Lactation curve modelling in dairy production

Applications at cow and herd level

Yongyan Chen

Lactation curve modelling in dairy production

Applications at cow and herd level

Yongyan Chen

2024

Lactation curve modelling in dairy production

Applications at cow and herd level

Yongyan Chen

PhD thesis, Faculty of Veterinary Medicine, Utrecht University, the Netherlands - with summaries in English, Dutch and Chinese

ISBN: 978-90-393-7643-0

DOI: https://doi.org/10.33540/2148

Cover and layout design: Yongyan Chen

Printing: Proefschriftmaken | www.proefschriftmaken.nl

Lactation curve modelling in dairy production

Applications at cow and herd level

Het modelleren van lactatiecurves in de melkveehouderij

Toepassingen bij koe en kudde

(met een samenvatting in het Nederlands)

泌乳曲线模型

在奶牛个体和牧场方向的应用

（中文摘要）

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

maandag 26 Februari 2024 des middags te 4.15 uur

door

Yongyan Chen

geboren op 15 februari 1995
te Guangdong, China

Promotor:

Prof. dr. M. Nielen

Copromotoren:

Dr. W. Steeneveld
Dr. M.M. Hostens

Beoordelingscommissie:

Prof. dr. ir. A. de Vries
Prof. dr. ir. H. Hogeveen
Prof. dr. T.J.G.M. Lam
Prof. dr. H. Soyeurt
Dr. ir. A.T.M. van Knegsel

This thesis was accomplished with financial support from the Oversea Study Program of the Guangzhou Elite Project under grant agreement No. JY201937.

Dit proefschrift werd mede mogelijk gemaakt met financiële steun van het Oversea Study Program van het Guangzhou Elite Project onder subsidieovereenkomst nr. JY201937.

Contents

Chapter 1

General Introduction

1.1. Dairy industry development

Dairy products are rich sources of essential nutrients like calcium, protein, and vitamins, which are important for maintaining good health (FAO, 2013; Guetouache et al., 2014; Pereira, 2014). The increasing global population has increased the demand for dairy products, prompting growth and innovation in the dairy industry. Among the many major dairy-producing nations, the Netherlands was the 4th biggest suppliers of milk on global markets in 2022, contributing 8.2% of the total milk exports (Workman, 2023). The Dutch dairy sector experienced changes in total milk production and number of cows from 2000 to 2022 (**Figure 1.1**). Notably, between 2000 and 2016, both total milk production and the cow population surged by 28% and 16% respectively. However, this growth trend reversed between 2016 and 2018, with both indicators declining by 3% and 7%, respectively. This contraction was attributed to the introduction of a new manure policy to limit phosphate emissions (Klootwijk et al., 2016; Jongeneel et al., 2017). Subsequently, the dairy industry stabilized, resulting in 1.571 million cows producing a total of 13.8 billion kg of milk in 2022.

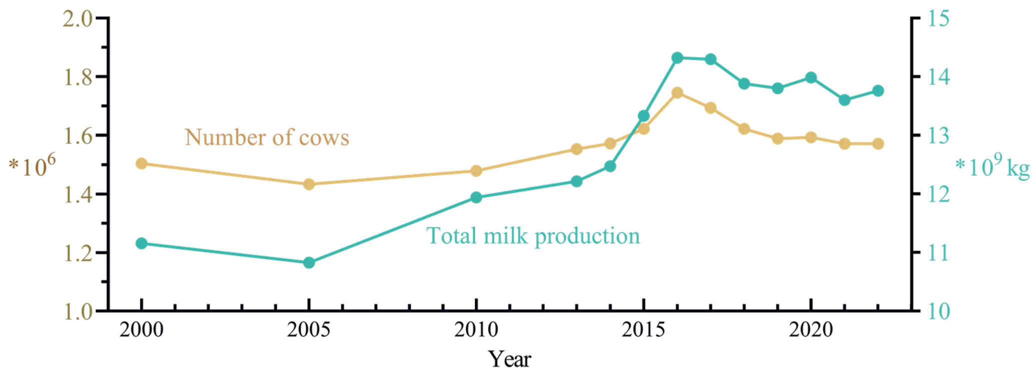


Figure 1.1 Overview of number of cows and total milk production in the Netherlands from 2000 to 2022 (adapted from annual report of CRV, 2022)

Behind this growing production scene, there has been a noticeable trend of increasing herd sizes accompanied by a reduction in the number of herds (**Figure 1.2**). Large farms are characterized by the investment in modernization and technology-driven solutions to replace expensive labour, and therefore becoming more competitive in terms of supply chain management, quality control and cost effectiveness (Oleggini et al., 2001; Gargiulo et al., 2018). Importantly, the average milk production per cow per year experienced a notable 14% increase between 2005 and 2018, remaining stable thereafter. This upward trend reflects the increased productivity of dairy cows.

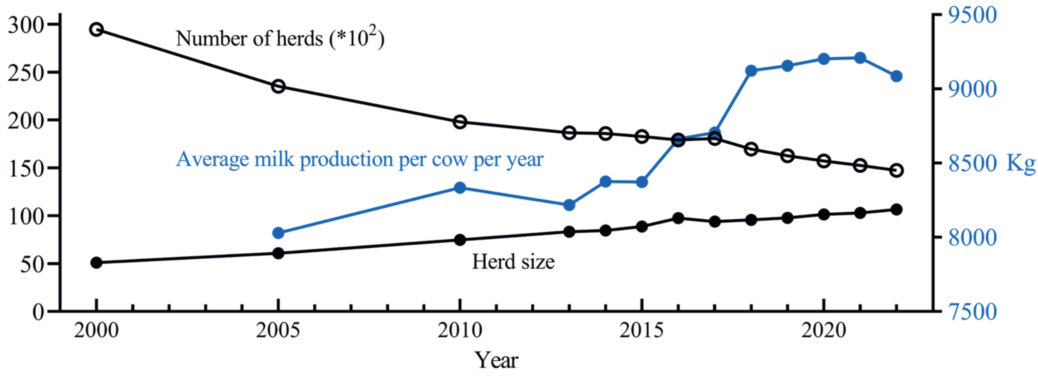


Figure 1.2 Overview of number of herds, herd size and average milk production per cow per year in the Netherlands from 2000 to 2022 (adapted from annual report of CRV, 2022)

The Netherlands has been a leader in adopting advanced technologies for dairy farming, such as robotic milking systems, automated feeding systems, data recording and data analytics (Rutten et al., 2013; Steeneveld and Hogeveen, 2015) with an increasing percentage of cows and herds undergoing milk recording over the past two decades (**Figure 1.3**), providing us with insights into quantity and quality of milk production. This recording system serves as a tool for frequent monitoring of milk production, thereby opening the possibility for better management in health, reproduction, feed and nutrition (Rutten et al., 2017; Deng et al., 2020; Hut et al., 2021).

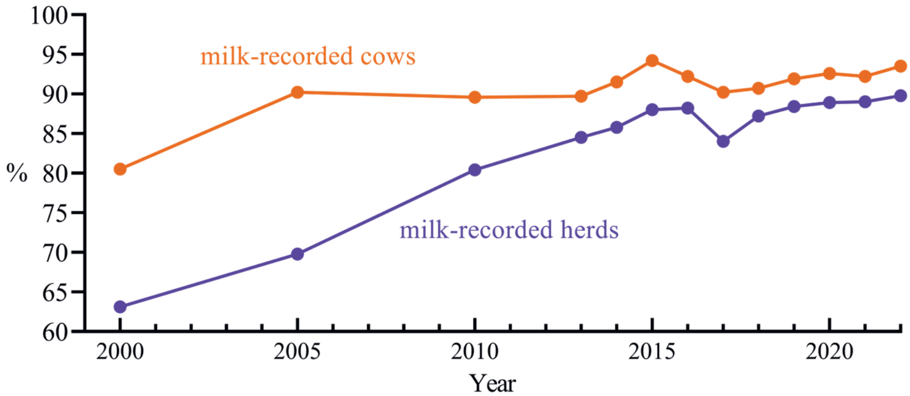


Figure 1.3 Overview of the percentage of milk-recorded cows and herd in the Netherlands from 2000 to 2022 (adapted from annual report of CRV, 2022)

1.1.1. Metrics to evaluate milk production

As milk production becomes increasingly accessible to record, the question arises: how can we effectively evaluate milk production performance and capture the difference between individual cows? Various metrics have been proposed to evaluate milk production, like cumulative milk production over a specific period (e.g., 305 days, 365 days, or an entire lactation) and milk yield per day within a certain period (such as lactating days or calving intervals) (Cole et al., 2012; Atashi et al., 2020; Burgers et al., 2021). These metrics are based on simple calculations on raw data. However, they only provide an overview of milk production performance, without capturing changes in milk production over the entire lactation period. These changes, or patterns in milk production offer more information (e.g., peak yield, peak time, persistency) about the lactation, which can be useful for breeding and selection, health monitoring, and other applications. For instance, persistency can be used to describe the cow's ability to maintain a slow rate of decline in production after the peak (Wood, 1967). Previous studies showed that persistent cows can have positive effects in a herd such as improved milk yield, improved conception rates, extended productive lifetimes or reduced culling rates (Dekkers et al., 1998; Hadley et al., 2006; Togashi et al., 2016). Therefore, farmers are eager to identify and selectively breed cows with high persistency (Dekkers et al., 1996; Cole and VanRaden, 2006; Togashi and Lin, 2009).

Changes in milk production over the entire lactation period can be visualized through lactation curves, which depict the relationship between milk yields and days in milk during lactation (Wood, 1974; Ehrlich, 2011). For example, consider two cows with the same 305d milk yield (**M305**). Their paths to achieving the same M305 can differ (**Figure 1.4**). Cow A displays a lower peak and a slower rate of milk production decline after the peak, in contrast to cow B. Potentially, cow A can be considered superior to cow B for several reasons. In the early lactation, cow A produce less milk and therefore it will experience less stress from negative energy balance (Butler, 2005; Wathes et al., 2007b; Lehmann et al., 2016). Moreover, cow A's ability to maintain a stable milk production rate makes her a strong candidate for extended lactation strategies, further optimizing milk production efficiency (Sorensen et al., 2008a; Kok et al., 2019; Sehested et al., 2019). However, evaluation by traditional metrics might fail to capture this difference between these two cows and therefore other metrics are needed. This is where lactation curve modelling becomes essential.

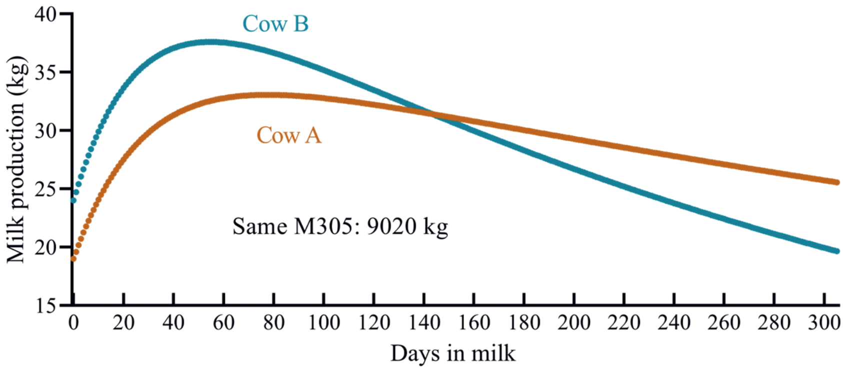


Figure 1.4 Lactation curves of two cows with the same 305d milk production (M305)

1.2. Lactation curve modeling

A lactation curve model is a mathematical representation used to describe the shape of the lactation curve based on measured milk production data. It can extrapolate and quantify a lactation curve and estimate actual production from incomplete data sets.

1.2.1. Difference between models

Various mathematical equation models have been proposed to fit the lactation data. According to their curve-fitting algorithms, different empirical models derive various lactation curve characteristics (**LCC**), quantifying the shape of the curve in different ways (**Table 1**). The first lactation model was introduced in 1923 (Brody et al., 1923). Initially, this model employed an exponential function to depict the declining phase of the lactation curve. A year later, the model was improved by incorporating the modelling of the ascending phase (Brody et al., 1924). While this pioneering effort marked a significant step, it was noted that the model tended to underestimate milk yield in the middle of lactation and overestimate the milk yield at the end of lactation (Cobby and Le Du, 1978). Nonetheless, it served as an important foundation for lactation curve modelling. Later, the Sikka model (Sikka, 1950) was built which provided a better fit for primiparous cows than multiparous cows (Gahlot et al., 1988). Building upon the Brody model, the Fischer model (Fischer, 1958) was introduced. However, this model demonstrated shortcomings, including underestimating maximum milk yield and estimating the peak date relatively early (Rowlands et al., 1982). To improve the fitting, Nelder (1966) proposed the inverse function to overcome the disadvantages of the ordinary polynomial (unbounded, built-in symmetry). On the other hand, Wood (1967) proposed an adjustment encompassing the entire curve through an incomplete gamma-type function, recognizing the limitations of the aforementioned exponential models in accurately capturing the ascending phase of the lactation curve. This classic Wood model emerged as the most widely recognized lactation curve model and served as an inspiration for subsequent improvements and innovations (Cobby and Le Du, 1978; Dhanoa, 1981; Jenkins and Ferrell, 1984). While these innovative models enhanced the Wood model in various ways, they also exhibited certain limitations. Among these improved models, the Wilmink model garnered notable recognition and adoption (Wilmink, 1987). The Wilmink model can make adjustment for herd, age at calving and the stage of lactation to improve prediction (Wilmink, 1987). However, it's important to note that the performance of the Wilmink model may be compromised when the first test day occurs after DIM 60 (Silvestre et al., 2006). Consequently, the effectiveness of the Wilmink model in making accurate predictions is significantly influenced by the interval between calving and the first test day. However, the previously mentioned models have been commonly used and have generally met people's expectations. In recent years, there has been a growing interest in extending lactations. While the other models has limited ability to describe the shape of the lactation curve beyond 305 days (Vargas et al., 2000; VanRaden et al., 2006; Dematawewa et al., 2007), the MilkBot model adjusts to extended lactations (Ehrlich, 2011). This model also offers greater flexibility for accounting for the influence of diseases and management practices, potentially leading to more accurate daily milk yield estimates by incorporating prior information (Ehrlich, 2011).According to the curve-fitting algorithms, different models derive various LCC, quantifying the shape of the curve in different ways. For example, the classic Wood model consists of the scale a (representing the level of production), the ramp b (representing the rising rate of milk to the peak production level) and the declining slope c (Wood, 1967). In MilkBot, scale and ramp are similar to the Wood model. The other characteristics are the time of maximal creation of productive capacity (offset, c) and the loss of productive capacity (decay, d), which can be easily transformed into a measure of persistency using the formula $Persistency = \frac{0.693}{decay}$ (Ehrlich, 2013). Persistency from Milkbot refers to the number of days it takes for the milk production to decrease by half during the declining stage of lactation. It can be thought of as the "half-life" of milk production. For instance, if a cow has a persistency of 300 days and reaches its peak yield of 40 kg at DIM 100, it means that this cow will attain a milk yield of 20kg at DIM 400.

Table 1. A list of empirical lactation curve models developed with their formula and main features.

Model	**Formula ($Y(t) =$)[1]**	**Features[2]**
Brody et al (1923)	ae^{-k1t}	Only depict the declining phase.
Brody et al (1924)	$ae^{-k1t} - be^{-k2t}$	Incorporating the modelling of the ascending phase.
Sikka (1950)	ae^{bt-ct^2}	The parabolic exponential function is effective for primiparous cows but less suitable for multiparous cows.
Fischer (1958)	$a - bt - ae^{-ct}$	Improve the Brody model by replacing the exponential decline by a linear decline .
Nelder (1966)	$\frac{t}{a + bt + ct^2}$	The inverse function is non-negative, bounded, and has no built-in symmetry.
Wood (1967)	$at^b e^{-ct}$	Incomplete gamma function; most popular; generates the standard form of the lactation curve; the direct relation of its parameters with the main elements of the shape of the lactation curve.
Cobby and Le Du (1978)	$a - bt - ae^{-ct}$	Declines in milk production are modelled exponentially; Allow a better adjustment of the initial phase of the lactation curve with a good estimate of peak production.
Dhanoa (1980)	$at^{bc} e^{-ct}$	Less correlations between characteristics; characteristic b can directly estimate the peak lactation date.
Wilmink (1987)	$a + be^{-kt} - ct$	Combines exponential and linear decline function; It can make adjustment of test-day milk, fat and protein yield for other factors such as age and season.
Milkbot (Ehrlich, 2011)	$a\left(1 - \frac{e^{\frac{c-t}{b}}}{2}\right)e^{-dt}$	The model is derived from a theoretical-mechanistic hypothesis; characteristics can be interpreted from two perspectives: their impact on the curve's shape and their association with mechanistic hypotheses.

[1] Y(t) is the milk yield; a, b, c and d are lactation curve characteristics that define the shape of the lactation curve; t is the days in milk; k1, k2 and k are the fixed value derived from a preliminary analysis of average production for different breed or parity.
[2] Main features are summarized by the original reference introducing the model and one review paper (Bouallègue and M'hamdi, 2019).

1.2.2. Lactation curve modeling at cow level

LCC serve as a crucial metric for evaluating milk production performance at the cow level and have diverse applications in various dairy research fields. They are intensely used in research related to feed composition and feeding systems (Chen et al., 2016; van Hoeij et al., 2017; Różańska-Zawieja et al., 2021). For instance, the effect of dietary energy source and dietary energy level on LCC were studied (van Hoeij et al., 2017). Moreover, LCC serve as a tool for identifying cows with a specific lactational phenotype (Yamazaki et al., 2011a; Ehrlich, 2013). For example, it's suggested that cows exhibiting high daily milk yield and long milking intervals are more efficient and thus suited for being milked with an automated milking system (Masía et al., 2020). Additionally, LCC can characterize perturbations of milk and offer insights about how cows respond to challenges during the lactation (Abdelkrim et al., 2021; Adriaens et al., 2021; Wang et al., 2022). For disease detection, lactation curve analysis can significantly contribute to the assessment of both short- and long-term effects of metabolic diseases on milk production (Yamazaki et al., 2009; Hostens et al., 2012; Masia et al., 2022). For calculation of economic disease impacts, lactation curve analysis are used to determine milk production losses due to disease (Steeneveld et al., 2007, 2011; Andersen et al., 2011). Disease-related impacts on milk production and economic impact might not have become apparent when total milk production alone (like M305) was analysed. For reproduction management, some studies have established associations between LCC and age at first calving (Elahi Torshizi, 2016; Atashi et al., 2021). Similarly, explorations have been conducted into the potential associations between LCC and parameters such as lactation length and dry period length (Atashi et al., 2013; Chen et al., 2016). Some researchers have utilized LCC to calculate cow persistency, allowing for comprehensive investigations into how persistency relates to reproduction (Inchaisri et al., 2011a; van Hoeij et al., 2017; Atashi et al., 2020). For instance, persistency metrics derived from LCC aids in the determination of the economically optimal voluntary waiting period (**VWP**) for first insemination (Inchaisri et al., 2011a).

Moreover, persistency metrics derived from the LCC of the Wilmink lactation model have gained recognition and significance as a critical breeding trait since 2001 in the Netherlands (personal communication with Gerben de Jong from the Cattle Improvement Cooperative CRV).

Beyond the aforementioned research and applications, some important research topics remained with little attention.

First, the shape of the lactation curve has been used as an argument to extend the lactation in dairy cows (Sorensen et al., 2008b; Lehmann et al., 2016; Burgers et al., 2021). Cows with flatter lactation curves, referred to as high persistent cows, can yield economic benefits when their lactation is extended (Dekkers et al., 1998; De Vries, 2006), which resulted in lower feed costs, fewer calvings, lower incidence of postpartum metabolic diseases and thus reduced veterinary costs (Van Amburgh et al., 1997; Kok et al., 2019; Lehmann et al., 2019a). Maintaining milk production in late lactation is a prerequisite for extended lactation (Stefanon et al., 2002; De Vries, 2006; Niozas et al., 2019). Therefore, when deciding on the VWP of an individual cow, it is useful to be aware of the persistency for the remainder of that lactation, especially for farmers who consider persistency in their reproduction management. Predictions of persistency for the current lactation could thus provide additional information to optimize the VWP. Currently, breeding values for persistency are calculated for dairy cows (Cole and VanRaden, 2006a; Togashi and Lin, 2008; Cole and Null, 2009). However, no studies have focused on predicting the lactation persistency and identifying cows eligible for an extended lactation based on readily available cow and herd data. Lactation curve modelling could allow for the prediction of persistency at any future timepoint (i.e. insemination moment) within the lactation period.

Furthermore, deciding on an optimal VWP for individual cows within a herd is complex because of a multitude of cow and management factors and becomes even more complicated when taking into account the direct effect of pregnancy on the level of milk production (Yamazaki et al., 2016; Lu and Bovenhuis, 2020). Compared to non-pregnant cows, pregnant cows experience an additional decline in milk yield during the gestation period (Roche, 2003; Bohmanova et al., 2009; Loker et al., 2009). The pregnancy effect is initially limited but becomes more pronounced later during the pregnancy due to increased nutrient requirements for the growing foetus (Loker et al., 2009) and hormonal changes, leading to an increased regression of the mammary gland (Bormann et al., 2002; Zhao et al., 2019). A significant decline in the milk yield of pregnant cows has been reported from months 4 or 5 of gestation onwards (Roche, 2003; Penasa et al., 2016). While the effect of days post conception on an absolute milk yield reduction has been reported in several studies (Bohmanova et al., 2009; Loker et al., 2009), the association between days post conception and lactation persistency based on a lactation curve model has not yet been quantified, i.e., if and how the persistency changes during pregnancy and whether or not these changes are related to the days in milk at conception. A better understanding of how the days post conception, and hence the moment of insemination during lactation, influence lactation persistency, enabling farmers to make more informed, evidence-based decisions when taking persistency into account in their reproduction management.

1.2.3. Lactation curve modeling within animal health economics

At its core, a dairy herd is a business and profitability is an essential part of a sustainable dairy farm (van Calker, 2005). A dairy herd has several variable cost factors, such as feed costs, labour costs and veterinary costs. Meanwhile, the revenues from a dairy farm originate mainly from milk production, and other revenues include the sale of cows, calves and surplus forage. A dairy farm can be characterized by several management areas, such as grassland management, youngstock management, and cow management. The latter one includes optimizing milk production, and making decisions on health and reproduction. Within the research field of animal health economics (AHE), theories, concepts, procedures and methodologies are developed to support the decision-making process for improving animal health and reproduction (McInerney et al., 1992; Dijkhuizen and Morris, 1996; Jonathan Rushton, 2009). In recent years several studies were performed to gain economic insights into health (Shim et al., 2004; Steeneveld et al., 2007; van den Borne et al., 2010) and reproduction management of dairy cows (Inchaisri et al., 2011b; Burgers et al., 2021; Vredenberg et al., 2021). The reproduction studies have focused on topics such as optimizing VWP and calving interval, and the timing of ceasing insemination (Sørensen and Østergaard, 2003; Inchaisri et al., 2011b, 2012). As it is generally known that the shape of lactation curve has a direct impact on both milk production and, consequently, milk revenue (e.g., Togashi and Lin, 2009; Němečková et al., 2015), AHE studies on reproduction management took into account the shape of the lactation curve, and thus included LCC in their decision support models. This approach provides opportunities for more informed decision making on reproductive decisions for individual cows.

Besides the cow-level decision support, within AHE also herd-level studies are performed focusing on health and reproduction factors associated with the profitability of herds (Vredenberg et al., 2021; Walsh et al., 2021; Yue et al., 2022). As it is generally known that individual cows with a persistent milk production are more profitable (e.g., Dekkers et al., 1998; Němečková et al., 2015), it might be that herds with relatively more persistent cows are associated with higher profitability. Therefore, a herd lactation curve is needed to reflect the lactation curves of all individual cows in the herd. As previously mentioned, lactation curves can differ between herds, since the environment and management of a dairy herd influence their shapes (Val-Arreola et al., 2004; Ehrlich, 2013). To develop herd lactation curves, a method to summarize individual lactation curves into a group lactation curve is needed. Previously, aggregating lactation curves has been a common approach, with primarily two ways applied. First, within a group (e.g., parity), either the average milk yield of cows at each DIM (VanRaden et al., 2006; Ehrlich, 2011) or the average milk yield and average DIM within a specific interval of the lactation were calculated (Vargas et al., 2000; Pietersma et al., 2001). Subsequently, the aggregated data for each group would be fitted to a lactation curve model. Alternatively, the pooled milk production data from all cows within a group was fitted as if the entire milk production came from one cow (Scott et al., 1996; Val-Arreola et al., 2004; Dematawewa et al., 2007). Both ways are relatively easy and straightforward but limit the ability to explore differences in cow level lactation curve characteristics within the group.

Economic evaluations of herd lactation curve characteristics **(HLCC)** need a valid aggregation of cow lactation curve characteristics as economic data are generally expressed at a calendar year basis. Such a herd lactation curve needs to be summarized on a year basis as well. This will be challenging as individual cow lactation curves often belong to multiple calendar years. Aggregating methods from cow to herd level lactation curve on a calendar year basis have not previously been described. The annual HLCC might open possibilities to explore economic differences between herds. In previous studies, the herd 305-d milk production (**HM305**) was often used to describe the herd's production performance (Pinedo et al., 2010; Nor et al., 2014; Shahid et al., 2015), and used in economic analyses as well (Ferguson et al., 2000; Green et al., 2002; Ferguson and Skidmore, 2013). However, it is unknown whether the HLCC is better or worse than the absolute volume of milk production at explaining the economic variation between herds.

1.3. Scope and outline of this thesis

The scope of this thesis was to explore the application of lactation curve modelling based on farm data collected on commercial dairy farms in the Netherlands and Belgium (**Figure 1.5**).

To achieve this overall scope, four objectives were formulated:
1. to predict lactation persistency for DIM 305 at different insemination moments
2. to investigate the association between days post conception and persistency

3. to summarize cow lactation curves into HLCC and illustrate a field application of HLCC
4. to compare whether HLCC or HM305 is better able to explain herd economic performance.

Milkbot lactation model was used throughout these applications for several reasons. Firstly, the MilkBot model is capable of accurately modelling extended lactations (Ehrlich, 2011). Other models, such as the Wood function, have limited capacity to accurately describe the shape of the lactation curve beyond DIM 305 (Dekkers et al., 1998; Bouallègue and M'hamdi, 2019). Secondly, the MilkBot model utilizes Bayesian statistics to provide a consistent fitting of individual cow lactation data. Even in cases where the data is sparse and noisy, the incorporation of prior information (i.e., the population mean LCC) can provide an LCC estimate.

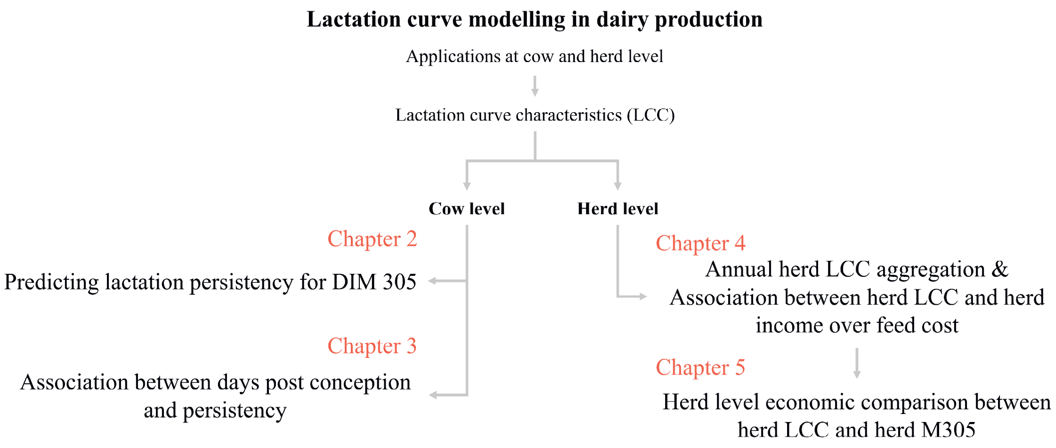


Figure 1.5 The application of lactation curve modelling in dairy production in the thesis

The outline of the thesis is visualized in **Figure 1.5**, and consisted of two kind of applications. The first application is at cow level (**Chapter 2 and 3**). **Chapter 2** aimed to predict lactation persistency for DIM 305 at different insemination moments (DIM 50, 75, 100 and 125). **Chapter 3** presented the association between days post conception and persistency, with an additional focus on the potential influence of DIM at conception on persistency. The second application is at herd level (**Chapter 4 and 5**). **Chapter 4** presented a procedure to summarize cow lactation curves into HLCC on a calendar year basis. Subsequently, a field application of HLCC is illustrated and this includes the association between HLCC and IOFC. In **Chapter 5**, we further investigated the application by determining whether HLCC or HM305 is better at explaining the variation in economic performance between herds.

Finally, in **Chapter 6**, the main results of this thesis are discussed. Additionally, the datasets used, cow and herd level clustering issues, the definition of persistency and the generalizability of our study results to China are discussed.

2

Chapter 2

Prediction of persistency for day 305 of lactation at the moment of the insemination decision

Yongyan Chen

Wilma Steeneveld

Mirjam Nielen

Miel Hostens

Published in Frontiers in Veterinary Science, doi:10.3389/fvets.2023.1264048.

Abstract

When deciding on the voluntary waiting period of an individual cow, it might be useful to have insight into the persistency for the remainder of that lactation at the moment of the insemination decision, especially for farmers who consider persistency in their reproduction management. Currently, breeding values for persistency are calculated for dairy cows but, to our knowledge, prediction models to accurately predict persistency at different moments of insemination are lacking. This study aimed to predict lactation persistency for DIM 305 at different insemination moments (DIM 50, 75, 100 and 125). Available cow and herd level data from 2005–2022 were collected for a total of 16,980 cows from 84 herds located in the Netherlands and Belgium. Lactation curve characteristics were estimated for every daily record using the data up to and including that day. Persistency was defined as the number of days it takes for the milk production to decrease by half during the declining stage of lactation, and calculated from the estimated lactation curve characteristic ‘decay’. Four linear regression models for each of the selected insemination moment were built separately to predict decay at DIM 305 (**decay-305**). Independent variables included the lactation curve characteristics at the selected insemination moment, daily milk yield, age, calving season, parity group and other herd variables. The average decay-305 of primiparous cows was lower than that of multiparous cows ($1.55 *10^{-3}$ vs $2.41*10^{-3}$, equivalent to a persistency of 447 vs 288 days, respectively). Results showed that our models had limitations in accurately predicting persistency, although predictions improved slightly at later insemination moments, with R^2 values ranging between 0.27 and 0.41. It can thus be concluded that, based only on cow and herd milk production information, accurate prediction of persistency for DIM 305 is not feasible.

Keywords

persistency, dairy, prediction model, milk production, insemination moment

2.1. Introduction

Traditionally, 12 to 13 months has been considered to be the economically optimal calving interval for dairy cows (Sørensen and Østergaard, 2003; Inchaisri et al., 2011b). Such a calving interval can maximize milk yield per cow per year, making use of peak production at the beginning of every lactation (M.J. et al., 2007; Kok et al., 2019). However, whether this yearly calving interval is the most optimal choice for every cow is now being questioned in the literature. First, cows can suffer from a negative energy balance at early lactation, especially high-producing cows (Butler, 2005; Kawashima et al., 2012). Subsequent conception rates might therefore be low as cows may not have recovered from the metabolic problems caused by this negative energy balance (Ingvartsen et al., 2003; LeBlanc et al., 2006). Second, a yearly calving interval can result in cows being dried off with a relatively high milk yield at the end of the lactation. This has been described as a risk factor for poor udder health in subsequent lactations (Rajala-Schultz et al., 2005; Odensten et al., 2007). Third, a yearly calving interval might be an indication for more metabolic disease treatments per year (Burgers et al., 2022). More costs (labour, veterinarian and insemination) may then be incurred and the cow's health, welfare and lifespan may be impaired (Bertulat et al., 2013; Zobel et al., 2015).

Extending lactation has been proposed as a solution to solve the above-mentioned issues. By extending lactation, farmers deliberately delay the first insemination moment. Several advantages of extended lactation have been identified (Österman and Bertilsson, 2003; Niozas et al., 2019; van Knegsel et al., 2022). Extended lactation could benefit cow health and production efficiency due to fewer transition periods in the lifespan of the cow. Extending the voluntary waiting period (**VWP**) for some cows has resulted in higher milk yield per day of calving interval (Arbel et al., 2001; Österman and Bertilsson, 2003; Burgers et al., 2021). In addition, extending the VWP can lower milk yield during the last six weeks before dry-off and benefit udder health in the following dry period and the next lactation (Rajala-

Schultz et al., 2005; Niozas et al., 2019; Burgers et al., 2021). Other advantages of extending lactation are that it may reduce greenhouse gas emissions per kg of milk produced, increase profitability and improve cow welfare (Wall et al., 2012; Lehmann et al., 2014; Browne et al., 2015). However, not all cows are suitable for extended lactation and the optimal VWP may vary per cow. It is therefore important to select the right cow for extended lactation (Lehmann et al., 2017; Sehested et al., 2019).

Maintaining milk production in late lactation is a prerequisite for extended lactation (Stefanon et al., 2002; De Vries, 2006; Niozas et al., 2019). Persistent cows decrease their milk yield at a lower rate after the peak day, resulting in a flatter lactation curve than non-persistent cows. Persistency is one of the factors that affect body condition scores at the end of the lactation, thus avoiding the risk of parturition diseases after the subsequent calving (Roche et al., 2007; Pires et al., 2013). From an economic perspective, extending lactation of persistent cows could increase the net partial cash flow at herd level (Kok et al., 2019). Extended lactations will be more beneficial, especially in herds with more persistent cows (Steeneveld and Hogeveen, 2012). Definitions of lactation persistency differ between previous studies. Persistency was defined as the milk yield difference at selected DIMs or the declining slope of milk yield within selected intervals after peak yield (Togashi et al., 2016; Grayaa et al., 2019; Burgers et al., 2021). Persistency can also be determined by using lactation curve models which quantify the lactation curve based on all available milk yield data (Wood, 1974; Ehrlich, 2011). One of the lactation curve characteristics that defines the curve is the decay, a lactation curve characteristic that can easily be transformed into other measures of persistency as the number of days it takes to halve milk production in the declining stage of lactation (Ehrlich, 2011).

When deciding on the VWP of an individual cow, it is useful to be aware of the persistency for the remainder of that lactation, especially for farmers who consider persistency in their reproduction management. Predictions of persistency for the current lactation could thus provide additional information to optimize the VWP. Currently, breeding values for persistency are calculated for dairy cows (Cole and VanRaden, 2006a; Togashi and Lin, 2008; Cole and Null, 2009) but, to our knowledge, prediction models to accurately predict persistency at the moment of insemination are lacking.

