fitting less sensitive to drops in MY.

This study aimed at analyzing the methods proposed by Adriaens et al. (2021), Ben Abdelkrim et al. (2021), and Poppe et al. (2020) considering their mathematical complexity, and gathering insights into their strengths and weaknesses. The end-user (e.g., farmer, veterinarian, researcher) can decide to utilize the method that best fits his specific application, to monitor the herd or to optimize breeding schemes and culling decisions. Moreover, the model by Ben Abdelkrim et al. (2021) was used as an application example to test the differences in cows' average lactation performances across breeds, parities, and calving seasons.

MATERIALS AND METHODS

Data collection and pre-processing

Back-up files of the management software of 14 AMS were collected at 10 farms with Lely AMS (Lely Industries N.V., Maassluis, the Netherlands) and 4 farms with DeLaval AMS (DeLaval, Tumba, Sweden). Half of the farms housed Italian Simmental dairy cows in the north of Italy, whereas the remaining were Holstein dairy farms from Belgium and the south of the Netherlands. The data tables containing daily historical MY data, together with cow and lactation identifiers with calving dates, were extracted from the AMS software back-up files of each farm using an automated pipeline in Python (Gote et al., 2022). The MY data in DeLaval software back-ups differed from the Lely ones: DeLaval takes the sum of the milkings of a day (24-hour period) as the daily value, Lely corrects for the varying number of milkings between days by taking a 3-day moving average of the MY time-series. Daily MY data from the DeLaval farms were therefore corrected by replacing each value with the average of the current and previous days (Adriaens et al., 2021).

The maximum time-period covered by each farm data ranged between 2015 and 2022. Lactations were selected based on the following criteria: (1) MY data were available from before 5 DIM to at least 200 DIM (Adriaens et al., 2020; Ben Abdelkrim et al., 2021), and (2) no more than 10% missing daily MY records were present. For each of the selected lactations, data up to 305 DIM were included in the analysis. After 305 DIM, indeed, the lactation dynamics can be influenced by the gestation stage and feed changes towards dry-off (Dematawewa et al., 2007), which was not of interest to this study.

The data processing and further modelization were performed with RStudio software (R version 4.2.3; Posit Software PBC, Boston, MA).

Iterative Wood model

The iterative Wood model (IW) presented hereafter was based on the work by Adriaens et al. (2021).

The unperturbed lactation curve (ULC), meaning the estimated expected milk production in the absence of perturbations, is calculated for each lactation using the nonlinear Wood model:

$$MY = a \cdot DIM^b \cdot e^{-c \cdot DIM} + \varepsilon, \qquad \text{(Eq.1)}$$

where ε is the error term, and a, b, c are positive parameters that define the shape of the lactation curve. Parameter a mainly determines the scaling of the curve, b and c determine, respectively, the moment of the peak production and the slope. An iterative fitting procedure is applied to gradually remove MY data during perturbations, following the steps below:

(1) iteration $i = 1$: fit the Wood's model (ULC_1) on all MY data of the lactation ('nlsLM' function of the 'minpack.lm' package; Elzhov et al., 2023);

(2) remove MY data below $ULC_1 - 1.6 \times$ standard deviation (SD_1) of the residuals of the model estimated at $i = 1$;

(3) iteration $i > 1$: fit Wood's model (ULC_i) on the filtered MY data resulting from the previous iteration $i - 1$;

(4) remove MY data below $ULC_i - 1.6 \times SD_i$ of the residuals of the model estimated at i;

(5) repeat (3) to (4) until the improvement in the root mean squared error of the model estimated at i compared with the previous iteration $i - 1$ is smaller than 0.1 kg, or after 20 iterations.

To increase the fitting stability in the first part of the lactation, MY values below $ULC_i - 1.6 \times SD_i$ are not removed when DIM < 30. The parameters a, b, c estimated at the previous iteration are used as the starting values for the next iteration (first iteration: a = 5; b = 0.2; c = 0.004).

Starting from the ULC, it is possible to identify the perturbations in the actual milk production. After computing the residuals by subtracting ULC from the original MY data, a perturbation is defined as a period of at least 5 successive days of negative residuals with at least one day in which MY is below 80% of ULC. The start and end DIM of each perturbation correspond to the first and last residual below zero, respectively. The days before the largest negative residual are identified as the development phase of the perturbation, while the days afterwards represent the recovery phase.

Perturbed lactation model

The perturbed lactation model (PLM) presented below was based on the work by Ben Abdelkrim et al. (2021).

The formula of PLM for a lactation with P individual perturbations is given by:

$$MY = a \cdot DIM^b \cdot e^{-c \cdot DIM} \cdot \prod_{p=1}^{P} \left(1 - \frac{k_{0p} \cdot k_{1p}}{k_{1p} - k_{2p}} \cdot \left(e^{-k_{2p} \cdot \Delta_p(DIM)} - e^{-k_{1p} \cdot \Delta_p(DIM)} \right) \right) + \varepsilon,$$

where ε is the error term. The model is composed of an unperturbed lactation model corresponding to Wood's equation (Eq.1) and a perturbation model that represents the global proportion of milk affected by the P perturbations. The parameter $k_{0p} \in (0,1)$ is the intensity of the pth perturbation, k_{1p} the collapse speed of the pth perturbation, and k_{2p} the recovery speed of the same perturbation; $\Delta_p(DIM)$ is the elapsed time since the beginning of the pth perturbation and is given by:

$$\Delta_p(DIM) = \begin{cases} 0 & if\ DIM < DIM_p \\ DIM - DIM_p & if\ DIM \geq DIM_p \end{cases},$$

where DIM_p is the time of start of the pth perturbation. The total number of parameters of the model is then equal to $4 + 4 \times P$ (i.e., a, b, c, P, $\{k_{0p}$, k_{1p}, k_{2p}, $DIM_p\} \times P$). The fitting strategy consists of 2 steps: (1) repeated fittings to estimate the most frequent number of detected perturbations, and (2) fix the number of perturbations to the value determined in the first step and estimate the remaining parameters of the model. Arbitrary values can be set in step (1) for the number of fittings to be performed and the maximum number of perturbations to be detected, as well as for the maximum number of iterations to be used in step (2) to reach the convergence of the estimation procedure. We performed 50 fittings in step (1) and set the maximum number of perturbations and the maximum number of iterations at 10 and 1000, respectively. The PLM provides an explicit representation of the ULC and the perturbed lactation curve where perturbations can occur one inside another.

Quantile regression

The approach of using quantile regression (QR) to estimate the ULC was based on the work by Poppe et al. (2020).

The QR estimates the conditional median or other quantiles of the outcome, instead of the conditional mean as for classical linear regression (Koenker, 2005). By using a quantile higher than 0.5, low values of the outcome have less impact on the estimated curve than high values. To estimate the ULC, Poppe et al. (2020) chose a 4th order polynomial quantile regression using a 0.7 quantile ('quantreg' package; Koenker, 2023):

$$MY = \beta_0 + \beta_1 \cdot DIM + \beta_2 \cdot DIM^2 + \beta_3 \cdot DIM^3 + \beta_4 \cdot DIM^4 + \varepsilon,$$

where ε is the error term and β contains the regression coefficients.

Comparison of methods

To compare IW, PLM, and QR, the cumulative 305-DIM unperturbed milk production (ULC_{305}) and the percentage of total milk loss (ML) were estimated for each lactation of the dataset. The ULC_{305} was calculated as the sum of the predicted unperturbed daily production until DIM 305 of the lactation. The ML was computed using the following formula:

$$ML = 100 \cdot \left(1 - \frac{\sum_{DIM} ULC}{\sum_{DIM} MY}\right),$$

where $\sum_{DIM} ULC$ is the total estimated unperturbed production for that lactation and $\sum_{\mathrm{DIM}} \mathrm{MY}$ the respective observed total MY ($200 \leq \max(\mathrm{DIM}) \leq 305$). Furthermore, every lactation is divided into 'early' (DIM ≤ 60), 'mid' (60 < DIM ≤ 150), and 'late' (150 < DIM ≤ 305), and the sum of the estimated unperturbed daily production and percentage of ML were computed at each lactation stage (ULC_{early} and ML_{early}, ULC_{mid} and ML_{mid}, and ULC_{late} and ML_{late}).

The PLM estimates an ULC as well as a perturbed lactation curve representing overlapping deviations from the ULC. The IW, on the other hand, comes with a criterion to identify individual perturbations. When the relative milk loss of an overlapping PLM perturbation was lower than 5%, that perturbation was considered as the same perturbation as the one it overlaps with and was not taken into account to determine the final number of perturbations. The IW and PLM were thus compared in terms of number of detected perturbations per lactation (N_P), and number of detected perturbations per lactation stage ($N_{P,early}$, $N_{P,mid}$, and $N_{P,late}$).

The outcomes of interest from each lactation of each farm (i.e., ULC_{305} and ML, also divided by lactation stage, estimated with all the methods; N_P, also divided by lactation stage, estimated with IW and PLM) were merged into one global table, keeping the information on cow's breed, parity, and calving dates. Pearson correlation coefficients were computed for ULC_{305} and ML to detect the degree of linear relationship between each pair of methods. A Kruskal-Wallis statistical test (Dodge, 2008) was applied to compare the outcomes estimated with the 3 methods, and a post-hoc Wilcoxon test with Holm correction (Holm, 1979) was performed if the median results differed among methods.

Application example

After selecting PLM as estimation method, a multivariate mixed-effects model (Fitzmaurice et al., 2004) was used to evaluate the effect of breed, parity, and calving season on ULC_{305}, ML, and N_P. The models were specified as follows ('lme4' package; Bates and Maechler, 2017):

$$y_{ghjlmq} = \mu + parity_h + calving\ season_j + breed_l + parity_h \cdot calving\ season_j + parity_h \cdot breed_l + calving\ season_j \cdot breed_l + cow_m + farm_q + \varepsilon_{ghjlmq}, \quad \text{(Eq.2)}$$

where y_{ghjlmq} was the outcome variable (i.e., ULC_{305}, ML, or N_P) related to the gth lactation of the mth cow in the qth farm of the lth breed, that calved in season j within parity h. The ULC_{305} and ML were continuous variables that were log-transformed if they did not have a normal distribution; N_P was a counting variable with a Poisson distribution. The fixed effects of the models corresponded to the variables 'parity' (3 categories: '1', '2', '≥3'), 'calving season' (4 categories: 'summer' when the cow calved between June and August, 'autumn' when the cow calved between September and November, 'winter' when the cow calved between December and February, 'spring' otherwise), 'breed' (2 categories: 'Holstein', 'Italian Simmental'), and the interaction terms between couples of those variables. The random effects corresponded to the 'cow' and the respective 'farm' variables (nested random effects). First, a likelihood ratio test (Silvey, 1975) was used to test the significance of including the random effects, then the effect of each fixed term on each response variable was investigated by examining the ANOVA tables. When a term was not significant, a reduced version of the model with respect to Eq.2 was built and evaluated.

RESULTS

The final data consisted of 2,250 cows (65% Holstein; 35% Italian Simmental) and 4,441 individual lactations with parity ranging from 1 to 12 (33% parity 1; 26% parity 2; 41% parity ≥3). Descriptive statistics over farms are given in Table 1.

Table 1. *Descriptive statistics of the dataset (farms n = 14).*

	Mean ± SD over farms	**Range over farms** [min; max]
Time period covered (yr)	6.2 ± 1.0	[4.3; 7.2]
Lactations (n)	317.2 ± 177.4	[61; 674]
Parity 1	104.9 ± 68.4	[0; 254]
Parity 2	82.4 ± 47.6	[25; 193]
Parity ≥ 3	129.9 ± 66.1	[31; 247]
Average daily MY[1] in first 305 d (kg)	30.6 ± 4.6	[24.1; 38.4]
Total sum of MY in first 305 d (kg)	9,119.3 ± 2,253.9	[1,447.1; 16,704.8]

[1] Milk yield.

Comparison of methods

Unperturbed lactation curve

First, the 3 mathematical methods were used to estimate the unperturbed MY for each lactation of the dataset. An example of ULC obtained with each method for one randomly selected lactation is shown in Figure 1. Second, starting from the ULC, ULC_{305} and ML were computed for every lactation. Table 2 reports Pearson correlation coefficients calculated for ULC_{305} and ML between pairs of estimation methods, while Figures 2A and 2B compare the distributions of ULC_{305} and ML (respectively) across estimation methods. The median ULC_{305} were 9,632 kg, 9,693 kg, and 9,814 kg when estimated with IW, PLM, and QR, respectively. The median ML was 3.5% using IW, 4.3% using PLM, and 4.8% using QR. Based on Kruskal-Wallis statistical tests, the medians ULC_{305} and ML resulted different across methods ($p < 0.001$, respectively). Afterwards, by performing post-hoc Wilcoxon tests with Holm correction between pairs of methods, only the ULC_{305} median estimated with IW and with PLM did not result statistically different ($p = 0.108$), whereas all the pairs of ML comparisons were statistically different ($p < 0.001$ in all scenarios).

Figure 1. *Unperturbed lactation curves estimated using iterative Wood model (pink), perturbed lactation model (green), and quantile regression (blue) for one lactation randomly selected.*

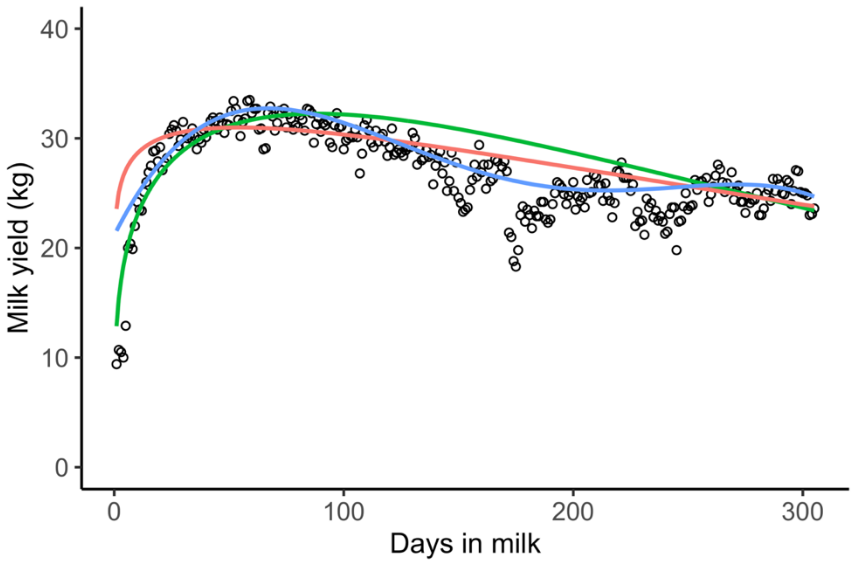


Figure 2. *(A) Distribution (boxplot and violin plot) of the estimated total unperturbed milk yield ('tot. unp. MY') across different estimation methods (iterative Wood model – IW, perturbed lactation model – PLM, quantile regression – QR); distributions with different letters differ ($p < 0.05$). (**B**) Distribution (boxplot and violin plot) of the estimated total milk loss across different estimation methods (top limit of the y-axis is reduced); distributions with different letters differ ($p < 0.05$).*

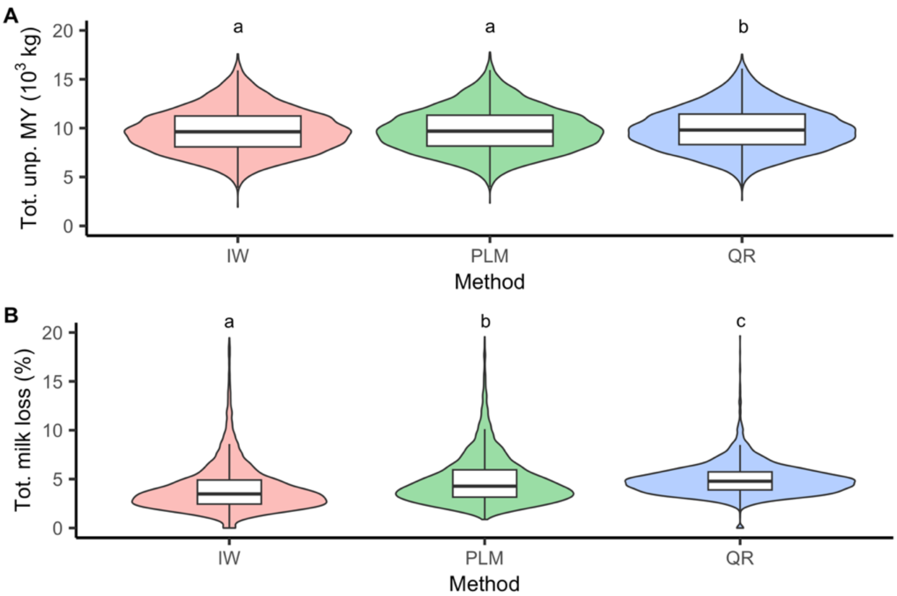
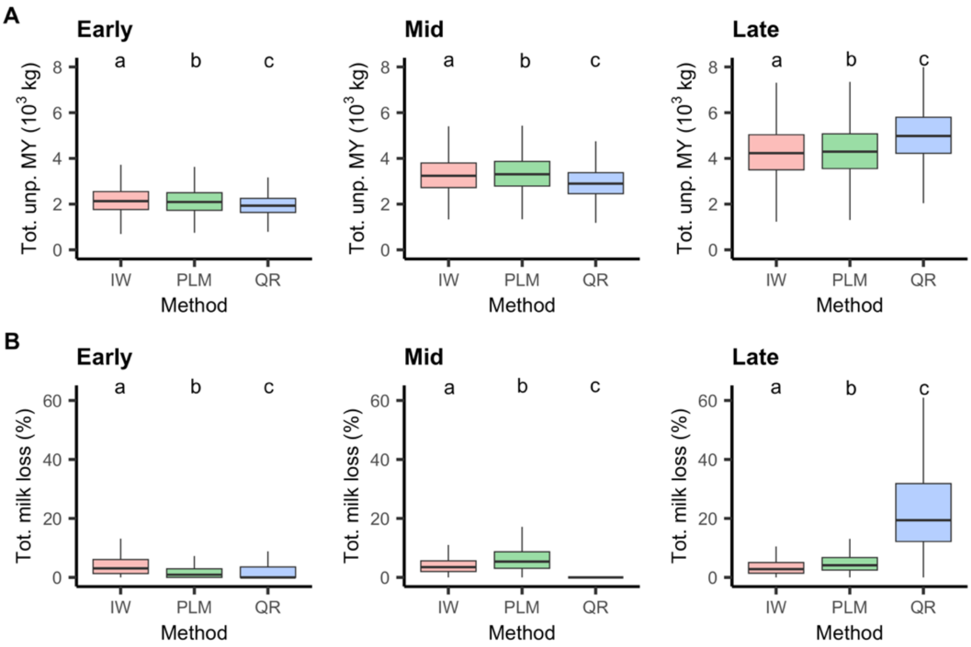


Table 2. *Pearson correlation coefficients for the total unperturbed milk yield (ULC_{305}) and the total milk loss (ML) between couples of estimation methods (iterative Wood model – IW, perturbed lactation model – PLM, quantile regression – QR).*

ULC_{305}	IW	PLM	QR
IW	1	0.992	0.988
PLM	0.992	1	0.984
QR	0.988	0.984	1
ML	IW	PLM	QR
IW	1	0.543	0.769
PLM	0.543	1	0.400
QR	0.769	0.400	1

To analyze more in detail IW, PLM, and QR performances, the total unperturbed MY and ML were estimated for each stage of each lactation, obtaining ULC_{early} and ML_{early}, ULC_{mid} and ML_{mid}, and ULC_{late} and ML_{late}. Figures 3A and 3B show the boxplots by stage of lactation for the total unperturbed MY and ML, respectively. The median ULC_{early} was 2,132 kg when estimated with IW, 2,092 kg when estimated with PLM, and 1,931 kg when estimated with QR; the median ULC_{mid} were 3,239 kg, 3,305 kg, and 2,896 kg, respectively; the median ULC_{late} were 4,229 kg, 4,294 kg, and 4,987 kg, respectively. The median ML_{early} was 3.1% using IW, 0.9% using PLM, and 0.0% using QR; the median ML_{mid} were 3.5%, 5.4%, and 0.0%, respectively; the median ML_{late} were 2.8%, 4.1%, and 19.4%, respectively. The medians of the total estimated unperturbed MY and ML at each lactation stage were significantly different across estimation methods (Kruskal-Wallis tests; p < 0.001 in all scenarios), and also between all pairs of methods (post-hoc Wilcoxon tests; $p < 0.05$ in all scenarios).

Figure 3. *(**A**) Boxplot of the estimated total unperturbed milk yield ('tot. unp. MY') per lactation stage (early: DIM ≤ 60, mid: 60 < DIM ≤ 150, late: 150 < DIM ≤ 305) using different estimation methods (iterative Wood model – IW, perturbed lactation model – PLM, quantile regression – QR); distributions with different letters within lactation stage differ* (p *< 0.05). (**B**) Boxplot of the estimated total milk loss per lactation stage using different estimation methods; distributions with different letters within lactation stage differ* (p *< 0.05).*



Perturbations of the lactation curve

The IW and PLM were used to detect the perturbations of every lactation of the dataset. Using the same MY data plotted in Figure 1, Figure 4A shows the ULC and the perturbations estimated with IW, whereas Figure 4B shows the ULC and the perturbed lactation curve estimated with PLM. The distribution of N_P across the 2 methods is represented in Figure 5. The mode of N_P was 3 when applying IW and 4 when applying PLM.

To better compare IW and PLM performances, Figure 6 shows the distribution of the number of perturbations identified by the 2 methods at each lactation stage. The modes of $N_{P,early}$, $N_{P,mid}$, and $N_{P,late}$ were 1 at each lactation stage using both IW and PLM.

Figure 4. *(**A**) Unperturbed lactation curve (pink line) estimated with iterative Wood model and perturbations identified (red points). (**B**) Unperturbed (green line) and perturbed (red line) lactation curves estimated with perturbed lactation model.*

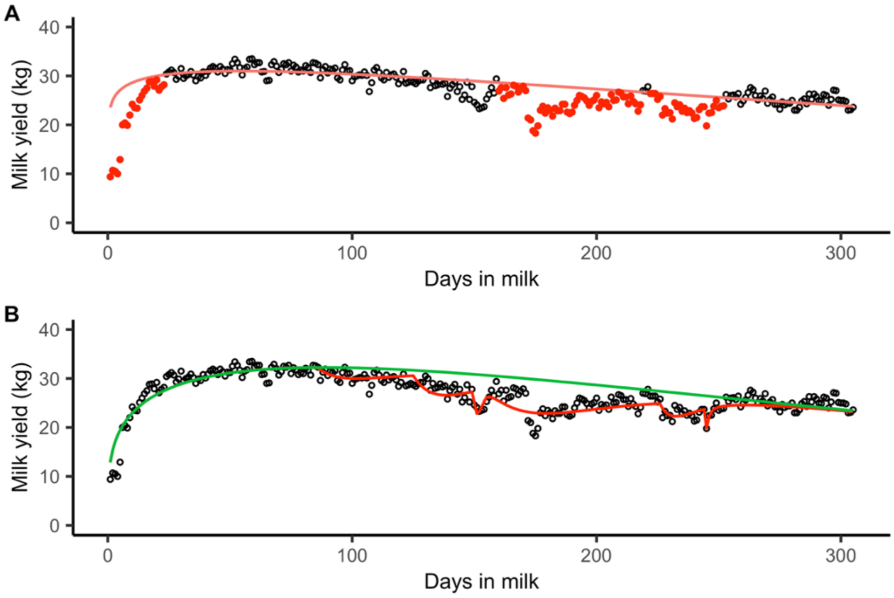


Figure 5. *Histogram of the number of perturbations identified across lactations using iterative Wood model (pink) and perturbed lactation model (green).*

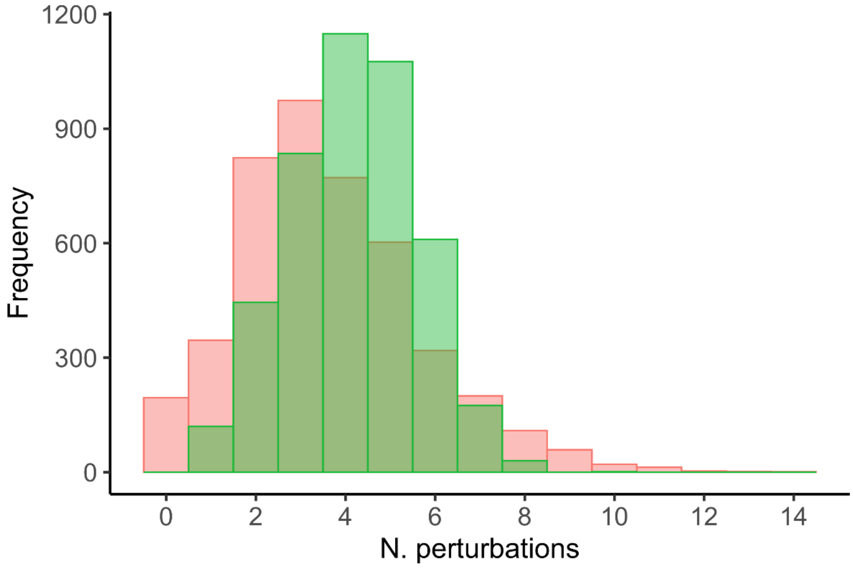
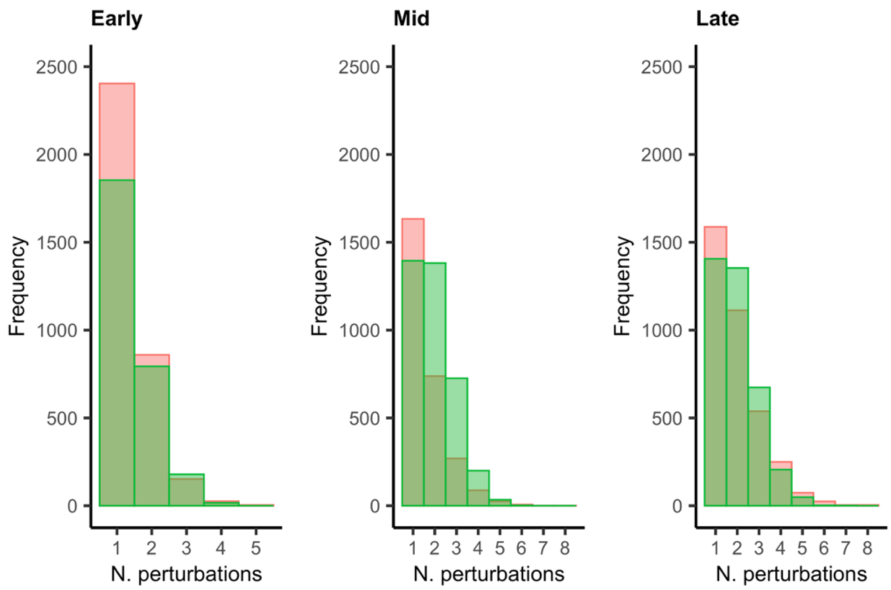


Figure 6. *Histogram of the number of perturbations identified across lactation stages (early: DIM ≤ 60, mid: 60 < DIM ≤ 150, late: 150 < DIM ≤ 305) with iterative Wood model (pink) and perturbed lactation model (green).*



Application example: effect of parity, calving season, and breed on lactation performances

The PLM was chosen to analyze the effect of parity, calving season, and breed on cows' lactation performances (i.e., ULC_{305}, ML, and N_P). Visual inspection of the residual plots of the linear regression on ULC_{305}, the linear regression on the logarithmic transformation of ML, and the Poisson regression on N_P did not reveal any important deviations from homoscedasticity and normality. As measured by the likelihood ratio tests, the inclusion of farm and cow as random effects improved the fitting of the models on ULC_{305} and ML, whereas it was sufficient to include only the farm random variable for the model on N_P. The ANOVA revealed that the parity and the season of calving had a highly significant effect on cows' lactation performances, whereas the breed had a significant effect only on ULC_{305} and N_P. The interaction parity × breed had a highly significant effect on ULC_{305} and ML. Milk losses were also influenced by the combined effect of calving season and breed ($p < 0.05$), while the interaction between parity and calving season highly affected ULC_{305} ($p < 0.001$).