This study aims to determine whether it is possible to predict lactation persistency for DIM 305 at different insemination moments (DIM 50, 75, 100 and 125) based on available cow and herd data (excluding breeding values).

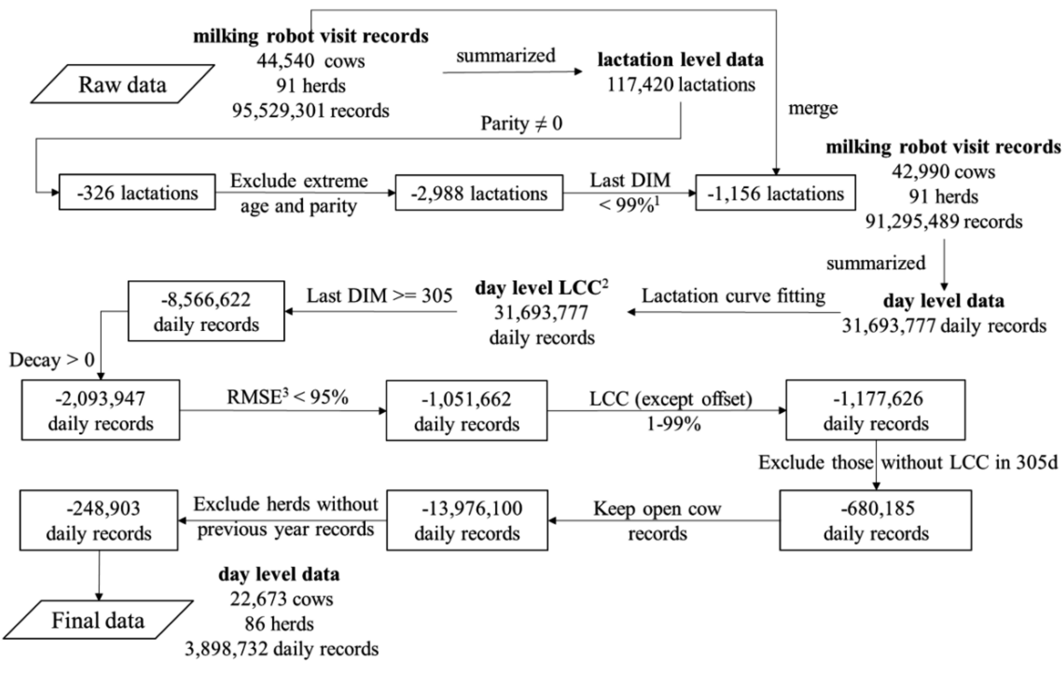
2.2. Materials and methods

2.2.1. Available data

Daily milk production and cow data were obtained for the years 2005–2022 from the MmmooOgle programme (Puurs, Belgium). Originally, the dataset included 95,529,301 milking robot visit records for 44,540 cows in 91 herds located throughout the Netherlands and Belgium. Milking robot visit records refer to detailed records generated by automated milking robots during milking of cows. All robot visits included general cow information (e.g., birth date, calving date, age in days and parity) and milk yield (kg). The number of lactating cows per herd varied between 26 (1%) and 394 (99%) per year, with a mean of 174 cows.

2.2.2. Preliminary data editing

The data editing diagram is shown in **Figure 2.1**. All exact calculations are shown in the Github repository mentioned at the end of this section. First, the milking robot visit records were summarized into 117,420 lactations from 44,540 cows. Subsequently, 326 lactations without parity information were excluded. Percentiles of age in days were calculated within every parity and 2,988 lactations with extreme age in days per parity (>99% percentile or <1% percentiles) were excluded. In addition, 1,156 lactations with extremely long lactation lengths (>99% percentile) were excluded. Applying these lactation level filters resulted in 91,295,489 milking robot visit records for 42,990 cows in 91 herds. Subsequently, a method (Lazenby et al., 2003) from the International Committee for Animal Recording (**ICAR**) was used to calculate a 24-hour milk yield using the 12 previous milkings for every milking robot visit record. The 24-hour milk yield of the last milking robot visit record on a given day was considered as the daily milk yield for that specific day. Afterwards, 31,693,777 daily records were summarized in 112,949 lactations. Among these, 34,646 lactations were from primiparous cows while 78,303 lactations were from multiparous cows.



[1] %: the percentile.
[2] LCC: lactation curve characteristics (magnitude, time to peak yield, offset and decay).
[3] RMSE: root mean squared error of the lactation curve fitting.

Figure 2.1 Diagram on data editing of the dataset on milk production per visit to an automatic milking system. The numbers in the boxes represent the excluded numbers.

2.2.3. Lactation curve modelling

A lactation curve was fitted for each daily record using the MilkBot model (Ehrlich, 2011) through the MilkBot lactation API (https://api.milkbot.com/). No records were dropped during the fitting process. The full MilkBot equation is shown as:

$$Y(t) = a\left(1 - \frac{e^{\frac{c-t}{b}}}{2}\right)e^{-dt} \quad (1)$$

where Y(t) is the estimated milk production when DIM is t, and scale a, ramp b, offset c and decay d are the lactation curve characteristics (**LCC**) describing the lactation curve. LCC are estimated for every daily record by fitting a lactation curve using the data up to and including that day. For example, LCC at DIM 50 are estimated after a lactation curve was fitted for the daily milk records up to and including DIM 50. Based on Bayesian statistics, the specific population mean lactation curve characteristics were used as prior information, the priors were previously adjusted to the population of Dutch dairy farms (Chen et al., 2022a). The prior was used to a greater extent when the fitted lactation had fewer daily records.

In the current study, the a (scale) was renamed magnitude of milk production (in kg/day) and the b (ramp) was renamed time to peak yield (in days). The d (decay) was transformed into a measure of persistency (in days) using the equation (Ehrlich, 2011):

$$Persistency = \frac{0.693}{d} \quad (2)$$

Persistency refers to the number of days it takes for the daily milk production to decrease by half during the declining stage of lactation. It can be thought of as the "half-life" of milk production. For instance, if a cow has a persistency of 300 days and reaches its peak yield of 40 kg at DIM 100, it means that this cow will attain a milk yield of 20kg at DIM 400.

The 305-day milk production (**M305**, in kg) can be estimated using the equation:

$$\mathrm{M305} = \frac{(a - ae^{-305d})}{d} + \frac{\left(abe^{\frac{c}{b}}\right)\times\left(-1+e^{-305\left(\frac{1}{b}+d\right)}\right)}{2+2bd} \quad (3)$$

2.2.4.Further data editing

After fitting the lactation curve model, 31,693,777 daily records with LCC from 42,990 cows in 91 herds remained. Daily records with LCC from lactations ending before DIM 305 (n = 8,566,622) were excluded because they did not have LCC for DIM 305. Daily records with negative decay (n = 2,093,947) and an extremely bad fitting (root mean squared error (**RMSE**) of lactation curve fitting >95% percentile, n = 1,051,662) were also excluded, as were extreme values for magnitude, time to peak yield and decay (>99% percentile or <1% percentiles, n = 1,177,626). In cases where lactations did not have LCC at DIM 305, LCC at DIM 304 was used as a substitute. This was determined based on the 90th percentile of the closest day to DIM 305. Following this, daily records from lactations without LCC at DIM 305 or 304 were excluded (n = 680,185). For every lactation, the calculated conception date was calculated by subtracting 282 days (Fitzgerald et al., 2015; Zamorano-Algandar et al., 2021) from the subsequent calving date. If no subsequent calving date was present, the breeding status was defined as ‘Never’. The breeding status was defined as ‘Bred’ if the calculated conception date was earlier than the date of the daily record; in all other cases the breeding status was defined as ‘Open’. Only daily records with an ‘Open’ breeding status were further included (excluding ‘Bred’ and ‘Never’ daily records, n = 13,976,100). To account for the herd effect, we aggregated herd level lactation curve characteristics (**HLCC** - herd magnitude, herd time to peak yield, herd offset and herd decay) and herd average 305-day milk production (**HM305**) from the previous year data, following the method described by Chen et al. (Chen et al., 2022a). In short, we aggregated individual lactations to the calendar year in which the lactation ended. Since LCC differ between primiparous cows and multiparous cows (Wood, 1969; Horan et al., 2005; Ehrlich, 2013), we divided herd lactations into two parity groups: primiparous cows and multiparous cows. HLCC was then calculated as the mean of the LCC per parity group per herd for each calendar year, while HM305 was calculated as the mean of M305 per herd for each calendar year. Daily records from lactations without HLCC and HM305 from the previous year were excluded (n = 248,903). In addition, age in months was calculated from age in days. The calving season was defined based on the calving month (3-5: Spring; 6-8: Summer; 9-11: Autumn; 12-2: Winter) (Steeneveld et al., 2014; Rutten et al., 2016). Two parity groups were defined (primiparous cows and multiparous cows). This method resulted in final dataset with 3,898,732 daily records from 43,430 lactations, 22,673 cows and 86 herds.

From the final dataset of daily records with breeding status ‘Open’, we constructed four datasets. The dataset for DIM 50 included daily records at DIM 50 from cows that was not yet conceived at DIM 50. Likewise, datasets were constructed for DIM 75, 100, and 125, which were considered as potential insemination moments. For lactations where LCC was not available on the exact selected insemination moments, we selected the closest day within the 90th percentile of the corresponding DIM (48, 74, 98, 122). After this selection, we have 99,593 daily records from all selected insemination moments from 37,021 lactations, 20,508 cows and 85 herds.

2.2.5. Model building

The model building for each selected insemination moment was carried out separately (**Figure 2.2**). In every selected insemination moment, cow-parity records were randomly split into two parts; 80% for the training set and 20% for the test set (**Table 1**). The training set was used for model training and validation (10-fold cross-validation). The test set was used for model evaluation.

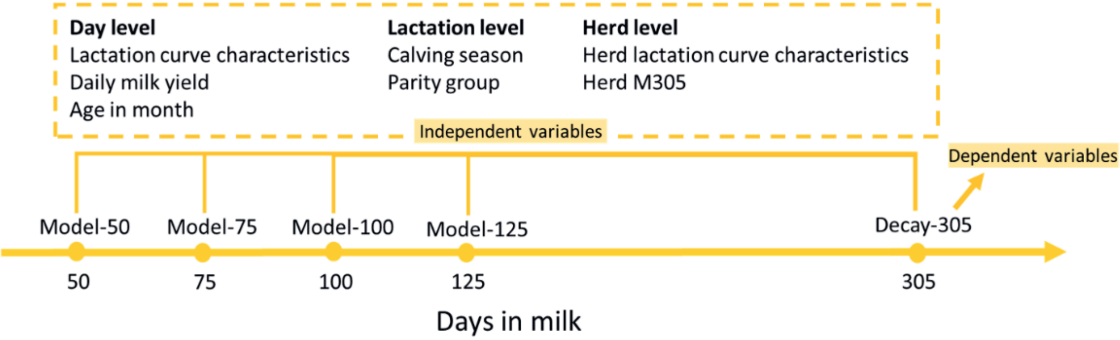


Figure 2.2 Diagram illustrating the model building procedure.

Table 1 Number of cow-parity records, cows and herds in training and test set for different selected insemination moments used for model training, validation and evaluation

Insemination moment (day)	Number of cow-parity records			Number of cows	Number of herds
	Training set	Test set	Total		
50	17,902	4,521	22,423	14,536	83
75	22,006	5,456	27,462	16,764	84
100	19,159	4,752	23,911	15,544	85
125	14,701	3,693	18,394	13,024	84

Due to the right-skewed distribution of persistency and the normal distribution of decay, decay was preferred for statistical analysis and converted to persistency afterwards for a more straightforward interpretation (Ehrlich, 2013). Decay at DIM 305 (**decay-305**) was therefore defined as the dependent variable. In total, four linear regression models for every selected insemination moment were built to predict decay-305 (**Figure 2.2**). The available details at every selected insemination moment were used as independent variables. These included the following cow level variables: LCC, daily milk yield (kg), age in months, calving season and parity group; and herd level variables: HLCC and HM305 from the year preceding the selected insemination moments. HLCC and HM305 were expected to explain herd variance since we could not add herd as the random effect in prediction models. To compare the strength of the effect of each independent variable to the dependent variable, we standardized all continuous independent variables. Funnel graphs were generated to visualize the ranking of the effect size for all continuous independent variables. To validate our method, we also used the same set of data and independent variables to predict M305 and assess the validity of our prediction model approach. The model is shown as:

$$y_{\mathrm{ijkl}} = \mu + LCC_i + Daily\ milk\ yield_i + Age_i + Calving\ season_j + Parity\ group_k + HLCC_l + HM305_l + \varepsilon_{\mathrm{ijkl}} \quad (4)$$

where y represents the dependent variables (decay-305 or M305), μ represents the overall mean, i represents the insemination moments (i = DIM 50, 75, 100 or 125), j represents the calving season class (j = spring, summer, autumn or winter), k represents the parity group class (k = primiparous cows or multiparous cows), l represents the previous year, and $\varepsilon_{\mathrm{ijkl}}$ represents the random residual term from a normal distribution.

2.2.6. Model evaluation

Model evaluation was carried out on test data with four metrics frequently used in similar research: coefficient of determination (**R^2**), RMSE, the mean absolute error (**MAE**) and the mean absolute percentage error (**MAPE**) (Liseune et al., 2020, 2021; Salamone et al., 2022). R^2 indicates the proportion of the variance of decay-305 explained by the independent variables. RMSE and MAE indicate the differences between predicted and observed decay-305, with MAE being less sensitive to extreme values in the prediction errors. MAPE measures how much the model's predictions deviate from the corresponding true value on average, ranging between 0 and 1. We used these four metrics to evaluate all decay prediction models while we only used R^2 and MAPE to evaluate all M305 prediction models, in order to compare them with the decay models. Data editing and analysis were carried out using the Python API for the Spark platform (**PySpark**). Visualization were conducted using GraphPad Prism version 8.0. Code scripts for the data editing steps and statistical analyses can be downloaded at (https://github.com/Bovi-analytics/Chen-et-al-2023a).

2.3. Results

Over all lactations the average M305 of primiparous cows (n = 11,562) varied between 6,253 (5%) and 11,390 (95%), with a mean of 8,809 kg. The average M305 of multiparous cows (n = 15,195) varied between 7,833 (5%) and 13,786 (95%), with a mean of 10,813 kg. The average decay-305 of primiparous cows was lower than that of multiparous cows ($1.55 *10^{-3}$ vs $2.41*10^{-3}$, equivalent to a persistency of 447 vs 288 days, respectively). Descriptive statistics for the independent variables are only shown for DIM 75 (**Table 2**); the statistics for the other insemination moments (DIM 50, 100 and 125) can be found in GitHub.

Table 2 Descriptive statistics of the dependent and independent variables used in the model predicting decay at DIM 305 (decay-305) at insemination moment DIM 75 based on milk production data from 16,764 cows in 84 Dutch and Belgium herds.

	Primiparous cows				Multiparous cows			
Variables	Mean	SD	5%[a]	95%	Mean	SD	5%	95%
Dependent variable								
Decay-305 ($*10^3$, day^{-1})	1.6 [b]	0.7	0.5	2.9	2.5	0.8	1.2	3.9
Independent variables [c]								
Cow level variables								
Magnitude (kg)	38.2	6.0	28.2	47.9	51.2	7.5	38.6	63.4
Time to peak yield (day)	28.2	2.4	24.1	31.9	21.0	3.8	13.3	26.4
Offset (day)	-0.50	$2.5*10^{-5}$	-0.50	-0.50	-0.53	0.36	-0.78	0.01
Decay ($*10^3$, day^{-1})	1.4 [b]	0.9	0.2	3.0	1.9 [b]	1.0	0.3	3.7
Daily milk yield (kg)	32.4	5.6	23.2	41.4	42.7	6.7	31.3	53.6
Age in months	28.2	2.5	25.2	33.1	56.6	18.0	38.0	91.9
Herd level variables								
Herd magnitude (kg)	37.0	3.5	30.9	42.5	49.9	3.8	43.6	56.0
Herd time to peak yield	28.5	1.2	26.7	30.6	21.6	1.3	20.1	23.6
Herd offset (day)	-0.50	$1.11*10^{-5}$	-0.50	-0.50	-0.54	0.09	-0.69	-0.40
Herd decay ($*10^3$, day^{-1})	1.5 [b]	0.4	0.9	2.1	2.2 [b]	0.3	1.7	2.7
Herd M305 (kg)	9,888	1,001	8,304	11,397	10,002	937	8,449	11,409

[a] 5% and 95%: the 5% and 95% percentile.
[b] Persistency was calculated based on the aforementioned decay using the equation 0.693/decay (Ehrlich, 2011). Persistency was not used in the prediction model because of non-normality; decay was used instead. A decay of 1.4, 1.5, 1.6, 1.9, 2.2 and 2.5 $*10^{-3}$ is equivalent to a persistency of 495, 462, 433, 365, 315 and 277 days, respectively.
[c] All values for the independent variables represent the value at DIM75. Herd variables were aggregated from the day level data of the previous year following the method described by Chen et al. (2022).

The model performance indicators of the prediction models for decay-305 at all selected insemination moments are summarized in **Table 3A**. Among all models, we found higher R^2 and lower RMSE, MAE and MAPE at later insemination moments. The R^2 of models for decay-305 range from 0.266 to 0.407, while RMSE, MAE and MAPE were slightly improved along the selected insemination moments.

Table 3A Model performance indicators[1] of prediction models on test set for decay at DIM 305 at different selected insemination moments (DIM 50, 75, 100 and 125).

Insemination moment (day)	R^2	RMSE	MAE	MAPE
50	0.266	$7.73*10^{-4}$	$6.16*10^{-4}$	0.391
75	0.270	$7.51*10^{-4}$	$5.98*10^{-4}$	0.400
100	0.325	$7.22*10^{-4}$	$5.72*10^{-4}$	0.371
125	0.407	$6.60*10^{-4}$	$5.22*10^{-4}$	0.370

[1] Model performance indicators: R^2: coefficient of determination; RMSE: root mean squared error; MAE: mean absolute error; MAPE: mean absolute percentage error.

Standardized coefficients of the model predicting decay-305 at all potential insemination moments are shown in **Figure 2.3**. Among all potential insemination moments, all variables had similar effects on the models. The three most influential variables affecting decay-305 were calving in autumn, daily milk yield and magnitude. However, the specific order of these variables varied across different models. Take model at DIM 75 for example, cows calving in autumn had on average 3.24 (SE=0.14) lower decay ($*10^4$) respectively than calving in winter. Increasing one unit of daily milk yield (7.90 kg) corresponded to an average 2.99 (SE=0.17) decrease in decay ($*10^4$). Increasing one unit of magnitude (9.22 kg/day) corresponded to an average 2.62 (SE=0.18) increase in decay ($*10^4$). **Table 3B** shows the results of the prediction models for M305, which showed much higher R^2.

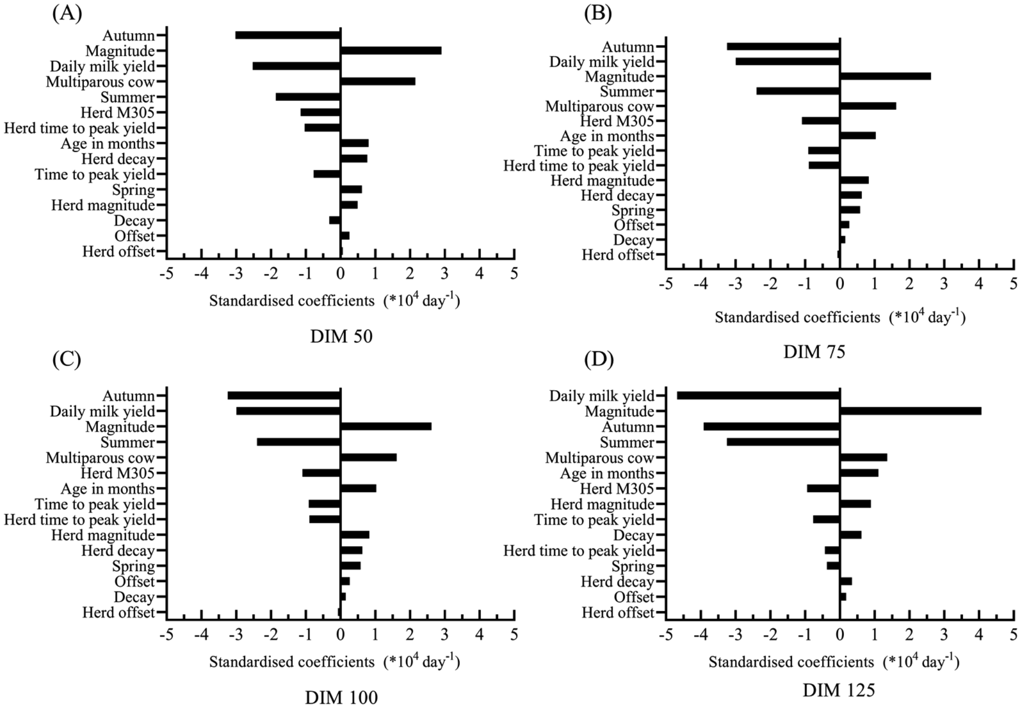


Figure 2.3 Standardized coefficients of the independent variables used to predict decay at DIM 305 at all potential insemination moments (DIM 50, 75, 100 and 125).

Table 3B Model performance indicators[1] of prediction models on test set for M305 at different selected insemination moments (DIM 50, 75, 100 and 125).

Insemination moment (day)	R^2	MAPE
50	0.785	0.073
75	0.850	0.061
100	0.889	0.051
125	0.921	0.043

[1] Model performance indicators: R^2: coefficient of determination; MAPE: mean absolute percentage error.

2.4. Discussion

This study aimed to predict lactation persistency for DIM 305 at insemination moments DIM 50, 75, 100 and 125. Our models have low prediction accuracy, although predictions improved at later insemination moments.

In our study, we used decay to measure persistency. The R^2 of all decay models was under 0.407, suggesting the bad predictive power of the model with all available information included (Olori et al., 1999). Using the same methodology, we were able to predict M305 much more accurately, thus confirming our prediction methodology to be valid for M305. Similar to previous studies (Gorgulu, 2012; Liseune et al., 2020, 2021), M305 was predictable at all insemination moments with R^2 values ranging between 0.79 and 0.92 for the different insemination moments. Other methodological approaches were explored to improve the prediction performance of the decay-305 models. First, we explored building prediction models for two parity groups separately. The results were similar (results shown in GitHub). Next to the linear regression, we built models using random forest, lasso regression and ridge regression but results were similar (results shown in GitHub). Models from lasso regression and ridge regression showed the same results to linear regression, indicating that penalization did not improve our models. In addition, adding LCC from the previous lactation did not improve the models in our study (results shown in GitHub).

In the current study, we only included cow and herd information in the prediction models that was available through the MmmooOgle herd management software. As persistency is a heritable trait and could be a target for selection (Cole and VanRaden, 2006a; Yamazaki et al., 2014; Torshizi et al., 2019) others have tried adding its breeding value to prediction models, though with little success (Kjeldsen et al., 2022). It's worth noting that the heritability of persistency varies, influenced by factors like the definition of persistency, the breed, and the parity of the cows, with heritability values spanning the range of 0.01 to 0.33 (Linde et al., 2000; Cole and Null, 2009; Torshizi et al., 2019). Breeding values were not available in our dataset. Persistency is furthermore influenced by feed management in herd (Sorensen et al., 2008a; Gaillard et al., 2016). We took into account this herd-level factor by including HLCC and HM305 into all of our models, rather than including herd as a random effect. This approach allows us to apply our prediction model to unknown farms and effectively consider the impact of herd-level factors on the study outcomes.

There is little existing literature to predict persistency for the mid-late lactation based on data from the beginning of lactation and herd information. We chose to predict persistency for DIM 305 because this is a classic time point for measurements like M305. Other studies chose to predict different parameters to help make insemination decisions. For example, Kjeldsen et al. predicted energy-corrected milk per day of calving interval at DIM 40 for primiparous and multiparous cows separately (Kjeldsen et al., 2022). They included the calving interval in the model while the future calving interval is actually unknown at the moment of making the insemination decision. We assumed that, in their research, predicting milk yield per day of calving interval was equivalent to predicting the milk yield. Another example, Manca et al. (Manca et al., 2020) used the threshold of daily milk yield at DIM 305 to determine whether a cow is persistent, and defining persistent cows as those with a daily milk yield at DIM 305 greater than 20 kg. Essentially, they used the lactation curve characteristics of the first DIM 90, 120, and 150 to predict the future daily milk yield at DIM 305. The results of Manca et al. (Manca et al., 2020) correspond with our results on predicting M305 as both achieved a high accuracy. It is important to note that persistency in our study primarily focuses on the slope or rate of decline in milk production over time. Consequently, persistency cannot be directly translated into the exact amount of milk that drops per day without knowledge of the initial peak milk production level. This consideration should be kept in mind when interpreting the findings and conclusions of this study.

Our prediction models could predict M305 well but could not predict persistency for DIM 305 accurately. We hypothesized that M305 is highly predictable due to its association with peak yield (Ehrlich, 2013; Atashi et al., 2020). Peak yield estimation was commonly established at our insemination moments from DIM 50 onwards (Kay et al., 2005; Peiter et al., 2021). In contrast, persistency was not highly correlated with information in early lactation. Additionally, the low prediction accuracy observed in our study may be attributed to other factors that influence persistency between the insemination moments and DIM 305. One potential factor that could impact persistency is pregnancy. However, we were unable to account for the pregnancy effect in our prediction model due to several reasons. Firstly, the exact timing of pregnancy is unknown at the time of making predictions for open cows. Secondly, the quantification of the pregnancy effect on persistency is lacking in previous studies, making it difficult to incorporate it into the model. As a result, we were unable to correct for the pregnancy effect in our prediction model.

There are multiple measures of persistency, and all these measures require the transformation from raw milk data (Togashi and Lin, 2009; Yamazaki et al., 2011b; Burgers et al., 2021). Simple measures of persistency are typically fixed at two time points in lactation (Togashi and Lin, 2009; Yamazaki et al., 2011b; Chen et al., 2016), limiting the ability to observe persistency changes throughout the lactation. To overcome this limitation, we employed lactation curve modelling using the MilkBot model, which allowed to assess persistency at any timepoint within the lactation period. This so called continuous measurement provides insights into the changes of persistency during lactation.

Our data were obtained from AMS farms and we therefore had access to milk production data for each robot visit. Such detailed data did not, however, result in high prediction values. The average M305 of the involved farms was higher than that of average dairy farms in the Netherlands and Belgium (Gelder, 2022a; b). Higher milk production can be explained by more frequent milking on AMS farms than on conventional farms (de Koning, 2010; Hogenboom et al., 2019). In our study, we deliberately only included cows with an 'Open' breeding status at the selected insemination moments. 'Open' was defined as cows which were not pregnant at the insemination moment but which could be pregnant in the future. Those open cows were the target object of our study since their insemination decisions were yet to be made. In our study, we only included lactations over 305 days, a period commonly accepted by the global standard for livestock data (ICAR, 2017).

2.5. Conclusion

Our results showed that based only on cow and herd milk production information, predicting persistency for DIM 305 at different insemination moments (DIM 50, 75, 100 and 125) is challenging. The accuracy of the predictions was found to be low in our models. In order to target decision-support at the insemination moment, other information is needed to improve the accuracy in predicting persistency.

Acknowledgements

We gratefully acknowledge the MmmooOgle programme (Puurs, Belgium) for providing the data. Author Yongyan Chen is financially supported by the Oversea Study Program of Guangzhou Elite Project.

Supplementary Figure

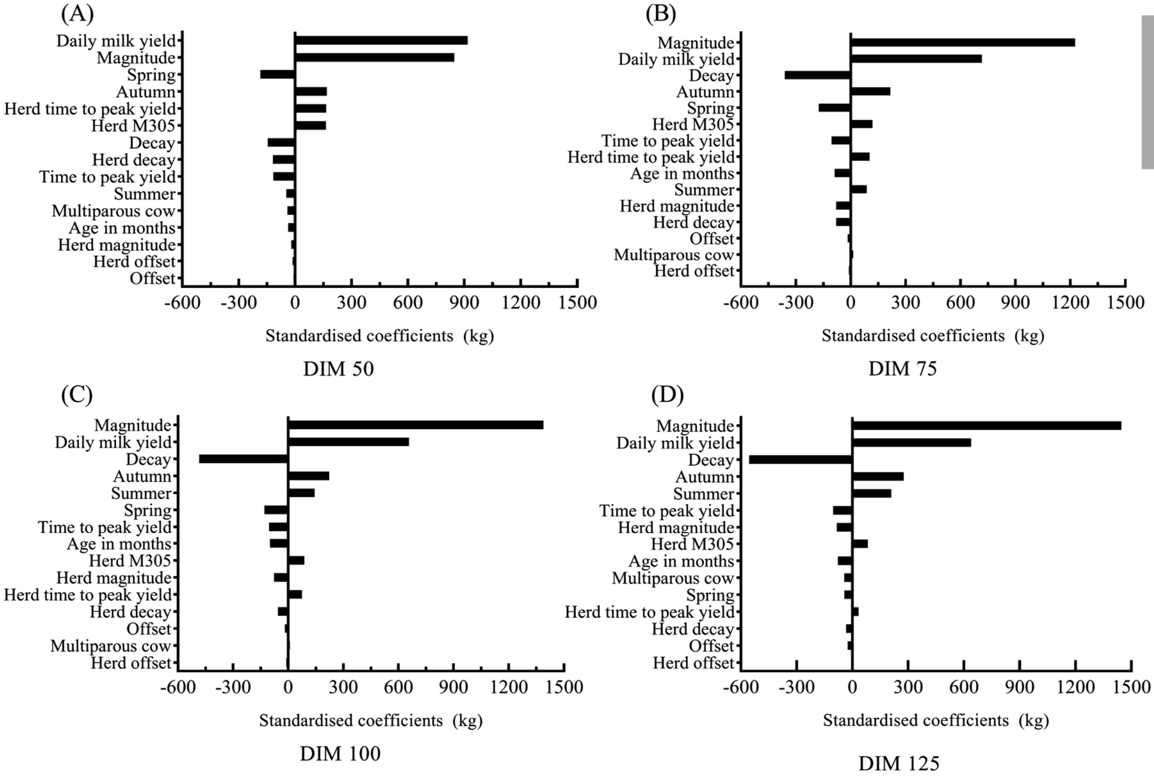


Figure 2.S1 Standardized coefficients of the independent variables used to predict M305 at DIM 305 at all potential insemination moments (DIM 50, 75, 100 and 125).

Chapter 3

Association between days post conception and lactation persistency in dairy cattle

Yongyan Chen

Wilma Steeneveld

Klaas Frankena

Imke Leemans

Hilde Aardema

Peter L.A.M. Vos

Mirjam Nielen

Miel Hostens

submitted, first revision

Abstract

Determining the optimal insemination moment for individual cows is complex, particularly when considering the impact of pregnancy on milk production. The effect of pregnancy on the absolute milk yield has already been reported in several studies. Currently, there is limited quantitative knowledge about the association between days post conception (**DPC**) and lactation persistency, based on a lactation curve model, and, specifically, how persistency changes during pregnancy and relates to the days in milk at conception (**DIMc**). Understanding this association might provide valuable insights to determine the optimal insemination moment. This study, therefore, aimed to investigate the association between DPC and lactation persistency, with an additional focus on the influence of DIMc. Available milk production data from 2005–2022 were available for 23,908 cows from 87 herds located throughout the Netherlands and Belgium. Persistency was measured by a lactation curve characteristic decay, representing the time taken to halve milk production after peak yield. Decay was calculated for eight DPC (0, 30, 60, 90, 120, 150, 180 and 210 days after DIMc) and served as the dependent variable. Independent variables included DPC, DIMc (<=60, 61-90, 91-120, 121-150, 151-180, 181-210, >210), parity group, DPC × parity group, DPC × DIMc and variables from 30 days before DIMc as covariates. The results showed an increase in decay, i.e., a decrease in persistency, during pregnancy for both parity groups, albeit in different ways. Specifically, from DPC 150 to DPC 210, multiparous cows showed a higher decline in persistency compared to primiparous cows. Furthermore, a later DIMc (cows conceiving later) was associated with higher persistency. Except for the early DIMc groups (DIMc<90), DIMc does not impact the change in persistency by gestation. The findings from this study contribute to a better understanding of how DPC and DIMc during lactation influence lactation persistency, enabling more informed decision-making by farmers who wish to take persistency into account in their reproduction management.

Keywords

Days post conception, lactation persistency, days in milk at conception, dairy

3.1. Introduction

Traditionally, 12 to 13 months is considered to be the economically optimal calving interval for dairy cows (Sørensen and Østergaard, 2003; Inchaisri et al., 2011b). Such a calving interval can optimize milk yield per cow per year (e.g., Auldist et al., 2007; Kok et al., 2019), making use of peak production at the beginning of every lactation. However, whether this yearly calving interval is the most optimal choice for every cow is now being questioned and different calving interval lengths, and thus different insemination moments during the lactation, have been studied (Lehmann et al., 2017; Burgers et al., 2021). The optimal insemination moment could be different for individual cows in a herd as it is influenced by multiple intrinsic and extrinsic factors (milk yield, parity, lactation persistency etc.,). For instance, the recommendation has been to inseminate high-producing cows later in lactation since no adverse effects on milk production, involuntary culling, udder health and body condition score were found (Niozas et al., 2019; Burgers et al., 2020).

Besides the level of milk production, the persistency of milk production during the lactation influences the optimal insemination moment (Inchaisri et al., 2011b; Burgers et al., 2021). Delaying the insemination moment for high persistent cows has been demonstrated to be profitable due to several the advantages it offers including reduced feed costs, decreased number of calvings, lower greenhouse gas emissions, decreased incidence of postpartum metabolic diseases and subsequently reduced veterinary expenses (Van Amburgh et al., 1997; Kok et al., 2019; Lehmann et al., 2019a). Persistency of milk production is the cow's ability to maintain a slow decline in milk production after peak production (Cole and VanRaden, 2006b; Togashi and Lin, 2009). However, the definition of persistency differs between studies (Togashi and Lin, 2009; Ehrlich, 2011; Burgers et al., 2021). Persistency has been defined as the milk yield difference at selected DIM timepoints on the lactation curve (Yamazaki et al., 2011b; Togashi et al., 2016) or as the declining slope of milk yield within selected intervals after peak milk yield (Chen et al., 2016; Burgers et al., 2021). Persistency has also frequently been defined using lactation curve models which quantify the slope of the lactation curve based on all available milk yield data, but persistency definitions differ between different lactation curve models (Wood, 1967; Wilmink, 1987; Ehrlich, 2011). For instance, in the MilkBot lactation curve model, as used in this study, persistency is expressed as the number of days it takes to halve the milk production in the declining stage of lactation (Ehrlich, 2011).