The non-significant terms resulted from the ANOVA tables were removed from the models on ULC_{305}, ML, and N_P, and the final estimated fixed effects are reported in Table 3. Multiparous cows of the same breed that calved in the same season potentially produce, on average, 1,500 kg (combined effect of main effect and interaction terms) of milk more during a whole lactation with respect to primiparous ones. High parity cows (parity ≥3) calving in winter produce 700 kg (combined effect of main effect and interaction term) of milk more than the ones calving in summer in the absence of perturbations of the lactation curve; high parity Holstein cows potentially produce around 2,000 kg (combined effect of main effect and interaction term) of milk more than high-parity Italian Simmental ones. Multiparous cows have more relative milk losses with respect to the unperturbed production than primiparous ones ($e^{0.13}\% = 1.1\% \times$ primiparous ML); the cows calving in winter have fewer relative milk losses than the cows calving in summer ($e^{-0.09}\% = 0.9\% \times$ summer ML). Last, summer-calving cows have fewer perturbations in the lactation curve than spring-calving ones (0.25 perturbations less), whereas the animals at parity 2 have fewer perturbations than the primiparous ones (0.25 perturbations less).

DISCUSSION

Comparison of methods

This work compared 3 existing mathematical methods (i.e., IW, PLM, and QR) to estimate dairy cows' lactation performances. The IW, PLM, and QR were found to be valuable methods to obtain cows' expected production and ML, each one with advantages and disadvantages depending on the specific application (see Table 4 for a summary).

From a mathematical and computational point of view, PLM is quite complex. Depending on the number of iterations imposed and the computational power of the machine, the estimation procedure of $4 + 4 \times P$ parameters can take several hours to fit the individual curves of an entire set of lactation data. For instance, setting 1000 iterations and using a simple machine with a 1.8 GHz Intel Core i5 dual-core processor and a RAM of 8 GB, it takes around 5 hours for a set of 300 lactations, compared to about 20 seconds for IW and 10 seconds for QR. In addition, PLM requires a maximum number of perturbations to be set, with the computational burden of the algorithm raising as this number increases. We set a maximum of 10 perturbations to be identified per lactation, knowing that higher numbers may occur especially in case of severe metabolic disorders or chronic mastitis (Hostens et al., 2012). Nonetheless, the resulting unperturbed and perturbed lactation curves obtained with PLM always fit the data well, as shown in the example of Figure 4B. The IW and QR are easier in terms of mathematical complexity. Yet, also these methods depend on meta-parameters that can significantly change the performances based on their values. The IW relies on the threshold for removing low MY values at each iteration of the algorithm, whereas QR relies on the reference quantile of MY data points. Both Adriaens et al. (2021) and Poppe et al. (2020) selected those meta-parameters upon thorough visual observation of numerous fitted lactation curves.

Table 3. *Estimates of the fixed effects of the mixed models on the total unperturbed milk yield (ULC_{305}), the logarithm of the total milk loss (ML), and the number of perturbations (N_P), with relative standard errors (SE) and observed levels of significance (p).*

	ULC_{305} (kg)			log(**ML**) (%)			**N_P**		
Fixed effect	Estimate	SE	*p*	Estimate	SE	*p*	Estimate	SE	*p*
Intercept	9088	412	***	1.4	0.07	***	4.2	0.09	***
Parity 2	1329	108	***	0.06	0.02	**	-0.25	0.08	**
Parity ≥3	1777	104	***	0.13	0.02	***	-0.09	0.07	n.s.
Calving season autumn	227	103	*	0.002	0.02	n.s.	-0.20	0.08	*
Calving season winter	176	105	n.s.	-0.09	0.03	***	0.06	0.08	n.s.
Calving season spring	127	105	n.s.	0.002	0.03	n.s.	0.25	0.09	**
Breed Italian Simmental	-1516	578	*	0.15	0.10	n.s.	0.21	0.09	*
Parity 2 × calving autumn	53	140	n.s.	-	-	-	-	-	-
Parity ≥3 × calving autumn	431	134	**	-	-	-	-	-	-
Parity 2 × calving winter	375	142	**	-	-	-	-	-	-
Parity ≥3 × calving winter	580	133	***	-	-	-	-	-	-
Parity 2 × calving spring	318	147	*	-	-	-	-	-	-
Parity ≥3 × calving spring	377	141	**	-	-	-	-	-	-
Parity 2 × breed I.S.[1]	-137	97.7	n.s.	-0.03	0.04	n.s.	-	-	-
Parity ≥3 × breed I.S.	-543	98.8	***	-0.10	0.03	**	-	-	-
Calving autumn × breed I.S.	-	-	-	-0.08	0.04	*	-	-	-
Calving winter × breed I.S.	-	-	-	0.02	0.04	n.s.	-	-	-
Calving spring × breed I.S.	-	-	-	0.02	0.04	n.s.	-	-	-

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, n.s. $p \geq 0.05$. Reference categories of the variables (parity 1; calving season summer; breed Holstein) are not reported.

[1] Italian Simmental.

The correlation coefficients between the different methods revealed that they were perfectly linearly related in the case of ULC_{305} estimation, whereas they were only moderately correlated in the case of ML estimation. The distribution of ULC_{305} was very similar among methods, but the same did not apply for ML especially when using QR compared to the other 2 methods. The QR, indeed, produces less variable milk losses estimates compared to IW and PLM because it tends to 'follow' the observed MY data closer. This is due both to the intrinsically higher sensibility of QR to very large perturbations from the expected lactation curve, and to the 4^{th} order polynomial regression that can take more variable shapes (Lever et al., 2016). The 3 methods performed differently according to the lactation stage. The IW estimated the ML to be higher during the early stage of the lactation with respect to the other 2 methods. As shown in Figure 1, indeed, IW is less performant in capturing the ascending phase of the ULC, although we set to never exclude MY values below $ULC - 1.6 \times SD$ when DIM < 30 in the estimation process. The PLM, on the other hand, produced higher ML estimates during the mid-stage of the lactation, whereas QR 'underestimated' ML in the mid-stage and highly 'overestimated' it in the last stage of the lactation due to too much fitting flexibility. The IW and PLM were compared also in terms of number of estimated MY perturbations. Both the distributions of N_P estimated using the 2 methods were right-skewed, but PLM detected on average more perturbations than IW. The IW detected more perturbations during the early stage of the lactation compared to PLM, which is mostly due to the poor adaptation of ULC estimated with IW on the first days after calving. Contrarily, PLM tended to find more perturbations during the mid and late stages of the MY curve, mainly because multiple perturbations can overlap, also affecting all the previous perturbations in the estimation process.

Both IW and PLM are Wood model-based, producing ULC with better shape stability compared to QR. The Wood model is one of the easiest and most used mathematical models for estimating a lactation curve (Ben Abdelkrim et al., 2021) because it describes its shape very well during the first 305 DIM; more complex models are preferred only when it comes to estimate the curve after 305 DIM (Dematawewa et al., 2007). On the other hand, polynomial QR is very flexible, fast, and easy to implement. The QR, as well as IW, can also operate with less frequent than daily MY data, while this is not yet verified for PLM. Finally, IW and PLM enable to detect perturbations of the lactation curve, within the estimation process in the case of PLM and using a characterization criterion in the case of IW. The PLM, in particular, allows to capture multiple (overlapping) perturbations with contrasted features (e.g., due to gestation, drying off, disease) and to produce metrics to compare the effect of perturbations on MY (i.e., parameters of scale and shape of each perturbation).

Possible improvements

The QR could also be used for the detection of perturbations using a criterion like the one of IW (i.e., at least 5 days of negative residuals with at least one day of MY below 80% of ULC), bearing in mind that QR does not produce ULC with robust shape, running the risk of not identifying relevant perturbations.

Specific adaptations of the 3 methods could further improve their performances, besides further tuning of their meta-parameters. Optimizing the weights of specific phases of the lactation during the model fitting could enhance the performances of both IW and QR. For example, the conditional quantile of QR could change based on DIM, in order to avoid the 'overestimation' or 'underestimation' of ULC during specific stages. As an alternative to the 4th order polynomial, a linearized Wood model could be used in combination with QR, to keep at the same time the ease of a regression and the stability of shape produced by the Wood model. Finally, PLM and the iterative approach of Adriaens et al. (2021) could be tested by substituting the Wood model with other typical lactation models, such as the one developed by Wilmink (Wilmink, 1987).

Practical implications

Dairy practitioners and researchers can choose the mathematical method that best fits their specific needs. Farmers, veterinarians, or technicians may need a tool for phenotyping purposes; they generally require a fast and 'rough' approach on large datasets with minimal computational effort, putting QR forward as the most suitable method. Farm technology suppliers may be interested in implementing one of these methods for cow monitoring; then IW or PLM would be preferable, considering that the former is less efficient during the pre-peak lactation phase than in the post-peak one, whereas the latter is very robust but requires significant computational power. If dairy practitioners are interested in perturbation detection, then IW is a fast and rather precise method for identifying individual perturbations of the herd lactation curves. Researchers that need a precise fitting of the expected production or want to characterize (i.e., parametrize) perturbations to link them with external factors or farm management practices, might be more interested in PLM. When high-frequency milk meter data are not available (e.g., a farm is not equipped with an AMS), then PLM should probably be avoided because it has not yet been tested in combination with less frequent than daily MY data.

Thanks to the estimation of ULC, it is possible to compare the cows based on their potential of MY and rank them according to the production level they would have achieved in a non-perturbed environment (Ben Abdelkrim et al., 2021). With this information, farmers could optimize culling decisions and breeding schemes, identify the animals that have both a high production potential and ability to cope with their environment, or understand which are the most resilient animals, namely the ones that are able to recover fast after a given challenge. Studying the characteristics of perturbations throughout many lactations and connecting them to genomic information could open the opportunity to evaluate their heritability and genetic impact (Ben Abdelkrim et al., 2021). Linking perturbations with other information on cows or farm environment, could help to detect sensitive periods where perturbations are more likely to occur or could assist the farmers in identifying the animals with greater adaptive capacities during the same stress phases (Ranzato et al., 2023). With a better understanding of environmental effects on animal production, on-farm preventive measures could be optimized (Ben Abdelkrim et al., 2021).

Application example: effect of parity, calving season, and breed on lactation performances

The results from the mixed-effects models on cows' lactation performances estimated with PLM, chosen for its robustness, were consistent with the literature. The expected production was higher as parities progressed. The heifers, in fact, do not yet have reached adult body weight and fully developed udders, producing less milk than multiparous cows (Wathes et al., 2007; Siewert et al., 2019). Non-summer calving led to a higher MY potential compared to summer calving, especially for multiparous cows. It has been widely demonstrated that lactations that start during the hot season usually have lower-than-average production performances as an effect of thermal stress on the animals (Torshizi, 2016; Li et al., 2022). For the same reason, ML was higher for the cows calving in summer compared to the ones calving in winter. Moreover, ML was higher for multiparous cows than for primiparous ones. The same result was found by Carvalho et al. (2019) and Adriaens et al. (2021), and it could be explained by a higher incidence rate of diseases during higher parities (Lee and Kim, 2006). The number of perturbations was larger in first than second lactations, which could be explained by a higher susceptibility to stressors due to management changes (e.g., regrouping or milking) of first-parity cows (Proudfoot and Huzzey, 2022). The number of perturbations increased also for spring calvings compared to summer calvings and this could be linked with heat stress episodes the animals calving in spring have to face during the first critical months of the lactation (McNamara, 2002). Last, as expected (Knob et al., 2023), Holstein cows had higher production potential than Italian Simmental ones, especially at high parities (parity $\geq$3).

CONCLUSIONS

This study compared 3 existing mathematical methods to estimate dairy cows' lactation performances using high-frequency MY data from on-farm AMS. The IW, PLM, and QR were all valuable methods to obtain cows' expected production and milk losses, each one with advantages and disadvantages that need to be considered depending on the specific application. The outcome of this study can help dairy practitioners in choosing the best decision-support method, or researchers in attempting to study MY dynamics.

ACKNOWLEDGMENTS

The Authors would like to thank 'Livestock Technology' group of KU Leuven (Geel, Belgium) and ANAPRI – Associazione Nazionale Allevatori Pezzata Rossa Italiana (Udine, Italy) for providing AMS data.

CHAPTER 5


Artificially generated image (DALL·E 2, OpenAI)

Modelling of milk fatty acids profiles in Holstein dairy cows during early and mid-lactation

G. Ranzato[1], M.C. Galli[1], G. Cozzi[1]

[1] *Department of Animal Medicine, Production and Health (MAPS), University of Padova, Viale dell'Università 16, 35020 Legnaro (PD), Italy*

Oral communication at the *24th Congress ASPA (Animal Science and Production Association)*

Summary

Detection of cows' health problems has become important especially during early lactation, when animals experience a stressful period that often impairs their welfare status and productive response. Milk composition analysis has been recognized as a tool to detect cows that are exposed to critical feeding or management situations. In particular, the proportions of different milk fatty acids (FA) groups (i.e., de novo, preformed and mixed) are considered promising biomarkers of cows' welfare and proper feeding. Understanding the trend of these groups of milk FA in healthy cows could be useful to identify animals in critical health conditions, even before clinical signs are visible. To define reference trends and confidence intervals for the three groups of FA in Holstein dairy cows according to days in milk (DIM), 300 individual milk samples were collected from 10 different herds belonging to Grana Padano (GP; 6 herds) and Parmigiano Reggiano (PR; 4 herds) dairy chains. In each farm, 10 multiparous cows (6 with DIM $\leq$ 45, 4 with 45 < DIM $\leq$ 175) in good health status were monthly sampled from August to October 2020. Milk samples were analyzed for milk composition and FA profiles with a mid-infrared instrument. Analytical data were statistically pre-processed and modelled including the effects of DIM and type of milk. To model the trend of FA, linear mixed models with nested random effects (i.e., cows and farms) were used to account for repeated measures. The percentage of de novo FA had an increasing overall trend according to DIM, while the trend was opposite for preformed FA. No significant differences in de novo, mixed, and preformed FA trends were detected between GP and PR. However, the percentage of de novo in PR milk tended to be always lower (p = 0.11) than GP, as a possible sign of a greater risk of subacute ruminal acidosis promoted by PR diets through a potential increase of cows' feed selection activity. Reference intervals for the three groups of FA estimated according to DIM and type of milk could help farmers assessing cows' state of well-being during early and mid-lactation.