Deciding on an optimal insemination moment for individual cows within a herd is complex because of a multitude of cow and management factors and becomes even more complicated when taking into account the direct effect of pregnancy on the level of milk production (Yamazaki et al., 2016; Lu and Bovenhuis, 2020). Compared to non-pregnant cows, pregnant cows experience an additional decline in milk yield during the gestation period (Roche, 2003; Bohmanova et al., 2009; Loker et al., 2009). The pregnancy effect is initially limited but becomes more pronounced later during the pregnancy due to increased nutrient requirements for the growing foetus (Loker et al., 2009) and hormonal changes, leading to an increased regression of the mammary gland (Bormann et al., 2002; Zhao et al., 2019). A significant decline in the milk yield of pregnant cows has been reported from months 4 or 5 of gestation onwards (Roche, 2003; Penasa et al., 2016). For instance, cows pregnant from DIM 121 to 210 and those pregnant from DIM 211 to 310 produced 0.9 and 3.1 kg/d less milk than those pregnant from DIM 1 to 120, respectively (Penasa et al., 2016).

The effect of days post conception (**DPC**) on an absolute milk yield reduction has been reported in several studies (Bohmanova et al., 2009; Loker et al., 2009). To our knowledge, the association between DPC and lactation persistency based on a lactation curve model has not yet been quantified, i.e., if and how the persistency changes during pregnancy and whether or not these changes are related to the days in milk at conception (**DIMc**). Insight into this association might support the determination of the optimal insemination moment for farmers who wish to take persistency into account in their reproduction management. Therefore, the goal of this study was to investigate the association between DPC and lactation persistency, based on a large observational dataset. Additionally, we aimed to explore whether this relationship was influenced by the days in milk at conception.

3.2. Materials and methods

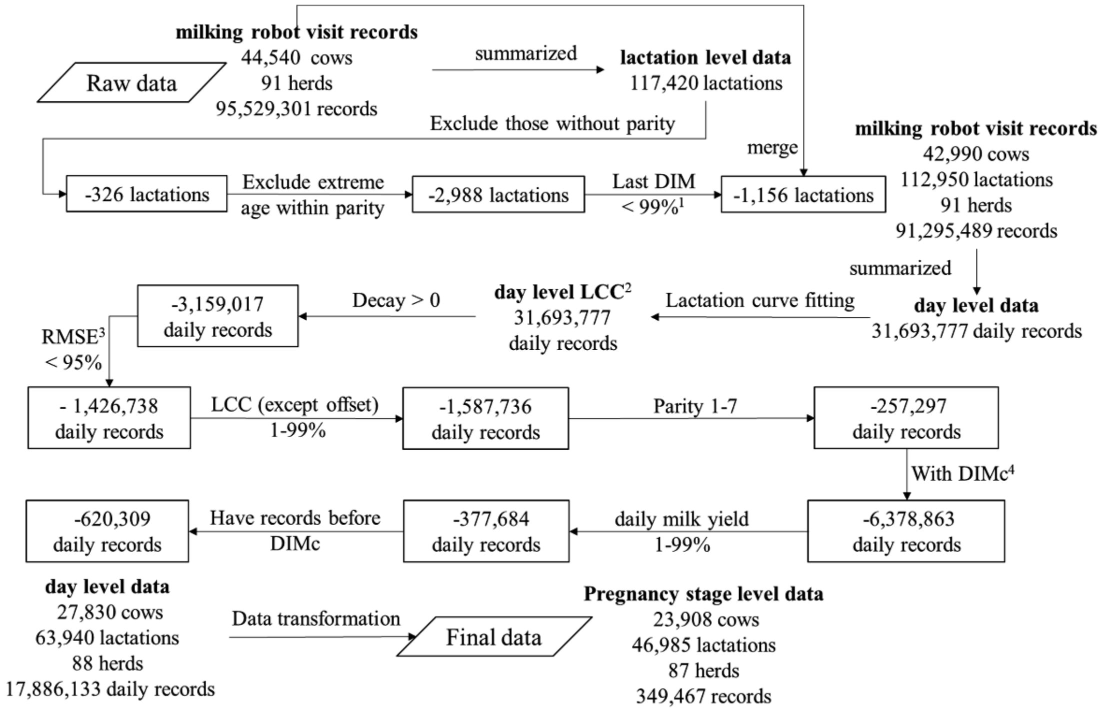
3.2.1. Available data

Anonymized daily milk production and cow data were obtained for the years 2005–2022 from the MmmooOgle programme (Puurs, Belgium). Originally, the dataset included 95,529,301 milking robot visit records for 44,540 cows (mainly Holstein-Friesian) in 91 herds located throughout the Netherlands and Belgium. All records included general cow information (e.g., birth date, calving date, age in days and parity) and milk yield. The number of lactating cows per herd per year varied between 71 (5% percentile) and 320 (95% percentile), with a mean of 174 cows (SD = 77, median = 163).

3.2.2. Data editing

The data editing diagram is shown in **Figure 3.1**. All exact calculations are shown in the Github repository mentioned at the end of this section. First, the milking robot visit records were summarized into 117,420 lactations from 44,540 cows. Subsequently, 326 lactations without parity information were excluded. Percentiles of age in days were calculated within every parity and 2,988 lactations with extreme age per parity (<1% percentile or >99% percentile) were excluded. In addition, 1,156 lactations with extremely long lactation lengths (>99% percentile) were also excluded. Applying these lactation level filters resulted in 91,295,489 milking robot visit records for 42,990 cows in 91 herds. Subsequently, a method (Lazenby et al., 2003) from the International Committee for Animal Recording (**ICAR**) was used to calculate a 24-hour milk yield using the 12 previous milkings for every milking robot visit record. The 24-hour milk yield on the last milking robot visit record on a given day was considered as the daily milk yield for that specific day. Afterwards, 31,693,777 daily records were summarized in 112,949 lactations. Out of these, 34,646 lactations belonged to primiparous cows, while 78,303 lactations belonged to multiparous cows.

The MilkBot lactation curve model (Ehrlich, 2011) was utilized to fit a lactation curve for each daily record, using the MilkBot lactation API (https://api.milkbot.com/). The model utilized prior information, specifically the population mean lactation curve characteristics, which were adjusted to the population of Dutch dairy farms based on a previous study (Chen et al., 2022a). Lactation curve characteristics (magnitude, time to peak yield, offset and decay) were estimated for every available daily record using the data up to and including that day. For example, lactation curve characteristics at DIM 50 were estimated after a lactation curve was fitted for the available daily milk records up to and including DIM 50. See supplementary material in Github for further description. No records were dropped during the fitting process. Persistency was calculated from the decay using the formula $Persistency = \frac{0.693}{decay}$ (Ehrlich, 2013), expressed as the number of days it takes for the milk production to decrease by half during the declining stage of lactation. It can be thought of as the "half-life" of milk production. For instance, if a cow has a persistency of 300 days and reaches its peak yield of 40kg at DIM100, it means that this cow will attain a milk yield of 20kg at DIM400. It is important to note that persistency in our study primarily focuses on the slope or rate of decline in milk production over time. Consequently, persistency cannot be directly translated into the exact amount of milk that drops per day without knowledge of the initial milk production level.



1 %: the percentile.
2 LCC: lactation curve characteristics (magnitude, time to peak yield, offset and decay).
3 RMSE: root mean squared error of the lactation curve fitting.
4 DIMc: days in milk at conception, calculated by subtracting 282 days from the subsequent calving date.

Figure 3.1 Diagram on data editing of the dataset on milk production per visit to the automatic milking system. The numbers in the boxes represent the numbers excluded.

Daily records with a negative decay (n = 3,159,017) or an extremely poor fit on the lactation curve (based on >95% percentile of the root mean squared error (**RMSE**), n = 1,426,738) were excluded, as were extreme values for magnitude, time to peak yield and decay (<1% percentile or >99% percentile, n = 1,587,736). The daily records of cows with parity above 7 were also excluded (n = 257,297). For every lactation, DIMc was calculated by subtracting 282 days (Fitzgerald et al., 2015; Zamorano-Algandar et al., 2021) from the subsequent calving date (**Figure 3.2**). Daily records of 40,026 lactations without a subsequent calving date were excluded (n = 6,378,863). Daily records with extreme daily milk yield were also excluded (<1%

percentile or >99% percentile, n = 377,684). All records of 5,223 lactations without records before DIMc were excluded (n = 620,309) as well. This data reduction procedure resulted in 17,886,133 daily records from 63,940 lactations, 27,830 cows and 88 herds. Of these, 22,135 lactations were from primiparous cows and 41,805 lactations were from multiparous cows.

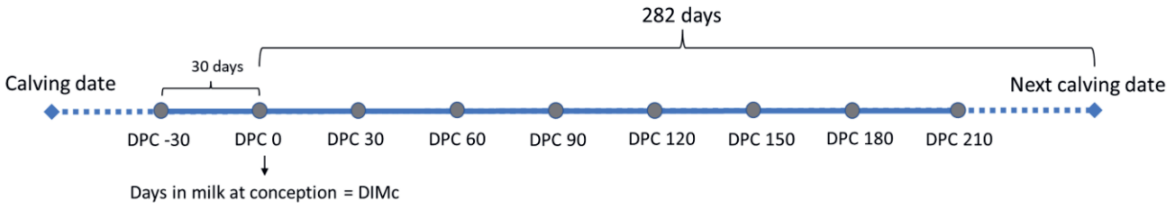


Figure 3.2 Illustration for selected days post conception (DPC) before and after conception defined at DIMc as 282 days before next calving date. Data from 46,985 lactations (15,328 for primiparous cows; 31,657 for multiparous cows), 23,908 cows and 87 herds located throughout the Netherlands and Belgium from 2005-2022.

3.2.3.Defining days post conception

For every lactation, we selected the decay at different DPC (-30, 0, 30, 60, 90, 120, 150, 180 and 210 days related to DIMc) (**Figure 3.2**). To calculate the decay at those selected DPC, we selected an interval of 10 days before and 10 days after. For example, decay at DPC 30 is the average of the decay from DPC 20 to 40 days (with a minimum requirement of at least 5 days of data available for calculation). The interval average decay calculation was used instead of the decay at a single day to exclude the effects of daily fluctuations in milk yield on persistency measures over time. This approach allowed for the inclusion of lactations with less frequent daily records, maximizing the dataset used for analysis. Peak yield and decay at 30 days before DIMc (peak yield-30) were calculated in the same way. The birth year of each cow was incorporated into the analysis. Cows born before 2000 were grouped into the '2000' category due to their limited numbers. Furthermore, cows born in 2019 were combined with those born in 2018 to increase the sample size. The calving age in months for each cow in every lactation was determined by calculating the difference between the calving date and the birth date. The number of lactations, cows and herds per DPC are shown in **Table 1**. Two parity groups were defined (primiparous cows and multiparous cows). The final dataset included 349,467 records from 46,985 lactations, 23,908 cows and 87 herds. Of these, 15,328 lactations were from primiparous cows and 31,657 lactations were from multiparous cows.

Table 1 Number of lactations, cows and herds available in the final dataset for different days post conception related to days in milk at conception (DIMc), based on data from 46,985 lactations (15,328 for primiparous cows; 31,657 for multiparous cows), 23,908 cows and 87 herds located throughout the Netherlands and Belgium from 2005-2022.

Days post conception	Number of lactations	Number of cows	Number of herds
-30	47,200	23,987	87
0	51,938	24,937	88
30	56,205	25,933	88
60	58,228	26,481	88
90	59,027	26,736	88
120	59,208	26,800	88
150	58,802	26,726	88
180	56,984	26,426	88
210	52,069	25,551	88

3.2.4. Statistical analysis

Due to the right-skewed distribution of persistency and the normal distribution of decay, decay was preferred for statistical analysis and converted to persistency afterwards for a more straightforward interpretation (Ehrlich, 2013). A linear mixed model was used to analyse the association between DPC and decay (dependent variable, multiplied by 1,000). Independent variables, based on an expected association with decay, were DPC, i.e., days after DIMc (0, 30, 60, 90, 120, 150, 180 and 210), parity group and DIMc. Decay at 30 days before DIMc (decay-30, multiplied by 1,000) and peak yield-30 were added as covariates to adjust for cow production and decay level before conception. The biologically relevant interaction terms DPC × parity group and DPC × DIMc were included. Herd, cow and birth year were entered into the model as random effects to account for unobserved herd-related (e.g., environment, feed management) and cow-related (e.g., individual characteristics, genetic improvement) heterogeneity. The linearity of the continuous predictors was assessed by adding their quadratic terms to the model; for all these predictors the linearity assumption was met as the quadratic terms were not significant ($P>0.05$). The conditional R^2 and the marginal R^2 were calculated to describe the variance explained by the entire model and the fixed effects, respectively. The normality of residuals was checked by a Q-Q plot. The interaction term was graphically presented for interpretation.

To visualize lactation curves for different DIMc in both parity groups, three conception groups (early, mean and late conception) were defined based on the population mean of DIMc (128 days) as well as one standard deviation below and above the mean value (63 and 193 days, respectively). For all conception and parity groups, the decay for each DPC was calculated using the estimated marginal means from the model. Other lactation curve characteristics (magnitude, time to peak yield and offset) were set at their population mean per parity group. The lactation curves were estimated for a duration of 240 days following conception.

Data editing, analysis and visualization were performed using Python API for the Spark platform (PySpark) and R version 3.6.3 (R Core Team, 2020), including R packages 'dplyr' (Wickham et al., 2023), 'lme4' (Bates et al., 2015), 'ggplot2' (Wickham, 2016), 'emmeans' (Lenth et al., 2020). Parts of the visualization were conducted using Excel and GraphPad Prism version 8.0. Code scripts for the data editing steps and statistical analyses can be downloaded from Github (https://github.com/Bovi-analytics/Chen-et-al-2023b).

3.3. Results

The distribution of DIMc for primiparous and multiparous cows is presented in **Figure 3.3**. Primiparous cows conceived on average earlier in lactation than multiparous cows (121 vs 133 days). In total, 47.5% of the primiparous cows and 37.5% of the multiparous cows conceived within 100 days of lactation. Primiparous cows had lower peak yield-30 (31.9 vs 43.3kg) than multiparous cows and showed a lower decay-30 (persistency of 517 vs 341 days) (**Table 2**). Over all DPC, primiparous cows had higher persistency than multiparous cows (median of 442 and 295 days, **Figure 3.4**).

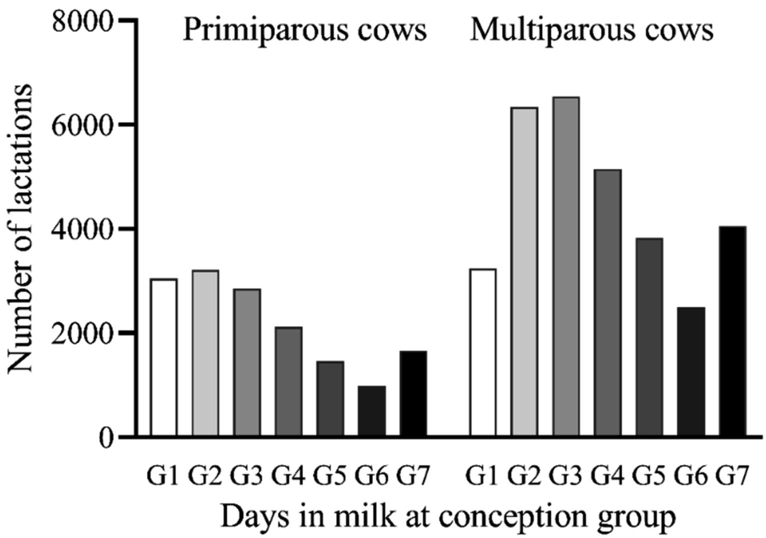


Figure 3.3 Frequency distribution of lactations of primiparous (n =15,328) and multiparous cows (n =31,657) over days in milk at conception groups in the final dataset of 23,908 cows and 87 herds. Days in milk at conception were categorized into seven groups: G1 (<=60), G2 (61-90), G3 (91-120), G4 (121-150), G5 (151-180), G6 (181-210), and G7 (>210). Data from 46,985 lactations (15,328 for primiparous cows; 31,657 for multiparous cows), 23,908 cows and 87 herds located throughout the Netherlands and Belgium from 2005-2022.

Table 2 Descriptive statistics of continuous variables based on data from 46,985 lactations (15,328 for primiparous cows; 31,657 for multiparous cows), 23,908 cows and 87 herds located throughout the Netherlands and Belgium from 2005-2022.

Variable	Parity group[1]	Mean	SD	Median	25%-75% IQR
Days in milk at conception (day)	P1	121	68	104	67-156
	P2+	133	66	119	83-168
Peak yield-30 (kg) [2]	P1	31.9	6.0	32.1	27.7-36.1
	P2+	43.3	7.0	43.8	39.1-48.2
Decay-30 ($*10^3$, day^{-1}) [3]	P1	1.34[4]	0.76	1.19	0.79-1.73
	P2+	2.03[4]	0.95	1.98	1.30-2.66

[1] P1: primiparous cows; P2+: multiparous cows.
[2] Peak yield-30: peak yield was defined based on the average peak yield estimations between 40 and 20 days before DIMc.
[3] Decay-30: average decay between 40 and 20 days before DIMc.
[4] A decay of $1.34*10^{-3}$ and $2.03*10^{-3}$ is equivalent to a persistency of 517 and 341 days, respectively.

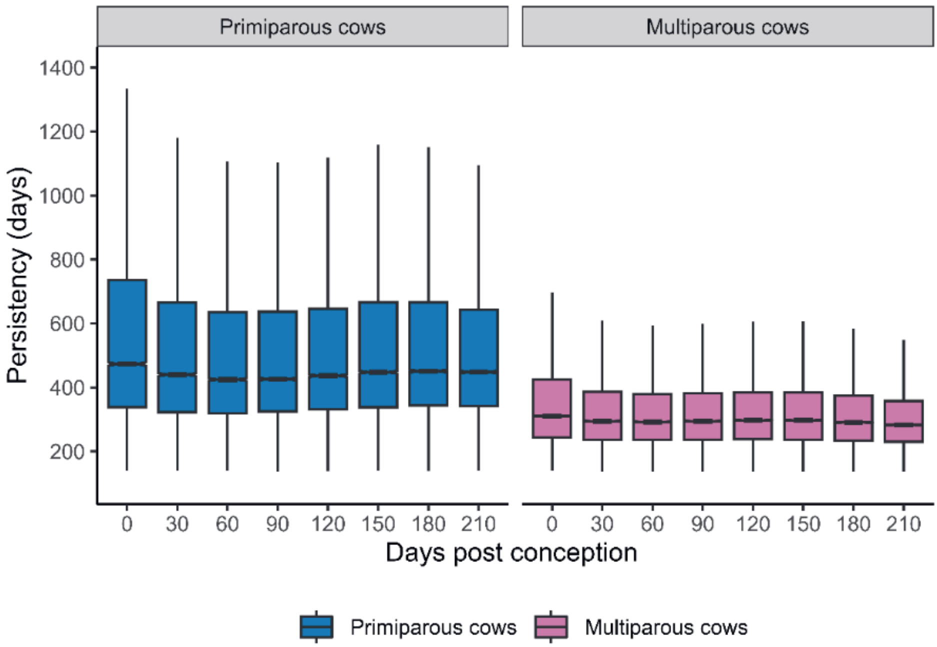


Figure 3.4. Distribution of persistency at different days post conception for primiparous and multiparous cows. The central mark is the median. The bottom and top edges of each box are the 25th (Q1) and 75th (Q3) percentiles. The whiskers extend to the extreme data points that are within the range of [Q1 – 1.5 × (Q3 – Q1), Q3 + 1.5 × (Q3 – Q1)]. Data from 46,985 lactations (15,328 for primiparous cows; 31,657 for multiparous cows), 23,908 cows and 87 herds located throughout the Netherlands and Belgium from 2005-2022.

The estimated associations between decay and the DPC are presented in **Table 3**. Since an interaction term is included, the interpretation of the other estimates is described together with the interaction (**Figure 3.5)**. In this model, the conditional R^2 and the marginal R^2 were 66.9 and 25.6%, respectively and model residuals followed a normal distribution.

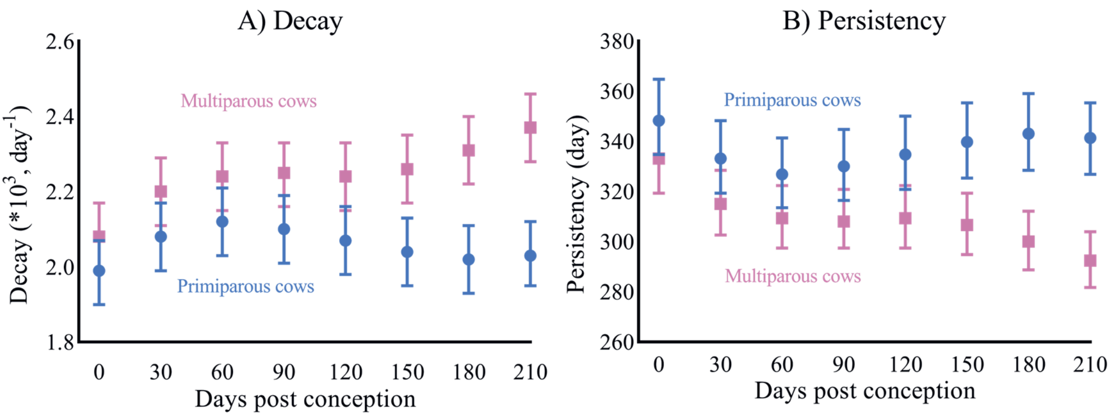


Figure 3.5 Visualization of interaction between days post conception and parity group on A) decay (*10^3) based on estimated marginal means from the model; B) persistency based on the direct transformation (0.693/decay) from decay in A). Data from 46,985 lactations (15,328 for primiparous cows; 31,657 for multiparous cows), 23,908 cows and 87 herds located throughout the Netherlands and Belgium from 2005-2022.

The interaction of DPC × parity group is shown in **Figure 3.5A** based on estimated marginal means derived from the model. For a more intuitive illustration, we interpreted the results using persistency instead of decay (**Figure 3.5B**). For primiparous cows, the persistency first decreased slightly from DPC 0 to 60 and gradually increased from DPC 60 to 180. For multiparous cows, the persistency gradually decreased from DPC 0 to 60, then remained stable up to DPC 150 and decreased again from DPC 150 to 210. During pregnancy, both primiparous and multiparous cows had a lower persistency at DPC 210 compared to that at DPC 0. Multiparous cows showed a higher decline in persistency compared to primiparous cows (a 12.2% decrease from 333 to 292 days and a 2.0% decrease from 348 to 341 days).

Table 3 Results of the linear mixed model on the association between decay (*10^3, day^{-1}) and days post conception (and other cow variables) based on data from 46,985 lactations, 23,908 cows and 87 herds located throughout the Netherlands and Belgium from 2005-2022.

Variable		β[1]	S.E.	P-value
Intercept		-0.238	0.0465	<0.001
Days post conception	0	Ref [2]		
	30	0.398	0.0124	<0.001
	60	0.655	0.0121	<0.001
	90	0.754	0.0120	<0.001
	120	0.771	0.0120	<0.001
	150	0.758	0.0120	<0.001
	180	0.759	0.0120	<0.001
	210	0.799	0.0121	<0.001
Days in milk at conception [3]	G1	Ref [2]		
	G2	0.445	0.0108	<0.001
	G3	0.534	0.0108	<0.001
	G4	0.530	0.0113	<0.001
	G5	0.472	0.0121	<0.001
	G6	0.434	0.0134	<0.001
	G7	0.359	0.0120	<0.001
Days post conception × parity group [4]	0 × $P2^+$	0.093	0.0068	<0.001
	30 × $P2^+$	0.030	0.0079	<0.001
	60 × $P2^+$	0.032	0.0079	<0.001
	90 × $P2^+$	0.050	0.0079	<0.001
	120 × $P2^+$	0.078	0.0079	<0.001
	150 × $P2^+$	0.125	0.0079	<0.001
	180 × $P2^+$	0.192	0.0080	<0.001
	210 × $P2^+$	0.245	0.0081	<0.001
Days post conception × Days in milk at conception				<0.001
Calving age in month		-0.001	0.0001	<0.001
Peak yield-30 [5]		0.033	0.0003	<0.001
Decay-30 [6]		0.305	0.0016	<0.001

[1] β: corrected for interaction Days post conception × Days in milk at conception (not shown).
[2] Ref: used as a reference category.
[3] Days in milk at conception were categorized into seven groups: G1 (<=60), G2 (61-90), G3 (91-120), G4 (121-150), G5 (151-180), G6 (181-210), and G7 (>210).
[4] Parity group: P1: primiparous cows; P2+: multiparous cows. The interaction term "days post conception × parity group" indicates the difference in mean decay between the two parity groups (P1 vs P2+) within each days post conception.
[5] Peak yield-30: peak yield was defined based on the average peak yield estimations between 40 and 20 days before DIMc.
[6] Decay-30: average decay between 40 and 20 days before DIMc.

The interaction of DPC × DIMc is shown in **Figure 3.6A** based on estimated marginal means derived from the model. For a more intuitive illustration, we interpreted the results using persistency instead of decay (**Figure 3.6B**). Across all DIMc groups, there was a marginal change in persistency on average (with a maximum of 14 days) from DPC 0 to DPC 210, except for G1 and G2, which experienced a 36% and 14% decrease (152 and 47 days), respectively. Specifically, cows in G1 experienced an average decrease of 123 days (29%) from DPC 0 to DPC 60, remaining stable thereafter. In G2, cows experienced an average decrease of 40 days (12%) from DPC 0 to DPC 30, and maintained stability thereafter.

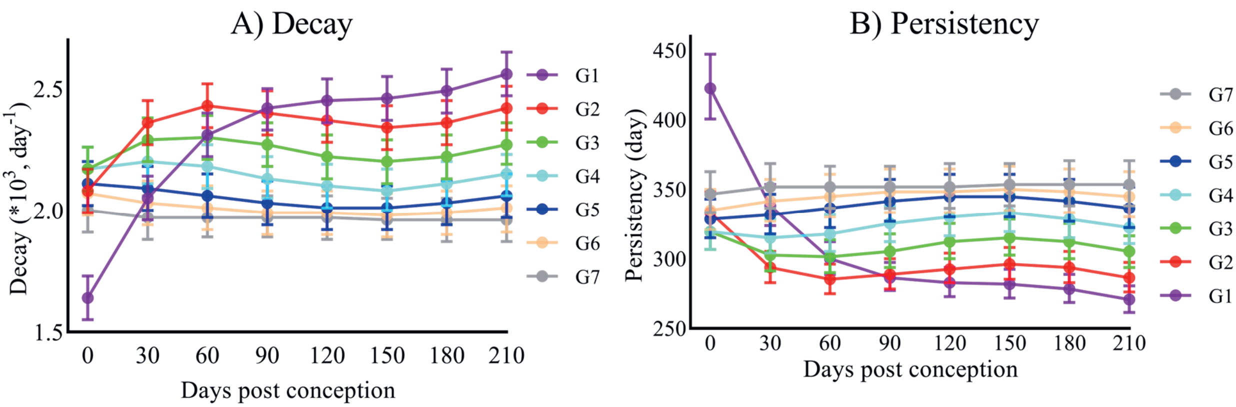


Figure 3.6. Visualization of interaction between days post conception and days in milk at conception on A) decay (*10^3) based on estimated marginal means from the model; B) persistency based on the direct transformation (0.693/decay) from decay in A). Days in milk at conception were categorized into seven groups: G1 (<=60), G2 (61-90), G3 (91-120), G4 (121-150), G5 (151-180), G6 (181-210), and G7 (>210). Data from 46,985 lactations (15,328 for primiparous cows; 31,657 for multiparous cows), 23,908 cows and 87 herds located throughout the Netherlands and Belgium from 2005-2022.

For both parity groups, a later DIMc was associated with higher persistency. Specifically, compared to G1, primiparous cows in G2 to G7 DIMc groups corresponded to an average of 18 to 100 additional days of persistency at DPC 210, reflecting a 6% to 34% increase. Similarly, compared to G1, multiparous cows in G2 to G7 DIMc groups corresponded to an average of 12 to 62 additional days of persistency at DPC 210, indicating a 5% to 26% increase. **Figure 3.7** presents the visualization of lactation curves for different DIMc groups (G2, G3 and G5). In both parity groups, a later DIMc corresponded to higher total milk production (as indicated by the area under the curve) and a lower dry-off milk yield.

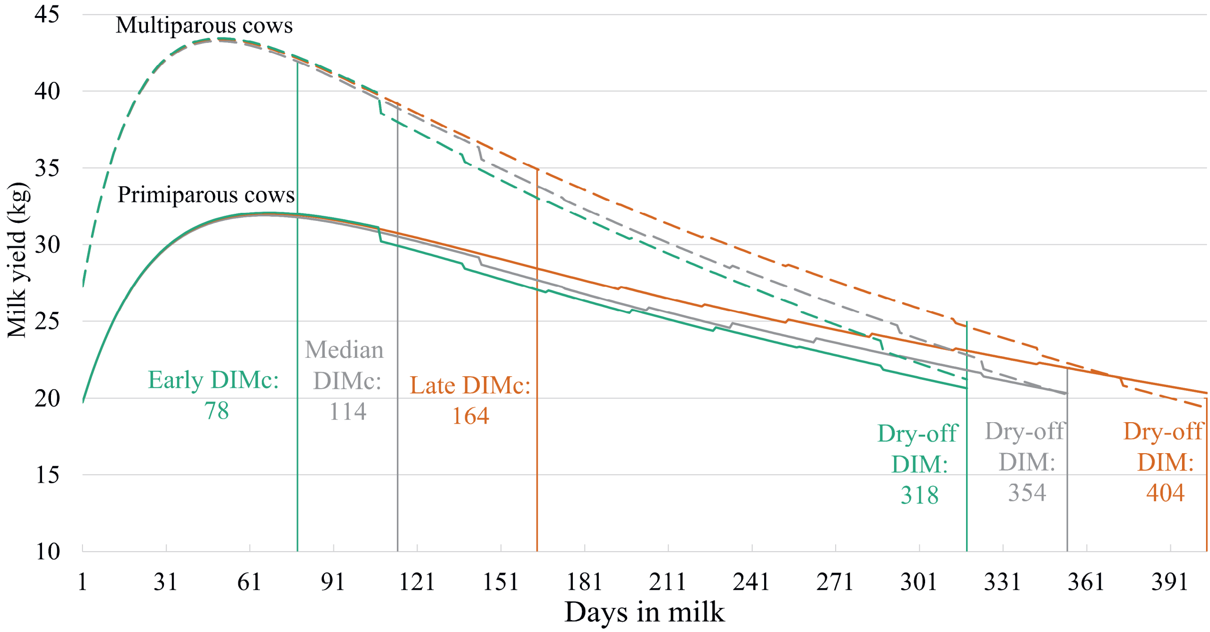


Figure 3.7. Lactation curve visualization of days post conception × parity group included in the model. Decay was based on the estimated marginal means from the model. Other lactation curve characteristics (magnitude, time to peak yield and offset) were set at their population mean per parity group. The conception groups are determined based on the first quartile, median, and third quartile of DIMc values (78, 114 and 164). The vertical line is the days in milk at conception. Each lactation curve is visualized for 240 days following conception. Data from 46,985 lactations (15,328 for primiparous cows; 31,657 for multiparous cows), 23,908 cows and 87 herds located throughout the Netherlands and Belgium from 2005-2022.

3.4. Discussion

This study aimed to investigate the association between DPC and lactation persistency, additionally exploring whether this relationship is influenced by the moment of conception (DIMc) during lactation. This association was quantified by a linear mixed model. The results showed a significant decrease in persistency during pregnancy for both parity groups, albeit in different ways. A later DIMc was only weakly associated with a higher persistency.

Generally, primiparous cows display higher persistency, but lower total milk production compared to multiparous cows (Niozas et al., 2019; Atashi et al., 2020; Marumo et al., 2022). Thus, we hypothesized that the association between DPC and lactation persistency might differ between parity groups. We therefore incorporated an interaction term between parity group and DPC in the model. A notable difference was indeed present during late pregnancy (i.e., from DPC 150 onwards, **Figure 3.5**). From DPC 150 to 210, multiparous cows experienced a decline in persistency by 4.6%, whereas primiparous cows remained stable or slightly inclined in that period.

Earlier DIMc groups (G1 and G2) exhibit a more pronounced decrease in persistency from DPC 0 to DPC 210, while the change in persistency on average is relatively small in other DIMc groups (**Figure 3.6**). The high persistency observed in the early DPC for the G1 and G2 groups may be a consequence of the lactation curve modelling. At that time (DIM before 90), the lactation curve is still ascending, indicating a continuous increase in daily milk yield. Consequently, the model might occasionally misinterpret this pattern as high persistency, potentially introducing inaccuracies. Previous research had indicated that multiparous cows exhibit a more rapid decline in milk production during the later stages of pregnancy compared to primiparous cows (Bormann et al., 2002; Leclerc et al., 2008; Yamazaki et al., 2016). This decline could potentially be attributed to the increased energy demands of foetal development during the third trimester of gestation (Roche, 2003; Brotherstone et al., 2004). Meanwhile, primiparous cows might be able to counterbalance this decline in a more controlled way because of the hyperactive maturation of the mammary glands (Jingar et al., 2014; Walter et al., 2022).

A later DIMc, indicating an increased number of days open, is associated with a higher persistency after conception. This finding may follow from: 1) biological factors: increasing days open leading to higher persistency or, 2) farm management practices: cows with higher persistency being more likely to be selected by the farmer for a postponed insemination moment and extended lactation (i.e., deliberately postponing the time of artificial insemination). Therefore, we cannot assert a purely biological causation. Previous research has established a positive relationship between days open and persistency (Niozas et al., 2019; Burgers et al., 2021). However, it is worth noting that primiparous cows, despite exhibiting a higher persistency, are generally inseminated earlier than multiparous cows (with an average DIMc of 121 days versus 133 days, as observed in our study), likely due to their lower total milk yield. This may suggest that farmers' economic considerations prioritize milk production levels above persistency. Different DIMc leads to variations in persistency patterns across different DPC, consequently impacting milk production performance as expected (**Figure 3.7**). Longer DIMc is generally associated with a lower dry-off milk yield and a higher M305 yield, aligning with findings from previous studies (Niozas et al., 2019; Burgers et al., 2021). To consider changes in lactation persistency throughout the DPC period when determining the optimal insemination timing after calving, one approach is to predict a lactation curve from conception to dry off, as illustrated in **Figure 3.7**. However, it is important to note that across the range of DIMc from G1 to G7, there is a difference in persistency of 100 days for primiparous cows and 62 days for multiparous cows. In this respect, it is important to note that our research focused specifically on Holstein-Friesian dairy cows that originate from the Netherlands and Belgium, where milk persistency has been recognized and emphasized as a breeding trait since 2001 (personal communication with Gerben de Jong from the Cattle Improvement

Cooperative CRV). This specific context and breeding focus may have influenced the results of our study.

To adjust the model for production parameters before conception, we incorporated covariates from the interval between 40 and 20 days before DIMc (decay-30 and peak yield-30). In our study, higher pre-conception persistency was associated with higher persistency after conception, while a higher peak yield-30 was linked to lower persistency. This observation was expected since higher peak milk yield has been shown to have a negative relationship with persistency (Hostens et al., 2012; Burgers et al., 2020; Marumo et al., 2022). These findings suggest clearly that both peak yield-30 and persistency before conception correlated to persistency after conception.

There are multiple measures of persistency, and all these measures require the transformation of milk production data (Togashi and Lin, 2009; Yamazaki et al., 2011b; Burgers et al., 2021). Simple measures of persistency are typically fixed at two time points in lactation (Togashi and Lin, 2009; Yamazaki et al., 2011b; Chen et al., 2016), limiting the ability to observe persistency changes throughout the lactation. To overcome this limitation, we employed lactation curve modelling using the MilkBot model, which allowed us to assess persistency at any timepoint within the lactation period. This so-called continuous measurement provides insights into the changes in persistency during lactation. There are a number of reasons why we selected the MilkBot model among all current models. Firstly, the MilkBot model is capable of accurately modelling extended lactations (Ehrlich, 2011). Other models, such as the Wood function, have limited capacity to accurately describe the shape of the lactation curve beyond DIM 305 (Dekkers et al., 1998; Bouallègue and M'hamdi, 2019). Secondly, the MilkBot model, utilizing Bayesian statistics, provides a consistent fitting of individual cow lactation data, even in cases where the data is sparse and noisy through the incorporation of prior information (i.e., the population mean lactation curve characteristics). However, the prior information should be selected with caution, as is the case with all models based on Bayesian statistics.

A number of limitations have to considered in our study. Firstly, we were unable to directly examine the effect of pregnancy on persistency since we did not have non-pregnant cows with complete lactation records. In this respect, non-pregnant cows that are milked for an extended period without experiencing a subsequent calving are not common. Such cows may provide valuable insights into the lactation effects on persistency in the absence of pregnancy, but as a farmer aims to get cows pregnant to ensure continuing milk production in continuing lactations, these cows are not present in high numbers on farms. Additionally, for a more comprehensive analysis, it is necessary to match pregnant and non-pregnant cows based on herd, parity and calving date. However, achieving such a matching with field data can be challenging as this requires a large experimental herd and even identical twins. These limitations have to be taken into consideration when interpreting the findings and conclusions of this study.

In the model, the fixed effects accounted for 25.6% of the variance explained. Within the unexplained variance, 46.1% was attributed to individual animal factors, 7.7% to herd components, and 3.9% to birth year. It is common that the individual animal contributes significantly more to the percentage of unexplained variance compared to a herd component. In the current study, the data were unbalanced (i.e., not all lactations had a decay for all DPC). However, results remained robust and consistent with the outcomes observed when using only lactations (n = 34,188) with decay for all DPC. This observation indicates that the analysis performed on the unbalanced dataset did not significantly affect the overall findings and conclusions of our study.

3.5. Conclusion

We have quantified the association between DPC and lactation persistency. During pregnancy, there was a decrease in persistency observed in both primiparous and multiparous cows. Notably, multiparous cows showed a higher decline in persistency compared to primiparous cows. Furthermore, a later DIMc (i.e., increased number of days open as cows conceiving later) was associated with a higher persistency. Except for the early DIMc groups (DIMc<90), DIMc does not impact the change in persistency by gestation. In conclusion, the outcomes from this study contribute to a better understanding of how the days post conception, and hence the moment of insemination during lactation, influence lactation persistency, enabling farmers to make more informed, evidence-based decisions when taking persistency into account in their reproduction management.

Acknowledgement

We gratefully acknowledge the MmmooOgle programme (Puurs, Belgium) and Dr. Bonifacius van Ranst for providing the data. The authors would like to thank Dr. Jan van den Broek and Dr. Hans Vernooij for statistical advice. Author Yongyan Chen is financially supported by the Oversea Study Program of the Guangzhou Elite Project. The authors reported no conflicts of interest.

Chapter 4

An empirical analysis of economic herd performance in relation to herd lactation curve characteristics

Yongyan Chen

Miel Hostens

Mirjam Nielen

Jim Ehrlich

Wilma Steeneveld

Published in Proceeding of European Conference on Precision Livestock Farming & International Conference on Precision Dairy Farming 2022. p. 728–736

Abstract

Lactation persistency gets increasing attention, and previous studies stated persistent cows are more profitable. These studies were however at cow level, and associations might differ from herd level as other herd factors are interfering with herd economic performance. Additionally, for other lactation curve characteristics (magnitude, time to peak yield) no economic evaluation is performed yet. Our objectives were to 1) present a procedure to aggregate cow lactation curves into herd lactation curves (herd magnitude, herd time to peak yield and herd persistency); 2) investigate the association between herd lactation curve characteristics and herd economic performance. Longitudinal Dutch data (8 years) on milk production and accounting of 1,673 herds were evaluated. Cow lactation curve characteristics were summarized to weighted median herd lactation curve characteristics on a calendar year basis, for primiparous and multiparous cows (**P1** and **P2+**). Data was analyzed using linear mixed modelling, with income over feed cost (**IOFC**) per cow as dependent variable, herd lactation curve characteristics and other herd variables as independent variables. Results indicated all herd lactation curve characteristics were associated with IOFC, except for time to peak yield for P1. All were positively associated with IOFC, except for the negative association with time to peak yield for P2+. In conclusion, we defined herd production patterns by aggregating the cow lactation curves into annual herd lactation curves for P1 and P2+. Associations between IOFC and the various herd lactation curve characteristics were deemed logical and interpretable, suggesting that the herd level aggregation was valid. More research is required to determine when herd economic analysis can be based on simple peak production or M305, or in which circumstances the more computationally challenging herd lactation curve characteristics are better suited.

Keywords

Lactation curve, dairy cow, herd economics

4.1. Introduction

A lactation curve model can quantify the lactation curve shape for a single lactation of a dairy cow and consists of various lactation curve characteristics which describe the curve in different ways. The classic Wood model (Wood, 1967) is the most common lactation curve model, inspiring certain improvements and innovations (e.g., Wilmink, 1987). The Wood model consists of the scale (representing the level of production), the ramp (representing the rising rate of milk production up to the peak level) and the declining slope (Wood, 1967). The MilkBot model adjusts the Wood model with extended lactations (Ehrlich, 2011). In MilkBot, scale and ramp are similar to the Wood model but include other characteristics - the estimated time between the start of milk synthesis and calving (offset) and the rate of late lactation decline (decay) - which can be transformed into a measure of persistency (Ehrlich, 2011).

Persistency is an important lactation curve characteristic describing the cow's ability to maintain a slow rate of decline in production after the peak (Wood, 1967). Persistent cows have increased milk yields, improved conception rates, extended productive lifetimes and decreased culling rates (Dekkers et al., 1998; Hadley et al., 2006; Togashi et al., 2016). The economic consequences of persistency have mainly been evaluated with bio-economic simulation models (Dekkers et al., 1996, 1998). Empirically, economic analyses have only included feed costs (Sölkner and Fuchs, 1987). Both normative and empirical studies have shown that cows with higher persistency are more profitable (e.g., Dekkers et al., 1998; Němečková et al., 2015).

To our knowledge, only Němečková et al. (2015) have presented empirical economic evaluations of other lactation curve characteristics (ramp, scale) besides persistency. Their study evaluated only 80 dairy cows from one herd. As lactation curve characteristics are only available at the cow level, economic evaluations (both empirically and normatively) were all performed at the cow level (Sölkner and Fuchs, 1987; Dekkers et al., 1998; Němečková et al., 2015). However, associations found at the cow level might differ at the herd level as other herd factors (e.g., management, herd size) can interfere with the herd's economic performance. Economic evaluations of lactation curve characteristics at the herd level need a valid aggregation of cow lactation curve characteristics. Such a herd lactation curve needs however to be summarized on a calendar year basis as only then it is possible to combine it with economic data, which is often expressed at a calendar year. This will, however, be challenging as individual cow lactation curves often belong to multiple calendar years. Aggregating methods from cow to herd level lactation curves, on a calendar year basis, have not previously been described.

This study aimed to 1) present a procedure to aggregate cow lactation curves into herd lactation curves (herd magnitude, herd time to peak yield and herd persistency); 2) investigate the association between the herd lactation curve characteristics and the economic performance of dairy herds.

4.2. Material and methods

4.2.1. Available data

Milk production data at the test-day level and herd level performance data for the years 2007 to 2016 were obtained from the Dutch Cattle Improvement Cooperative (CRV, Arnhem, The Netherlands). Originally, the cow test-day data included 159,173,868 test-day records from 6,710,117 cows in 20,760 herds. All test-day records included general cow information (e.g., birth date, calving date, parity, health status), milk yield (kg) and milk component (protein and fat percentage). At the cow level, days in milk (**DIM**), age in days and calving intervals were calculated for every lactation. Herd level performance data contained annual averages of somatic cell counts (**SCC**), calving intervals, age in days and the 305-day milk production level (**M305**).

Herd accounting data from a Dutch accounting agency (Flynth, Arnhem, The Netherlands) was obtained. The data represented 2,058 herds with 18,108 yearly records from 2008 to 2015. The herd accounting data included annual information on all revenues (e.g., milk, livestock), fixed costs (e.g., depreciation, maintenance costs) and variable costs (e.g., feed costs, breeding costs, health costs), as well as on general herd characteristics (e.g., soil type, herd size, milking system).

4.2.2. Development of herd lactation curve characteristics

We used the cow test-day data to calculate herd lactation curve characteristics. First, we fitted a lactation curve for each lactation with the MilkBot model using a proprietary maximum likelihood fitting algorithm by the DairySight fitting engine (Ehrlich, 2011). The MilkBot equation is shown as:

$$Y(t) = a\left(1 - \frac{e^{\frac{-t}{b}}}{2}\right)e^{-dt} \qquad (1)$$

in which Y(t) is the estimated milk production when DIM is t, and a (scale), b (ramp), and d (decay) are lactation curve characteristics describing the lactation curve. As c (offset) is practically undetectable without daily milk production records at the beginning of lactation we decided not to use offset. In the current study, a (scale) was renamed magnitude of milk production (in kg day^{-1}), b (ramp) was renamed time to peak yield (in days), and d (decay) was transformed into a measure of persistency using the equation (Ehrlich, 2011):

$$Persistency = \frac{0.693}{\mathrm{d}} \qquad (2)$$

Persistency (in days) is the time needed for milk production to drop by half after the peak.

After fitting, every lactation had a set of three lactation curve characteristics (magnitude, time to peak yield and persistency). Two parity groups were defined: primiparous and multiparous cows. To summarise herd lactation curve characteristics on a calendar year basis, we used a weighted method as the partitioning method to deal with lactations in multiple calendar years (**Figure 4.1**). Lactations belong to every calendar year with a specific weight relative to the number of test-day records. Using the number of test-days as weight, the contribution of the lactation for different calendar years was calculated. For example, cow A started a lactation in 2008 and finished in 2009. This lactation had 5 test-day records in 2008 and 3 in 2009. Suppose there were n and m test-day records in total from all lactations in 2008 and 2009 in the herd, in which cow A belonged. Then cow A's lactation curve characteristics would contribute 5 / n to the herd lactation curve characteristics in 2008 and 3 / m in 2009. Using the number of test-days as weight, weighted medians were calculated per parity group per herd for each calendar year.

To include only complete lactations for aggregation to herd level, we excluded herd level calculations for the first record year (2007) and last record year (2016), resulting in 273,322 records from 20,000 herds.

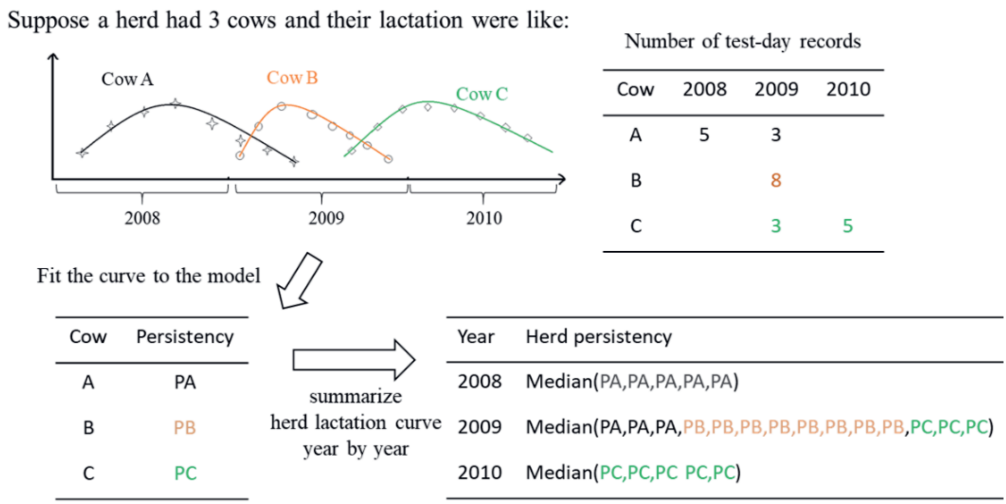


Figure 4.1 Example of how to aggregate herd lactation curve characteristics from individual cow lactation curve characteristics illustrated for persistency.

4.2.3. Data management

We defined several additional variables based on the accounting dataset. First, income over feed cost (**IOFC**) was calculated as total milk revenue minus total feed costs (Wolf, 2010), and was expressed per cow. Secondly, the relative yearly herd milk price was calculated as the difference between herd milk price and the Dutch raw milk price for the corresponding year. Finally, the equity ratio was calculated as the total equity divided by the total assets. The expansion rate was calculated as:

$$\text{Expansion rate} = \left[\frac{(\text{herd size in m year} - \text{herd size in n year})}{\text{herd size in n year}}\right] / (m - n) \quad (3)$$

The yearly herd accounting data of 2,058 herds were merged with herd lactation curve characteristics (n = 20,000 herds) and herd performance data (n = 20,760 herds) for the corresponding years. This merging was possible for 1,887 herds and resulted in a dataset of 12,849 yearly records from 2008-2015. We first excluded 184 yearly records as they were not consecutive (< 2 years consecutive). Secondly, we excluded herds selling milk products on farm (direct sellers) and organic herds (153 yearly records). We also excluded extremely small herds (herd size < 1% percentiles; 126 yearly records). Finally, extreme outliers and records with missing values were excluded (1,578 yearly records). The final dataset included 1,673 herds with 10,808 yearly records.

4

4.2.4. Statistical analysis

Using IOFC per cow as the dependent variable, we developed a linear mixed model to analyse the association between herd economic performance and herd lactation curve characteristics. Apart from the lactation curve characteristics, other variables (e.g., soil type, equity ratio, milking system) were selected as independent variables based on an expected association with IOFC per cow. Multicollinearity between several variables was checked using variance inflation factors. A year variable was forced into the model as a fixed effect to account for potential year effects (e.g., absolute milk price differences). A herd variable was entered into the model as a random effect to account for unobserved herd-related heterogeneity (e.g., environment, feed management). In order to compare the strength of the effect of each individual independent variable to the dependent variable, we standardised continuous independent variables. Akaike information criterion (**AIC**) and backward selection were used to find the best model. The conditional R^2, the marginal R^2 and the part R^2 were calculated to describe the variance explained by the entire model, the fixed effects and a single variable, respectively. Data editing and analysis were performed using the Python API for the Spark platform (PySpark) and R version 3.6.3 (R Core Team, 2020), respectively.

4.3. Results and Discussion

This study presented a procedure to summarize individual cow lactation curves into herd level lactation curve characteristics per calendar year. Aggregating lactation curves fitted from data aggregated by DIM or pooled milk production data from all cows within groups has been applied previously (Vargas et al., 2000; Dematawewa et al., 2007). This method was, however, not applied in our present study, since a lactation curve from aggregated milk recording data would neglect individual cow variation (Ehrlich, 2013). Therefore, we fitted every lactation curve first and subsequently summarized the cow lactation curve characteristics to create annual herd level lactation curve characteristics. In this case, all information of cow lactation curves is available, which allows analysis of both inter-lactation and intra-lactation variability. It makes a better understanding of the variance of cow level lactation curve characteristics possible, opening possibilities to explore differences within and between herds.

Table 1 Distribution of annual herd lactation curve characteristics from weighted median method for primiparous cows (P1) and multiparous cows (P2+).

Lactation curve characteristics	P1			P2+		
	Mean (SD)	Q1[1]	Q3	Mean (SD)	Q1	Q3
Herd magnitude (kg/day)	34.3 (4.08)	31.8	37.1	45.9 (5.85)	42.5	49.8
Herd time to peak yield (day)	29.6 (0.44)	29.4	29.9	22.0 (1.25)	21.9	22.5
Herd persistency (day)	373 (82.3)	316	417	255 (38.8)	229	277

[1]Q1 and Q3: The first and the third quartile

Average herd lactation curve characteristics are presented in **Table 1**. Herd persistency was on average 373 and 255 days for primiparous and multiparous cows, respectively. As in previous studies, primiparous cows tend to have higher persistency than multiparous cows (Gengler, 1996) and we found the variance of herd persistency was higher in primiparous cows as well (**Table 1**). Mean milk weights calculated for each DIM were often used when making aggregate lactation curves (VanRaden et al., 2006). Previously, median and mean milk weights were aggregated for each DIM to describe aggregated curves for different dairy breeds and parities. These median and means were however based on normally distributed data, and therefore mean and median curves were similar in all cases (Ehrlich, 2011). We demonstrated that for a skewed distributed variable, like persistency, using median was an appropriate way to aggregate to herd level lactation curve characteristics as the variance was smaller.

The results of the final reduced linear mixed model to estimate the associations between IOFC per cow and herd lactation curve characteristics are presented in **Table 2**. All herd lactation curve characteristics were associated ($P < 0.01$) with IOFC per cow, except for the time to peak yield for primiparous cows. Apart from the negative association with time to peak yield for multiparous cows, all estimated coefficients were positive, indicating that an increased lactation curve characteristic was associated with an increased IOFC per cow. The standardised coefficients indicated that for multiparous cows, herd magnitude had a larger effect on IOFC per cow than it did for primiparous cows. Increasing one unit of magnitude for multiparous and primiparous cows corresponded to a €152.9 and €48.0 increase in IOFC per cow, respectively. Of this model, the conditional R^2 and the marginal R^2 were 88.9% and 76.6%, respectively. Herd lactation curve characteristics explained 14.0% variance of IOFC per cow, 88.9% of which was explained by multiparous cows.

Table 2 Results of the final reduced linear mixed model on the association between herd lactation curve characteristics and income over feed cost per cow (€).

Variable		β	S.E.	P value
Intercept		2,436.7	5.93	< 0.001
Primiparous cows	Magnitude	48.0	3.13	< 0.001
	Time to peak yield	1.1	1.85	0.600
	Persistency	14.1	2.44	< 0.001
Multiparous cows	Magnitude	152.9	3.77	< 0.001
	Time to peak yield	-5.2	1.98	0.010
	Persistency	66.4	2.87	< 0.001
Year	2008	Ref[1]		
	2009	-593.1	5.41	< 0.001
	2010	-225.0	6.45	< 0.001
	2011	80.7	5.86	< 0.001
	2012	-248.4	6.46	< 0.001
	2013	119.6	7.06	< 0.001
	2014	188.1	6.45	< 0.001
	2015	-492.6	7.20	< 0.001
Milking system	Conventional	Ref[1]		
	Automatic	14.2		0.030
Herd size		-16.1	3.44	< 0.001
SCC		-23.1	2.10	< 0.001
Equity ratio		6.4	2.55	< 0.001
Herd intensity		-13.9	2.82	< 0.001
Calving interval		-22.1	2.14	< 0.001
Relative herd milk price		149.0	2.32	< 0.001

[1]Ref: This category is used as a reference category in the regression analysis.

Magnitude was most strongly associated with IOFC per cow among the herd lactation curve characteristics of both parity groups. This was expected, as, of all lactation curve characteristics, herd magnitude has the highest correlation with M305 (Ehrlich, 2013) and M305 explains most milk revenues (Demeter et al., 2011). Herd persistency of both parity groups was positively associated with IOFC per cow. These results correspond with earlier findings (Sölkner and Fuchs, 1987; Dekkers et al., 1998; Němečková et al., 2015), with previous studies also mentioning persistency as an important economic parameter (De Vries, 2006; Togashi and Lin, 2009). Time to peak yield was least associated with IOFC per cow in our study. This was expected because of the weak phenotypic correlation between the rising rate of milk to the peak yield and M305 (Elahi Torshizi, 2016; Atashi et al., 2020).

Lactation curve characteristics for multiparous cows were more strongly associated with herd economics than those for primiparous cows. The herd magnitude of multiparous cows was positively associated with IOFC per cow. Multiparous cows have a higher milk production compared to primiparous cows (Cole et al., 2012); they generally make up 60-70% of the dairy herd and are thus the main milk suppliers of the herd.

4.4. Conclusions

In this study, we defined herd production patterns by aggregating the individual cow level lactation curve characteristics to a yearly herd level for primiparous and multiparous cows separately. The associations between IOFC per cow and the various herd lactation curve characteristics were deemed logical and interpretable, suggesting that the herd level aggregation was valid. More research is required to determine when herd level economic analysis can be based on simple peak production or M305, or in which circumstances the more computationally challenging herd lactation curve characteristics are better suited.

Acknowledgements

We gratefully acknowledge the Cattle Improvement Cooperative (CRV, Arnhem, The Netherlands) and Flynth (Arnhem, The Netherlands) for providing the data. Author Yongyan Chen is financially supported by the Oversea Study Program of Guangzhou Elite Project.

Chapter 5

Herd level economic comparison between the shape of the lactation curve and 305d milk production

Yongyan Chen

Miel Hostens

Mirjam Nielen

Jim Ehrlich

Wilma Steeneveld

Published in Frontiers in Veterinary Science, doi:10.3389/fvets.2022.997962.

Abstract

Herd milk production performance is generally evaluated using the herd's average 305-day milk production (**HM305**). Economic comparisons between herds are also often made using HM305. Comparing herds is thus based on summarized milk production, and not on the form of the lactation curves of the cows within the herd. Cow lactation curve characteristics can be aggregated on a calendar year basis to herd lactation curve characteristics (**HLCC**) (herd magnitude, herd time to peak yield and herd persistency). Thus far, no literature has evaluated whether the shape of the lactation curve (described by HLCC) is better able to explain the economic variation of herds than summarized milk production such as HM305 does. This study aims to determine whether HM305 or HLCC is better able to explain the variation in economic performance between herds. To do so, we evaluated eight years of Dutch longitudinal data on milk production and the financial accounts of 1,664 herds. Cow lactation curve characteristics were calculated through lactation curve modelling and aggregated to HLCC on a calendar year basis for two parity groups (primiparous cows and multiparous cows). Using income over feed cost per cow (**IOFC-cow**) or per 100kg milk (**IOFC-milk**) as the dependent variable separately, we developed four linear mixed models. Two models were used to analyse the association between herd economic performance and HLCC; the other two models were used to analyse the association between herd economic performance and HM305. A Cox test and J test were used to compare two non-nested models to investigate whether HM305 or HLCC better explain IOFC. The average IOFC-cow was €2,305 (SD = 408) per year, while the average IOFC-milk was €32.1 (SD = 4.6). Results showed that HLCC and HM305 explain the same amount of variance of IOFC-cow or IOFC-milk. IOFC-cow was associated with HM305 and HLCC (except herd time to peak yield for primiparous cows). Herd magnitude was most strongly associated with IOFC-cow, followed by herd persistency and herd time to peak yield of multiparous cows. IOFC-milk was not associated with HM305 or HLCC (except for a weak negative association with herd persistency for primiparous cows). IOFC-cow and IOFC-milk were driven most by time effects. In conclusion, HLCC and HM305 explain the same amount of variance in IOFC-cow or IOFC-milk. HLCC is more computationally expensive, while HM305 is more readily available.

Keywords
lactation curve; milk production; dairy; economics; herd aggregation

5.1. Introduction

The milk production performance of cows is generally described by 305-d milk production (**M305**) (Fleischer et al., 2001; Buckley et al., 2003), which is an indicator of absolute milk production. Additionally, lactation curve characteristics (**LCC**), describing the lactation curve in different ways, can also be used to evaluate the milk production performance of cows. LCCs are derived from a lactation curve model such as the classic Wood model (Wood, 1967), the Wilmink model (Wilmink, 1987) and the Milkbot model (Ehrlich, 2011). The MilkBot model, for example, consists of the scale (representing the level of production), the ramp (representing the rising rate of milk production up to the peak level), the estimated time between the start of milk synthesis and calving (offset) and the rate of late lactation decline (decay). The latter can be easily transformed into a measure of persistency (Ehrlich, 2011). Both M305 and LCC are commonly used to compare the milk production performance of cows (Mellado et al., 2011; Hostens et al., 2012) as well as economic performance (Shim et al., 2004; Němečková et al., 2015). Results show that cows with a higher M305 have lower costs per kg of milk and produce a higher IOFC-cow (Němečková et al., 2015).

The milk production performance of the herd is generally evaluated using the herd's average 305-day milk production (**HM305**) (Pinedo et al., 2010; Nor et al., 2014; Shahid et al., 2015) along with some other variables such as average milk production per cow per year (Oleggini et al., 2001; Kristensen et al., 2008). Economic comparisons between herds are also generally made using HM305 (corrected for milk price) (Ferguson et al., 2000; Green et al., 2002). Comparing herds is thus based on the absolute volume of milk production rather than on the form of the lactation curves of the cows within the herd. Comparing herds based on LCC is challenging as LCCs are at cow level. Chen et al. (Chen et al., 2022b) have already presented a procedure to aggregate the individual cow level LCC to the annual herd level for primiparous and multiparous cows separately. The annual herd lactation curve characteristics (**HLCC**) open possibilities to explore differences between herds. Potentially, HLCC can be an additional herd performance indicator. It differs between herds since the environment, management and cow genetics of a dairy herd influence individual cow's LCCs and hence HLCCs (Val-Arreola et al., 2004; Ehrlich, 2013). Persistency is one of the lactation curve characteristics that was shown to increase profitability at cow level, where more persistent cows were more profitable (Dekkers et al., 1996, 1998; Němečková et al., 2015). This association was not studied at herd level, where HLCC might be associated with herd level economic results. It is therefore not known whether the shape of the curve is better or worse than the absolute volume of milk production at explaining the economic variation of herds. The herd's economic performance can be expressed in many ways, depending on data availability ant the aim of the research. A herd's economic performance includes revenues, fixed costs and variable costs, which are difficult data to gather precisely. When the value of farm assets is not well-known, partial measure of farm profitability can be used (Kristensen et al., 2008; Vredenberg et al., 2021), such as gross margin, income over feed cost and milk-to-feed price ratio. Gross margin states the difference between total revenues and total variable costs. If only milk revenue and feed costs data are available, economic calculations, such as income over feed cost and milk-to-feed price ratio, can be used (Bailey et al., 2009; Wolf, 2010; Atzori et al., 2021). Income over feed cost is often used to monitor whether the feed cost is in line for the milk production or whether the feed management is successful (Bailey et al., 2005; Buza et al., 2014a; Cowley et al., 2020). Milk-to-feed price ratio indicates the convenience of transforming feed into milk in terms of market opportunity (Atzori et al., 2021). However, when the price of milk and feed are volatile income over feed cost is a better measure of profitability than milk-to-feed price (Wolf, 2010).

This study aims to determine whether HM305 or HLCC is better at explaining the economic performance variation between herds.

5.2. Materials and methods

5.2.1. Available data

For this study, we obtained milk production data at the test-day level and herd level performance data for the years 2007–2016 from the Dutch Cattle Improvement Cooperative (CRV, Arnhem, The Netherlands). Originally, the cow test-day data included 159,173,868 test-day records from 6,710,117 cows in 20,760 herds. All test-day records included general cow information (e.g., birth date, calving date, parity, health status), milk yield (kg) and milk component (protein and fat percentage). At the cow level, days in milk, age in days and calving intervals were calculated for every lactation. Herd level performance data contained annual averages of somatic cell counts (SCC), calving intervals, age in days and HM305.

We retrieved herd accounting data from a Dutch accounting agency (Flynth, Arnhem, The Netherlands). The data represented 2,058 commercial herds with 18,108 yearly records from 2008–2015,herd size varied between 5 and 1075. The herd accounting data included annual information on all revenues (e.g., milk, livestock), fixed costs (e.g., depreciation, maintenance costs) and variable costs (e.g., feed costs, breeding costs, health costs), as well as on general herd characteristics (e.g., soil type, herd size, milking system).

5.2.2. Development of HLCC

The development of HLCC was described in detail in our previous study (Chen et al., 2022b). In short, we used the cow test-day data to calculate HLCC. First, we fitted a lactation curve for each whole lactation with the MilkBot model using a proprietary maximum likelihood fitting algorithm of the DairySight fitting engine (Ehrlich, 2011). The full MilkBot equation is shown as:

$$Y(t) = a\left(1 - \frac{e^{\frac{c-t}{b}}}{2}\right)e^{-dt} \qquad (1)$$

in which Y(t) is the estimated milk production when days in milk is t, and a (scale), b (ramp), c (offset) and d (decay) are LCC describing the lactation curve. As offset is practically undetectable without daily milk production records at the beginning of lactation, we decided not to use that measurement, resulting in a simplified equation:

$$Y(t) = a\left(1 - \frac{e^{\frac{-t}{b}}}{2}\right)e^{-dt} \qquad (2)$$

In the current study, a (scale) was renamed magnitude of milk production (in kg/day) and b (ramp) was renamed time to peak yield (in days). d (decay) was transformed into a measure of persistency using the equation (Ehrlich, 2011):

$$Persistency = \frac{0.693}{d} \qquad (3)$$

Persistency (in days) is the time needed for milk production to drop by half after the peak.

After fitting, every lactation had a set of three LCCs (magnitude, time to peak yield and persistency). Two parity groups were defined: primiparous cows and multiparous cows. To summarise HLCC on a calendar year basis, we used a weighted method (Chen et al., 2022b) as the partitioning method to deal with lactations in multiple calendar years. Lactations belong to every calendar year with a specific weight relative to the number of test-day records. Using the number of test-days as weight, the contribution of the lactation for different calendar years was calculated. For example, cow A started a lactation in 2008 and finished in 2009. This lactation had 5 test-day records in 2008 and 3 in 2009. Suppose there were n and m test-day records in total from all lactations in 2008 and 2009 in the herd. Cow A's lactation curve characteristics would contribute 5/n to the herd lactation curve characteristics in 2008 and 3/m in 2009. Using the number of test days per year as weight, the median HLCCs were defined as the annual HLCC per parity group, per herd and per year. As described in Chen et al. (Chen et al., 2022b), we only included complete lactations to aggregate at herd level and thus excluded herd level calculations for the first record year (2007) and last record year (2016), resulting in 273,322 records from 20,000 herds. The lactation lengths varied between 56 (5%) and 495 (95%) days, with a mean of 336 days.

5.2.3. Data management

The definition of all variables is shown in **Table 1**. We defined several additional variables based on the accounting dataset. First, income over feed cost (**IOFC**) was calculated as total milk revenue minus total feed costs (Bailey and Ishler, 2008; Wolf, 2010). Total annual milk revenue was available in the dataset and total annual feed costs were calculated by adding up the annual costs for concentrates, vitamins, minerals, wet by-products and roughage. We calculated two variables for IOFC, one expressed per cow (**IOFC-cow**) and the other expressed per 100kg milk (**IOFC-milk**).

Secondly, we calculated annual herd milk prices by dividing the total kg of milk delivered to the factory by milk revenue. In addition, we looked up average Dutch yearly raw milk prices (European Commission, 2021) and calculated the relative annual herd milk price as the difference between herd milk price and the Dutch raw milk price for the corresponding year. Thirdly, as an indication of the level of economic leverage, we calculated the equity ratio per herd per year as follows:

$$\text{Equity ratio} = \frac{(\text{total assets} - \text{total liabilities})}{\text{total assets}}$$

Finally, we calculated the expansion rate from year n to year m as follows:

$$\text{Expansion rate} = \left[\frac{(\text{herd size in m year} - \text{herd size in n year})}{\text{herd size in n year}}\right] / (m - n)$$

An overview of all datasets and defined variables included is shown in **Figure 5.1**. The yearly herd accounting data of 2,058 herds were merged with calculated HLCC (n = 20,000 herds) and herd performance data (n = 20,760 herds) for the corresponding years. This merging was possible for 1,887 herds and resulted in a dataset of 12,849 yearly records from 2008-2015 for further analysis.