INTRODUCTION

Among the several applications of PLF, the monitoring of animal health and welfare is one of the most crucial, especially under intensive production systems (Berckmans, 2014). Within the dairy sector, health and welfare monitoring is particularly important at the onset of the lactation, when the cows are more susceptible to production disorders and diseases such as milk fever, ketosis, retained placenta, displaced abomasum, metritis, mastitis, and lameness (Mulligan et al., 2006; Roche et al., 2013). According to Cardoso et al. (2020), the peak of disease incidence after parturition corresponds with the time of greatest negative energy balance, the peak in blood concentrations of non-esterified fatty acids (FA), and the greatest acceleration of MY. Several technologies like cameras or sound devices have been used for dairy cows monitoring; however, many of them are still too expensive to be adopted by many farms, and in particular by the small-scale ones (Lora et al., 2020).

Milk composition has been recognized as a good source of health and metabolic information, given that there is interaction between circulating blood and milk synthesis outcome (Gengler et al., 2016). Barbano et al. (2014) introduced the application of Fourier-transform mid-infrared spectroscopy (FTIR) for rapid milk FA analysis. Several studies have tested the use of FTIR estimates of milk β-hydroxybutyrate and milk FA as a method for assessing cow's health in early lactation (Denis-Robichaud et al., 2014; Santschi et al., 2016). A large and systematic change in FA composition of milk fat occurs, in fact, with the progress of the lactation (Lynch et al., 1992). De novo FA (from C_4 to C_{14}), synthesized in the cow's mammary gland from acetate and butyrate, start out as low proportion of total FA during early lactation and increase when cows reach a positive energy balance. An opposite trend is observed for preformed FA ($\geq C_{18}$), which enter the mammary cells from the blood stream and originate from the cow's intestinal, liver, or adipose tissues (Woolpert et al., 2017).

The use of individual milk FA profiles obtained with FTIR represents a non-invasive, economic, and ready-to-use method to detect cows at risk of metabolic dysfunctions. To this end, the aim of this pilot study was to calculate reference intervals for de novo, mixed (C_{16}, $C_{16:1}$, and C_{17}), and preformed FA from individual milk samples collected from healthy Holstein dairy cows during early and mid-lactation.

MATERIALS AND METHODS

Ten Holstein dairy farms were enrolled based on their willingness to participate in the study. Six farms belonged to the dairy chain of Grana Padano (GP) cheese and fed cows according to the P.D.O. specifications which allow the use of ensiled feedstuffs in the ration of the lactating cows (Grana Padano, 2019). Four farms were part of Parmigiano Reggiano (PR) cheese chain, which restricts to hay and pasture the roughage to be included in the diets of the lactating cows (Parmigiano Reggiano, 2018). All farms fed the cows with total mixed rations that were delivered in the morning for ad libitum intake. In every farm, cows were milked twice a day, in the morning (around 6:00 AM) and in late afternoon (after 5:30 PM).

Individual milk samples were collected through an automatic milk sampler from 10 multiparous Holstein cows of each farm during the evening milking, for 3 following sampling sessions from August to October 2020. Sampled cows were considered healthy by the farm veterinarian according to the lack of any peripartum disorder. Each sampling session considered 6 cows in early lactation (DIM $\leq$ 45), and 4 cows in mid-lactation (45 < DIM $\leq$ 175). Ear tag ID, DIM, and daily MY of the selected cows were recorded at every sampling. From each milk sample, 120 ml aliquots were placed on ice and stored at 4°C and transported on the next morning to Rumilab (Nutristar S.p.A., Reggio Emilia, Italy) for milk composition analysis. The content of milk protein and fat was predicted with a Lactoscope FT-A (PerkinElmer, Waltham, USA), and the somatic cells count was assessed using a Fossomatic FC (Foss Electric A/S, Hillerød, Denmark). De novo, mixed, and preformed FA were predicted by FTIR using the PLS prediction model described by Woolpert et al. (2017).

A total of 300 milk samples from 214 multiparous cows were collected and analyzed. The editing criteria included restrictions to discriminate between ‘healthy’ and ‘non-healthy’ cows. All samples with more than 400,000 n/ml somatic cells were discarded from the dataset (Agriculture and Horticulture Development Board, 2023). Furthermore, de novo, mixed, and preformed concentrations expressed in mass percentage that exceeded their mean value $\pm$ 3 SD were considered outliers and excluded from the analysis. After assessing the normality assumption of the outcomes, one linear mixed-effects model was used to estimate the trend of the FA according to DIM, while another one was used to evaluate the association between the type of dairy chain (GP vs. PR) and the FA percentages in time. The models were specified as follows:

$$y_{hijl} = \mu + \mathrm{ns}(DIM_i, k = 45) + cow_j + farm_l + \varepsilon_{hijl}, \qquad \text{(Eq.1)}$$

$$y_{hijlm} = \mu + \mathrm{ns}(DIM_i, k = 45) + milk_m + \mathrm{ns}(DIM_i, k = 45) \cdot milk_m + cow_j + farm_l + \varepsilon_{hijlm}, \qquad \text{(Eq.2)}$$

where y was the outcome variable (i.e., de novo, mixed, or preformed FA in mass percentages) related to the h milk sample at DIM i from cow j of farm l (belonging to the dairy chain m). The fixed effect was associated to DIM modelled with a natural cubic spline with a knot at DIM 45 to distinguish the trends between early and mid-lactation (and to the type of dairy chain $milk$ and the interaction between $milk$ and DIM); the random effects were associated to the 'cow' (n = 214) and the respective 'farm' (n = 10) variables.

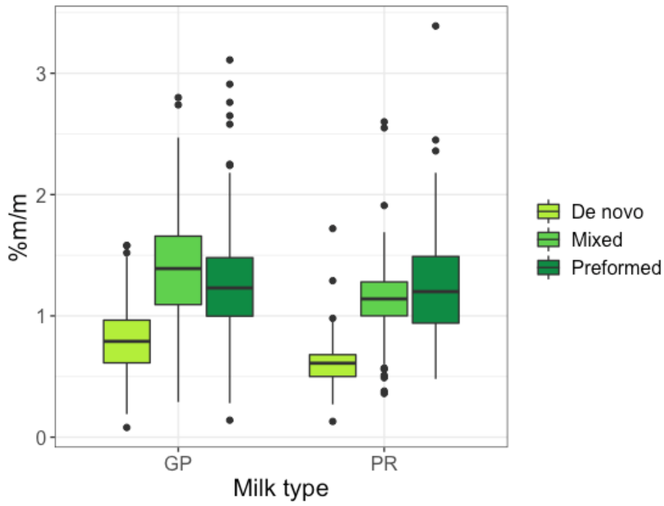
RESULTS AND DISCUSSION

The final dataset considered 288 individual records of 214 healthy cows (58% in early lactation, 42% in mid-lactation) from 10 different dairy farms (59% of GP dairy chain, 41% of PR dairy chain). To have a general overview, Table 1 reports the means (±SD) of MY and milk content across lactation stages and dairy chains. The only statistically significant difference ($p < 0.05$) was the milk protein content in the two dairy chains: PR milk had a higher protein content compared to GP milk. Focusing on milk FA concentrations, PR tended to have lower de novo percentages compared to GP (p = 0.11), but no significant differences ($p < 0.05$) were detected for average concentrations between the two types of milk (Figure 1).

Table 1. *Effect of lactation stage (early: DIM ≤ 45, mid: 45 < DIM ≤ 175) and type of dairy chain (Grana Padano – GP, Parmigiano Reggiano – PR) on milk yield and milk protein and fat content.*

	Lactation stage			**Dairy chain**		
	Early	Mid	*p*	GP	PR	*p*
Milk yield (kg/d)	37.4 (±2.04)	38.6 (±2.09)	0.15	38.4 (±2.74)	37.0 (±3.35)	0.77
Milk protein (%m/m)	3.09 (±0.04)	3.09 (±0.04)	0.88	3.03 (±0.04)	3.18 (±0.04)	0.02
Milk fat (%m/m)	3.73 (±0.24)	3.51 (±0.25)	0.08	3.82 (±0.31)	3.40 (±0.38)	0.41

Figure 1. *Boxplots of fatty acids percentages for Grana Padano (GP) and Parmigiano Reggiano (PR) milks.*



The monitoring of individual cows' health in the opening of lactation is a target for dairy farmers to reduce milk losses and medical treatments. During early lactation, dietary intake is unable to meet the demand of energy and nutrients for milk production. Therefore, cows enter a period of negative balance, which requires the mobilization of body reserves to balance the deficit between intake and milk production (Bauman and Bruce Currie, 1980). The process of mobilization affects cow's welfare and other biological pathways; the reproduction status, for example, can be impaired by a too negative energy balance post-calving (Butler and Smith, 1989). To identify a simple system for assessing the animal's health status, milk has been used as an alternative to determinations based on blood parameters, which additionally eliminates the problem of variation in the values observed for blood parameters due to the timing of the sampling. It has been observed that milk de novo FA and the concentration of non-esterified blood FA, the main marker of cows' energy deficit, have an inversely proportional trend in cows in good health and correctly fed. In particular, in the opening of the lactation, de novo FA are low while NEFA are high (Bach et al., 2019).

In the present study, results from Eq.1 confirmed the increase of de novo and mixed FA, and an opposite trend for preformed FA especially during the early lactation. Specifically, de novo FA significantly increased during mid-lactation (p = 0.02), while mixed FA during both early (p = 0.03) and mid-lactation (p = 0.001); preformed FA significantly decreased during early lactation ($p < 0.001$). References intervals were obtained starting from the estimation of the fixed effect, and they are reported in Table 2, Table 3, and Table 4 for de novo, mixed, and preformed FA, respectively. These would represent the range of 'standard' values for milk FA at different moments of cows' early and mid-lactation. The trends and reference intervals of de novo, mixed, and preformed FA are also shown in Figure 2, Figure 3, and Figure 4, respectively.

Results from Eq.2 did not reveal any differences of FA trends between GP and PR milk, except for a slightly significant effect of the interaction term in the case of de novo FA (p = 0.055). As it is also visible in Figure 5, de novo FA were always lower in PR milk, and were more stable during the early lactation for GP milk compared to PR milk. Specific limitations of feeds, described by the Parmigiano Reggiano cheese consortium (Parmigiano Reggiano, 2018), restrict to dry hay (primarily alfalfa and grass) and pasture the roughage to be included in the diets of the lactating cows. This could promote a more selective feed intake by the cows, leading to a greater risk of subacute ruminal acidosis (Fustini et al., 2016). The differences between the two types of milk were, instead, less pronounced for mixed and preformed FA as shown in Figure 6 and Figure 7, respectively.

Table 2. *95% confidence intervals (CI) for de novo fatty acids at specific DIM.*

DIM	**95% CI** (%m/m)
5	[0.25, 1.12]
31	[0.27, 1.13]
56	[0.29, 1.16]
78	[0.31, 1.18]
104	[0.34, 1.21]

Figure 2. *De novo fatty acids trend with confidence intervals during early and mid-lactation.*

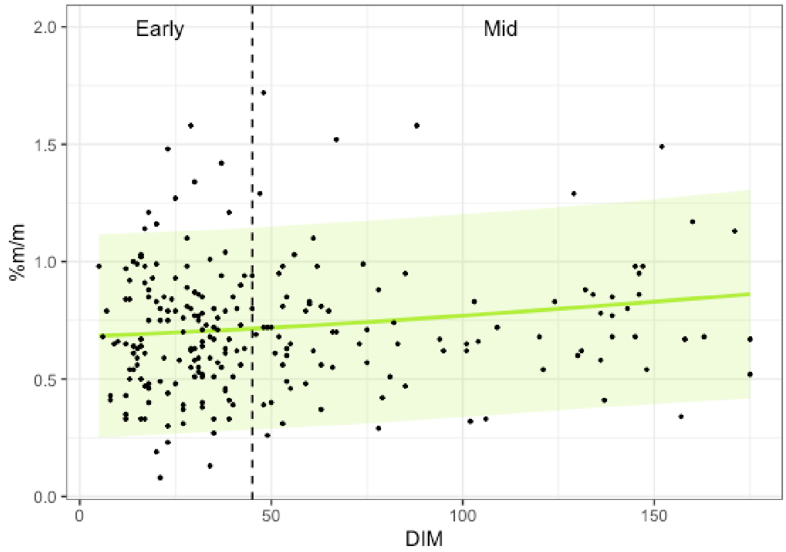


Table 3. *95% confidence intervals (CI) for mixed fatty acids at specific DIM.*

DIM	**95% CI** (%m/m)
5	[0.57, 1.86]
31	[0.61, 1.89]
56	[0.65, 1.93]
78	[0.69, 1.97]
104	[0.75, 2.03]

Figure 3. *Mixed fatty acids trend with confidence intervals during early and mid-lactation.*

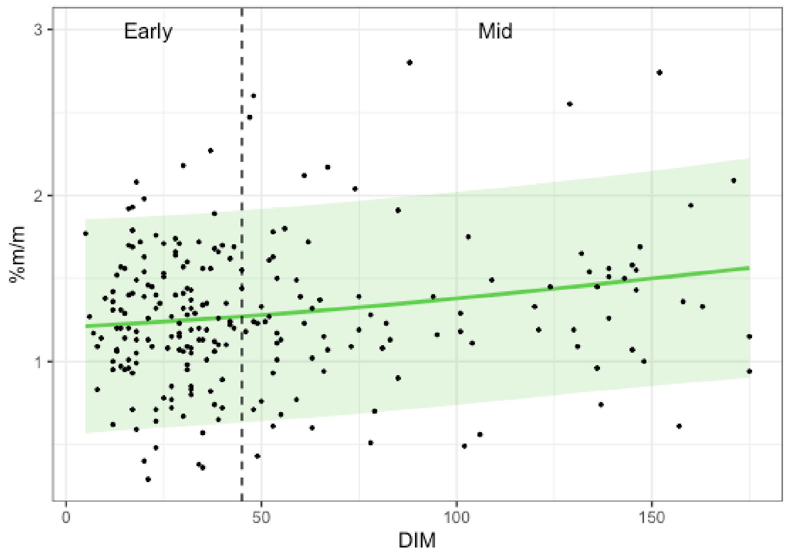


Table 4. *95% confidence intervals (CI) for preformed fatty acids at specific DIM.*