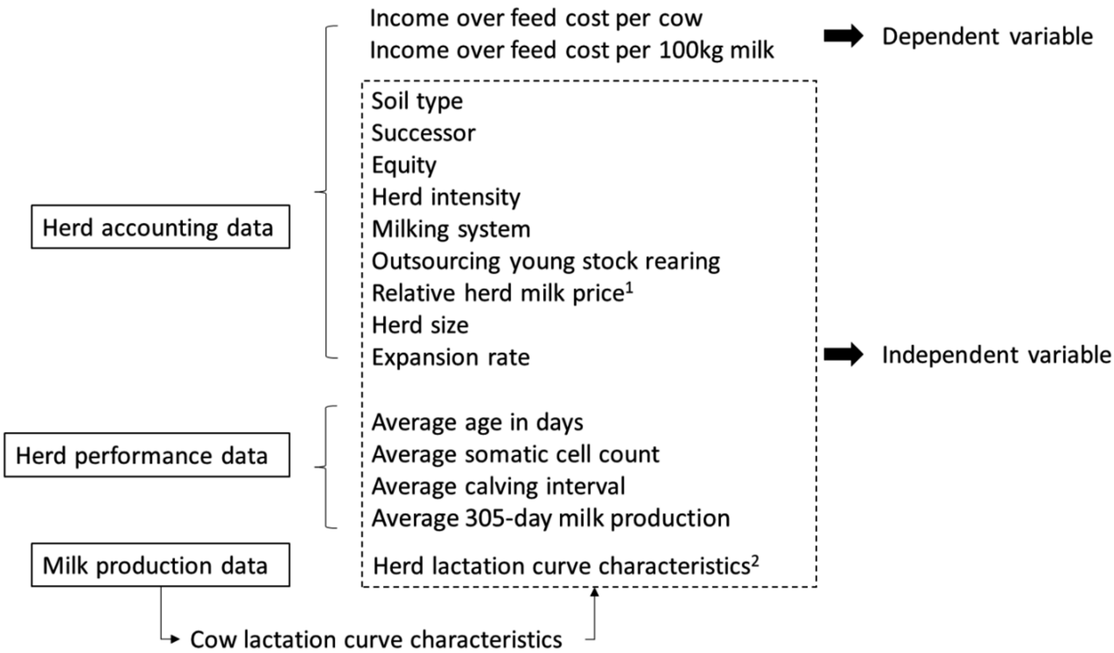


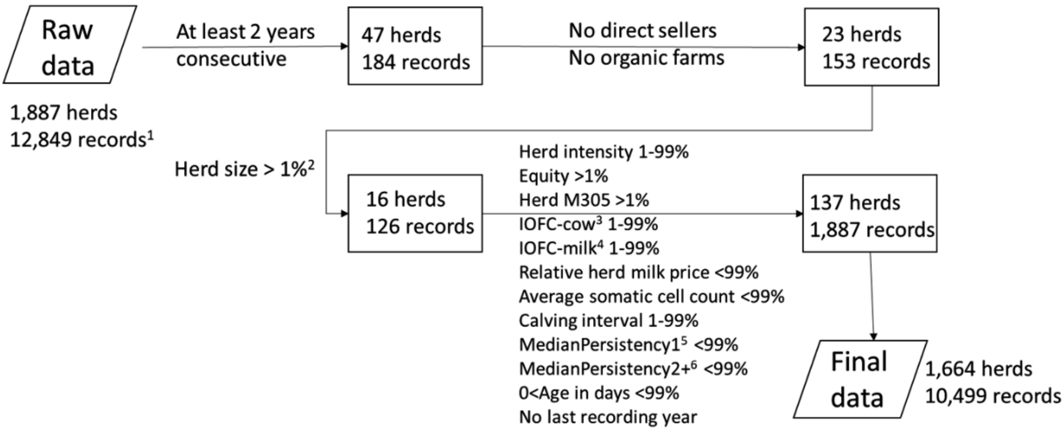
Figure 5.1 Overview of variables used in the statistical analyses and the dataset they originate from.

The data editing flow diagram is presented in **Figure 5.2**. We first excluded 184 yearly records as they were not consecutive (<2 years consecutive). Secondly, we excluded herds selling milk products on farms (direct sellers) and organic herds since their milk prices differed too greatly from those of conventional herds (153 yearly records). We also excluded extremely small herds (herd size <1% percentiles; 126 yearly records). In addition, we calculated percentiles for IOFC- cow, IOFC-milk, herd intensity, equity ratio, HM305, relative herd milk price, SCC, calving interval, herd persistency for primiparous cows, herd persistency for multiparous cows and age in days. Of these variables, extreme outliers and records with missing values were excluded (1,887 yearly records). The final dataset included 1,664 herds with 10,499 yearly records.

Table 1 Descriptive statistics of continuous variables over 1,664 Dutch herds for the years 2008–2015.

	Description (unit)	Mean	SD	5%[a]	95%[a]
IOFC-cow [b]	(Milk revenue - feed cost)/herd size (€)	2,305	408	1,609	2,961
IOFC-milk [c]	100*(Milk revenue - feed cost)/milk delivered to factory (€)	32.1	4.6	24.0	39.3
HM305	Average 305-day milk production in the herd (kg)	8,686	899	7,099	10,107
Equity ratio	(total assets - total liabilities) / total assets	0.45	0.31	-0.11	0.93
Herd intensity	Milk production per ha (kg of milk/ha)	15,129	3,841	9,564	22,159
Relative herd milk price	The price difference in relation to national raw milk price[d] (€/100kg)	2.52	2.08	-0.86	5.71
Herd size	Number of cows present in the herd	85.3	43.2	39.0	151.0
Expansion rate	((herd size – last year's herd size)/ last year herd size)/year difference	0.03	0.06	-0.06	0.14
Age in days	Average age in days of cows in the herd	1,716	160	1,483	1,998
Somatic cell counts	Average somatic cell counts of cows in the herd ($*10^3$ cells/ml)	193	59	105	301
Calving interval	Average calving interval of cows in the herd	414	23	384	457
Herd magnitude1[e]	Weighted median magnitude of primiparous cows (kg/day)	34.8	3.7	28.3	40.5
Herd time to peak yield1	Weighted median time to peak yield of primiparous cows (day)	29.6	0.4	28.9	30.2
Herd persistency1	Weighted median persistency of primiparous cows (day)	358	70	263	492
Herd magnitude2+[f]	Weighted median magnitude of multiparous cows (kg/day)	47.7	5.3	38.0	55.8
Herd time to peak yield2+	Weighted median time to peak yield of multiparous cows (day)	22.1	1.3	20.3	23.3
Herd persistency2+	Weighted median persistency of multiparous cows (day)	240	33	194	304

[a] 5% and 95%: the 5% and 95% percentile.
[b] IOFC-cow: income over feed cost per cow.
[c] IOFC-milk: income over feed cost per 100kg.
[d] Average yearly raw milk price aggregated by monthly raw milk price from official milk market observatory (European Commission, 2021).
[e] 1: primiparous cows.
[f] 2+: multiparous cows.



[1]: the difference in milk price and the Dutch raw milk price for the corresponding year.
[2]: herd magnitude, herd time to peak yield and herd persistency for primiparous cows and multiparous cows (Chen et al., 2022b).

Figure 5.2 Diagram on data editing of the combined production and accounting dataset. The numbers in the boxes represent the excluded numbers.

5.2.4. Statistical analysis

Using IOFC-cow or IOFC-milk separately as the dependent variable, we developed four linear mixed models. Two models were used to analyse the association between herd economic performance and HLCC; the other two models were used to analyse the association between herd economic performance and HM305. We selected other herd variables as independent variables based on an expected association with IOFC. Those selected herd variables were soil type (sand/other), successor availability (yes/no), equity ratio, herd intensity (kg of milk/ha), milking system (automatic/conventional), use of outsourced heifer rearing (yes/no), relative herd milk price, herd size, expansion rate, SCC and calving interval. We used variance inflation factors to check for multicollinearity between several variables. A year variable was forced onto all models as a fixed effect to account for potential year effects (e.g., absolute milk price differences). A herd variable was entered into the models as a random effect to account for unobserved herd-related heterogeneity (e.g., environment, feed management). To compare the strength of the effect of each independent variable to the dependent variable, we standardised continuous independent variables. Akaike information criterion and backward selection were used to find the best models, which were eventually presented in the results. The conditional R^2, the marginal R^2 and the part R^2 were calculated to describe the variance explained by the entire model, the fixed effects and a single variable, respectively. A Cox test and a J test (Davidson and MacKinnon, 1981) were used to compare the two non-nested models to investigate whether HM305 or HLCC better explain IOFC. Both tests are used for non-nested hypothesis testing. For example, models A and B are two non-nested models with the same dependent variable. In the non-nested hypothesis testing, model A would have a null hypothesis that the regressors from model B cannot improve the performance of model A. If the null hypothesis of model A is rejected, model B is the 'true' model, having an additional explanatory power beyond that contributed by model A. If the null hypothesis of model A is not rejected, model A is the 'true' model. The same test can be done for model B to determine whether the regressors from model A can improve the performance of model B.

Data editing and analysis were performed using the Python API for the Spark platform (PySpark) and R version 3.6.3 (R Core Team, 2020), respectively. Code scripts for the data editing steps, statistical analyses and figure visualizing average herd lactation curve for primiparous cow and multiparous cow can be downloaded at https://github.com/Bovi-analytics/Chen-et-al-2022b.

5.3. Results

Total feed costs over all farms varied between €20,345 (5%) and €132,519 (95%) per year, with a mean of €63,320. Total revenues likewise varied between €114,589 (5%) and €540,270 (95%) per year, with a mean of €287,787. The descriptive statistics of the continuous variables over all herds and all years are shown in **Table 1**. The average IOFC-cow was €2,305 (SD = 408) per year, while the average IOFC-milk was €32.1 (SD = 4.6). The same patterns were found in both IOFCs for the years 2008-2015, with the lowest value in the year 2009 and the highest value in the year 2013 (**Figure 5.3**). Average herd magnitude, herd time to peak and herd persistency were 34.8kg (SD = 3.7), 29.6 days (SD = 0.4) and 358 days (SD = 70) for primiparous cows, respectively. Average herd magnitude, herd time to peak and herd persistency were 47.7kg (SD = 5.3), 22.1 days (SD = 1.3) and 240 days (SD = 33) for multiparous cows, respectively. The average HM305 was 8,686kg (SD = 899). The average herd intensity was 15,129kg of milk/ha (SD = 3,841) and the average herd size was 85.3 cows (SD = 43.2).

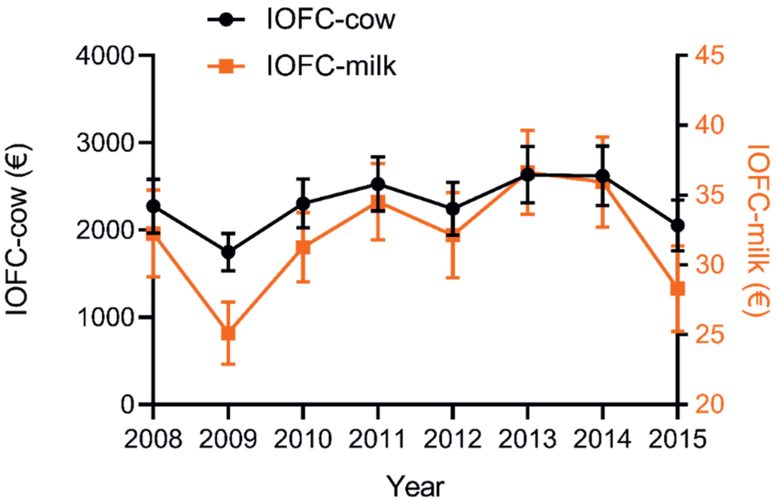


Figure 5.3 Average income over feed cost per cow (IOFC-cow) and per 100 kg milk (IOFC-milk) for the years 2008 to 2015.

The results of the final reduced linear mixed models to estimate the associations between the two IOFC definitions and HLCC are presented in **Tables 2 and 3** respectively.

All HLCCs were associated (P<0.01) with IOFC-cow, except for the herd time to peak yield for primiparous cows (**Table 2**). Apart from the negative association with herd time to peak yield for multiparous cows, all estimated coefficients were positive, indicating that an increased herd lactation curve characteristic was associated with an increased IOFC-cow. The standardised coefficients indicated that herd magnitude had a larger effect on IOFC-cow for multiparous cows than it did for primiparous cows. Increasing one unit of herd magnitude for multiparous cows and primiparous cows corresponded to a €154.3 and €48.0 increase in IOFC-cow, respectively (**Table 2**). In this model, the conditional R^2 and the marginal R^2 were 88.9% and 76.9%, respectively. HLCC explained 12.3% variance of IOFC-cow, 86.6% of which was explained by multiparous cows. The top three variables explaining the variance of IOFC-cow were the year, herd magnitude for multiparous cows and relative herd milk price, at 38.2%, 10.7% and 7.0% part R^2, respectively.

Table 2 Results of the final reduced linear mixed model on the association between income over feed cost per cow (€) and herd lactation curve characteristics (and other herd variables) based on data from 1,664 Dutch herds.

Variable		β	S.E.	P value
Intercept		2,437.1	5.82	< 0.001
Primiparous cows	Magnitude	48.0	3.09	< 0.001
	Time to peak yield	1.8	1.85	0.336
	Persistency	13.3	2.44	< 0.001
Multiparous cows	Magnitude	154.3	3.72	< 0.001
	Time to peak yield	-4.4	1.98	0.027
	Persistency	69.0	2.87	< 0.001
Year	2008	Ref[1]		
	2009	-586.3	5.44	< 0.001
	2010	-222.2	6.43	< 0.001
	2011	85.1	5.78	< 0.001
	2012	-241.2	6.32	< 0.001
	2013	129.7	6.88	< 0.001
	2014	195.2	6.21	< 0.001
	2015	-481.8	6.95	< 0.001
Herd size		-16.1	3.40	< 0.001
Somatic cell counts		-22.9	2.10	< 0.001
Herd intensity		-14.7	2.81	< 0.001
Calving interval		-21.6	2.12	< 0.001
Relative herd milk price		146.4	2.31	< 0.001

[1] Ref: used as a reference category.

5

Table 3 Results of the final reduced linear mixed model on the association between income over feed cost per 100kg milk (€) and herd lactation curve characteristics (and other herd variables) based on data from 1,664 Dutch herds.

Variable		β	S.E.	P value
Intercept		33.60	0.091	< 0.001
Primiparous cows	Magnitude	-0.05	0.036	0.1343
	Time to peak yield	-0.01	0.021	0.5026
	Persistency	-0.13	0.028	< 0.001
Multiparous cows	Magnitude	0.07	0.043	0.1174
	Time to peak yield	-0.03	0.023	0.1500
	Persistency	-0.03	0.033	0.3987
Year	2008	Ref[1]		
	2009	-7.54	0.064	< 0.001
	2010	-3.58	0.075	< 0.001
	2011	0.74	0.069	< 0.001
	2012	-2.53	0.074	< 0.001
	2013	1.64	0.080	< 0.001
	2014	2.07	0.074	< 0.001
	2015	-6.51	0.081	< 0.001
Soil type	Other soil	Ref		
	Sand soil	0.56	0.085	< 0.001
Somatic cell counts		-0.07	0.024	0.003
Equity ratio		0.08	0.028	0.005
Outsourcing heifer rearing	No	Ref		
	Yes	0.61	0.088	< 0.001
Herd intensity		-1.21	0.031	< 0.001
Calving interval		-0.11	0.025	< 0.001
Relative herd milk price		1.88	0.027	< 0.001
Expansion rate		0.11	0.018	< 0.001
Age in days		0.06	0.025	0.011

[1] Ref: used as a reference category.

The IOFC-milk was negatively associated with herd persistency for primiparous cows **(Table 3)**. A one-unit increase in herd persistency for primiparous cows decreased IOFC-milk by €0.13 ($P<0.01$) on average. In this model, the conditional R^2 and the marginal R^2 were 88.7% and 78.9%, respectively. HLCC only explained

0.20% variance of IOFC-milk. The top three variables explaining the variance of IOFC-milk were the year, relative herd milk price and herd intensity, at 53.2%, 9.3% and 4.5% part R^2, respectively.

The results of the final reduced linear mixed models to estimate the associations between the two IOFC definitions and HM305 are presented in **Tables 4 and 5** respectively.

Table 4 Results of the final reduced linear mixed model on the association between income over feed cost per cow (€) and average herd 305-day milk production (and other herd variables) based on data from 1,664 Dutch herds.

Variable		β	S.E.	P value
Intercept		2,435.2	5.70	< 0.001
Average herd 305-day milk production		206.6	2.95	< 0.001
Year	2008	Ref[1]		
	2009	-584.5	5.34	< 0.001
	2010	-224.3	6.33	< 0.001
	2011	86.2	5.73	< 0.001
	2012	-248.9	6.25	< 0.001
	2013	129.5	6.80	< 0.001
	2014	190.92	6.22	< 0.001
	2015	-478.1	6.89	< 0.001
Herd size		-14.2	3.31	< 0.001
Milking system	Conventional	Ref		
	Automatic	21.0	6.31	< 0.001
Somatic cell counts		-22.4	2.10	< 0.001
Herd intensity		-24.0	2.80	< 0.001
Calving interval		-17.8	2.10	< 0.001
Relative herd milk price		148.6	2.26	< 0.001
Expansion rate		4.84	1.50	0.001

[1] Ref: used as a reference category.

The final reduced linear mixed model on the association between HM305 and IOFC-cow is shown in **Table 4**. Increasing one unit of HM305 corresponded to a €206.6 increase in IOFC-cow. In this model, the conditional R^2 and the marginal R^2 were

5

89.6% and 78.7%, respectively. HM305 explained 18.9% variance of IOFC-cow. The top three variables explaining the variance of IOFC-cow were the year, HM305 and relative herd milk price, at 39.5%, 18.9% and 7.5% part R^2, respectively.

Table 5 Results of the final reduced linear mixed model on the association between income over feed cost per 100kg milk (€) and average herd 305-day milk production (and other herd variables) based on data from 1,664 Dutch herds.

Variable		β	S.E.	P value
Intercept		33.60	0.091	< 0.001
Average herd 305-day milk production		-0.01	0.033	0.700
Year	2008	Ref[1]		
	2009	-7.52	0.064	< 0.001
	2010	-3.57	0.075	< 0.001
	2011	0.76	0.069	< 0.001
	2012	-2.49	0.074	< 0.001
	2013	1.65	0.080	< 0.001
	2014	2.09	0.073	< 0.001
	2015	-6.52	0.081	< 0.001
Soil type	Other soil	Ref		
	Sand soil	0.61	0.085	< 0.001
Somatic cell counts		-0.08	0.024	0.001
Equity ratio		0.08	0.028	0.008
Outsourcing heifer rearing	No	Ref		
	Yes	0.62	0.088	< 0.001
Herd intensity		-1.21	0.031	< 0.001
Calving interval		-0.14	0.024	< 0.001
Relative herd milk price		1.88	0.027	< 0.001
Expansion rate		0.12	0.018	< 0.001
Age in days		0.07	0.025	0.006

[1] Ref: used as a reference category.

The final reduced linear mixed model on the association between HM305 and IOFC-milk is shown in **Table 5**. HM305 was not associated with IOFC-milk. In this model, the conditional R^2 and the marginal R^2 were 88.7% and 78.8%, respectively. HM305 explained 0.03% variance of IOFC-cow. The top three variables explaining the variance of IOFC-milk were again the year, relative herd milk price and herd intensity, at 53.3%, 9.1% and 4.2% part R^2, respectively.

The results of the J test and Cox test are shown in **Table 6**. For IOFC-cow, there is no difference between the model including HM305 and the model including HLCC. For IOFC-milk, the model including HLCC is significantly better at explaining the variance of IOFC-milk than the model including HM305. However, both HM305 and HLCC variables explained almost no variance at all.

Table 6 Results of non-nested hypothesis testing from Cox test and J test.

Test	Comparison[1]		Estimate	Std	Value[2]	P value	Interpretation
Cox test	IOFC-cow	HLCC ~ HM305	-576	22.3	-25.8	< 0.001	No difference
		HM305 ~ HLCC	-141	25.6	-6.0	< 0.001	
	IOFC-milk	HLCC ~ HM305	-0.9	0.67	-1.38	0.166	HLCC is better than HM305
		HM305 ~ HLCC	-32.2	0.72	-44.5	< 0.001	
J test	IOFC-cow	HLCC ~ HM305	0.9	0.04	24.5	< 0.001	No difference
		HM305 ~ HLCC	0.3	0.03	8.8	< 0.001	
	IOFC-milk	HLCC ~ HM305	1.8	1.34	1.3	0.190	HLCC is better than HM305
		HM305 ~ HLCC	1.0	0.12	8.0	< 0.001	

[1] IOFC-cow: models for income over feed cost per cow; IOFC-milk: models for income over feed cost per 100kg milk; HLCC: models include herd lactation curve characteristics; HM305: models include average herd 305-day milk production.
[2] Value: z value for cox test and t value for J test.

All four final multivariable models included variables that showed expected associations with the IOFC outcomes (**Tables 2-5**). For both IOFC variants, SCC, herd intensity and calving interval were negatively associated, while relative herd milk price was positively associated (P<0.01).

Outsourcing heifer rearing, expansion rate, equity ratio and age in days were positively associated with IOFC-milk (P<0.05). Herd size was negatively associated in both IOFC-cow models, while the milking system was only associated with IOFC-cow when HM305 was present in the model (P<0.01).

5.4. Discussion

The goal of this empirical study was to investigate how HM305 or HLCC are associated with economic performance at herd level, defined as IOFC. We used a unique dataset incorporating eight years of milk production and accounting data for 1,664 Dutch herds. Accounting data is rarely available on such a large scale (Steeneveld and Hogeveen, 2012; Steeneveld et al., 2015) and having access to it provided new opportunities to evaluate dairy herd economic performance. In our study, both HM305 and HLCC were associated with IOFC-cow, but they explained approximately the same amount of variance. HLCC is significantly better in explaining the variance of IOFC-milk than HM305. However, both HM305 and HLCC variables explained almost no variance in IOFC-milk at all.

IOFC was chosen as the herd economic performance indicator as the lactation curve is most closely related to milk production and thus milk revenue. In addition, feed costs are between 40% to 60% of the total costs of producing milk (Bailey and Ishler, 2008; Alqaisi et al., 2011). Therefore, milk revenues and feed costs seem to be the two economic components that could be most influenced by variations in lactation curves between herds when ignoring other variable costs (such as health and breeding costs). Other studies have, for instance, evaluated gross margin and the milk-to-feed price ratio (Hadley et al., 2006; Vredenberg et al., 2021). We chose to focus on IOFC because it is a better measure of profitability in periods of volatility (e.g., fluctuations in milk price) compared, for instance, to the milk-to-feed price ratio (Wolf, 2010).

The average IOFC-cow was €2,305 per year, equivalent to €6.22 per day. This value corresponds with previous research on IOFC-cow from similar time periods (Hardie et al., 2014; Wu et al., 2019). IOFC-cow was associated with HM305, HLCC (except herd time to peak yield for primiparous cows), year and other herd characteristics (such as relative milk price) (**Tables 2 and 3**). Our findings on the association between IOFC-cow and HM305 correspond with existing literature, as a higher milk yield per cow resulted in a higher IOFC-cow (Buza et al., 2014b). Previously, Laroche et al. (Laroche et al., 2020) had explained that the IOFC-cow depends mainly on milk production per cow. HLCC and HM305 are both indicators that could reflect the herd's production level. That is why they were both highly associated with IOFC-cow. In the current study, HM305 could explain 18.9% variance of IOFC-cow, similar to findings from other studies (Demeter et al., 2011). In the same way, we could explain the HLCCs' association with IOFC-cow by their correlation with HM305. In the HLCC model, herd magnitude was most strongly associated with IOFC-cow among the HLCCs of both parity groups. This was expected, as, of all LCCs, the magnitude has the highest correlation with M305 (Ehrlich, 2013). Herd persistency of both parity groups was positively associated with IOFC-cow although their relative contribution was 2.2-3.6 times smaller than the magnitude. These results correspond with earlier findings (Sölkner and Fuchs, 1987; Dekkers et al., 1998; Němečková et al., 2015) and with previous studies also mentioning persistency as an important economic parameter (De Vries, 2006; Togashi and Lin, 2009). Time to peak yield was least associated with IOFC-cow in our study, supported by a weak phenotypic correlation between the rising rate of milk to the peak yield and M305 (Elahi Torshizi, 2016; Atashi et al., 2020).

HLCCs for multiparous cows were more strongly associated with IOFC-cow than those for primiparous cows. We expected this finding, since multiparous cows have higher milk production than primiparous cows (Cole et al., 2012). As multiparous cows generally make up 60-70% of the dairy herd they are thus the main milk suppliers of the herd.

The average IOFC-milk was €33.6, which is in line with previous studies (Wolf, 2010; Bozic et al., 2012). IOFC-milk was not associated with HM305 and HLCC (except for a weak association with herd persistency for primiparous cows). Again, we were not surprised by this finding, as IOFC-milk depends primarily on milk quality payment characteristics (e.g., milk fat and protein) and the cost of concentrates (Laroche et al., 2020). The weak negative association with herd persistency for primiparous cows found in the HLCC model can be explained by the

fact that primiparous cows are still growing and need more feed than multiparous cows to produce the same amount of milk (Santos et al., 2001; Van Knegsel et al., 2007; Wathes et al., 2007a). However, this association was so weak that it only has a small effect compared, for example, to year and relative herd milk price. Other studies using accounting data have illustrated similar challenges in finding economic effects; the hypothesis is that this is due to large heterogeneity between farms and years (Steeneveld et al., 2015; Vredenberg et al., 2021).
Our results indicate that the year effect is most strongly associated with IOFC. The year effect of course reflects the milk price in the Netherlands and we therefore expected, for instance, to see the lowest year effect in 2009 because in that year the milk price was lowest (European Commission, 2021). We also found that the relative herd milk price (the price difference in relation to the national raw milk price) was strongly associated with IOFC. This indicates that herds selling milk with a relatively higher milk price due to better components (fat and protein) achieve better economic performance, which is in agreement with previous studies (Bailey et al., 2005; Rodrigues et al., 2005). Herd intensity was negatively associated with IOFC, again corresponding with an earlier study (Vibart et al., 2012).
In our study, we defined HLCC by aggregating the individual cow level LCC to a yearly herd level for primiparous and multiparous cows separately. The associations between IOFC and the various HLCCs were deemed logical and interpretable, suggesting that the herd level aggregation was valid. We had expected HLCC to be able to explain more variance in IOFC than HM305 in herd economics since persistent cows are proven to be more profitable in cow level studies (Dekkers et al., 1996, 1998; Němečková et al., 2015). However, in the current herd level study, HLCC was not better associated with IOFC than HM305, a finding that we did not expect. There might, however, be logical explanations for this finding. First, the absolute volume of milk production (HM305) is basically the area under the lactation curve. This area consists mainly of the magnitude and the persistency of milk production, and, to a lesser extent, of the time to peak yield. This means that the

shape of the curve might essentially be another way to describe the absolute volume of milk production, which is equally captured by M305. A second potential explanation lies in the way LCC is aggregated at herd level. Aggregating HLCC on a calendar year basis is challenging, as individual cow lactation curves often belong to multiple calendar years (Chen et al., 2022b). In our current study, we used the weighted median aggregation method to aggregate HLCC. More sophisticated aggregation methods could probably be used in future studies to improve the aggregation of HLCC. This may result in a more precise HLCC explaining more variance of IOFC than HM305. Potentially, such an improved HLCC might be able to reflect economic variation between herds, irrespective of whether this is defined by IOFC.

In our study, HLCC and HM305 explained a similar variance of IOFC. HLCC is more computationally expensive, while HM305 is more readily available. Potentially, HLCC can be an additional herd indicator, helping farmers and their advisors to evaluate herd lactation when making specific decisions and/or analyses. For instance, when comparing the HLCC of a single herd over several years, the HLCC trends over time may illustrate the genetic improvement of dairy cows for persistency.

Acknowledgements

We gratefully acknowledge the Cattle Improvement Cooperative (CRV, Arnhem, The Netherlands) and Flynth (Arnhem, The Netherlands) for providing the data. Author Yongyan Chen is financially supported by the Oversea Study Program of Guangzhou Elite Project (GEP).

Chapter 6

General Discussion

6

This thesis was conducted to explore the application of lactation curve modelling based on routinely collected farm data on commercial dairy farms in the Netherlands and Belgium. Applications in cow reproduction performance and herd economic performance were shown. Four objectives were formulated: (1) to predict lactation persistency for DIM 305 at different insemination moments (2) to investigate the association between days post conception (**DPC**) and persistency (3) to summarize cow lactation curves into herd lactation curve characteristics **(HLCC** - herd magnitude, herd time to peak yield and herd decay**)** and illustrate a field application of HLCC (4) to compare whether HLCC or the herd 305-d milk production (**HM305**) is better able to explain herd economic performance. In this chapter, the main results of this thesis are discussed. Additionally, the datasets used, the definition of persistency, cow and herd level clustering issues and the generalizability of our study results to China are discussed.

6.1. Application in cow reproduction performance

In **Chapter 2**, we found that it was not possible to predict future persistency at DIM 305, although the prediction power was increasing at later insemination moment based on the cow and herd data available up to the insemination moments. The low prediction accuracy observed in our study may be attributed to other factors that influence persistency between the insemination moments and DIM 305. One potential factor that could impact persistency is pregnancy. However, we were unable to account for the pregnancy effect in our prediction model due to several reasons. Firstly, the exact timing of pregnancy is unknown at the time of making predictions for open cows. Secondly, the quantification of the pregnancy effect on persistency (e.g., the specific mechanisms or causal relationships) is lacking in previous studies, making it difficult to incorporate it into the model. As a result, we were unable to correct for the pregnancy effect in our prediction model, sparking our curiosity to explore the relationship between DPC and persistency in **Chapter 3**.

Chapter 3 summarized how DPC and days in milk at conception (**DIMc**) during lactation influence lactation persistency. We found that persistency decreases during pregnancy, but differently for primiparous and multiparous cows. Specifically, from DPC 150 to DPC 210, multiparous cows showed a larger decline in persistency compared to primiparous cows. Furthermore, later DIMc was weakly associated with higher persistency. This weak association suggests that, in general, DIMc does not impact the change in persistency.

Based on the insights gained from **Chapters 2 and 3**, I wondered which steps can be implemented to enhance the accuracy of persistency prediction. Firstly, the pregnancy effect need to be examined. In **Chapter 3** we did not directly examine the pregnancy effect on persistency since we did not use non-pregnant cows with complete lactation records. Non-pregnant cows that are milked for an extended period without experiencing a subsequent calving are uncommon. Such cows may provide valuable insights into the lactation effects on persistency in the absence of pregnancy, but as a farmer aims to get cows pregnant to ensure continuing milk production in continuing lactations, these cows could only be present in high numbers under experimental settings. Additionally, for a more comprehensive analysis, it would be advisable to match pregnant and non-pregnant cows based on herd, parity and calving date. However, achieving such a matching would be challenging as this requires a large experimental herd and even identical twins. Finally, the integration of the association between pregnancy and persistency into the prediction model relies on identifying the relevant pregnancy-related features (such as changes of hormones levels, body condition score, body weight) for predicting persistency. Body condition score and body weight are now easily measurable with the advancement of technology (such as three dimensional cameras and ultrasound measurements) (Yukun et al., 2019; Martins et al., 2020; Albornoz et al., 2022). Additional fundamental and perhaps experimental research into the biological aspects of this association may be necessary to uncover these specific features.

6.2. Application in animal health economics at herd level

Chapter 4 presented various procedures to aggregate individual cow lactation curves into annual HLCC and illustrate a field application of HLCC. The HLCC was then used in **Chapter 5** to study the association between herd economic performance and HLCC or HM305. HLCC and HM305 explained the same amount of variance of income over feed cost per cow or per 100 kg milk (**IOFC-cow** or **IOFC-milk**).

We had expected HLCC to be able to explain more variance in IOFC than HM305 in herd economics since persistent cows are proven to be more profitable in cow level studies (Dekkers et al., 1996, 1998; Němečková et al., 2015). However, in the current herd level study, HLCC was not better associated with IOFC than HM305. There might, however, be logical explanations for this finding. Firstly, IOFC is most strongly associated with the year effect. The year effect reflects the milk price and feed price in the Netherlands. The milk price and feed price are changing every year due to lots of reasons (international and regional policy, the supply and demand) (Dong et al., 2011; Alqaisi et al., 2019). While within a farm, the farm management is relatively stable between years, leading to relatively stable HM305 and HLCC between years. Therefore, most of the variance of IOFC were explained by year effect. Secondly, the absolute volume of milk production, often represented by HM305, is basically the area under the lactation curve. This area consists mainly of the magnitude and the persistency of milk production, and to a lesser extent of the time to peak yield. This means that the shape of the curve might essentially be another way to describe the absolute volume of milk production, which is equally captured by M305. Thirdly, we applied the weighted median method discussed in **Chapter 4**, which we believe to be the most robust and fair aggregation method among all the methods we introduced in that chapter. Aggregating HLCC on a calendar year basis is challenging, as individual cow lactation curves often belong to multiple calendar years. More sophisticated aggregation methods for summarizing the distribution of cow variance could probably be used in future studies to improve the aggregation of HLCC. This may result in a more precise HLCC explaining more variance of IOFC than HM305.

In **Chapters 4 and 5**, we aggregated the HLCC on a calendar year basis to align with the herd accounting data, which also operates on a calendar year basis. However, it's worth noting that the aggregation of lactation curves may vary depending on the specific objectives. Our aggregation method is versatile and can be applied for group aggregation, not limited to herd-level aggregation. For instance, when monitoring population-level trends for genetic purposes, individual lactations can easily be aggregated to the year in which the lactation started or ended.

6.3. Data used in this thesis

Regarding milk production data in this thesis, we had access to two extensive historical datasets: test-day records over the years 2007–2019 from the Dutch Cattle Improvement Cooperative (CRV, Arnhem, the Netherlands) and milking robot visit records over the years 2005–2022 from the MmmooOgle programme (Puurs, Belgium). In addition, we included herd accounting data over 2008–2015 from a Dutch accounting agency (Flynth, Arnhem, The Netherlands). All of these comprehensive historical datasets are unique, and comparable datasets in size do not exist (Bijl et al., 2007; Steeneveld et al., 2012; 2015). Our longitudinal data, collected over the span of several years, provided substantial analytical potential. In **Chapter 2**, longitudinal data were used for prediction modelling, where past observations were used to forecast future outcomes. In **Chapter 3**, by repeatedly measuring the cow persistency before and after conception in longitudinal data, we could capture persistency changes within a lactation. Longitudinal data is a robust approach when handling variables that experience temporal fluctuations. In **Chapter 4 and 5**, variables such as milk prices and feed costs fluctuated over time, potentially impacting associations between IOFC and HLCC or HM305. We gained comprehensive understanding of the associations between IOFC and HLCC, particularly when including the year effect within our model.

However, it should be noted that we used existing data collected in the field which were not specifically designed for our research purposes. Consequently, these extensive datasets required editing before they could effectively be used in our research. The data editing process can be subjective, influenced by personal preferences (including research background and habits) and research objectives. Even when individuals are collaborating on the same project, different filters may be applied during the data editing process, resulting in variations in the final dataset. It is possible that during the data editing process, excluding what we perceived as extreme values could result in the loss of true observations. While working with this existing field data, time and efforts are needed to carefully observe the data and determine a sensible data editing procedure. It's undeniable that these extensive datasets hold substantial analytical potential. However, it's crucial to be careful during the data editing phase to avoid over-editing and potential loss of valuable information.

Nowadays, data is generated relatively inexpensive and should be seen as a by-product of current animal husbandry offering opportunities for dairy researchers. Working on existing dataset can save lots of time, labour and money to design an experiment for a specific research purpose. Our application of lactation curve modelling on existing datasets is a nice step in effectively applying real-world examples in the interdisciplinary domain of data and dairy science. It's important to note that we are actively embracing more open science practices. We have made all our methods, including data editing and model building, available in a public GitHub repository. While our data is not openly accessible, it can be obtained through a request to the data owner. All of these open science practices align with the principles of making data Findable, Accessible, Interoperable, and Reusable (**FAIR**) (Wilkinson et al., 2016). These practices contribute to the broader goal of enhancing the accessibility and transparency of research data and methodologies, which is a fundamental aspect of open science.

6.4. The definition of persistency

There are multiple measures of persistency (**Table 1**), aiming to quantify the declining rate of the lactation curve after the peak yield from different perspectives. The absence of a standardized measurement method for persistency may hinder comparability, reliability, reproducibility, and generalizability of research findings. All these measures require the transformation of raw milk data (Togashi and Lin, 2009; Yamazaki et al., 2011b; Burgers et al., 2021).

Table 1 List of common definition of persistency in previous studies.

Measure of persistency	Formular	Reference
Difference in milk (kg)	$MY240 - MY60$	Yamazaki et al., 2011b
	$MY240 - MY60 + 100$	Togashi et al., 2016
	$MY305 - MY60$	Lehmann et al., 2019
Daily milk reduction (kg/day)	$\frac{MY260 - MY100}{160}$	Chen et al., 2016
	$\frac{MY200 - MY100}{100}$	Burgers et al., 2021
	$\frac{MY_{7d\ before\ dry-off} - MY100}{DIM_{7d\ before\ dry-off} - 100}$	Burgers et al., 2021
	$\frac{MY60 - MY270}{210}$	Adediran et al., 2012
	$\frac{MY280 - MY90}{190}$	Inchaisri et al., 2010
Dimensionless quantity	$c^{-(b+1)}$	Wood, 1967
Half-life of milk production (days)	$\frac{0.693}{decay}$	Ehrlich, 2011

The selection of a specific method for measuring persistency depends on the research objectives. Simple measures of persistency (difference in milk and the daily milk reduction) are calculating the absolute difference and the slope of milk reduction, typically fixed at two selected timepoints in lactation (Togashi and Lin, 2009; Yamazaki et al., 2011b; Chen et al., 2016). These measures are straightforward and easy to interpret. These simpler methods involve either using the moving average of daily milk yield or estimating milk yield through lactation curve modelling, depending on the frequency of milk recording. In **Chapter 3**, we aimed to assess persistency at various timepoints before and after DIMc, utilizing daily milk yield data up to that specific timepoint. The earliest measurement was taken 30 days before

DIMc, which could potentially precede the peak day. Similarly, in **Chapter 2**, our earliest measurement for persistency was at DIM 50. In both cases, calculating the absolute milk difference or slope using the moving average of daily milk yield was not suitable for capturing the declining rate.

An alternative approach we could have used in our studies is employing lactation curve modelling to estimate milk yield in the declining stage of lactation (Yamazaki et al., 2011b; Chen et al., 2016), such as milk yield at DIM 100 and 200, followed by calculating the absolute milk difference or slope. In our studies, we opted not to calculate the absolute milk difference or slope using the estimated milk yield generated by MilkBot. This decision was primarily driven by the advantages offered by the MilkBot model. Notably, this model allows for a direct calculation of persistency using one of the LCC known as "decay." What made this approach particularly appealing is that the decay variable is already normally distributed, eliminating the necessity of transforming the right-skewed persistency into a new variable for further statistical analysis (Ehrlich, 2011, 2013). This streamlined the analysis and provided a more direct and suitable method for assessing persistency in our study.

For achieving a more accurate fit when dealing with a limited number of records in early lactation, the Milkbot model stands out as a strong choice. This model employs Bayesian statistics, offering a relatively reliable fitting of individual cow lactation data, even in situations where the data is sparse and subject to noise (Ehrlich, 2013). This robustness is achieved through the incorporation of prior information (i.e., the population mean lactation curve characteristics). However, the utilization of prior information is sensitive to the available data, particularly when dealing with a limited number of milking records. In our analysis conducted in **Chapters 4 and 5**, we relied on test-day data; however, not all lactations had complete test-day records spanning from the beginning to the end of the lactation period. Our investigation revealed that fitting results for lactations with sparse test-day records aligned with the prior information, which is expected given the scarcity of data points. Hence, it is important to carefully consider whether you want to include the fitting results from those lactations with only one or two milking records in your analysis.

It should also be noticed that the prior information mentioned above should be selected with caution, as is the case with all models based on Bayesian statistics (van de Schoot et al., 2021). We initiated the process by fitting our data to the MilkBot model using the default prior designed for US dairy cows. Subsequently, we calculated our own priors tailored to cows in the Netherlands. For future users considering the utilization of the MilkBot model, it is important to be mindful of the prior information, as customizing it for your specific dataset is a critical step to achieve accurate and meaningful results.

Furthermore, we came across instances of test-day records that exhibited deviations from the anticipated pattern. For instance, there were situations in which cows exhibited an extremely early peak day or milk yield that appeared to increase towards the end of the lactation, resulting in unexpected fitting outcomes (see **Figure 6.1**). Such records pose a challenge for fitting with conventional lactation models. It is important to acknowledge that these anomalous records can yield fitting results that appear unconventional. In such scenarios, the process of data editing plays a critical role in recognizing and mitigating the influence of these unusual fitting results to the greatest extent possible.

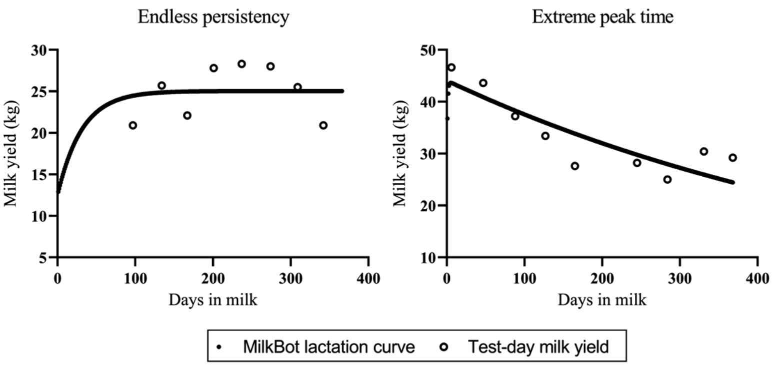


Figure 6.1 Example of lactations with anomalous milk records and the corresponding fitted lactation curve from MilkBot model.

6.5. Cow and herd level clustering in the data

Longitudinal data allowed us to observe several lactations from individual cows over the years. To account for the distinct lactation curve characteristics patterns between multiparous and primiparous cows (Wood, 1969; Horan et al., 2005; Ehrlich, 2013), we introduced the parity group (multiparous cows and primiparous cows) as a categorical variable in our applications. However, applying a parity group introduced an imbalance in our dataset, as some cows could have multiple lactations as multiparous cows, while they would only have one lactation as a primiparous cow. This imbalance raised concerns regarding potential bias in our results, as the group with more observations (multiparous cows) could dominate the model, potentially overshadowing the characteristics of the smaller group (primiparous cows). In this case, results might become less generalizable, as they may primarily reflect the larger group's characteristics and offer limited insights into the smaller group. We employed various methods to address this issue in each chapter. In **Chapter 2**, we included HLCC for each parity group (from the year preceding the selected insemination moments) and also added parity group (multiparous cows and primiparous cows) as a categorical variable in the prediction model. Additionally, we built separate prediction models for two parity groups to assess whether this imbalance affects our model's predictive capability. Nevertheless, the results yielded similar outcomes, indicating that this imbalance may not substantially impact our prediction model. In **Chapter 3**, we included a biologically relevant interaction term between DPC and the parity group into our association model. This interaction term enabled us to assess whether the relationship between DPC and persistency differs between multiparous and primiparous cows. By doing so, we aimed to mitigate potential bias arising from the dataset's imbalance and obtain a more accurate understanding of the association between DPC and persistency for both parity groups. In **Chapter 4**, we separately summarized the HLCC for each parity group to ensure that the HLCC values were not influenced by the imbalance in the data. In **Chapter 5**, we therefore included HLCC for both parity group into the association model and this enables a more nuanced understanding of how different parity groups of HLCC contributed to the association with IOFC.

Considering that the four main chapters of our thesis utilize longitudinal data collected from a large number of herds, with indicators derived from cows that are nested within these herds, it becomes essential to account for the herd effect. This is crucial for mitigating unobserved variations related to herds, including factors like environmental conditions and feed management, which are not directly measured but can significantly impact the outcomes of our analysis. The approach to correct for this herd effect varies depending on the research objective. In the association models presented in **Chapters 3 and 5**, we incorporated a herd variable as a random effect within the model to account for unobserved herd-related factors and inherent within-herd correlation. Associations found from models with random effects are not limited to the herds in our studies but can be generalized to a broader population of herds. In the prediction model described in **Chapter 2**, we handled the herd effect differently by including the variables HLCC and HM305 in all of our prediction models, rather than treating the herd as a random effect. This approach enabled us to apply our prediction model to unknown farms and effectively consider the influence of herd-level factors on the study outcomes.

6.6. Generalizability of our study to China

Throughout this thesis, we are using the data from the Netherlands and Belgium for the application of lactation curve modelling. I have consistently reminded myself of the importance of applying the knowledge and skills I acquire to benefit China in the future. Now, the question arises: Can the application of lactation curve modelling also be beneficial for the Chinese dairy industry? Before addressing this question, it is essential to gain understanding of the dairy industry in China.

The number of cows and total milk production both increased from 2000 to 2015 (**Figure 6.2**). During the adjustment period 2015 to 2017, the number of cows in

China was down from around 15.1 million in 2015 to about 10.8 million in 2017. The decline in the cattle population can be attributed to two primary factors: 1) strict environmental regulations on animal manure management and 2) farmers leaving the industry as a result of supply and demand imbalances, stemming from a significant surge in milk powder imports (Li et al., 2016; Wang et al., 2021). Given the dominant position of dairy processors in China's dairy sector, they have shifted their milk sourcing strategy from domestically produced raw milk to imported powder. This shift was driven by the economic advantage of using imported milk powder for manufacturing dairy products compared to using domestically produced raw milk. This change has led to a reduced demand for domestically sourced raw milk, which, in turn, has resulted in issues such as dumping raw milk and culling of dairy cows.

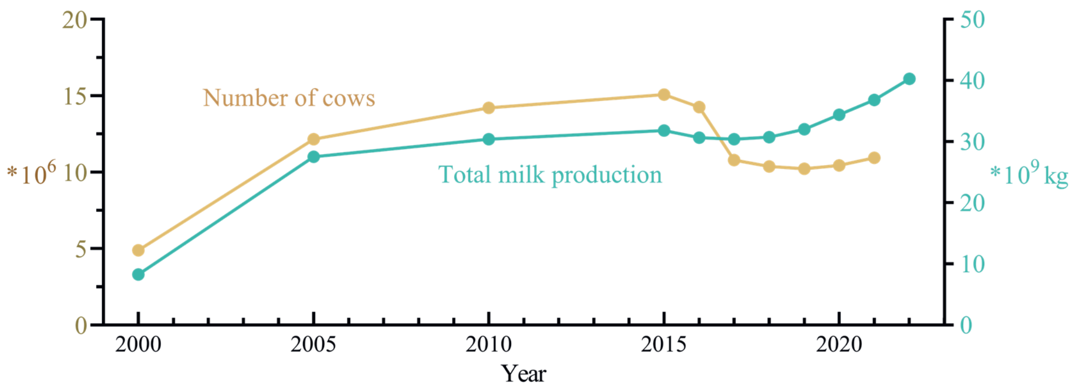


Figure 6.2 Overview of number of cows and total milk production in China from 2000 to 2022 (adapted from National Bureau of Statistics)

Although the number of dairy cattle did not increase or even declined since 2015, the total milk production was steady and supply has been guaranteed due to the improvement of the milk yield per cow and the efficiency of farming (**Figure 6.3**). The proportion of large scale dairy farms (more than 100 cows per herd) in China has increased year by year. It was 11.2% in 2005 and 72% in 2022. The average milk yield per cow per year was 2605 kg in 2000, 4760 kg in 2010, 8300 kg in 2020 and reached 9200 kg in 2022. This rate of increase was much higher than that observed in the Netherlands over the same period (**Chapter 1**). The Netherlands, as a leader in the dairy industry, has seen well-established and stable growth. In contrast, the Chinese dairy industry, starting later, possesses significant untapped potential. As a latecomer, the Chinese dairy industry has experienced rapid advancement, facilitated by factors such as genetic improvement (importing high-quality semen and cows), advanced equipment (TMR, milking parlors, milking robots), and enhanced management knowledge in areas like feeding, reproduction, and herd management (Li et al., 2016).

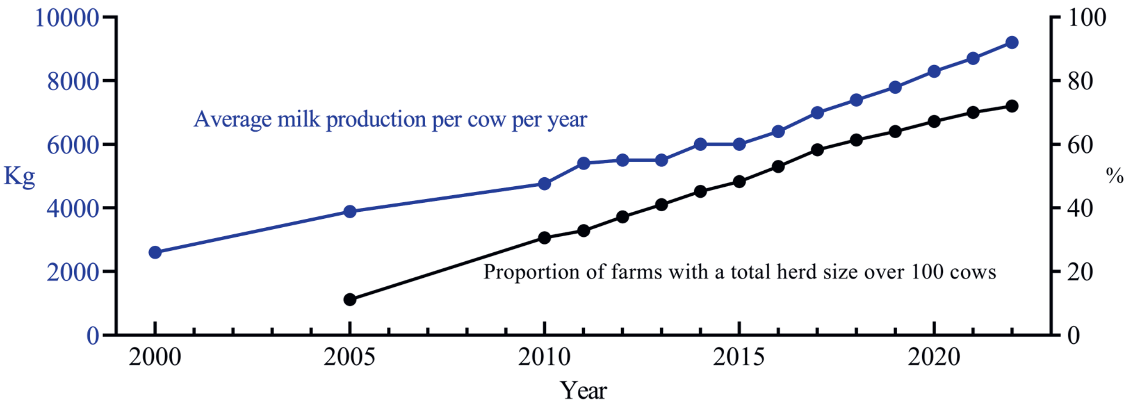


Figure 6.3 Overview of proportion of farms with over 100 cows and average milk production per cow per year in China from 2000 to 2022 (adapted from National Bureau of Statistics)

Milk production data is typically recorded by the milking system or through a dairy herd improvement system, providing the necessary data for modelling lactation curves. The method used to summarize HLCC and associate it with herd economics can be adapted for use with data from China. Given that milk revenue is the main income and feed cost represents a major cost in farms of China, similar to the situation in the Netherlands, it is reasonable to assume that the findings from **Chapter 5** can be generalized to the Chinese context.

China recognizes its role in promoting a more sustainable dairy industry and is actively pursuing this goal. Sustainable dairy farming involves optimizing resource utilization, enhancing resource efficiency, and reducing environmental impact. Currently, the requirements and pressure of environmental protection are increasing on China's dairy industry (PRC, 2017, 2020). From my perspective, persistency can be a valuable trait that may contribute to a more sustainable dairy industry in China. Persistent cows can contribute to more efficient milk production due to their extended lactation periods, which reduces the need for frequent calving and the associated resource inputs such as feed, land, and labour for raising replacement heifers (Dekkers et al., 1998; Hadley et al., 2006; Togashi et al., 2016). Additionally, persistent cows require fewer concentrates and consume a higher proportion of roughage in their diet, resulting in reduced feed costs (Sölkner and Fuchs, 1987). However, regarding the concept of persistency, it is relatively new in the Chinese dairy industry and has not yet been incorporated into the breeding system. Currently, the breeding system in China primarily focuses on factors such as milk production, fat and protein percentages, fat and protein yields, feet and legs score, body condition score, somatic cell score, and udder conformation score (National Animal Husbandry Central Station, 2022). I would therefore recommend the inclusion of persistency in the breeding system. By selectively breeding for more persistent cows, we might improve the milk production efficiency of the herd. Additionally, I suggest incorporating persistency as a key milk trait in individual cow management. This approach can lead to more informed breeding decisions, ultimately contributing to the advancement of precision dairy farming practices. By having more persistent cows and monitoring the cow persistency, it becomes possible to achieve the same level of milk production with fewer calvings. This has the potential to reduce the overall environmental footprint of the dairy industry in China.

Overall, I firmly believe that the application of lactation curve modelling holds significant promise for the Chinese dairy industry, offering numerous advantages and potential enhancements. An individual cow lactation curve modelling approach empowers dairy farmers with a deeper understanding of individual cows' milk production patterns, paving the way for data-driven decision-making. For instance, utilizing lactation curve modelling to assess cow persistency and optimize insemination decisions. However, the reliability of such applications is likely to be significantly bolstered as advanced technologies become more prevalent in dairy farming. These technologies may include robotic milking systems, automated feeding systems, and comprehensive data recording and analytics. With these advancements, we can look forward to an era of precision dairy cow farming in China. This transition will further amplify the benefits of lactation curve modelling and data-driven approaches, positioning the Chinese dairy industry for increased efficiency, improved sustainability, and enhanced productivity.

Reference list

Abdelkrim, A.B., L. Puillet, P. Gomes, and O. Martin. 2021. Lactation curve model with explicit representation of perturbations as a phenotyping tool for dairy livestock precision farming. Animal 15:100074. doi:10.1016/j.animal.2020.100074.

Adediran, S.A., D.A. Ratkowsky, D.J. Donaghy, and A.E.O. Malau-Aduli. 2012. Comparative evaluation of a new lactation curve model for pasture-based Holstein-Friesian dairy cows. J. Dairy Sci. 95:5344–5356. doi:10.3168/jds.2011-4663.

Adriaens, I., I. van den Brulle, L. D'Anvers, J.M.E. Statham, K. Geerinckx, S. De Vliegher, S. Piepers, and B. Aernouts. 2021. Milk losses and dynamics during perturbations in dairy cows differ with parity and lactation stage. J. Dairy Sci. 104:405–418. doi:10.3168/jds.2020-19195.

Albornoz, R.I., K. Giri, M.C. Hannah, and W.J. Wales. 2022. An improved approach to automated measurement of body condition score in dairy cows using a three-dimensional camera system. Animals 12. doi:10.3390/ani12010072.

Alqaisi, O., O. Asaah, and T. Hemme. 2011. Global View on Feed Cost and Feed Efficiency on Dairy Farms. Accessed September 5, 2022. https://www.allaboutfeed.net/animal-feed/feed-processing/global-view-on-feed-cost-and-feed-efficiency-on-dairy-farms/.

Alqaisi, O., L.E. Moraes, O.A. Ndambi, and R.B. Williams. 2019. Optimal dairy feed input selection under alternative feeds availability and relative prices. Inf. Process. Agric. 6:438–453. doi:10.1016/j.inpa.2019.03.004.

Van Amburgh, M.E., D.M. Galton, D.E. Bauman, and R.W. Everett. 1997. Management and economics of extended calving intervals with use of bovine somatotropin. Livest. Prod. Sci. 50:15–28. doi:10.1016/S0301-6226(97)00069-9.

Andersen, F., O. Østers, O. Reksen, and Y.T. Gröhn. 2011. Mastitis and the shape of the lactation curve in Norwegian dairy cows. J. Dairy Res. 78:23–31. doi:10.1017/S0022029910000749.

Arbel, R., Y. Bigun, E. Ezra, H. Sturman, and D. Hojman. 2001. The Effect of Extended Calving Intervals in High Lactating Cows on Milk Production and Profitability. J. Dairy Sci. 84:600–608. doi:10.3168/jds.S0022-0302(01)74513-4.

Atashi, H., A. Asaadi, and M. Hostens. 2021. Association between age at first calving and lactation performance, lactation curve, calving interval, calf birth weight, and dystocia in Holstein dairy cows. PLoS One 16:1–13. doi:10.1371/journal.pone.0244825.

Atashi, H., M. Salavati, J. De Koster, J. Ehrlich, M. Crowe, G. Opsomer, M. Hostens, N. McLoughlin, A. Fahey, E. Matthews, A. Santoro, C. Byrne, P. Rudd, R. O'Flaherty, S. Hallinan, C. Wathes, Z. Cheng, A. Fouladi, G. Pollott, D. Werling, B.S. Bernardo, A. Wylie, M. Bell, M. Vaneetvelde, K. Hermans, S. Moerman, H. Bogaert, J. Vandepitte, L. Vandevelde, B. Vanranst, J. Hoglund, S. Dahl, S. Ostergaard, J. Rothmann, M. Krogh, E. Meyer, C. Gaillard, J. Ettema, T. Rousing, F. Signorelli, F. Napolitano, B. Moioli, A. Crisà, L. Buttazzoni, J. McClure, D. Matthews, F. Kearney, A. Cromie, M. McClure, S. Zhang, X. Chen, H. Chen, J. Zhao, L. Yang, G. Hua, C. Tan, G. Wang, M. Bonneau, A. Pompozzi, A. Pearn, A. Evertson, L. Kosten, A. Fogh, T. Andersen, M. Lucey, C. Elsik, G. Conant, J. Taylor, N. Gengler, M. Georges, F. Colinet, M.R. Pamplona, H. Hammami, C. Bastin, H. Takeda, A. Laine, A.S. Van Laere, M. Schulze, S.P. Vera, C. Ferris, and C. Marchitelli. 2020. Genome-wide association for milk production and lactation curve parameters in Holstein dairy cows. J. Anim. Breed. Genet. 137:292–304. doi:10.1111/jbg.12442.

Atashi, H., M.J. Zamiri, and M. Dadpasand. 2013. Association between dry period length and lactation performance, lactation curve, calf birth weight, and dystocia in holstein dairy cows in iran. J. Dairy Sci. 96:3632–3638. doi:10.3168/jds.2012-5943.

Atzori, A.S., C. Valsecchi, E. Manca, F. Masoero, A. Cannas, and A. Gallo. 2021. Assessment of feed and economic efficiency of dairy farms based on multivariate aggregation of partial indicators measured on field. J. Dairy Sci. 104:12679–12692. doi:10.3168/jds.2020-19764.

Auldist, M.J., G. O'Brien, D. Cole, K.L. Macmillan, and C. Grainger. 2007. Effects of varying lactation length on milk production capacity of cows in pasture-based dairying systems. J. Dairy Sci. 90:3234–3241. doi:10.3168/jds.2006-683.

Bailey, K., T. Beck, E. Cowan, and V. Ishler. 2009. Managing income over feed cost.. Penn State Ext. 1–12.

Bailey, K., and V.A. Ishler. 2008. Tracking Milk Prices and Feed Costs. Accessed September 5, 2022. https://extension.psu.edu/tracking-milk-prices-and-feed-costs.

Bailey, K.W., C.M. Jones, and A.J. Heinrichs. 2005. Economic returns to Holstein and Jersey herds under multiple component pricing. J. Dairy Sci. 88:2269–2280. doi:10.3168/jds.S0022-0302(05)72903-9.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using {lme4}. J. Stat. Softw. 67:1–48. doi:10.18637/jss.v067.i01.

Bertulat, S., C. Fischer-Tenhagen, V. Suthar, E. Möstl, N. Isaka, and W. Heuwieser. 2013. Measurement of fecal glucocorticoid metabolites and evaluation of udder characteristics to estimate stress after sudden dry-off in dairy cows with different milk yields. J. Dairy Sci. 96:3774–3787. doi:10.3168/jds.2012-6425.

Bijl, R., S.R. Kooistra, and H. Hogeveen. 2007. The profitability of automatic milking on Dutch dairy farms. J. Dairy Sci. 90:239–248. doi:10.3168/jds.S0022-0302(07)72625-5.

Bohmanova, J., J. Jamrozik, and F. Miglior. 2009. Effect of pregnancy on production traits of Canadian Holstein cows. J. Dairy Sci. 92:2947–2959. doi:10.3168/jds.2008-1782.

Bormann, J., G.R. Wiggans, T. Druet, and N. Gengler. 2002. Estimating effects of permanent environment, lactation stage, age, and pregnancy on test-day yield. J. Dairy Sci. 85:263.e1-263.e21. doi:10.3168/jds.S0022-0302(02)74076-9.

van den Borne, B.H.P., T. Halasa, G. van Schaik, H. Hogeveen, and M. Nielen. 2010. Bioeconomic modeling of lactational antimicrobial treatment of new bovine subclinical intramammary infections caused by contagious pathogens. J. Dairy Sci. 93:4034–4044. doi:10.3168/jds.2009-3030.

Bouallègue, M., and N. M'hamdi. 2019. Mathematical Modeling of Lactation Curves: A Critical Review of Used Models.

Bozic, M., J. Newton, C.S. Thraen, and B.W. Gould. 2012. Mean-reversion in income over feed cost margins: Evidence and implications for managing margin risk by US dairy producers. J. Dairy Sci. 95:7417–7428. doi:10.3168/jds.2012-5818.

Brody, S., A.C. Ragsdale, and C.W. Turner. 1923. The rate of decline of milk secretion with the advance of the period of lactation. J. Gen. Physiol. 5:441–444. doi:10.1085/jgp.5.4.441.

Brody, S., C.W. Turner, and A.C. Ragsdale. 1924. The relation between the initial rise and the subsequent decline of milk secretion following parturition. J. Gen. Physiol. 6:541–545. doi:10.1085/jgp.6.5.541.

Brotherstone, S., R. Thompson, and I.M.S. White. 2004. Effects of pregnancy on daily milk yield of Holstein-Friesian dairy cattle. Livest. Prod. Sci. 87:265–269. doi:10.1016/j.livprodsci.2003.07.014.

Browne, N.A., R. Behrendt, R.S. Kingwell, and R.J. Eckard. 2015. Does producing more product over a lifetime reduce greenhouse gas emissions and increase profitability in dairy and wool enterprises?. Anim. Prod. Sci. 55:49–55. doi:10.1071/AN13188.

Buckley, F., K. O'Sullivan, J.F. Mee, R.D. Evans, and P. Dillon. 2003. Relationships among milk yield, body condition, cow weight, and reproduction in spring-calved Holstein-Friesians. J. Dairy Sci. 86:2308–2319. doi:10.3168/jds.S0022-0302(03)73823-5.

Burgers, E.E.A., A. Kok, R.M.A. Goselink, H. Hogeveen, B. Kemp, and A.T.M. van Knegsel. 2020. Fertility and milk production on commercial dairy farms with customized lactation lengths. J. Dairy Sci.. doi:10.3168/jds.2019-17947.

Burgers, E.E.A., A. Kok, R.M.A. Goselink, H. Hogeveen, B. Kemp, and A.T.M. van Knegsel. 2021. Effects of extended voluntary waiting period from calving until first insemination on body condition, milk yield, and lactation persistency. J. Dairy Sci.. doi:10.3168/jds.2020-19914.

Burgers, E.E.A., A. Kok, R.M.A. Goselink, H. Hogeveen, B. Kemp, and A.T.M. van Knegsel. 2022. Revenues and costs of dairy cows with different voluntary waiting periods based on data of a randomized control trial.. J. Dairy Sci.. doi:10.3168/jds.2021-20707.

Butler, W.R. 2005. Nutrition, negative energy balance and fertility in the postpartum dairy cow. Cattle Pract. 13:13–18.

Buza, M.H., L.A. Holden, R.A. White, and V.A. Ishler. 2014a. Evaluating the effect of ration composition on income over feed cost and milk yield. J. Dairy Sci. 97:3073–3080. doi:10.3168/jds.2013-7622.

Buza, M.H., L.A. Holden, R.A. White, and V.A. Ishler. 2014b. Evaluating the effect of ration composition on income over feed cost and milk yield. J. Dairy Sci. 97:3073–3080. doi:10.3168/jds.2013-7622.

van Calker, K.J. 2005. Sustainability of Dutch Dairy Farming Systems : A Modelling Approach.

Chen, J., A. Kok, G.J. Remmelink, J.J. Gross, R.M. Bruckmaier, B. Kemp, and A.T.M. van Knegsel. 2016. Effects of dry period length and dietary energy source on lactation curve characteristics over 2 subsequent lactations. J. Dairy Sci. 99:9287–9299. doi:10.3168/jds.2016-11253.

Chen, Y., M. Hostens, M. Nielen, J. Ehrlich, and W. Steeneveld. 2022a. Herd level economic comparison between the shape of the lactation curve and 305 d milk production. Front. Vet. Sci. 9. doi:10.3389/fvets.2022.997962.

Chen, Y., M. Hostens, M. Nielen, J. Ehrlich, and W. Steeneveld. 2022b. An empirical analysis of economic herd performance in relation to herd lactation curve characteristics. Pages 728–736 in European Conference on Precision Livestock Farming & International Conference on Precision Dairy Farming 2022.

Cobby, J.M., and Y.L.P. Le Du. 1978. On fitting curves to lactation data. Anim. Prod. 26:127–133. doi:10.1017/S0003356100039532.

Cole, J.B., J.L. Ehrlich, and D.J. Null. 2012. Short communication: Projecting milk yield using best prediction and the MilkBot lactation model. J. Dairy Sci. 95:4041–4044. doi:10.3168/jds.2011-4905.

Cole, J.B., and D.J. Null. 2009. Genetic evaluation of lactation persistency for five breeds of dairy cattle. J. Dairy Sci. 92:2248–2258. doi:10.3168/jds.2008-1825.

Cole, J.B., and P.M. VanRaden. 2006a. Genetic evaluation and best prediction of lactation persistency. J. Dairy Sci. 89:2722–2728. doi:10.3168/jds.S0022-0302(06)72348-7.

Cole, J.B., and P.M. VanRaden. 2006b. Genetic evaluation and best prediction of lactation persistency. J. Dairy Sci. 89:2722–2728. doi:10.3168/jds.S0022-0302(06)72348-7.

Cowley, F.C., T.M. Syahniar, D. Ratnawati, D.E. Mayberry, Marsetyo, D. Pamungkas, and D.P. Poppi. 2020. Greater farmer investment in well-formulated diets can increase liveweight gain and smallholder gross margins from cattle fattening. Livest. Sci. 242:104297. doi:10.1016/j.livsci.2020.104297.

CRV. 2022. Jaarstatistieken 2022. Arnhem,the Netherlands.

Davidson, R., and J.G. MacKinnon. 1981. Several Tests for Model Specification in the Presence of Alternative Hypotheses. Econometrica 49:781. doi:10.2307/1911522.

Dekkers, J.C.M., J.H. Ten Hag, and A. Weersink. 1998. Economic aspects of persistency of lactation in dairy cattle. Livest. Prod. Sci. 53:237–252. doi:10.1016/S0301-6226(97)00124-3.

Dekkers, J.C.M., J. Jamrozik, J.H. ten Hag, L.R. Schaeffer, and A. Weersink. 1996. Genetic and economic evaluation of persistency in dairy cattle. Interbull Bull. 12:97–102.

Dematawewa, C.M.B., R.E. Pearson, and P.M. VanRaden. 2007. Modeling extended lactations of Holsteins. J. Dairy Sci. 90:3924–3936. doi:10.3168/jds.2006-790.

Demeter, R.M., A.R. Kristensen, J. Dijkstra, A.G.J.M. Oude Lansink, M.P.M. Meuwissen, and J.A.M. van Arendonk. 2011. A multi-level hierarchic Markov process with Bayesian updating for herd optimization and simulation in dairy cattle. J. Dairy Sci. 94:5938–5962. doi:10.3168/jds.2011-4258.

Deng, Z., H. Hogeveen, T.J.G.M. Lam, R. van der Tol, and G. Koop. 2020. Performance of Online Somatic Cell Count Estimation in Automatic Milking Systems. Front. Vet. Sci. 7. doi:10.3389/fvets.2020.00221.

Dhanoa, M.S. 1981. A note on an alternative form of the lactation model of Wood. Anim. Sci. TA - TT - 32:349–351. doi:10.1017/S0003356100027276 LK - https://utrechtuniversity.on.worldcat.org/oclc/4669310650.

Dijkhuizen, A.A., and R.S. Morris. 1996. Animal Health Economics: Principles and Applications 265–273.

Dong, F., X. Du, and B.W. Gould. 2011. Milk price volatility and its determinants. AAEA NAREA Jt. Annu. Meet. Pittsbg. 24–26.

Ehrlich, J.L. 2011. Quantifying shape of lactation curves , and benchmark curves for common dairy breeds and parities. Bov. Pract. 45:88–96. doi:https://doi.org/10.21423/bovine-vol45no1p88-95.

Ehrlich, J.L. 2013. Quantifying inter-group variability in lactation curve shape and magnitude with the MilkBot® lactation model. PeerJ 2013:1–18. doi:10.7717/peerj.54.

Elahi Torshizi, M. 2016. Effects of season and age at first calving on genetic and phenotypic characteristics of lactation curve parameters in Holstein cows. J. Anim. Sci. Technol. 58:1–14. doi:10.1186/s40781-016-0089-1.

European Commission. 2021. Milk Market Observatory. Accessed September 5, 2022. https://ec.europa.eu/info/food-farming-fisheries/farming/facts-and-figures/markets/overviews/market-observatories/milk_en#globan-dropdown-6kovo54wpha.

FAO. 2013. Milk and Dairy Products in Human Nutrition. E. Muehlhof, A. Bennett, and D. McMahon, ed. Food and agriculture organization of the united nations, Rome.

Ferguson, J.D., D.K. Beede, R.D. Shaver, C.E. Polan, J.T. Huber, and P.T. Chandler. 2000. Effects of inclusion of a blended protein product in 35 dairy herds in five regions of the country. J. Dairy Sci. 83:1813–1828. doi:10.3168/jds.S0022-0302(00)75052-1.

Ferguson, J.D., and A. Skidmore. 2013. Reproductive performance in a select sample of dairy herds. J. Dairy Sci. 96:1269–1289. doi:10.3168/jds.2012-5805.

Fischer, A. 1958. Research with Württemberg spotted mountain cows on the shape ofthe lactation curve and how it may be influenced by non-genetic factors. Ziichtungskunde 30:296–304.

Fitzgerald, A.M., D.P. Ryan, and D.P. Berry. 2015. Factors associated with the differential in actual gestational age and gestational age predicted from transrectal ultrasonography in pregnant dairy cows. Theriogenology 84:358–364. doi:10.1016/j.theriogenology.2015.03.023.

Fleischer, P., M. Metzner, M. Beyerbach, M. Hoedemaker, and W. Klee. 2001. The relationship between milk yield and the incidence of some diseases in dairy cows. J. Dairy Sci. 84:2025–2035. doi:10.3168/jds.S0022-0302(01)74646-2.

Gaillard, C., N.C. Friggens, M. Taghipoor, M.R. Weisbjerg, J.O. Lehmann, and J. Sehested. 2016. Effects of an individual weight-adjusted feeding strategy in early lactation on milk production of Holstein cows during extended lactation. J. Dairy Sci. 99:2221–2236. doi:10.3168/jds.2015-10359.

Gargiulo, J.I., C.R. Eastwood, S.C. Garcia, and N.A. Lyons. 2018. Dairy farmers with larger herd sizes adopt more precision dairy technologies. J. Dairy Sci. 101:5466–5473. doi:10.3168/jds.2017-13324.