DIM	95% CI (%m/m)
5	[0.76, 2.39]
31	[0.53, 2.14]
56	[0.35, 1.97]
78	[0.28, 1.90]
104	[0.26, 1.89]

Figure 4. *Preformed fatty acids trend with confidence intervals during early and mid-lactation.*

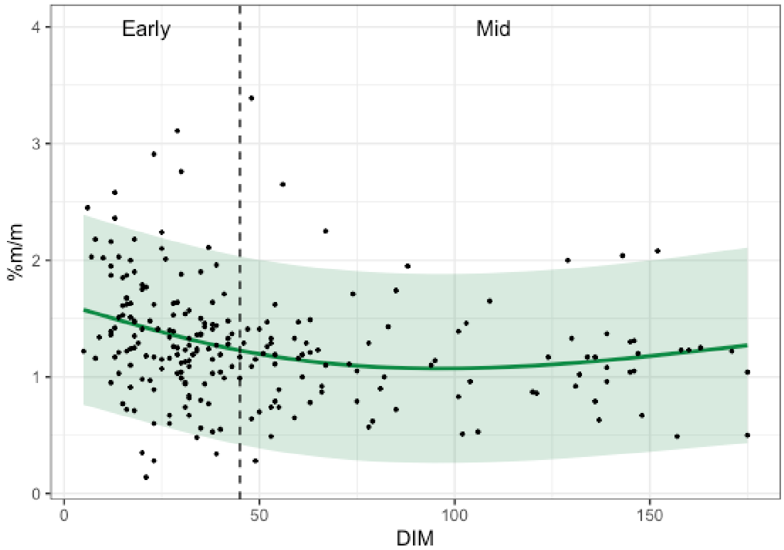


Figure 5. *De novo fatty acids trends with confidence intervals during early and mid-lactation for Grana Padano (GP) and Parmigiano Reggiano (PR) milk.*

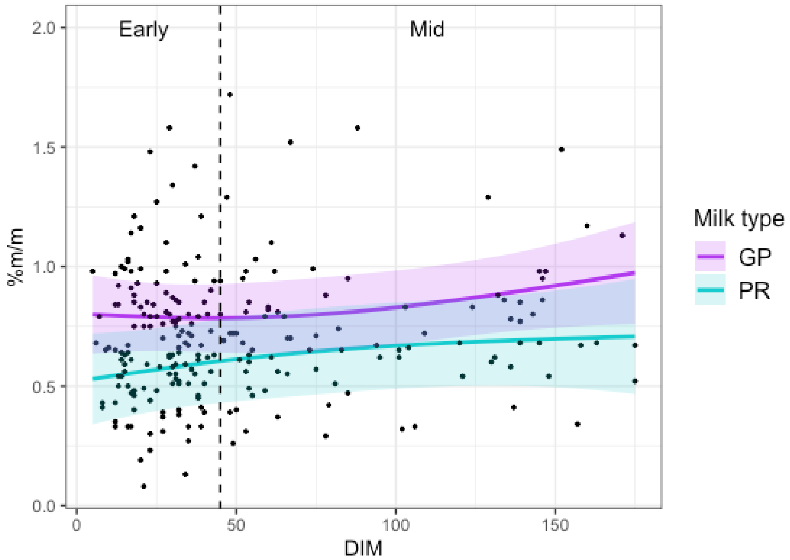


Figure 6. *Mixed fatty acids trends with confidence intervals during early and mid-lactation for Grana Padano (GP) and Parmigiano Reggiano (PR) milk.*

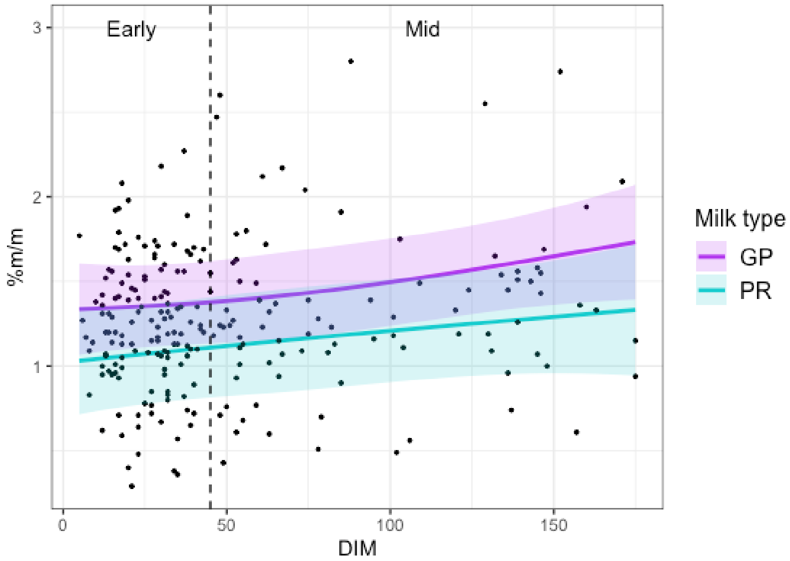
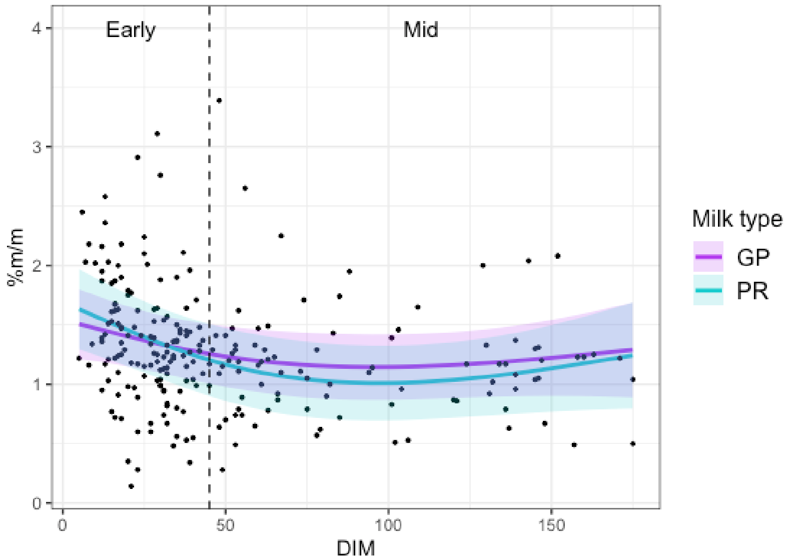


Figure 7. *Preformed fatty acids trends with confidence intervals during early and mid-lactation for Grana Padano (GP) and Parmigiano Reggiano (PR) milk.*



Through the analysis of a greater number of milk samples from a larger number of farms, and considering other potential factors such as calving season and parity, these reference ranges could help farmers to screen cows at risk for a specific health disorder (i.e., subclinical ketosis), even before its clinical signs are visible. A further step should consider the inclusion of this screening system into a decision support tool for the end-user. The best option would be its installation within the software operating in the milking parlor.

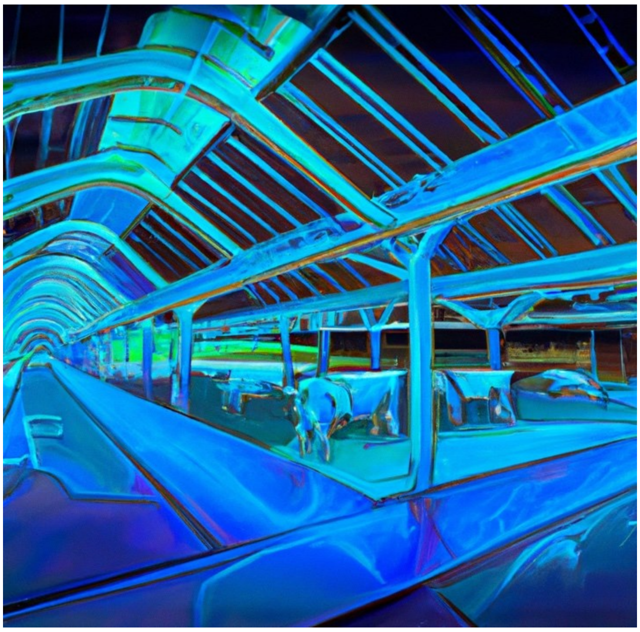
CONCLUSIONS

The outcomes of this pilot study showed the promising feasibility of the use of milk FA trends according to the progress of the lactation as biomarkers for the early detection of cows at risk of specific health problems. The trends observed with the progress of the lactation for de novo, mixed, and preformed FA were consistent with their expected pattern in healthy cows. Their reference intervals according to DIM may help farmers to detect cows having too marked energy deficits. However, it must be pointed out that the reference intervals values of this study must be considered as preliminary and need to be reinforced by further research with more cows and farms. No statistically significant differences in de novo, mixed, and preformed FA reference intervals were detected between GP and PR dairy chains.

ACKNOWLEDGMENTS

This research was funded by the project n.PG-2019-916776 within POR FESR Emilia Romagna with Rumilab Nutristar S.p.A.

CHAPTER 6



Artificially generated image (DALL·E 2, OpenAI)

General conclusions

This thesis presented four studies regarding different PLF applications to the dairy sector. We explored the development and feasibility of different decision support tools addressed to optimize farm management strategies and cows' welfare. In particular, specific solutions have been tested to identify the cows that better cope with the housing conditions and the surrounding environment or to early identify animals at risk of specific metabolic disorders. Dairy farmers and dairy farm advisers are, in fact, increasingly using decision support tools for better decision-making, which is translated into reduced health problems and culling rates, lower environmental impact, and enhanced performance and profitability. For example, based on survival predictions at farm level, dairy farmers could decide to select the cows that better cope with the pathological and environmental challenges of their specific farms, optimizing breeding schemes and culling decisions. Furthermore, sensor data could assist farmers in the early identification of cows for which personalized interventions to alleviate heat stress are needed. Mathematical modelling of high-frequency milk yield data could support them in improving milk production through continuous monitoring of cow health and reproduction, reducing losses in milk yield. In addition, reference ranges of milk contents could help them in screening the cows at risk of specific health disorders, even before clinical signs are visible.

This work highlighted the potential of PLF in assisting dairy farmers to make better choices about the sustainability of their production system, by providing more objective information about health and productivity of their animals. Precision livestock farming introduced a degree of management control over the component processes that was previously impossible. The basis of this control is the detailed knowledge, provided by technology, of individual animals or herds. But as stated by Norton et al. (2019), PLF is much more than '*farming by numbers':* technologies are used to support the farmers, not to substitute them. Farmers, in fact, can collect relevant information about the animals in a continuous manner and thereby build more in-depth insights into their needs, making choices that are not only driven on profits. However, PLF systems do not guarantee improved herd performance unless they are managed by skilled personnel who can professionally translate in practice the impressive amount of information generated by sensors.

REFERENCES

Abeni, F., & Galli, A. (2017). Monitoring cow activity and rumination time for an early detection of heat stress in dairy cow. *International Journal of Biometeorology*, *61*(3), 417–425. https://doi.org/10.1007/s00484-016-1222-z

Abeni, F., Petrera, F., & Galli, A. (2019). A Survey of Italian Dairy Farmers' Propensity for Precision Livestock Farming Tools. *Animals*, *9*(5), 202. https://doi.org/10.3390/ani9050202

Achour, B., Belkadi, M., Saddaoui, R., Filali, I., Aoudjit, R., & Laghrouche, M. (2022). High-accuracy and energy-efficient wearable device for dairy cows' localization and activity detection using low-cost IMU/RFID sensors. *Microsystem Technologies*, *28*(5), 1241–1251. https://doi.org/10.1007/s00542-022-05288-7

Adriaens, I., Friggens, N. C., Ouweltjes, W., Scott, H., Aernouts, B., & Statham, J. (2020). Productive life span and resilience rank can be predicted from on-farm first-parity sensor time series but not using a common equation across farms. *Journal of Dairy Science*, *103*(8), 7155–7171. https://doi.org/10.3168/jds.2019-17826

Adriaens, I., van den Brulle, I., D'Anvers, L., Statham, J. M. E., Geerinckx, K., De Vliegher, S., Piepers, S., & Aernouts, B. (2021). Milk losses and dynamics during perturbations in dairy cows differ with parity and lactation stage. *Journal of Dairy Science*, *104*(1), 405–418. https://doi.org/10.3168/JDS.2020-19195

Agenzia Regionale Protezione Ambiente. (2022). *Meteorologia*. ARPA Lombardia. https://www.arpalombardia.it:443/Pages/Meteorologia/Richiesta-dati-misurati.aspx

Agriculture and Horticulture Development Board. (2023). *Somatic Cell Count, an indicator of milk quality*. AHDB. https://ahdb.org.uk/somatic-cell-count-milk-quality-indicator#:~:text=Milk%20with%20an%20SCC%20of,from%20inside%20the%20animal's%20udder.

Allen, J. D., Hall, L. W., Collier, R. J., & Smith, J. F. (2015). Effect of core body temperature, time of day, and climate conditions on behavioral patterns of lactating dairy cows experiencing mild to moderate heat stress. *Journal of Dairy Science*, *98*(1), 118–127. https://doi.org/10.3168/jds.2013-7704

Alonso, R. S., Sittón-Candanedo, I., García, Ó., Prieto, J., & Rodríguez-González, S. (2020). An intelligent Edge-IoT platform for monitoring livestock and crops in a dairy farming scenario. *Ad Hoc Networks*, *98*, 102047. https://doi.org/10.1016/j.adhoc.2019.102047

Azizi, O., Hasselmann, L., & Kaufmann, O. (2010). Variations in feeding behaviour of high-yielding dairy cows in relation to parity during early to peak lactation. *Archives*

Animal Breeding, *53*(2), 130–140. https://doi.org/10.5194/aab-53-130-2010

Bach, A. (2011). Associations between several aspects of heifer development and dairy cow survivability to second lactation. *Journal of Dairy Science*, *94*(2), 1052–1057. https://doi.org/10.3168/JDS.2010-3633

Bach, K. D., Barbano, D. M., & McArt, J. A. A. (2019). Association of mid-infrared-predicted milk and blood constituents with earlylactation disease, removal, and production outcomes in Holstein cows. *Journal of Dairy Science*, *102*(11), 10129–10139. https://doi.org/10.3168/jds .2019-16926

Bach, A., Valls, N., Solans, A., & Torrent, T. (2008). Associations Between Nondietary Factors and Dairy Herd Performance. *Journal of Dairy Science*, *91*(8), 3259–3267. https://doi.org/10.3168/jds.2008-1030

Balaine, L., Dillon, E. J., Läpple, D., & Lynch, J. (2020). Can technology help achieve sustainable intensification? Evidence from milk recording on Irish dairy farms. *Land Use Policy*, *92*, 104437. https://doi.org/10.1016/j.landusepol.2019.104437

Banhazi, T. M., Lehr, H., Black, J. L., Crabtree, H., Schofield, P., Tscharke, M., & Berckmans, D. (2012). Precision Livestock Farming: An international review of scientific and commercial aspects. *International Journal of Agricultural and Biological Engineering*, *5*(3). https://doi.org/10.25165/ijabe.v5i3.599

Barbano, D. M., Melilli, C., & Overton, T. R. (2014). *Advanced Use of FTIR Spectra of Milk for Feeding and Health Management*. Cornell Nutrition Conference for Feed Manufacturers, Ithaca, NY. https://ecommons.cornell.edu/handle/1813/37978

Barriuso, A. L., Villarrubia González, G., De Paz, J. F., Lozano, Á., & Bajo, J. (2018). Combination of Multi-Agent Systems and Wireless Sensor Networks for the Monitoring of Cattle. *Sensors*, *18*(1). https://doi.org/10.3390/s18010108

Bauman, D. E., & Bruce Currie, W. (1980). Partitioning of nutrients during pregnancy and lactation: a review of mechanisms involving homeostasis and homeorhesis. *Journal of Dairy Science*, *63*(9), 1514–1529. https://doi.org/10.3168/jds.S0022-0302(80)83111-0

Bates, D., & Maechler, M. (2017). *Linear Mixed-Effects Models using «Eigen» and S4*. http://cran.nexr.com/web/packages/lme4/index.html

Beauchemin, K. A. (2018). Invited review: Current perspectives on eating and rumination activity in dairy cows. *Journal of Dairy Science*, *101*(6), 4762–4784. https://doi.org/10.3168/jds.2017-13706

Beauchemin, K. A., Yang, W. Z., & Rode, L. M. (2003). Effects of Particle Size of Alfalfa-Based Dairy Cow Diets on Chewing Activity, Ruminal Fermentation, and Milk Production1. *Journal of Dairy Science*, *86*(2), 630–643. https://doi.org/10.3168/jds.S0022-0302(03)73641-8

Becker, C. A., Collier, R. J., & Stone, A. E. (2020). Invited review: Physiological and behavioral effects of heat stress in dairy cows. *Journal of Dairy Science*, *103*(8), 6751–6770. https://doi.org/10.3168/jds.2019-17929

Ben Abdelkrim, A., Puillet, L., Gomes, P., & Martin, O. (2021). Lactation curve model with explicit representation of perturbations as a phenotyping tool for dairy livestock precision farming. *Animal*, *15*(1), 100074. https://doi.org/10.1016/j.animal.2020.100074

Beniston, M., Stephenson, D. B., Christensen, O. B., Ferro, C. A. T., Frei, C., Goyette, S., Halsnaes, K., Holt, T., Jylhä, K., Koffi, B., Palutikof, J., Schöll, R., Semmler, T., & Woth, K. (2007). Future extreme events in European climate: An exploration of regional climate model projections. *Climatic Change*, *81*(1), 71–95. https://doi.org/10.1007/s10584-006-9226-z

Berckmans, D. (2014). Precision livestock farming technologies for welfare management in intensive livestock systems. *Revue Scientifique Et Technique (International Office of Epizootics)*, *33*(1), 189–196. https://doi.org/10.20506/rst.33.1.2273

Bianchi, M. C., Bava, L., Sandrucci, A., Tangorra, F. M., Tamburini, A., Gislon, G., & Zucali, M. (2022). Diffusion of precision livestock farming technologies in dairy cattle farms. *Animal*, *16*(11), 100650. https://doi.org/10.1016/j.animal.2022.100650

Bohmanova, J., Misztal, I., & Cole, J. B. (2007). Temperature-Humidity Indices as Indicators of Milk Production Losses due to Heat Stress. *Journal of Dairy Science*, *90*(4), 1947–1956. https://doi.org/10.3168/JDS.2006-513

Borchers, M. R., Chang, Y. M., Tsai, I. C., Wadsworth, B. A., & Bewley, J. M. (2016). A validation of technologies monitoring dairy cow feeding, ruminating, and lying behaviors. *Journal of Dairy Science*, *99*(9), 7458–7466. https://doi.org/10.3168/jds.2015-10843

Brown-Brandl, T. M., Nienaber, J. A., Eigenberg, R. A., Hahn, G. L., & Freetly, H. (2003). Thermoregulatory responses of feeder cattle. *Journal of Thermal Biology*, *28*(2), 149–157. https://doi.org/10.1016/S0306-4565(02)00052-9

Bruijnis, M. R. N., Meijboom, F. L. B., & Stassen, E. N. (2013). Longevity as an Animal Welfare Issue Applied to the Case of Foot Disorders in Dairy Cattle. *Journal of Agricultural and Environmental Ethics*, *26*(1), 191–205. https://doi.org/10.1007/S10806-012-9376-0/TABLES/1

Brzozowska, A., Łukaszewicz, M., Sender, G., Kolasińska, D., & Oprządek, J. (2014). Locomotor activity of dairy cows in relation to season and lactation. *Applied Animal Behaviour Science*, *156*, 6–11. https://doi.org/10.1016/j.applanim.2014.04.009

Butler, W. R., & Smith, R. D. (1989). Interrelationships between energy balance and postpartum reproductive function in dairy cattle. *Journal of Dairy Science*, *72*(3),

767–783. https://doi.org/10.3168/jds.S0022-0302(89)79169-4
Cabrera, V. E. (2018). Invited review: Helping dairy farmers to improve economic performance utilizing data-driving decision support tools. *Animal*, *12*(1), 134–144. https://doi.org/10.1017/S1751731117001665
Cardoso, F. C., Kalscheur, K. F., & Drackley, J. K. (2020). Symposium review: Nutrition strategies for improved health, production, and fertility during the transition period. *Journal of Dairy Science*, *103*(6), 5684–5693. https://doi.org/10.3168/jds.2019-17271
Carpentier, L., Berckmans, D., Youssef, A., Berckmans, D., van Waterschoot, T., Johnston, D., Ferguson, N., Earley, B., Fontana, I., Tullo, E., Guarino, M., Vranken, E., & Norton, T. (2018). Automatic cough detection for bovine respiratory disease in a calf house. *Biosystems Engineering*, *173*, 45–56. https://doi.org/10.1016/j.biosystemseng.2018.06.018
Chen, S., Wang, J., Peng, D., Li, G., Chen, J., & Gu, X. (2018). Exposure to heat-stress environment affects the physiology, circulation levels of cytokines, and microbiome in dairy cows. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-32886-1
CLAL. (2022). *Italia: Patrimonio zootecnico (Bovini da Latte)*. https://teseo.clal.it/clal20/index.php?section=vacche_italia
Coimbra, P. A. D., Machado Filho, L. C. P., & Hötzel, M. J. (2012). Effects of social dominance, water trough location and shade availability on drinking behaviour of cows on pasture. *Applied Animal Behaviour Science*, *139*(3), 175–182. https://doi.org/10.1016/j.applanim.2012.04.009
Cook, N. B., Mentink, R. L., Bennett, T. B., & Burgi, K. (2007). The Effect of Heat Stress and Lameness on Time Budgets of Lactating Dairy Cows. *Journal of Dairy Science*, *90*(4), 1674–1682. https://doi.org/10.3168/jds.2006-634
CREA – Research Centre for Agricultural Policies and Bioeconomy. (2022). *Italian Agriculture in Figures*. https://www.crea.gov.it/web/politiche-e-bioeconomia/-/agricoltura-italiana-conta
Dallago, G. M., Wade, K. M., Cue, R. I., McClure, J. T., Lacroix, R., Pellerin, D., & Vasseur, E. (2021). Keeping Dairy Cows for Longer: A Critical Literature Review on Dairy Cow Longevity in High Milk-Producing Countries. *Animals*, *11*(3). https://doi.org/10.3390/ani11030808
Davis, M. S., Mader, T. L., Holt, S. M., & Parkhurst, A. M. (2003). Strategies to reduce feedlot cattle heat stress: Effects on tympanic temperature. *Journal of Animal Science*, *81*(3), 649–661. https://doi.org/10.2527/2003.813649x
de Andrade Ferrazza, R., Mogollón Garcia, H. D., Vallejo Aristizábal, V. H., de Souza Nogueira, C., Veríssimo, C. J., Sartori, J. R., Sartori, R., & Pinheiro Ferreira, J. C.

(2017). Thermoregulatory responses of Holstein cows exposed to experimentally induced heat stress. *Journal of Thermal Biology*, *66*, 68–80. https://doi.org/10.1016/j.jtherbio.2017.03.014

de Koning, K. (2002). MILKING MACHINES | Robotic Milking. In *Encyclopedia of Dairy Sciences* (pp. 2038–2045). Elsevier. https://doi.org/10.1016/B0-12-227235-8/00346-1

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review*, *65*(3), 122–135. https://doi.org/10.1108/LR-06-2015-0061

Dematawewa, C. M. B., Pearson, R. E., & VanRaden, P. M. (2007). Modeling Extended Lactations of Holsteins. *Journal of Dairy Science*, *90*(8), 3924–3936. https://doi.org/10.3168/jds.2006-790

Denis-Robichaud, J., Dubuc, J., Lefebvre, D., & DesCôteaux, L. (2014). Accuracy of milk ketone bodies from flow-injection analysis for the diagnosis of hyperketonemia in dairy cows. *Journal of Dairy Science*, *97*(6), 3364–3370. https://doi.org/10.3168/jds.2013-6744

Dickinson, R. A., Morton, J. M., Beggs, D. S., Anderson, G. A., Pyman, M. F., Mansell, P. D., & Blackwood, C. B. (2013). An automated walk-over weighing system as a tool for measuring liveweight change in lactating dairy cows. *Journal of Dairy Science*, *96*(7), 4477–4486. https://doi.org/10.3168/jds.2012-6522

Dodge, Y. (2008). Kruskal-Wallis Test. In *The Concise Encyclopedia of Statistics* (pp. 288–290). Springer.

Elzhov, T. V., Mullen, K. M., Spiess, A.-N., & Bolker, B. (2023, gennaio 26). *minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds*. https://cran.r-project.org/web/packages/minpack.lm/minpack.lm.pdf

European Parliament. (2015). *The future of the EU dairy sector after the end of milk quotas. Briefing Document: EPRS_BRI(2015)564359.*

EUROSTAT. (2020). *Milk and milk product statistics*. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Milk_and_milk_product_statistics

Evangelista, C., Basiricò, L., & Bernabucci, U. (2021). An Overview on the Use of Near Infrared Spectroscopy (NIRS) on Farms for the Management of Dairy Cows. *Agriculture*, *11*(4). https://doi.org/10.3390/agriculture11040296

Fitzmaurice, G., Davidian, M., Verbeke, G., & Molenberghs, G. (2008). Generalized linear mixed-effects models. In *Longitudinal Data Analysis* (1st ed.). Chapman & Hall/CRC.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied Longitudinal Analysis*. John Wiley & Sons.

Frich, P., Alexander, L. V., Della-Marta, P., Gleason, B., Haylock, M., Tank, A. M. G. K., & Peterson, T. (2002). Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research*, *19*(3), 193–212. https://doi.org/10.3354/cr019193

Fustini, M., Heinrichs, A. J., Palmonari, A., & Formigoni, A. (2016). Case Study: Farm characteristics and total mixed ration particle size issues on Parmigiano Reggiano farms in northern Italy. *The Professional Animal Scientist*, *32*(6), 869-873. https://doi.org/10.15232/pas.2016-01550

Gao, X., & Giorgi, F. (2008). Increased aridity in the Mediterranean region under greenhouse gas forcing estimated from high resolution simulations with a regional climate model. *Global and Planetary Change*, *62*(3), 195–209. https://doi.org/10.1016/j.gloplacha.2008.02.002

García, R., Aguilar, J., Toro, M., Pinto, A., & Rodríguez, P. (2020). A systematic literature review on the use of machine learning in precision livestock farming. *Computers and Electronics in Agriculture*, *179*, 105826. https://doi.org/10.1016/j.compag.2020.105826

Gengler, N., Soyeurt, H., Dehareng, F., Bastin, C., Colinet, F., Hammami, H., Vanrobays, M.-L., Lainé, A., Vanderick, S., Grelet, C., Vanlierde, A., Froidmont, E., & Dardenne, P. (2016). Capitalizing on fine milk composition for breeding and management of dairy cows1. *Journal of Dairy Science*, *99*(5), 4071–4079. https://doi.org/10.3168/jds.2015-10140

GenTORE. (2017). *Genomic management Tools to Optimise Resilience and Efficiency*. Genomic management Tools to Optimise Resilience and Efficiency. https://www.gentore.eu/project.html

Gerds, T. A., & Schumacher, M. (2006). Consistent estimation of the expected Brier score in general survival models with right-censored event times. *Biometrical Journal*, *48*(6), 1029–1040. https://doi.org/10.1002/BIMJ.200610301

Goldstein-Greenwood, J. (2021). *A Brief on Brier Scores*. https://data.library.virginia.edu/a-brief-on-brier-scores/

Gote, M. J., D'Anvers, L., Meuwissen, D., Aernouts, B., & Adriaens, I. (2022). *An automated pipeline for real-time data integration, health monitoring, and validation for dairy cows*. *28*, 549.