Gelder, K. van. 2022a. Average Milk Yield per Cow in Belgium 2010-2021. Accessed February 21, 2023. https://www.statista.com/statistics/1097660/average-milk-yield-per-cow-in-belgium/.

Gelder, K. van. 2022b. Average Milk Yield per Cow in the Netherlands 2000-2021. Accessed February 21, 2023. https://www.statista.com/statistics/1097599/average-milk-yield-per-cow-in-the-netherlands/.

Gengler, N. 1996. Persistency of lactation yields : a review. Interbull Bull. 12:87–96.

Gorgulu, O. 2012. Prediction of 305-day milk yield in Brown Swiss cattle using artificial neural networks. S. Afr. J. Anim. Sci. 42. doi:10.4314/sajas.v42i3.10.

Grayaa, M., S. Vanderick, B. Rekik, A. Ben Gara, C. Hanzen, S. Grayaa, R. Reis Mota, H. Hammami, and N. Gengler. 2019. Linking first lactation survival to milk yield and components and lactation persistency in Tunisian Holstein cows. Arch. Anim. Breed. 62:153–160. doi:10.5194/aab-62-153-2019.

Green, L.E., V.J. Hedges, Y.H. Schukken, R.W. Blowey, and A.J. Packington. 2002. The impact of clinical lameness on the milk yield of dairy cows. J. Dairy Sci. 85:2250–2256. doi:10.3168/jds.S0022-0302(02)74304-X.

Guetouache, M., Guessas, Bettache, Medjekal, and Samir. 2014. Composition and nutritional value of raw milk. Issues Biol. Sci. Pharm. Res. 2:115–122.

Hadley, G.L., C.A. Wolf, and S.B. Harsh. 2006. Dairy cattle culling patterns, explanations, and implications. J. Dairy Sci. 89:2286–2296. doi:10.3168/jds.S0022-0302(06)72300-1.

Hardie, C.A., M. Wattiaux, M. Dutreuil, R. Gildersleeve, N.S. Keuler, and V.E. Cabrera. 2014. Feeding strategies on certified organic dairy farms in Wisconsin and their effect on milk production and income over feed costs. J. Dairy Sci. 97:4612–4623. doi:10.3168/jds.2013-7763.

van Hoeij, R.J., J. Dijkstra, R.M. Bruckmaier, J.J. Gross, T.J.G.M. Lam, G.J. Remmelink, B. Kemp, and A.T.M. van Knegsel. 2017. Consequences of dietary energy source and energy level on energy balance, lactogenic hormones, and lactation curve characteristics of cows after a short or omitted dry period. J. Dairy Sci. 100:8544–8564. doi:10.3168/jds.2017-12855.

Hogenboom, J.A., L. Pellegrino, A. Sandrucci, V. Rosi, and P. D'Incecco. 2019. Invited review: Hygienic quality, composition, and technological performance of raw milk obtained by robotic milking of cows. J. Dairy Sci. 102:7640–7654. doi:10.3168/jds.2018-16013.

Horan, B., P. Dillon, D.P. Berry, P. O'Connor, and M. Rath. 2005. The effect of strain of Holstein-Friesian, feeding system and parity on lactation curves characteristics of spring-calving dairy cows. Livest. Prod. Sci. 95:231–241. doi:10.1016/j.livprodsci.2004.12.021.

Hostens, M., J. Ehrlich, B. Van Ranst, and G. Opsomer. 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. J. Dairy Sci. 95:2988–3007. doi:10.3168/jds.2011-4791.

Hut, P.R., M.M. Hostens, M.J. Beijaard, F.J.C.M. van Eerdenburg, J.H.J.L. Hulsen, G.A. Hooijer, E.N. Stassen, and M. Nielen. 2021. Associations between body condition score, locomotion score, and sensor-based time budgets of dairy cattle during the dry period and early lactation. J. Dairy Sci. 104:4746–4763. doi:10.3168/jds.2020-19200.

ICAR. 2017. Procedure 2 of Section 2 of ICAR Guidelines Computing of Accumulated Lactation Yield. ICAR Guidel. 5–15.

Inchaisri, C., H. Hogeveen, P.L.A.M. Vos, G.C. van der Weijden, and R. Jorritsma. 2010. Effect of milk yield characteristics, breed, and parity on success of the first insemination in Dutch dairy cows. J. Dairy Sci. 93:5179–5187. doi:10.3168/jds.2010-3234.

Inchaisri, C., R. Jorritsma, P.L.A.M. Vos, G.C. van der Weijden, and H. Hogeveen. 2011a. Analysis of the economically optimal voluntary waiting period for first insemination.. J. Dairy Sci. 94:3811–3823. doi:10.3168/jds.2010-3790.

Inchaisri, C., R. Jorritsma, P.L.A.M. Vos, G.C. van der Weijden, and H. Hogeveen. 2011b. Analysis of the economically optimal voluntary waiting period for first insemination. J. Dairy Sci. 94:3811–3823. doi:10.3168/jds.2010-3790.

Inchaisri, C., A. De Vries, R. Jorritsma, and H. Hogeveen. 2012. Improved knowledge about conception rates influences the decision to stop insemination in dairy cows. Reprod. Domest. Anim. 47:820–826. doi:10.1111/j.1439-0531.2011.01975.x.

Ingvartsen, K.L., R.J. Dewhurst, and N.C. Friggens. 2003. On the relationship between lactational performance and health: Is it yield or metabolic imbalance that cause production diseases in dairy cattle? A position paper. Livest. Prod. Sci. 83:277–308. doi:10.1016/S0301-6226(03)00110-6.

Jenkins, T.G., and C.L. Ferrell. 1984. A note on lactation curves of crossbred cows. Anim. Prod. 39:479–482. doi:10.1017/S0003356100032232.

Jingar, S., R.K. Mehla, M. Singh, and A.K. Roy. 2014. Lactation Curve Pattern and Prediction of Milk Production Performance in Crossbred Cows. J. Vet. Med. 2014:1–6. doi:10.1155/2014/814768.

Jonathan Rushton. 2009. The economics of animal health and production.pdf.

Jongeneel, R., C. Daatselaar, M. van Leeuwen, and H. Silvis. 2017. Phosphate Production Reduction Decree of the Netherlands.

Kawashima, C., M. Matsui, T. Shimizu, K. Kida, and A. Miyamoto. 2012. Nutritional factors that regulate ovulation of the dominant follicle during the first follicular wave postpartum in high-producing dairy cows. J. Reprod. Dev. 58:10–16. doi:10.1262/jrd.11-139N.

Kay, J.K., W.J. Weber, C.E. Moore, D.E. Bauman, L.B. Hansen, H. Chester-Jones, B.A. Crooker, and L.H. Baumgard. 2005. Effects of week of lactation and genetic selection for milk yield on milk fatty acid composition in Holstein cows. J. Dairy Sci. 88:3886–3893. doi:10.3168/jds.S0022-0302(05)73074-5.

Kjeldsen, A.M.H., V. Thorup, S. Østergaard, and L.A.H. Nielsen. 2022. Decision parameters for individual lactation length of dairy cows. Pages 737–743 in Precision livestock farming 2022. University of Veterinary Medicine Vienna, Vienna, Austria.

Klootwijk, C.W., C.E. Van Middelaar, P.B.M. Berentsen, and I.J.M. de Boer. 2016. Dutch dairy farms after milk quota abolition: Economic and environmental consequences of a new manure policy. J. Dairy Sci. 99:8384–8396. doi:10.3168/jds.2015-10781.

Van Knegsel, A.T.M., H. Van Den Brand, J. Dijkstra, W.M. Van Straalen, R. Jorritsma, S. Tamminga, and B. Kemp. 2007. Effect of glucogenic vs. lipogenic diets on energy balance, blood metabolites, and reproduction in primiparous and multiparous dairy cows in early lactation. J. Dairy Sci. 90:3397–3409. doi:10.3168/jds.2006-837.

van Knegsel, A.T.M., E.E.A. Burgers, J. Ma, R.M.A. Goselink, and A. Kok. 2022. Extending Lactation Length: Consequences for Cow, Calf and Farmer. J. Anim. Sci. 1–10. doi:10.1093/jas/skac220.

Kok, A., J.O. Lehmann, B. Kemp, H. Hogeveen, C.E. Van Middelaar, I.J.M. De Boer, and A.T.M. Van Knegsel. 2019. Production, partial cash flows and greenhouse gas emissions of simulated dairy herds with extended lactations. Animal 13:1074–1083. doi:10.1017/S1751731118002562.

de Koning, C.J.A.M. (Kees). 2010. Automatic Milking – Common Practice on Dairy Farms. Page 16 in The First North American Conference on Precision Dairy Management. Wageningen UR Livestock Research, Lelystad, the Netherlands.

Kristensen, E., S. Østergaard, M.A. Krogh, and C. Enevoldsen. 2008. Technical indicators of financial performance in the dairy herd. J. Dairy Sci. 91:620–631. doi:10.3168/jds.2007-0201.

Laroche, J.-P., R. Roy, and K. Bergeron. 2020. Per Cow, Per Hectolitre, or Per Kg Milk Fat: Which Is the Best Indicator of Margin? Accessed February 1, 2022. https://lactanet.ca/en/per-cow-per-hectolitre-or-per-kg-milk-fat-which-is-the-best-indicator-of-margin/.

Lazenby, D., E. Bohlsen, K.J. Hand, D.F. Kelton, F. Miglior, and K.D. Lissemore. 2003. Methods to estimate 24-hour yields for milk, fat and protein in robotic milking herds.. Perform. Rec. Anim. state art 2002. Proc. 33rd Bienn. Sess. ICAR, Interlaken, Switzerland, 26-31 May, 2002 65–71.

LeBlanc, S.J., K.D. Lissemore, D.F. Kelton, T.F. Duffield, and K.E. Leslie. 2006. Major advances in disease prevention in dairy cattle. J. Dairy Sci. 89:1267–1279. doi:10.3168/jds.S0022-0302(06)72195-6.

Leclerc, H., D. Duclos, A. Barbat, T. Druet, and V. Ducrocq. 2008. Environmental effects on lactation curves included in a test-day model genetic evaluation. Animal 2:344–353. doi:10.1017/S175173110700119X.

Lehmann, J.O., J.G. Fadel, L. Mogensen, T. Kristensen, C. Gaillard, and E. Kebreab. 2016. Effect of calving interval and parity on milk yield per feeding day in Danish commercial dairy herds. J. Dairy Sci. 99:621–633. doi:10.3168/jds.2015-9583.

Lehmann, J.O., L. Mogensen, and T. Kristensen. 2014. Extended lactations may improve cow health, productivity and reduce greenhouse gas emissions from organic dairy production. Org. Agric. 4:295–299. doi:10.1007/s13165-014-0070-6.

Lehmann, J.O., L. Mogensen, and T. Kristensen. 2017. Early lactation production, health, and welfare characteristics of cows selected for extended lactation. J. Dairy Sci. 100:1487–1501. doi:10.3168/jds.2016-11162.

Lehmann, J.O., L. Mogensen, and T. Kristensen. 2019a. Extended lactations in dairy production: Economic, productivity and climatic impact at herd, farm and sector level. Livest. Sci. 220:100–110. doi:10.1016/j.livsci.2018.12.014.

Lehmann, J.O., L. Mogensen, and T. Kristensen. 2019b. Extended lactations in dairy production: Economic, productivity and climatic impact at herd, farm and sector level. Livest. Sci. 220:100–110. doi:10.1016/j.livsci.2018.12.014.

Lenth, R., P. Buerkner, M. Herve, J. Love, H. Riebl, and H. Singmann. 2020. Emmeans: Estimated Marginal Means, Aka Least-Squares Means.. Aka Least-Squares Means. https://cran.r-project.org/package=emmeans.

Li, S., S. Li, Y. Li, C. Liu, X. Liu, C. Luan, Y. Wang, J. Xia, Y. Yang, K. Yao, and L. Zhang. 2016. White Paper on China Dairy (2016).

Linde, R. Van Der, A. Groen, and G. De Jong. 2000. Estimation of genetic parameters for persistency of milk production in dairy cattle. Page 113 in Interbull Bulletin.

Liseune, A., M. Salamone, D. Van Den Poel, and B. Van Ranst. 2021. Predicting the milk yield curve of dairy cows in the subsequent lactation period using deep learning. Comput. Electron. Agric. 180:105904. doi:10.1016/j.compag.2020.105904.

Liseune, A., M. Salamone, D. Van den Poel, B. Van Ranst, and M. Hostens. 2020. Leveraging latent representations for milk yield prediction and interpolation using deep learning. Comput. Electron. Agric. 175:105600. doi:10.1016/j.compag.2020.105600.

Loker, S., F. Miglior, J. Bohmanova, J. Jamrozik, and L.R. Schaeffer. 2009. Phenotypic analysis of pregnancy effect on milk, fat, and protein yields of Canadian Ayrshire, Jersey, Brown Swiss, and Guernsey breeds. J. Dairy Sci. 92:1300–1312. doi:10.3168/jds.2008-1425.

Lu, H., and H. Bovenhuis. 2020. Phenotypic and genetic effects of pregnancy on milk production traits in Holstein-Friesian cattle. J. Dairy Sci. 103:11597–11604. doi:10.3168/jds.2020-18561.

M.J., A., O. G., C. D., M. K.L., and G. C. 2007. Effects of varying lactation length on milk production capacity of cows in pasture-based dairying systems. J. Dairy Sci. 90:3234–3241.

Manca, E., A. Cesarani, N.P.P. Macciotta, A.S. Atzori, G. Pulina, and C. Dimauro. 2020. Use of discriminant statistical procedures for an early detection of persistent lactations in dairy cows. Comput. Electron. Agric. 176:105657. doi:10.1016/j.compag.2020.105657.

Martins, B.M., A.L.C. Mendes, L.F. Silva, T.R. Moreira, J.H.C. Costa, P.P. Rotta, M.L. Chizzotti, and M.I. Marcondes. 2020. Estimating body weight, body condition score, and type traits in dairy cows using three dimensional cameras and manual body measurements. Livest. Sci. 236. doi:10.1016/j.livsci.2020.104054.

Marumo, J.L., D. Lusseau, J.R. Speakman, M. Mackie, and C. Hambly. 2022. Influence of environmental factors and parity on milk yield dynamics in barn-housed dairy cattle. J. Dairy Sci.. doi:10.3168/jds.2021-20698.

Masia, F., G. Molina, C. Vissio, M. Balzarini, R.L. de la Sota, and M. Piccardi. 2022. Quantifying the negative impact of clinical diseases on productive and reproductive performance of dairy cows in central Argentina. Livest. Sci. 259. doi:10.1016/j.livsci.2022.104894.

Masía, F.M., N.A. Lyons, M. Piccardi, M. Balzarini, R.C. Hovey, and S.C. Garcia. 2020. Modeling variability of the lactation curves of cows in automated milking systems. J. Dairy Sci. 103:8189–8196. doi:10.3168/jds.2019-17962.

McInerney, J.P., K.S. Howe, and J.A. Schepers. 1992. A framework for the economic analysis of disease in farm livestock. Prev. Vet. Med. 13:137–154. doi:10.1016/0167-5877(92)90098-Z.

Mellado, M., E. Antonio-Chirino, C. Meza-Herrera, F.G. Veliz, J.R. Arevalo, J. Mellado, and A. de Santiago. 2011. Effect of lactation number, year, and season of initiation of lactation on milk yield of cows hormonally induced into lactation and treated with recombinant bovine somatotropin. J. Dairy Sci. 94:4524–4530. doi:10.3168/jds.2011-4152.

National Animal Husbandry Central Station. 2022. Genetic Evaluation Results of Chinese Dairy Bull. Beijing.

Nelder, A.J.A. 1966. Inverse Polynomials , a Useful Group of Multi-Factor Response Functions Published by : International Biometric Society Stable URL : https://www.jstor.org/stable/2528220 REFERENCES Linked references are available on JSTOR for this article : reference # ref 22:128–141.

Němečková, D., L. Stádník, and J. Čítek. 2015. Associations between milk production level, calving interval length, lactation curve parameters and economic results in Holstein cows. Mljekarstvo 65:243–250. doi:10.15567/mljekarstvo.2015.0404.

Niozas, G., G. Tsousis, C. Malesios, I. Steinhöfel, C. Boscos, H. Bollwein, and M. Kaske. 2019. Extended lactation in high-yielding dairy cows. II. Effects on milk production, udder health, and body measurements. J. Dairy Sci. 102:811–823. doi:10.3168/jds.2018-15117.

Nor, N.M., W. Steeneveld, and H. Hogeveen. 2014. The average culling rate of Dutch dairy herds over the years 2007 to 2010 and its association with herd reproduction, performance and health. J. Dairy Res. 81:1–8. doi:10.1017/S0022029913000460.

Odensten, M.O., B. Berglund, K.P. Waller, and K. Holtenius. 2007. Metabolism and udder health at dry-off in cows of different breeds and production levels. J. Dairy Sci. 90:1417–1428. doi:10.3168/jds.S0022-0302(07)71627-2.

Oleggini, G.H., L.O. Ely, and J.W. Smith. 2001. Effect of region and herd size on dairy herd performance parameters. J. Dairy Sci. 84:1044–1050. doi:10.3168/jds.S0022-0302(01)74564-X.

Olori, V.E., S. Brotherstone, W.G. Hill, and B.J. McGuirk. 1999. Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd. Livest. Prod. Sci. 58:55–63. doi:10.1016/S0301-6226(98)00194-8.

Österman, S., and J. Bertilsson. 2003. Extended calving interval in combination with milking two or three times per day: effects on milk production and milk composition. Livest. Prod. Sci. 82:139–149. doi:10.1016/S0301-6226(03)00036-8.

Peiter, M., H.N. Phillips, and M.I. Endres. 2021. Association between early postpartum rumination time and peak milk yield in dairy cows. J. Dairy Sci. 104:5898–5908. doi:10.3168/jds.2020-19698.

Penasa, M., M. De Marchi, and M. Cassandro. 2016. Short communication: Effects of pregnancy on milk yield, composition traits, and coagulation properties of Holstein cows. J. Dairy Sci. 99:4864–4869. doi:10.3168/jds.2015-10168.

Pereira, P.C. 2014. Milk nutritional composition and its role in human health. Nutrition 30:619–627. doi:10.1016/j.nut.2013.10.011.

Pietersma, D., R. Lacroix, D. Lefebvre, E. Block, and K.M. Wade. 2001. A case-acquisition and decision-support system for the analysis of group-average lactation curves. J. Dairy Sci. 84:730–739. doi:10.3168/jds.S0022-0302(01)74527-4.

Pinedo, P.J., A. De Vries, and D.W. Webb. 2010. Dynamics of culling risk with disposal codes reported by Dairy Herd Improvement dairy herds. J. Dairy Sci. 93:2250–2261. doi:10.3168/jds.2009-2572.

Pires, J.A.A., C. Delavaud, Y. Faulconnier, D. Pomiès, and Y. Chilliard. 2013. Effects of body condition score at calving on indicators of fat and protein mobilization of periparturient Holstein-Friesian cows. J. Dairy Sci. 96:6423–6439. doi:10.3168/jds.2013-6801.

PRC. 2017. Opinions of the General Office of the State Council on Accelerating the Promotion of Resource

Utilization of Livestock and Poultry Farming Waste. Accessed October 17, 2023. https://www.gov.cn/zhengce/content/2017-06/12/content_5201790.htm.

PRC. 2020. Opinions of the General Office of the State Council on Promoting High-Quality Development in Animal Husbandry. Accessed October 17, 2023. https://www.gov.cn/zhengce/content/2020-09/27/content_5547612.htm.

R Core Team. 2020. R: A Language and Environment for Statistical Computing.

Rajala-Schultz, P.J., J.S. Hogan, and K.L. Smith. 2005. Short communication: Association between milk yield at dry-off and probability of intramammary infections at calving. J. Dairy Sci. 88:577–579. doi:10.3168/jds.S0022-0302(05)72720-X.

Roche, J.R. 2003. Effect of pregnancy on milk production and bodyweight from identical twin study. J. Dairy Sci. 86:777–783. doi:10.3168/jds.S0022-0302(03)73659-5.

Roche, J.R., K.A. Macdonald, C.R. Burke, J.M. Lee, and D.P. Berry. 2007. Associations among body condition score, body weight, and reproductive performance in seasonal-calving dairy cattle. J. Dairy Sci. 90:376–391. doi:10.3168/jds.S0022-0302(07)72639-5.

Rodrigues, A.C.O., D.Z. Caraviello, and P.L. Ruegg. 2005. Management of Wisconsin dairy herds enrolled in milk quality teams. J. Dairy Sci. 88:2660–2671. doi:10.3168/jds.S0022-0302(05)72943-X.

Rowlands, G.J., S. Lucey, and A.M. Russell. 1982. A comparison of different models of the lactation curve in dairy cattle. Anim. Prod. 35:135–144. doi:10.1017/S0003356100000908.

Różańska-Zawieja, J., S. Winnicki, J. Zyprych-Walczak, A. Szabelska-Beręsewicz, I. Siatkowski, W. Nowak, B. Stefańska, R. Kujawiak, and Z. Sobek. 2021. Article the effect of feeding management and culling of cows on the lactation curves and milk production of primiparous dairy cows. Animals 11. doi:10.3390/ani11071959.

Rutten, C.J., C. Kamphuis, H. Hogeveen, K. Huijps, M. Nielen, and W. Steeneveld. 2017. Sensor data on cow activity, rumination, and ear temperature improve prediction of the start of calving in dairy cows. Comput. Electron. Agric. 132:108–118. doi:10.1016/j.compag.2016.11.009.

Rutten, C.J., W. Steeneveld, J.C.M. Vernooij, K. Huijps, M. Nielen, and H. Hogeveen. 2016. A prognostic model to predict the success of artificial insemination in dairy cows based on readily available data. J. Dairy Sci. 99:6764–6779. doi:10.3168/jds.2016-10935.

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

Salamone, M., I. Adriaens, A. Vervaet, G. Opsomer, H. Atashi, V. Fievez, B. Aernouts, and M. Hostens. 2022. Animal The international journal of animal biosciences Prediction of first test day milk yield using historical records in dairy cows. animal 16:100658. doi:10.1016/j.animal.2022.100658.

Santos, J.E.P., E.J. DePeters, P.W. Jardon, and J.T. Huber. 2001. Effect of prepartum dietary protein level on performance of primigravid and multiparous Holstein dairy cows. J. Dairy Sci. 84:213–224. doi:10.3168/jds.S0022-0302(01)74471-2.

van de Schoot, R., S. Depaoli, R. King, B. Kramer, K. Märtens, M.G. Tadesse, M. Vannucci, A. Gelman, D. Veen, J. Willemsen, and C. Yau. 2021. Bayesian statistics and modelling. Nat. Rev. Methods Prim. 1. doi:10.1038/s43586-020-00001-2.

Scott, T.A., B. Yandell, L. Zepeda, R.D. Shaver, and T.R. Smith. 1996. Use of Lactation Curves for Analysis of Milk Production Data. J. Dairy Sci.. doi:10.3168/jds.S0022-0302(96)76557-8.

Sehested, J., C. Gaillard, J.O. Lehmann, G.M. Maciel, M. Vestergaard, M.R. Weisbjerg, L. Mogensen, L.B. Larsen, N.A. Poulsen, and T. Kristensen. 2019. Review: extended lactation in dairy cattle. Animal 13:A65–A74. doi:10.1017/S1751731119000806.

Shahid, M.Q., J.K. Reneau, H. Chester-Jones, R.C. Chebel, and M.I. Endres. 2015. Cow- and herd-level risk factors for on-farm mortality in Midwest US dairy herds. J. Dairy Sci. 98:4401–4413.

doi:10.3168/jds.2014-8513.

Shim, E.H., R.D. Shanks, and D.E. Morin. 2004. Milk loss and treatment costs associated with two treatment protocols for clinical mastitis in dairy cows. J. Dairy Sci. 87:2702–2708. doi:10.3168/jds.S0022-0302(04)73397-4.

Sikka, L.C. 1950. A study of lactation as affected by heredity and environment. J. Dairy Res. 17:231–252. doi:10.1017/S0022029900005811.

Silvestre, A.M., F. Petim-Batista, and J. Colaço. 2006. The accuracy of seven mathematical functions in modeling dairy cattle lactation curves based on test-day records from varying sample schemes. J. Dairy Sci. 89:1813–1821. doi:10.3168/jds.S0022-0302(06)72250-0.

Sölkner, J., and W. Fuchs. 1987. A comparison of different measures of persistency with special respect to variation of test-day milk yields. Livest. Prod. Sci. 16:305–319. doi:10.1016/0301-6226(87)90001-7.

Sorensen, A., D.D. Muir, and C.H. Knight. 2008a. Extended lactation in dairy cows: Effects of milking frequency, calving season and nutrition on lactation persistency and milk quality. J. Dairy Res. 75:90–97. doi:10.1017/S0022029907002944.

Sorensen, A., D.D. Muir, and C.H. Knight. 2008b. Extended lactation in dairy cows: Effects of milking frequency, calving season and nutrition on lactation persistency and milk quality. J. Dairy Res. 75:90–97. doi:10.1017/S0022029907002944.

Sørensen, J.T., and S. Østergaard. 2003. Economic consequences of postponed first insemination of cows in a dairy cattle herd. Livest. Prod. Sci. 79:145–153. doi:10.1016/S0301-6226(02)00150-1.

Steeneveld, W., and H. Hogeveen. 2012. Economic consequences of immediate or delayed insemination of a cow in oestrus. Vet. Rec. 171:17. doi:10.1136/vr.100183.

Steeneveld, W., and H. Hogeveen. 2015. Characterization of Dutch dairy farms using sensor systems for cow management. J. Dairy Sci. 98:709–717. doi:10.3168/jds.2014-8595.

Steeneveld, W., H. Hogeveen, and A.G.J.M. Oude Lansink. 2015. Economic consequences of investing in sensor systems on dairy farms. Comput. Electron. Agric. 119:33–39. doi:10.1016/j.compag.2015.10.006.

Steeneveld, W., A.T.M. van Knegsel, G.J. Remmelink, B. Kemp, J.C.M. Vernooij, and H. Hogeveen. 2014. Cow characteristics and their association with production performance with different dry period lengths. J. Dairy Sci. 97:4922–4931. doi:10.3168/jds.2013-7859.

Steeneveld, W., J. Swinkels, and H. Hogeveen. 2007. Stochastic modelling to assess economic effects of treatment of chronic subclinical mastitis caused by Streptococcus uberis. J. Dairy Res. 74:459–467. doi:10.1017/S0022029907002828.

Steeneveld, W., L.W. Tauer, H. Hogeveen, and A.G.J.M. Oude Lansink. 2012. Comparing technical efficiency of farms with an automatic milking system and a conventional milking system. J. Dairy Sci. 95:7391–7398. doi:10.3168/jds.2012-5482.

Steeneveld, W., T. van Werven, H.W. Barkema, and H. Hogeveen. 2011. Cow-specific treatment of clinical mastitis: An economic approach. J. Dairy Sci. 94:174–188. doi:10.3168/jds.2010-3367.

Stefanon, B., M. Colitti, G. Gabai, C.H. Knight, and C.J. Wilde. 2002. Mammary apoptosis and lactation persistency in dairy animals. J. Dairy Res. 69:37–52. doi:10.1017/S0022029901005246.

Togashi, K., and C.Y. Lin. 2008. Genetic improvement of total milk yield and total lactation persistency of the first three lactations in dairy cattle. J. Dairy Sci. 91:2836–2843. doi:10.3168/jds.2007-0783.

Togashi, K., and C.Y. Lin. 2009. Economic weights for genetic improvement of lactation persistency and milk yield. J. Dairy Sci. 92:2915–2921. doi:10.3168/jds.2008-1647.

Togashi, K., K. Moribe, S. Iwama, S. Matsumoto, S. Yamaguchi, K. Adachi, T. Takahashi, S. Saito, T. Nobukuni, T. Yamazaki, and T. Ikeda. 2016. Genotype-by-environment interaction on genetic relationships between lactation persistency and conception measures in Japanese Holstein cows. Livest. Sci. 183:40–47.

doi:10.1016/j.livsci.2015.11.008.

Torshizi, M.E., M.H. Mashhadi, and H. Farhangfar. 2019. Different aspects of lactation persistency in dairy cows. Indian J. Anim. Sci. 89:607–614. doi:10.56093/ijans.v89i6.91098.

Val-Arreola, D., E. Kebreab, J. Dijkstra, and J. France. 2004. Study of the lactation curve in dairy cattle on farms in central Mexico. J. Dairy Sci. 87:3789–3799. doi:10.3168/jds.S0022-0302(04)73518-3.

VanRaden, P.M., C.M.B. Dematawewa, R.E. Pearson, and M.E. Tooker. 2006. Productive life including all lactations and longer lactations with diminishing credits. J. Dairy Sci. 89:3213–3220. doi:10.3168/jds.S0022-0302(06)72596-6.

Vargas, B., W.J. Koops, M. Herrero, and J.A.M. Van Arendonk. 2000. Modeling extended lactations of dairy cows. J. Dairy Sci. 83:1371–1380. doi:10.3168/jds.S0022-0302(00)75005-3.

Vibart, R.E., S.P. Washburn, J.T. Green, G.A. Benson, C.M. Williams, D. Pacheco, and N. Lopez-Villalobos. 2012. Effects of feeding strategy on milk production, reproduction, pasture utilization, and economics of autumn-calving dairy cows in eastern North Carolina. J. Dairy Sci. 95:997–1010. doi:10.3168/jds.2011-4755.

Vredenberg, I., R. Han, M. Mourits, H. Hogeveen, and W. Steeneveld. 2021. An Empirical Analysis on the Longevity of Dairy Cows in Relation to Economic Herd Performance. Front. Vet. Sci. 8:1–8. doi:10.3389/fvets.2021.646672.

De Vries, A. 2006. Economic value of pregnancy in dairy cattle. J. Dairy Sci. 89:3876–3885. doi:10.3168/jds.S0022-0302(06)72430-4.

Wall, E., M.P. Coffey, and G.E. Pollott. 2012. The effect of lactation length on greenhouse gas emissions from the national dairy herd. Animal 6:1857–1867. doi:10.1017/S1751731112000936.

Walsh, D.P., A.G. Fahey, F.J. Mulligan, and M. Wallace. 2021. Effects of herd fertility on the economics of sexed semen in a high-producing, pasture-based dairy production system. J. Dairy Sci. 104:3181–3196. doi:10.3168/jds.2020-18676.

Walter, L.L., T. Gärtner, E. Gernand, A. Wehrend, and K. Donat. 2022. Effects of Parity and Stage of Lactation on Trend and Variability of Metabolic Markers in Dairy Cows. Animals 12. doi:10.3390/ani12081008.

Wang, A., L.F. Brito, H. Zhang, R. Shi, L. Zhu, D. Liu, G. Guo, and Y. Wang. 2022. Exploring milk loss and variability during environmental perturbations across lactation stages as resilience indicators in Holstein cattle. Front. Genet. 13. doi:10.3389/fgene.2022.1031557.

Wang, Q., L. Wei, and W. Wang. 2021. REVIEW: Challenges and prospects for milk production in China after the 2008 milk scandal. Appl. Anim. Sci. 37:166–175. doi:10.15232/aas.2020-02074.

Wathes, D.C., Z. Cheng, N. Bourne, V.J. Taylor, M.P. Coffey, and S. Brotherstone. 2007a. Differences between primiparous and multiparous dairy cows in the inter-relationships between metabolic traits, milk yield and body condition score in the periparturient period. Domest. Anim. Endocrinol. 33:203–225. doi:10.1016/j.domaniend.2006.05.004.

Wathes, D.C., M. Fenwick, Z. Cheng, N. Bourne, S. Llewellyn, D.G. Morris, D. Kenny, J. Murphy, and R. Fitzpatrick. 2007b. Influence of negative energy balance on cyclicity and fertility in the high producing dairy cow. Theriogenology 68:232–241. doi:10.1016/j.theriogenology.2007.04.006.

Wickham, H. 2016. Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wickham, H., R. François, L. Henry, K. Müller, and D. Vaughan. 2023. dplyr: A Grammar of Data Manipulation.

Wilkinson, M., M. Dumontier, and I. Aalbersberg. 2016. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3 (2016). Sci. data 3:1–9.

Wilmink, J.B.M. 1987. Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. Livest. Prod. Sci. 16:335–348. doi:10.1016/0301-6226(87)90003-0.

Wolf, C.A. 2010. Understanding the milk-to-feed price ratio as a proxy for dairy farm profitability. J. Dairy Sci. 93:4942–4948. doi:10.3168/jds.2009-2998.

Wood, P.D.P. 1967. Algebraic model of the lactation curve in cattle. Nature 216:164–165. doi:10.1038/216164a0.

Wood, P.D.P. 1969. Factors affecting the shape of the lactation curve in cattle. Anim. Prod. 11:307–316. doi:10.1017/S0003356100026945.

Wood, P.D.P. 1974. A note on the estimation of total lactation yield from production on a single day. Anim. Prod. 19:393–396. doi:10.1017/S0003356100022972.

Workman, D. 2023. Top Milk Exporting Countries. Accessed August 31, 2023. https://www.worldstopexports.com/top-milk-exporting-countries/?expand_article=1.

Wu, Y., D. Liang, R.D. Shaver, and V.E. Cabrera. 2019. An income over feed cost nutritional grouping strategy. J. Dairy Sci. 102:4682–4693. doi:10.3168/jds.2018-15302.

Yamazaki, T., K. Hagiya, H. Takeda, T. Osawa, S. Yamaguchi, and Y. Nagamine. 2016. Effects of stage of pregnancy on variance components, daily milk yields and 305-day milk yield in Holstein cows, as estimated by using a test-day model. Animal 10:1263–1270. doi:10.1017/S1751731116000185.

Yamazaki, T., K. Hagiya, H. Takeda, S. Yamaguchi, T. Osawa, and Y. Nagamine. 2014. Genetic correlations among female fertility, 305-day milk yield and persistency during the first three lactations of Japanese Holstein cows. Livest. Sci. 168:26–31. doi:10.1016/j.livsci.2014.08.005.

Yamazaki, T., H. Takeda, A. Nishiura, Y. Sasai, N. Sugawara, and K. Togashi. 2011a. Phenotypic relationship between lactation persistency and change in body condition score in first-lactation Holstein cows. Asian-Australasian J. Anim. Sci. 24:610–615. doi:10.5713/ajas.2011.10271.

Yamazaki, T., H. Takeda, A. Nishiura, Y. Sasai, N. Sugawara, and K. Togashi. 2011b. Phenotypic relationship between lactation persistency and change in body condition score in first-lactation Holstein cows. Asian-Australasian J. Anim. Sci. 24:610–615. doi:10.5713/ajas.2011.10271.