Græsbøll, K., Kirkeby, C., Nielsen, S. S., Halasa, T., Toft, N., & Christiansen, L. E. (2016). Models to estimate lactation curves of milk yield and somatic cell count in dairy cows at the herd level for the use in simulations and predictive models. *Frontiers* *in Veterinary Science*, *3*(DEC), 115–115. https://doi.org/10.3389/fvets.2016.00115

Grana Padano. (2019, novembre). *Specification of Grana Padano P.D.O.* chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.granapadano.it/public/file/Production-Specification-RulesNov12-38345.pdf

Grandl, F., Furger, M., Kreuzer, M., & Zehetmeier, M. (2019). Impact of longevity on greenhouse gas emissions and profitability of individual dairy cows analysed with different system boundaries. *Animal*, *13*(1), 198–208. https://doi.org/10.1017/S175173111800112X

Grant, R. J., & Albright, J. L. (2001). Effect of Animal Grouping on Feeding Behavior and Intake of Dairy Cattle. *Journal of Dairy Science*, *84*, E156–E163. https://doi.org/10.3168/jds.S0022-0302(01)70210-X

Halekoh, U., Højsgaard, S., & Yan, J. (2006). The R Package geepack for Generalized Estimating Equations. *Journal of Statistical Software*, *15*(2), 1–11. https://doi.org/10.18637/JSS.V015.I02

Hardin, J. W., & Hilbe, J. M. (2013). *Generalized Estimating Equations* (2ª ed.). Chapman and Hall/CRC.

Heagerty, P. J., & Zheng, Y. (2005). Survival model predictive accuracy and ROC curves. *Biometrics*, *61*(1), 92–105. https://doi.org/10.1111/J.0006-341X.2005.030814.X

Heinicke, J., Hoffmann, G., Ammon, C., Amon, B., & Amon, T. (2018). Effects of the daily heat load duration exceeding determined heat load thresholds on activity traits of lactating dairy cows. *Journal of Thermal Biology*, *77*, 67–74. https://doi.org/10.1016/j.jtherbio.2018.08.012

Hendriks, S. J., Phyn, C. V. C., Huzzey, J. M., Mueller, K. R., Turner, S.-A., Donaghy, D. J., & Roche, J. R. (2020). Graduate Student Literature Review: Evaluating the appropriate use of wearable accelerometers in research to monitor lying behaviors of dairy cows. *Journal of Dairy Science*, *103*(12), 12140–12157. https://doi.org/10.3168/jds.2019-17887

Herbut, P., Angrecka, S., & Walczak, J. (2018). Environmental parameters to assessing of heat stress in dairy cattle—A review. *International Journal of Biometeorology*, *62*(12), 2089–2097. https://doi.org/10.1007/s00484-018-1629-9

Hertl, J. A., Schukken, Y. H., Welcome, F. L., Tauer, L. W., & Gröhn, Y. T. (2014). Pathogen-specific effects on milk yield in repeated clinical mastitis episodes in Holstein dairy cows. *Journal of Dairy Science*, *97*(3), 1465–1480. https://doi.org/10.3168/jds.2013-7266

Hin, L. Y., & Wang, Y. G. (2009). Working-correlation-structure identification in generalized estimating equations. *Statistics in Medicine*, *28*(4), 642–658. https://doi.org/10.1002/SIM.3489

Hoffman, J. M., & Valencak, T. G. (2020). A short life on the farm: Aging and longevity in agricultural, large-bodied mammals. *GeroScience*, *42*(3), 909–922. https://doi.org/10.1007/s11357-020-00190-4

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Hostens, M., Ehrlich, J., Van Ranst, B., & Opsomer, G. (2012). On-farm evaluation of the effect of metabolic diseases on the shape of the lactation curve in dairy cows through the MilkBot lactation model. *Journal of Dairy Science*, *95*(6), 2988–3007. https://doi.org/10.3168/jds.2011-4791

Hubbard, A. E., Ahern, J., Fleischer, N. L., Van der Laan, M., Satariano, S. A., Jewell, N., Bruckner, T., & Satariano, W. A. (2010). To GEE or Not to GEE: Comparing Population Average and Mixed Models for Estimating the Associations Between Neighborhood Risk Factors and Health. *Epidemiology*, *21*(4), 467–474.

Hut, P. R., Scheurwater, J., Nielen, M., van den Broek, J., & Hostens, M. M. (2022). Heat stress in a temperate climate leads to adapted sensor-based behavioral patterns of dairy cows. *Journal of Dairy Science*, *105*(8), 6909–6922. https://doi.org/10.3168/jds.2021-21756

Islam, M. A., Lomax, S., Doughty, A. K., Islam, M. R., Thomson, P. C., & Clark, C. E. F. (2021). Revealing the diversity in cattle behavioural response to high environmental heat using accelerometer-based ear tag sensors. *Computers and Electronics in Agriculture*, *191*, 106511. https://doi.org/10.1016/j.compag.2021.106511

Jiang, B., Tang, W., Cui, L., & Deng, X. (2023). Precision Livestock Farming Research: A Global Scientometric Review. *Animals*, *13*(13). https://doi.org/10.3390/ani13132096

John, A. J., Clark, C. E. F., Freeman, M. J., Kerrisk, K. L., Garcia, S. C., & Halachmi, I. (2016). Review: Milking robot utilization, a successful precision livestock farming evolution. *Animal*, *10*(9), 1484–1492. https://doi.org/10.1017/S1751731116000495

Kalbfleisch, J. D., & Prentice, R. L. (2002). Relative Risk (Cox) Regression Models. In *The statistical analysis of failure time data* (2nd ed., pp. 95–147). Wiley and Sons.

Karimi, M. T., Ghorbani, G. R., Kargar, S., & Drackley, J. K. (2015). Late-gestation heat stress abatement on performance and behavior of Holstein dairy cows. *Journal of Dairy Science*, *98*(10), 6865–6875. https://doi.org/10.3168/jds.2014-9281

Kelly, C. F., & Bond, T. E. (1971). Bioclimatic factors and their measurement. In *A guide to environmental research on animals* (pp. 7–92). National Academy of Sciences.

King, M. T. M., & DeVries, T. J. (2018). Graduate Student Literature Review: Detecting health disorders using data from automatic milking systems and associated technologies. *Journal of Dairy Science*, *101*(9), 8605–8614.

https://doi.org/10.3168/jds.2018-14521

Knob, D. A., Scholz, A. M., Perazzoli, L., Mendes, B. P. B., Kappes, R., Alessio, D. R. M., Rech, Â. F., & Thaler Neto, A. (2023). Feed Efficiency and Physiological Parameters of Holstein and Crossbred Holstein × Simmental Cows. *Animals*, *13*(10). https://doi.org/10.3390/ani13101668

Koenker, R. (2005). *Quantile Regression*. Cambridge University Press. https://doi.org/10.1017/CBO9780511754098

Koenker, R. (2023, aprile 8). *quantreg: Quantile Regression*. https://cran.r-project.org/web/packages/quantreg/index.html

Krieger, S., Oczak, M., Lidauer, L., Berger, A., Kickinger, F., Öhlschuster, M., Auer, W., Drillich, M., & Iwersen, M. (2019). An ear-attached accelerometer as an on-farm device to predict the onset of calving in dairy cows. *Biosystems Engineering*, *184*, 190–199. https://doi.org/10.1016/j.biosystemseng.2019.06.011

Lee, J. Y., & Kim, I. H. (2006). Advancing parity is associated with high milk production at the cost of body condition and increased periparturient disorders in dairy herds. *Journal of Veterinary Science*, *7*(2), 161–166. https://doi.org/10.4142/jvs.2006.7.2.161

Levene, H. (1960). Robust tests for equality of variances. In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (pp. 278–292). Stanford University Press.

Lever, J., Krzywinski, M., & Altman, N. (2016, agosto 30). *Model selection and overfitting*. https://www.nature.com/articles/nmeth.3968

Li, M., Rosa, G. J. M., Reed, K. F., & Cabrera, V. E. (2022). Investigating the effect of temporal, geographic, and management factors on US Holstein lactation curve parameters. *Journal of Dairy Science*, *105*(9), 7525–7538. https://doi.org/10.3168/jds.2022-21882

Lin, J., Shen, Z., Zhang, A., & Chai, Y. (2018). Blockchain and IoT based Food Traceability for Smart Agriculture. *Proceedings of the 3rd International Conference on Crowd Science and Engineering*, 1–6. https://doi.org/10.1145/3265689.3265692

Liu, Z., Ezernieks, V., Wang, J., Arachchillage, N. W., Garner, J. B., Wales, W. J., Cocks, B. G., & Rochfort, S. (2017). Heat Stress in Dairy Cattle Alters Lipid Composition of Milk. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-01120-9

Lokhorst, C., de Mol, R. M., & Kamphuis, C. (2019). Invited review: Big Data in precision dairy farming. *Animal*, *13*(7), 1519–1528. https://doi.org/10.1017/S1751731118003439

Lopes, S. I., Bexiga, R., Araújo, J. P., Cerqueira, J. L., Abreu, C., Paredes, C., & Alonso, J. M. (2016). Precision livestock farming for reproductive performance optimization:

A survey. In *Food futures: Ethics, science and culture* (pp. 587–592). Wageningen Academic Publishers. https://doi.org/10.3920/978-90-8686-834-6_90

Lora, I., Gottardo, F., Contiero, B., Zidi, A., Magrin, L., Cassandro, M., & Cozzi, G. (2020). A survey on sensor systems used in Italian dairy farms and comparison between performances of similar herds equipped or not equipped with sensors. *Journal of Dairy Science*, *103*(11), 10264–10272. https://doi.org/10.3168/jds.2019-17973

Lovarelli, D., Bacenetti, J., & Guarino, M. (2020). A review on dairy cattle farming: Is precision livestock farming the compromise for an environmental, economic and social sustainable production? *Journal of Cleaner Production*, *262*, 121409. https://doi.org/10.1016/j.jclepro.2020.121409

Machado, N. A. F., Da Costa, L. B. S., Barbosa-Filho, J. A. D., De Oliveira, K. P. L., De Sampaio, L. C., Peixoto, M. S. M., & Damasceno, F. A. (2021). Using infrared thermography to detect subclinical mastitis in dairy cows in compost barn systems. *Journal of Thermal Biology*, *97*, 102881. https://doi.org/10.1016/j.jtherbio.2021.102881

Maggiolino, A., Landi, V., Bartolomeo, N., Bernabucci, U., Santus, E., Bragaglio, A., & De Palo, P. (2022). Effect of Heat Waves on Some Italian Brown Swiss Dairy Cows' Production Patterns. *Frontiers in Animal Science*, *2*, 800680. https://doi.org/10.3389/fanim.2021.800680

Maia, G. G., Siqueira, L. G. B., Vasconcelos, C. O. de P., Tomich, T. R., Camargo, L. S. de A., Rodrigues, J. P. P., de Menezes, R. A., Gonçalves, L. C., Teixeira, B. F., Grando, R. de O., Nogueira, L. A. G., & Pereira, L. G. R. (2020). Effects of heat stress on rumination activity in Holstein-Gyr dry cows. *Livestock Science*, *239*, 104092. https://doi.org/10.1016/j.livsci.2020.104092

Maulfair, D. D., Zanton, G. I., Fustini, M., & Heinrichs, A. J. (2010). Effect of feed sorting on chewing behavior, production, and rumen fermentation in lactating dairy cows1. *Journal of Dairy Science*, *93*(10), 4791–4803. https://doi.org/10.3168/jds.2010-3278

McNamara, J. P. (2002). BODY CONDITION | Effects on Health, Milk Production and Reproduction. In *Encyclopedia of Dairy Sciences* (pp. 168–173). Academic Press. http://www.sciencedirect.com:5070/referencework/9780122272356/encyclopedia-of-dairy-sciences

Meen, G. H., Schellekens, M. A., Slegers, M. H. M., Leenders, N. L. G., van Erp-van der Kooij, E., & Noldus, L. P. J. J. (2015). Sound analysis in dairy cattle vocalisation as a potential welfare monitor. *Computers and Electronics in Agriculture*, *118*, 111–115. https://doi.org/10.1016/j.compag.2015.08.028

Morrone, S., Dimauro, C., Gambella, F., & Cappai, M. G. (2022). Industry 4.0 and Precision Livestock Farming (PLF): An up to Date Overview across Animal

Productions. *Sensors*, *22*(12), 4319. https://doi.org/10.3390/s22124319

Mottram, T., Lowe, J., McGowan, M., & Phillips, N. (2008). Technical note: A wireless telemetric method of monitoring clinical acidosis in dairy cows. *Computers and Electronics in Agriculture*, *64*(1), 45–48. https://doi.org/10.1016/j.compag.2008.05.015

Mulligan, F. J., O'Grady, L., Rice, D. A., & Doherty, M. L. (2006). A herd health approach to dairy cow nutrition and production diseases of the transition cow. *Animal Reproduction Science*, *96*(3), 331–353. https://doi.org/10.1016/j.anireprosci.2006.08.011

Neave, H. W., Lomb, J., von Keyserlingk, M. A. G., Behnam-Shabahang, A., & Weary, D. M. (2017). Parity differences in the behavior of transition dairy cows. *Journal of Dairy Science*, *100*(1), 548–561. https://doi.org/10.3168/jds.2016-10987

Neethirajan, S. (2017). Recent advances in wearable sensors for animal health management. *Sensing and Bio-Sensing Research*, *12*, 15–29. https://doi.org/10.1016/j.sbsr.2016.11.004

Neethirajan, S. (2020). The role of sensors, big data and machine learning in modern animal farming. *Sensing and Bio-Sensing Research*, *29*, 100367. https://doi.org/10.1016/j.sbsr.2020.100367

Neethirajan, S., & Kemp, B. (2021). Digital Livestock Farming. *Sensing and Bio-Sensing Research*, *32*, 100408. https://doi.org/10.1016/j.sbsr.2021.100408

Niozas, G., Tsousis, G., Malesios, C., Steinhöfel, I., Boscos, C., Bollwein, H., & Kaske, M. (2019). Extended lactation in high-yielding dairy cows. II. Effects on milk production, udder health, and body measurements. *Journal of Dairy Science*, *102*(1), 811–823. https://doi.org/10.3168/jds.2018-15117

Norton, T., & Berckmans, D. (2017). Developing precision livestock farming tools for precision dairy farming. *Animal Frontiers*, *7*(1), 18–23. https://doi.org/10.2527/af.2017.0104

Norton, T., Chen, C., Larsen, M. L. V., & Berckmans, D. (2019). Review: Precision livestock farming: building 'digital representations' to bring the animals closer to the farmer. *Animal*, *13*(12), 3009–3017. https://doi.org/10.1017/S175173111900199X

Oczak, M., Ismayilova, G., Costa, A., Viazzi, S., Sonoda, L. T., Fels, M., Bahr, C., Hartung, J., Guarino, M., Berckmans, D., & Vranken, E. (2013). Analysis of aggressive behaviours of pigs by automatic video recordings. *Computers and Electronics in Agriculture*, *99*, 209–217. https://doi.org/10.1016/j.compag.2013.09.015

Olechnowicz, J., Kneblewski, P., Jaśkowski, J. M., & Włodarek, J. (2016). Effect of selected factors on longevity in cattle: A review. *Journal of Animal and Plant*

Sciences, *26*(6), 1533–1541.

Ouweltjes, W., Spoelstra, M., Ducro, B., de Haas, Y., & Kamphuis, C. (2021). A data-driven prediction of lifetime resilience of dairy cows using commercial sensor data collected during first lactation. *Journal of Dairy Science*, *104*(11), 11759–11769. https://doi.org/10.3168/JDS.2021-20413

Parmigiano Reggiano. (2018). *Parmigiano Reggiano PDO Specifications*. https://www.parmigianoreggiano.com/consortium-specifications-and-legislation

Perdomo, M. C., Marsola, R. S., Favoreto, M. G., Adesogan, A., Staples, C. R., & Santos, J. E. P. (2020). Effects of feeding live yeast at 2 dosages on performance and feeding behavior of dairy cows under heat stress. *Journal of Dairy Science*, *103*(1), 325–339. https://doi.org/10.3168/jds.2019-17303

Picchi, V. V., de Castro, E. F. J., Marino, F. C. H., & Ribeiro, S. L. (2020). Increasing the Confidence of the Brazilian Livestock Production Chain using Blockchain. *Proceedings of the 2019 2nd International Conference on Blockchain Technology and Applications*, 93–98. https://doi.org/10.1145/3376044.3376058

Poppe, M., Veerkamp, R. F., van Pelt, M. L., & Mulder, H. A. (2020). Exploration of variance, autocorrelation, and skewness of deviations from lactation curves as resilience indicators for breeding. *Journal of Dairy Science*, *103*(2), 1667–1684. https://doi.org/10.3168/jds.2019-17290

Proudfoot, K. L., & Huzzey, J. M. (2022). A first time for everything: The influence of parity on the behavior of transition dairy cows. *JDS Communications*, *3*(6), 467–471. https://doi.org/10.3168/jdsc.2022-0290

Ranzato, G. (2023). Environmental and behavioral sensor data to identify heat stress in dairy cows ver 1. *Environmental Data Initiative*. https://doi.org/10.6073/pasta/4a8cbfc4011ff759a0b5a3f08d7fc872

Ranzato, G., Adriaens, I., Lora, I., Aernouts, B., Statham, J., Azzolina, D., Meuwissen, D., Prosepe, I., Zidi, A., & Cozzi, G. (2022). Joint Models to Predict Dairy Cow Survival from Sensor Data Recorded during the First Lactation. *Animals*, *12*(24), 3494. https://doi.org/10.3390/ani12243494

Ranzato, G., Lora, I., Aernouts, B., Adriaens, I., Gottardo, F., & Cozzi, G. (2023). Sensor-based behavioral patterns can identify heat-sensitive lactating dairy cows. *International Journal of Biometeorology*. Accepted for publication.

Ravagnolo, O., Misztal, I., & Hoogenboom, G. (2000). Genetic Component of Heat Stress in Dairy Cattle, Development of Heat Index Function. *Journal of Dairy Science*, *83*(9), 2120–2125. https://doi.org/10.3168/jds.S0022-0302(00)75094-6

Reiter, S., Sattlecker, G., Lidauer, L., Kickinger, F., Öhlschuster, M., Auer, W., Schweinzer, V., Klein-Jöbstl, D., Drillich, M., & Iwersen, M. (2018). Evaluation of an ear-tag-based accelerometer for monitoring rumination in dairy cows. *Journal of Dairy Science*, *101*(4), 3398–3411. https://doi.org/10.3168/jds.2017-12686

Rizopoulos, D. (2012). *Joint models for longitudinal and time-to-event data: With applications in R* (1st ed.). Chapman & Hall/CRC.

Rizopoulos, D., Papageorgiou, G., & Miranda Afonso, P. (2022). *JMbayes2: Extended Joint Models for Longitudinal and Time-to-Event Data*. https://cran.r-project.org/web/packages/JMbayes2/JMbayes2.pdf

Roche, J. R., Bell, A. W., Overton, T. R., & Loor, J. J. (2013). Nutritional management of the transition cow in the 21st century – a paradigm shift in thinking. *Animal Production Science*, *53*(9), 1000–1023. https://doi.org/10.1071/AN12293

Rojo-Gimeno, C., van der Voort, M., Niemi, J. K., Lauwers, L., Kristensen, A. R., & Wauters, E. (2019). Assessment of the value of information of precision livestock farming: A conceptual framework. *NJAS: Wageningen Journal of Life Sciences*, *90–91*(1), 1–9. https://doi.org/10.1016/j.njas.2019.100311

Roland, L., Schweinzer, V., Kanz, P., Sattlecker, G., Kickinger, F., Lidauer, L., Berger, A., Auer, W., Mayer, J., Sturm, V., Efrosinin, D., Breitenberger, S., Drillich, M., & Iwersen, M. (2018). Technical note: Evaluation of a triaxial accelerometer for monitoring selected behaviors in dairy calves. *Journal of Dairy Science*, *101*(11), 10421–10427. https://doi.org/10.3168/jds.2018-14720

Rutten, C. J., Velthuis, A. G. J., Steeneveld, W., & Hogeveen, H. (2013). Invited review: Sensors to support health management on dairy farms. *Journal of Dairy Science*, *96*(4), 1928–1952. https://doi.org/10.3168/JDS.2012-6107

Santschi, D. E., Lacroix, R., Durocher, J., Duplessis, M., Moore, R. K., & Lefebvre, D. M. (2016). Prevalence of elevated milk β-hydroxybutyrate concentrations in Holstein cows measured by Fourier-transform infrared analysis in Dairy Herd Improvement milk samples and association with milk yield and components. *Journal of Dairy Science*, *99*(11), 9263–9270. https://doi.org/10.3168/jds.2016-11128

Schillings, J., Bennett, R., & Rose, D. C. (2021). Animal welfare and other ethical implications of Precision Livestock Farming technology. *CABI Agriculture and Bioscience*, *2*(1), 17. https://doi.org/10.1186/s43170-021-00037-8

Schuster, J. C., Barkema, H. W., De Vries, A., Kelton, D. F., & Orsel, K. (2020). Invited review: Academic and applied approach to evaluating longevity in dairy cows. *Journal of Dairy Science*, *103*(12), 11008–11024. https://doi.org/10.3168/JDS.2020-19043

Segnalini, M., Bernabucci, U., Vitali, A., Nardone, A., & Lacetera, N. (2013). Temperature humidity index scenarios in the Mediterranean basin. *International Journal of Biometeorology*, *57*(3), 451–458. https://doi.org/10.1007/s00484-012-0571-5

Shahriar, Md. S., Smith, D., Rahman, A., Freeman, M., Hills, J., Rawnsley, R., Henry, D., & Bishop-Hurley, G. (2016). Detecting heat events in dairy cows using accelerometers and unsupervised learning. *Computers and Electronics in Agriculture*, *128*, 20–26. https://doi.org/10.1016/j.compag.2016.08.009

Siewert, J. M., Salfer, J. A., & Endres, M. I. (2019). Milk yield and milking station visits of primiparous versus multiparous cows on automatic milking system farms in the Upper Midwest United States. *Journal of Dairy Science*, *102*(4), 3523–3530. https://doi.org/10.3168/jds.2018-15382

Silvey, S. D. (1975). The Likelihood-Ratio Test and Alternative 'Large-Sample' Equivalents of it. In *Statistical Inference* (1ª ed., p. Chapter 7). Chapman & Hall/CRC.

Spiers, D. E., Spain, J. N., Sampson, J. D., & Rhoads, R. P. (2004). Use of physiological parameters to predict milk yield and feed intake in heat-stressed dairy cows. *Journal of Thermal Biology*, *29*(7), 759–764. https://doi.org/10.1016/j.jtherbio.2004.08.051

Steeneveld, W., & Hogeveen, H. (2015). Characterization of Dutch dairy farms using sensor systems for cow management. *Journal of Dairy Science*, *98*(1), 709–717. https://doi.org/10.3168/JDS.2014-8595

Steensels, M., Antler, A., Bahr, C., Berckmans, D., Maltz, E., & Halachmi, I. (2016). A decision-tree model to detect post-calving diseases based on rumination, activity, milk yield, BW and voluntary visits to the milking robot. *Animal*, *10*(9), 1493–1500. https://doi.org/10.1017/S1751731116000744

Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society. Series B (Methodological)*, *36*(2), 111–147. https://doi.org/10.1111/j.2517-6161.1974.tb00994.x

Stygar, A. H., Gómez, Y., Berteselli, G. V., Dalla Costa, E., Canali, E., Niemi, J. K., Llonch, P., & Pastell, M. (2021). A Systematic Review on Commercially Available and Validated Sensor Technologies for Welfare Assessment of Dairy Cattle. *Frontiers in Veterinary Science*, *8*, 634338. https://doi.org/10.3389/fvets.2021.634338

Torshizi, M. E. (2016). Effects of season and age at first calving on genetic and phenotypic characteristics of lactation curve parameters in Holstein cows. *Journal of Animal Science and Technology*, *58*. https://doi.org/10.1186/s40781-016-0089-1

Tse, C., Barkema, H. W., DeVries, T. J., Rushen, J., Vasseur, E., & Pajor, E. A. (2018). Producer experience with transitioning to automatic milking: Cow training, challenges, and effect on quality of life. *Journal of Dairy Science*, *101*(10), 9599–9607. https://doi.org/10.3168/jds.2018-14662

Tullo, E., Finzi, A., & Guarino, M. (2019). Review: Environmental impact of livestock farming and Precision Livestock Farming as a mitigation strategy. *Science of The*

Total Environment, *650*, 2751–2760. https://doi.org/10.1016/j.scitotenv.2018.10.018

van der Heide, E. M. M., Kamphuis, C., Veerkamp, R. F., Athanasiadis, I. N., Azzopardi, G., van Pelt, M. L., & Ducro, B. J. (2020). Improving predictive performance on survival in dairy cattle using an ensemble learning approach. *Computers and Electronics in Agriculture*, *177*, 105675–105675. https://doi.org/10.1016/J.COMPAG.2020.105675

van der Heide, E. M. M., Veerkamp, R. F., van Pelt, M. L., Kamphuis, C., Athanasiadis, I., & Ducro, B. J. (2019). Comparing regression, naive Bayes, and random forest methods in the prediction of individual survival to second lactation in Holstein cattle. *Journal of Dairy Science*, *102*(10), 9409–9421. https://doi.org/10.3168/JDS.2019-16295

van Pelt, M. L., Meuwissen, T. H. E., de Jong, G., & Veerkamp, R. F. (2015). Genetic analysis of longevity in Dutch dairy cattle using random regression. *Journal of Dairy Science*, *98*(6), 4117–4130. https://doi.org/10.3168/JDS.2014-9090

van Ravenzwaaij, D., Cassey, P., & Brown, S. D. (2018). A simple introduction to Markov Chain Monte–Carlo sampling. *Psychonomic Bulletin and Review*, *25*(1), 143–154. https://doi.org/10.3758/S13423-016-1015-8

VandeHaar, M. J., & St-Pierre, N. (2006). Major Advances in Nutrition: Relevance to the Sustainability of the Dairy Industry. *Journal of Dairy Science*, *89*(4), 1280–1291. https://doi.org/10.3168/JDS.S0022-0302(06)72196-8

Vitali, A., Felici, A., Esposito, S., Bernabucci, U., Bertocchi, L., Maresca, C., Nardone, A., & Lacetera, N. (2015). The effect of heat waves on dairy cow mortality. *Journal of Dairy Science*, *98*(7), 4572–4579. https://doi.org/10.3168/jds.2015-9331

Wathes, D. C., Cheng, Z., Bourne, N., Taylor, V. J., Coffey, M. P., & Brotherstone, S. (2007). Differences between primiparous and multiparous dairy cows in the inter-relationships between metabolic traits, milk yield and body condition score in the periparturient period. *Domestic Animal Endocrinology*, *33*(2), 203–225. https://doi.org/10.1016/j.domaniend.2006.05.004

West, J. W. (2003). Effects of Heat-Stress on Production in Dairy Cattle. *Journal of Dairy Science*, *86*(6), 2131–2144. https://doi.org/10.3168/jds.S0022-0302(03)73803-X

Wilmink, J. B. M. (1987). Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. *Livestock Production Science*, *16*(4), 335–348. https://doi.org/10.1016/0301-6226(87)90003-0

Wood, P. D. P. (1967). Algebraic Model of the Lactation Curve in Cattle. *Nature*, *216*(5111), 164–165. https://doi.org/10.1038/216164a0

Woolpert, M. E., Dann, H. M., Cotanch, K. W., Melilli, C., Chase, L. E., Grant, R. J., & Barbano, D. M. (2017). Management practices, physically effective fiber, and ether extract are related to bulk tank milk de novo fatty acid concentration on Holstein dairy farms. *Journal of Dairy Science*, *100*(6), 5097–5106. https://doi.org/10.3168/jds.2016-12046

Wurtz, K., Camerlink, I., D'Eath, R. B., Fernández, A. P., Norton, T., Steibel, J., & Siegford, J. (2019). Recording behaviour of indoor-housed farm animals automatically using machine vision technology: A systematic review. *PLOS ONE*, *14*(12). https://doi.org/10.1371/journal.pone.0226669

Zambelis, A., Wolfe, T., & Vasseur, E. (2019). Technical note: Validation of an ear-tag accelerometer to identify feeding and activity behaviors of tiestall-housed dairy cattle. *Journal of Dairy Science*, *102*(5), 4536–4540. https://doi.org/10.3168/jds.2018-15766

Zebari, H. M., Rutter, S. M., & Bleach, E. C. L. (2018). Characterizing changes in activity and feeding behaviour of lactating dairy cows during behavioural and silent oestrus. *Applied Animal Behaviour Science*, *206*, 12–17. https://doi.org/10.1016/j.applanim.2018.06.002