Yamazaki, T., H. Takeda, A. Nishiura, and K. Togashi. 2009. Relationship between the lactation curve and udder disease incidence in different lactation stages in first-lactation Holstein cows. Anim. Sci. J. 80:636–643. doi:10.1111/j.1740-0929.2009.00695.x.

Yue, X., J. Wu, M. van der Voort, W. Steeneveld, and H. Hogeveen. 2022. Estimating the Effect of a Bovine Viral Diarrhea Virus Control Program: An Empirical Study on the Performance of Dutch Dairy Herds. Front. Vet. Sci. 9. doi:10.3389/fvets.2022.892928.

Yukun, S., H. Pengju, W. Yujie, C. Ziqi, L. Yang, D. Baisheng, L. Runze, and Z. Yonggen. 2019. Automatic monitoring system for individual dairy cows based on a deep learning framework that provides identification via body parts and estimation of body condition score. J. Dairy Sci. 102:10140–10151. doi:10.3168/jds.2018-16164.

Zamorano-Algandar, R., M.A. Sánchez-Castro, A.I. Hernández-Cordero, R.M. Enns, S.E. Speidel, M.G. Thomas, J.F. Medrano, G. Rincón, J.C. Leyva-Corona, G. Luna-Nevárez, J.R. Reyna-Granados, and P. Luna-Nevárez. 2021. Molecular marker prediction for days open and pregnancy rate in Holstein cows managed in a warm climate. Livest. Sci. 250. doi:10.1016/j.livsci.2021.104536.

Zhao, X., B. Ponchon, S. Lanctôt, and P. Lacasse. 2019. Invited review: Accelerating mammary gland involution after drying-off in dairy cattle. J. Dairy Sci. 102:6701–6717. doi:10.3168/jds.2019-16377.

Zobel, G., D.M. Weary, K.E. Leslie, and M.A.G. von Keyserlingk. 2015. Invited review: Cessation of lactation: Effects on animal welfare. J. Dairy Sci. 98:8263–8277. doi:10.3168/jds.2015-9617.

Summary

Various metrics have been proposed to evaluate milk production of dairy cows, like cumulative milk production over a specific period (e.g., 305 days, 365 days, or an entire lactation) and milk yield per day within a certain period. These metrics are based on simple calculations on raw data. However, they provide an overview of milk production performance without capturing changes in milk production over the entire lactation period. These changes, or patterns in milk production offer more information (e.g., peak yield, peak time, persistency) about the lactation, which can be useful for breeding and selection, health monitoring, and other applications. Lactation curve models can extrapolate and quantify lactation curves and estimate actual production from incomplete data sets, generating various lactation curve characteristics (**LCC**) to describe the curve in different ways. LCC can serve as a metric to evaluate milk production performance at the cow level and have diverse applications in various dairy research fields. However, some important research topics have received insufficient attention.

This thesis was conducted to explore the application of lactation curve modelling based on farm data collected on commercial dairy farms in the Netherlands and Belgium. Applications in cow reproduction performance and herd economic performance were developed. Four objectives were formulated: (1) to predict lactation persistency for DIM 305 at different insemination moments, (2) to investigate the association between days post conception and persistency, (3) to summarize cow lactation curves into herd lactation curve characteristics (**HLCC**) and illustrate a field application of HLCC, and (4) to compare whether HLCC or the herd's average 305-day milk production (**HM305**) is better able to explain herd economic performance. Throughout this thesis, the MilkBot lactation model was chosen for its ability to accurately model extended lactations, leveraging Bayesian statistics for consistent fitting in the presence of sparse and noisy data through the incorporation of prior information. The MilkBot model consists of LCC: magnitude

(representing the level of production), time to peak yield (representing the rising rate of milk to the peak production level), offset (the time of maximal creation of productive capacity) and decay (the loss of productive capacity), which can be easily transformed into a measure of persistency. Persistency was defined as the number of days it takes for the milk production to decrease by half during the declining stage of lactation.

First, the shape of the lactation curve has been used as an argument to extend the lactation in dairy cows. Cows with flatter lactation curves, referred to as high persistent cows, can yield economic benefits when their lactation is extended. Therefore, when deciding on the optimal voluntary waiting period (**VWP**) of an individual cow, it is useful to be aware of the persistency for the remainder of that lactation, especially for farmers who consider persistency in their reproduction management. Currently, breeding values for persistency are calculated for dairy cows but no studies have focused on predicting the lactation persistency based on readily available cow and herd data. In **Chapter 2**, available cow and herd level data from 2005–2022 were used for a total of 16,980 cows from 84 herds. LCC were estimated for every daily record using the data up to and including that day. Due to the right-skewed distribution of persistency and the normal distribution of decay, decay was preferred for statistical analysis and converted to persistency afterwards for a more straightforward interpretation. Four linear regression models for each of the selected insemination moment (DIM 50, 75, 100 and 125) were built separately to predict decay at DIM 305 (**decay-305**). Results showed that our models had limitations in accurately predicting persistency, although predictions improved slightly at later insemination moments, with R^2 values ranging between 0.27 and 0.41.

The low prediction accuracy observed in our study may be attributed to other factors that influence persistency between the insemination moments and DIM 305. One potential factor that could impact persistency is pregnancy. However, we were unable to account for the pregnancy effect in our prediction model due to several reasons. Firstly, the exact timing of pregnancy is unknown at the time of making predictions for open cows. Secondly, the quantification of the pregnancy effect on persistency (the specific mechanisms or causal relationships) is lacking in previous studies, making it difficult to incorporate it into the model. As a result, we were unable to correct for the pregnancy effect in our prediction model, sparking our curiosity to explore the relationship between days post conception (**DPC**) and persistency in **Chapter 3**.

In **Chapter 3**, we investigated the association between DPC and persistency, with an additional focus on the potential influence of DIM at conception (**DIMc**) on persistency. Milk production data from 2005–2022 were available for 23,908 cows from 87 herds. Persistency was measured by the decay as abovementioned. Decay was calculated for eight DPC (0, 30, 60, 90, 120, 150, 180 and 210 days after DIMc) and served as the dependent variable in a linear model. Independent variables included DPC, DIMc (<=60, 61-90, 91-120, 121-150, 151-180, 181-210, >210) , parity group, DPC × parity group, DPC × DIMc and variables from 30 days before DIMc as covariates. Results showed an increase in decay, i.e., a decrease in persistency, during pregnancy for both parity groups, albeit in different ways. Specifically, from DPC 150 to DPC 210, multiparous cows showed a higher decline in persistency compared to primiparous cows. Furthermore, a later DIMc (cows conceiving later) was associated with higher persistency. Except for the early DIMc groups (DIMc<90), DIMc does not impact the change in persistency by gestation. The findings from this study contribute to a better understanding of how DPC and DIMc during lactation influence lactation persistency.

Within the research field of animal health economics, herd-level studies are performed focusing on health and reproduction factors associated with the profitability of herds. Lactation persistency gets increasing attention and previous studies stated persistent cows are more profitable. These studies were however at cow level, and associations might differ from herd level as other herd factors are interfering with herd economic performance. Additionally, for other LCC (magnitude, time to peak yield) no economic evaluation is performed yet. **Chapter 4** aimed to 1) present a procedure to aggregate cow lactation curves into HLCC; 2) illustrate a field application of HLCC by investigating its association with herd economic performance. Eight years of longitudinal Dutch data on milk production and accounting from 1,673 herds were evaluated. Cow level LCC were summarized to weighted median HLCC on a calendar year basis, for primiparous and multiparous cows (**P1 and P2+**). Data was analyzed using linear mixed modelling, with income over feed cost per cow (**IOFC-cow**) as dependent variable, and HLCC and other herd variables as independent variables. Results indicated that all HLCC were associated with IOFC-cow, except for herd time to peak yield for P1. Of those associated with IOFC-cow, all had positive association except for herd time to peak yield for P2+. In conclusion, we defined herd production patterns by aggregating the cow lactation curves into annual HLCC for P1 and P2+. Associations between IOFC-cow and the various HLCC were deemed logical and interpretable, suggesting that the HLCC aggregation was valid.

Herd milk production performance is generally evaluated using HM305. Economic comparisons between herds are also often made using HM305. Comparing herds is thus based on summarized milk production, and not on the form of the lactation curves of the cows within the herd. Thus far, no literature has evaluated whether the shape of the lactation curve (described by HLCC) is better able to explain the economic variation of herds than summarized milk production such as HM305 does. **Chapter 5** aimed to determine whether HM305 or HLCC is better able to explain the variation in economic performance between herds. To do so, we evaluated eight years of Dutch longitudinal data on milk production and the accounting of 1,664 herds. Cow level LCC were calculated through lactation curve modelling and aggregated to HLCC on a calendar year basis for P1 and P2+. Using income over feed cost per cow (**IOFC-cow**) or per 100kg milk (**IOFC-milk**) as the dependent variable separately, we developed four linear mixed models. Two models were used to analyse the association between herd economic performance and HLCC; the other

two models were used to analyse the association between herd economic performance and HM305. Results showed that HLCC and HM305 explain the same amount of variance of IOFC-cow or IOFC-milk. Both IOFC-cow and IOFC-milk were driven most by year effects (reflecting the milk price and feed price). IOFC-cow was associated with HM305 and HLCC (except herd time to peak yield for P1). Herd magnitude was most strongly associated with IOFC-cow, followed by herd persistency and herd time to peak yield of P2+. IOFC-milk was not associated with HM305 or HLCC (except for a weak negative association with herd persistency for P1).

In the general discussion (**Chapter 6**), the main results of this thesis are discussed. I discussed the limitations in predicting future persistency for DIM 305 and proposed steps for improving the accuracy of persistency prediction, drawing insights from **Chapters 2 and 3**. I provided an explanation for why HLCC exhibited the same association with IOFC compared to HM305 and suggested the potential application of aggregation methods. Additionally, the datasets used, the definition of persistency, cow and herd level clustering issues and the generalizability of our study results to China are discussed.

Samenvatting

Er bestaan verschillende indicatoren om de melkproductie van melkkoeien te evalueren, zoals cumulatieve melkproductie over een specifieke periode (bijvoorbeeld 305 dagen, 365 dagen of een hele lactatie) en melkgift per dag binnen een bepaalde periode. Deze indicatoren zijn gebaseerd op eenvoudige berekeningen van ruwe gegevens. Ze bieden inzicht in de melkproductie van koeien, maar ze geven geen inzicht in de veranderingen in melkproductie over een periode. Deze veranderingen, of patronen in melkproductie, kunnen meer informatie bieden (bijv. piekproductie, moment van piekproductie, persistentie) over de lactatie. Dit kan nuttig zijn voor fokkerij- en selectiedoeleinden, gezondheidsmonitoring en andere toepassingen. Lactatiecurve modellen zijn in staat om lactatiecurves van melkkoeien grafisch weer te geven en om de melkproductie te kwantificeren. Met lactatiecurve modellen kan op basis van onvolledige datasets, een lactatiecurve worden geschat, en kunnen er verschillende lactatiecurve kenmerken (**LCC**) worden gegenereerd (bijvoorbeeld persistentie) die de melkproductiecurve op verschillende manieren beschrijven. LCC kunnen dienen als een maatstaf om de melkproductie van melkkoeien te evalueren en hebben diverse toepassingen in verschillende onderzoeksvelden. Enkele belangrijke onderwerpen hebben echter onvoldoende aandacht gekregen.

Dit proefschrift had als doel het toepassen van lactatiecurve-modellering op basis van bedrijfsgegevens verzameld op commerciële melkveebedrijven in Nederland en België te onderzoeken. Toepassingen op het gebied van vruchtbaarheid van melkkoeien en economische prestaties van het bedrijf werden ontwikkeld. Vier doelstellingen werden geformuleerd: (1) het voorspellen van lactatiepersistentie voor dag in lactatie 305 bij verschillende inseminatiemomenten, (2) onderzoek naar de associatie tussen dagen na conceptie en persistentie, (3) het samenvatten van de lactatiecurves van individuele koeien tot bedrijfsniveau lactatiecurves (**HLCC**) en het illustreren van een praktische toepassing van HLCC, en (4) het vergelijken of

HLCC of de gemiddelde 305-dagen melkproductie van het bedrijf (**HM305**) beter in staat is om de economische prestaties van een melkveebedrijf te verklaren. In dit proefschrift is het MilkBot lactatiemodel gebruikt om nauwkeurig verlengde lactaties te modelleren. In het MilkBot lactatiemodel is gebruik gemaakt van Bayesiaanse statistiek en is dus in geval van ontbrekende data prior informatie gebruikt. Het MilkBot lactatiemodel genereert LCC, namelijk de hoogte van de start van de melkproductie, de snelheid van stijging van melkproductie, de hoogte van de piekproductie, tijd tot piekproductie, en de daling na piekproductie wat gemakkelijk kan worden omgezet in persistentie. Persistentie werd gedefinieerd als het aantal dagen dat nodig is voor de melkproductie om te halveren tijdens de fase van dalende melkproductie.

In eerder onderzoek is de persistentie van melkproductie genoemd als argument om de lactatie bij melkkoeien te verlengen. Er was gevonden dat voor melkkoeien met vlakkere lactatiecurves, aangeduid als persistente koeien, het economisch voordelig is om de lactatie te verlengen. Daarom is het handig om op het inseminatie moment te weten of de koe in het vervolg van de lactatie een persistente melkproductie zal hebben. Op deze manier kan de vrijwillige wachttijd (**VWP**) van een individuele koe geoptimaliseerd worden. Momenteel worden fokwaarden voor persistentie berekend voor melkkoeien, maar er zijn geen studies die zich hebben gericht op het voorspellen van de persistentie van melkproductie op basis routine matig beschikbare data van koeien en bedrijven. In **Hoofdstuk 2** zijn beschikbare data van in totaal 16.980 melkkoeien van 84 bedrijven gedurende de jaren van 2005-2022 gebruikt. LCC werden geschat voor elke dag in lactatie op basis van gegevens tot aan die dag. Vanwege de niet-normale verdeling van persistentie en de normale verdeling van de decay, is deze laatste gebruikt bij de statistische analyses. De decay is later omgezet naar persistentie. Vier lineaire regressiemodellen voor elk van de geselecteerde inseminatiemomenten (dag in lactatie 50, 75, 100 en 125) werden afzonderlijk gemaakt om decay bij 305 dagen in melk (**decay-305**) te voorspellen. De resultaten toonden aan dat onze modellen niet nauwkeurig de persistentie van melkproductie konden voorspellen. De voorspellingen verbeterden echter wel bij latere inseminatiemomenten, met R^2-waarden variërend tussen 0,27 en 0,41.

De lage nauwkeurigheid van voorspellingen in ons onderzoek kan worden toegeschreven aan andere factoren die tussen de inseminatiemomenten en 305 dagen in melk van invloed zijn op persistentie. Een mogelijke factor die van invloed kan zijn op persistentie is de dracht. We konden echter geen rekening houden met het effect van de dracht in ons voorspellingsmodel om verschillende redenen. Ten eerste is de exacte timing van de dracht niet bekend op het moment van voorspelling voor nog niet drachtige koeien. Ten tweede ontbreekt in eerdere studies de kwantificatie van het dracht effect op persistentie (de specifieke mechanismen of causale verbanden), waardoor het moeilijk is om het in het model op te nemen. Als gevolg daarvan konden we het niet corrigeren voor het effect van dracht in ons voorspellingsmodel, wat onze nieuwsgierigheid aanwakkerde om de relatie tussen dagen na conceptie (**DPC**) en persistentie in **Hoofdstuk 3** te verkennen.

In **Hoofdstuk 3** onderzochten we de associatie tussen DPC en persistentie, met een extra focus op de mogelijke invloed van dagen in lactatie bij conceptie (**DIMc**) op persistentie. Melkproductiegegevens van 2005-2022 waren beschikbaar voor 23.908 koeien van 87 bedrijven. Persistentie werd gedefinieerd als het aantal dagen dat nodig is voor de melkproductie om te halveren tijdens de fase van dalende melkproductie. Verval werd berekend voor acht DPC (0, 30, 60, 90, 120, 150, 180 en 210 dagen na DIMc) en diende als afhankelijke variabele in een lineair model. Onafhankelijke variabelen omvatten DPC, DIMc (<=60, 61-90, 91-120, 121-150, 151-180, 181-210, >210), pariteitsgroep, DPC × pariteitsgroep, DPC × DIMc en variabelen vanaf 30 dagen voor DIMc als covariaten. De resultaten toonden een afname van persistentie tijdens de dracht voor beide pariteitsgroepen, zij het op verschillende manieren. Specifiek lieten oudere koeien van DPC 150 tot DPC 210 een grotere afname in persistentie zien in vergelijking met koeien in de eerste lactatie. Bovendien was een latere DIMc (koeien die later drachtig worden) geassocieerd met een hogere persistentie. Behalve voor de vroege DIMc-groepen (DIMc <90) heeft

DIMc geen invloed op de verandering in persistentie door de dracht. De bevindingen van dit onderzoek dragen bij aan een beter begrip van hoe DPC en DIMc tijdens de lactatie de lactatiepersistentie beïnvloeden.

Binnen het onderzoeksveld van de economie van diergezondheid wordt er onderzoek verricht naar de associatie tussen gezondheids- en vruchtbaarheidsfactoren en de winstgevendheid van melkveebedrijven. De persistentie van melkproductie krijgt steeds meer aandacht en eerdere studies stelden dat koeien met een persistente melkproductie winstgevender zijn. Deze studies waren echter allemaal op koe-niveau, en associaties kunnen verschillen op bedrijfsniveau omdat andere bedrijfsfactoren interfereren met de economische prestaties van het bedrijf. Bovendien is er nog geen economische evaluatie uitgevoerd voor andere LCC dan persistentie (piekproductie en tijd tot piekproductie). **Hoofdstuk 4** had als doel om 1) een procedure te presenteren om de lactatiecurves van individuele koeien te aggregeren tot HLCC; 2) een toepassing van HLCC te illustreren door de associatie met de economische prestaties van melkveebedrijven te onderzoeken. Acht jaar aan longitudinale Nederlandse data over melkproductie en boekhouddata van 1.673 melkveebedrijven werden geanalyseerd. LCC op koe-niveau werden samengevat tot een gewogen mediane HLCC op jaarbasis, voor eerste kalfskoeien en oudere koeien (**P1 en P2+**). De gegevens werden geanalyseerd met linear mixed models, waarbij het voersaldo per koe (**IOFC-koe**) de afhankelijke variabele was, en HLCC en andere bedrijfsfactoren als onafhankelijke variabelen. De resultaten gaven aan dat alle HLCC geassocieerd waren met IOFC-koe, behalve de tijd tot piekproductie voor P1 op bedrijfsniveau. Van die variabelen die geassocieerd waren met IOFC-koe, hadden alle een positieve associatie behalve tijd tot piekproductie voor P2+ op bedrijfsniveau. Concluderend hebben we de productiepatronen van een melkveebedrijf gedefinieerd door de lactatiecurves van individuele koeien samen te voegen tot jaarlijkse HLCC voor P1 en P2+. De associaties tussen IOFC-koe en de verschillende HLCC waren logisch en te interpreteren, wat suggereert dat de aggregatie van HLCC op een goede manier uitgevoerd is.

De melkproductie op bedrijfsniveau wordt over het algemeen geanalyseerd met de HM305. Economische vergelijkingen tussen melkveebedrijven worden daarom ook vaak gemaakt met behulp van HM305. Het vergelijken van melkveebedrijven is dus gebaseerd op de totale melkproductie en niet op de vorm van de lactatiecurves van de koeien binnen het bedrijf. Tot nu toe zijn er geen studies die bepalen of de vorm van de lactatiecurve (beschreven door HLCC) beter in staat is om de economische variatie tussen bedrijven te verklaren dan de totale melkproductie zoals HM305 dat doet. **Hoofdstuk 5** had als doel om te bepalen of HM305 of HLCC beter in staat is om de variatie in economische prestaties tussen bedrijven te verklaren. Om dit te onderzoeken hebben we acht jaar aan Nederlandse longitudinale data over melkproductie en boekhouddata van 1.664 melkveebedrijven gebruikt. LCC op koe-niveau werden berekend door middel van lactatiecurve-modellering en samengevoegd tot HLCC op jaarbasis voor P1 en P2+. Met voersaldo per koe (**IOFC-koe**) of per 100 kg melk (**IOFC-melk**) als afzonderlijke afhankelijke variabele ontwikkelden we vier linear mixed models. Twee modellen werden gebruikt om de associatie tussen economische prestaties van de bedrijven en HLCC te analyseren; de andere twee modellen werden gebruikt om de associatie tussen economische prestaties van de bedrijven en HM305 te analyseren. Resultaten toonden aan dat zowel HLCC als HM305 dezelfde hoeveelheid variantie van IOFC-koe of IOFC-melk verklaren. Zowel IOFC-koe als IOFC-melk werden voornamelijk beïnvloed door jaareffecten (die de melkprijs en voerprijs weerspiegelen). IOFC-koe was geassocieerd met zowel HM305 als HLCC (behalve de tijd tot piekproductie op bedrijfsniveau voor P1). De bedrijfsgrootte was het sterkst geassocieerd met IOFC-koe, gevolgd door persistentie op bedrijfsniveau en tijd tot piekproductie van P2+ op bedrijfsniveau. IOFC-melk was niet geassocieerd met HM305 of HLCC (behalve een zwakke negatieve associatie met persistentie voor P1 op bedrijfsniveau).

In de algemene discussie (**Hoofdstuk 6**) worden de belangrijkste resultaten van dit proefschrift besproken. Ik heb de beperkingen besproken bij het voorspellen van toekomstige persistentie voor dag in lactatie 305 en stappen voorgesteld om de nauwkeurigheid van persistentie -voorspelling te verbeteren, gebruikmakend van inzichten uit **Hoofdstukken 2 en 3**. Ik heb een verklaring gegeven waarom HLCC dezelfde associatie met IOFC vertoonden in vergelijking met HM305 en de mogelijke toepassing van aggregatiemethoden voorgesteld. Daarnaast werden de gebruikte datasets, de definitie van persistentie, beperkingen met betrekking tot clustering op het niveau van koeien en bedrijven, en de generaliseerbaarheid van de resultaten van ons onderzoek naar China besproken.

中文简介

在奶牛养殖产业中，我们可使用不同的度量指标去评估奶牛的泌乳性能，例如在特定泌乳天数内的累积产奶量（305 天、365 天或全泌乳期）及其日均产奶量。此类指标仅需对原始的产奶数据进行简单计算即可获得。但它们仅提供了对泌乳性能的总概，无法捕捉泌乳期内的产奶量变化。而这些奶量的变化，抑或是当中的规律，蕴藏着大量信息（如产奶峰值、峰值时间、泌乳持续力等）。这些信息可为牧场的繁殖、选育和健康监测等工作提供重要价值。泌乳曲线模型可用于预测和量化泌乳曲线。在使用泌乳曲线模型对（无论完整与否的）产奶数据进行拟合后，我们会获得一系列的泌乳曲线参数（LCC），它们会从不同角度来描述泌乳曲线。根据特定公式，我们可通过 LCC 较精确地估算实际产奶量。因此，目前 LCC 亦是评估奶牛泌乳性能的实用度量指标之一，且于奶牛研究领域中有各种不同的应用。然而，一些重要的研究方向还未得到足够的关注。

本论文旨在探讨泌乳曲线模型在荷兰和比利时奶牛牧场数据上的应用，分别在个体奶牛繁殖表现和牧场经济效益表现方面进行了研究。论文共有四个目标：（1）在不同的配种时刻，我们能否预测泌乳天数第 305 日的泌乳持续力；（2）探索孕后天数与泌乳持续力之间的关系；（3）对牧场内个体奶牛泌乳曲线进行总结，计算出牧场泌乳曲线参数（HLCC），并展示 HLCC 的实际应用；以及（4）比较 HLCC 和牧场的平均 305 天产奶量（HM305）哪个更能解释牧场的经济效益表现。在整个论文中，我们选用了 MilkBot 泌乳模型，因它能较为准确地拟合相对较长的泌乳期，通过纳入先验信息和贝叶斯统计，即使在数据缺失较多和有噪音的情况下，该模型拟合表现相对稳定。MilkBot 模型会生成四个 LCC：幅度（反映产量水平）、峰值时间（反映泌乳高峰前的泌乳量的上升速度）、偏移（反映最初出现泌乳的时间）和衰减（反映泌乳高峰后泌乳量的下降速度），其

中我们可通过衰减变量计算出泌乳持续力。该泌乳持续力被定义为在泌乳高峰过后，日产奶量减半所需的天数。

泌乳曲线的形状被用作延长奶牛泌乳期的有效依据。泌乳曲线较为平坦的奶牛，即高持续力奶牛，在延长泌乳期时可以带来经济效益。因此，在考虑个体奶牛的最佳自愿等待期时，了解该泌乳期未来的泌乳持续力是非常重要的，尤其是对于那些相对关注持续力的牧场管理者而言。目前，奶牛的泌乳持续力有育种值作为参考，暂未研究使用较简单易得的奶牛和牧场数据去预测泌乳持续力。在第二章中，我们分析了来自 84 个牧场的 16,980 头奶牛在 2005 年至 2022 年间的产奶和牧场数据，使用它们的日产奶记录通过泌乳模型拟合后获得当日的LCC。由于持续力呈右偏分布，而衰减变量呈正态分布，因此我们先使用衰减变量进行统计分析，其后转换为持续力以便解释。我们于四个不同的配种时刻（泌乳第 50、75、100 和 125 天），建立了四个线性回归模型来预测泌乳第 305 天时的衰减（decay-305）。结果显示，模型在越迟的配种时刻，预测准确性越高，但 R^2 值在 0.27 和 0.41 之间，说明模型无法准确预测泌乳持续力。

所观察到的低预测准确性可能归因于那些影响配种时刻和泌乳第 305 天之间持续力的潜在因素，而其中潜在因素之一是妊娠。我们未能在预测模型中校正妊娠效应，原因如下：首先，在为未孕奶牛进行预测时，妊娠的发生时间是未知的。其次，暂无研究对妊娠效应（具体机制或因果关系）进行量化，这使其难以被纳入模型内。综上所述，我们未能在预测模型中校正妊娠效应，但因此萌生了对妊娠效应和泌乳持续力之间关系的好奇心。

在第三章中，我们研究了孕后天数（DPC）与泌乳持续力的关联，并探索了受精时的泌乳天数（DIMc）对该关联的影响。我们分析了来自 87 个牧场的 23,908 头奶牛在 2005 年至 2022 年间的产奶数据。持续力同样通过 LCC 中的衰减变量来计算获得。我们选择了八个 DPC（受精后的 0、30、60、90、120、150、

180 和 210 天）的衰减变量，并将其作为线性模型的因变量。自变量包括 DPC、DIMc（≤60, 61-90, 91-120, 121-150, 151-180, 181-210, ＞210）、胎次分组（初产牛 P1 和经产牛 P2+）、DPC × 胎次分组、DIMc × DPC 以及 DIMc 前 30 天的变量作为协变量。结果显示，在妊娠期间，两个胎次分组的持续力均有所下降，但规律不同。具体而言，在 DPC 150 到 DPC 210 之间，P2+的持续力下降幅度较 P1 更大。此外，越晚的 DIMc（怀孕较晚）与越高的持续力相关。除了早期的 DIMc 组（DIMc < 90），DIMc 不影响妊娠期间持续力的变化。此发现有助于大家更好地理解孕后天数 DPC 和 DIMc 如何影响泌乳持续力。

在动物卫生经济学研究领域，有许多研究者在牧场水平上进行关于与牧场收益相关的健康和生殖因素的研究。泌乳持续力越来越受到关注，有研究表明高持续力的奶牛更具盈利性。然而，这些研究是在个体奶牛水平进行的，它们的结论无法直接套用至牧场水平（即牧场若拥有越多高持续力的牛，其盈利情况越好）。因为在牧场水平，往往存在其他复杂的牧场层面的因素影响牧场经济效益。此外，尚未有研究对其他泌乳曲线参数（幅度、峰值时间）进行经济评估。第四章旨在 1）提出计算出牧场泌乳曲线参数（HLCC）的方法；2）分析 HLCC 与牧场经济效益的关联，从而展示 HLCC 在实际生产上的应用。我们分析了来自 1,673 个牧场共 8 年的产奶和牧场数据，以及财务报表。在每个日历年，我们对牧场内 P1 和 P2+分别进行统计，胎次组内所有个体奶牛 LCC 的加权中位数，为该胎次组于该日历年的 HLCC。随后，我们使用线性混合模型对数据进行分析，以平均每头奶牛的饲料成本收益（IOFC-cow）为因变量，HLCC 和其他牧场变量为自变量。结果显示，所有 HLCC 都与 IOFC-cow 相关，除了初产牛的牧场峰值时间。在所有与 IOFC-cow 相关的 HLCC 中，它们都与 IOFC-cow 呈正相关，除了经产牛的牧场峰值时间呈负相关外。总而言之，我们提出计算 HLCC 的方法。IOFC-cow 与各 HLCC 之间的关联也是合理的，证明 HLCC 的计算是有效和实用的。

中文简介

人们通常使用 HM305 对牧场的产奶表现进行评估，和比较牧场之间的经济效益。因此，牧场之间的比较是基于产奶总量，而不是基于牧场内每头牛的泌乳曲线形状。迄今为止，未文献评估牧场泌乳曲线的形状（由 HLCC 描述）是否能解释牧场经济效益，相比之下，HM305 等的累积产奶量是否更具有解释力。第五章的目标是比较 HM305 和 HLCC 哪个更能解释牧场内经济效益。为此，我们分析了来自 1,664 个牧场共 8 年的产奶和牧场数据，以及财务报表。通过泌乳曲线模型的拟合计算了每头牛在每个泌乳期的LCC，并在日历年的基础上计算出牧场 P1 和 P2+ 的 HLCC。以平均每头奶牛和平均每公斤奶的饲料成本收益（IOFC-cow 和 IOFC-milk）为独立的因变量，我们共建立了四个线性混合模型。当中两个模型用于分析牧场经济效益与 HLCC 之间的关系，另外两个模型用于分析牧场经济效益与 HM305 之间的关系。结果表明，HLCC 和 HM305 对 IOFC-cow 或 IOFC-milk 的具有相同的解释力。IOFC-cow 和 IOFC-milk 主要受所在年份（反映了牛奶价格和饲料价格）的影响。IOFC-cow 与 HM305 和 HLCC 相关（除 P1 的牧场峰值时间外）。牧场产幅度与 IOFC-cow 有最强的相关性，其次是牧场的持续力和P2+的牧场峰值时间。IOFC-milk 与 HM305 和 HLCC 均无关联（除了与 P1 牧场持续力存在微弱负相关）。在总结性讨论（第六章）中，我首先简述了论文的主要发现。基于第二、三章的研究结果，我探讨了无法预测未来泌乳持续力的原因，并提出了提高预测准确性的方案。基于第四、五章的研究结果，我探讨了为何 HLCC 和 HM305 对 IOFC 有相同解释力，提出 HLCC 的潜在应用价值。此外，就本论文使用的数据集、对泌乳持续力的定义、在个体牛和牧场水平上出现的聚类问题以及本研究结果在中国的普适性等方面，本人提出自己的一些浅知拙见，以此与大家共同探讨与研究，敬请广大同仁不吝赐教。

Acknowledgements

Oh yeah! Finally, I finished my thesis. Congratulations to myself! During these four years, I met lots of people and received a tremendous amount of love and support from them. Now, it's a great time to express my gratitude to all of you.

Firstly, I want to express my gratitude to my supervisor team. **Mirjam**, your energy, humor, care, straightforwardness, and sensitivity create a fantastic combination. I love to hear your laughter, and we all enjoy chatting with you. When we worked together, you could sense what would be unnecessary and helped us detect and avoid it immediately. I guess your experience has made you a great team leader. That wisdom and clarity are truly commendable. **Wilma**, you are remarkably sensitive and responsive. You are good at keeping everything on the right path and remind us to plan for everything. You are always the first to congratulate me when good things happen. I truly feel secure and warm under your guidance. **Miel**, you are passionate, creative, and always good at multitasking (in a positive way). Moreover, you are incredibly patient in guiding me from a beginner to a proficient user in all the coding and visualization programs. Thanks my supervisor team for encouraging me to attend conferences and present on stage, even with our limited budget. Though my publication progress may seem more time-consuming to others due to various reasons, I feel a bit anxious but confident with our work because I know you three are always behind my back. Many thanks for guiding me to be a more mature, persistent, confident researcher.

Thanks to all my co-authors in my works: **Jim** (Ehrlich), **Klaas** (Frankena), **Peter** (Vos), and **Hilde** (Aardema). A special thanks to **Klaas**. Although we've never had the opportunity to meet in person, I've always appreciated our insightful discussions and collaboration online. Your exceptional analytical skills and your willingness to dive into coding and modeling challenges with great patience have left a lasting impression on me.

I also extend my gratitude to my colleagues and the secretary in our department. Thanks **Jan** (van den Broek) and **Hans** (Vernooij) for statistical advice whenever we need. Thanks **Arjan** (Stegeman), **Egil** (Fischer), **Paul** (Dobbelaar), **Lindert** (Benedictus) valuable ideas and insights into my studies. Thanks PhD colleagues in our farm animal health group: **Zhaoju**, **Yao**, **Marina**, **Kitty**, **Jesse**, **Yara**, **Marloes**, **Imke**, **Iram**, **Pornsin**, **Pranav**, **Imke**, **Lucia**, **Iram**, **Afonso**, **Jerrold**, **Vivian**, **Maxime**, **Jasmine** and **Nynke**. Although I didn't visit the office frequently, I enjoyed the talks and fun during coffee breaks, lunch time, cake days, gift swapping days and other team building activities.

Thanks **Natcha** and **Ann** for being my paranymphs. Both of you have been very sweet and helpful. More importantly, you have witnessed different stages of my journey during these years. **Ann**, you saw the shy and introverted me when I was new to our department. You introduced me to new housing options, invited me to lunch, and even after you went to Wageningen, we still kept in touch. **Natcha**, you witnessed the later stage of my journey when we both were a bit nervous and anxious about the progress of our project. We encouraged each other and had some fun girls' talks. I remember the promise you made in Zwolle, and I'll keep waiting for that day!

To my friends outside our office building, I want to express my gratitude. I'm not a talkative person and, somehow, I'm a bit apprehensive about socializing with others. So, I don't have many friends, but fortunately, I've met some wonderful people from college, my neighborhood, and perhaps, by destiny. Thank you, **Xinwei Ye**, **Xiaole Tong**, **Nan Zhang**, **Danni Liu**, **Zonglin Tian**, **Wenrui Cao**, **Xing Li**, **Shanshan Bi** from Utrecht. Thanks also to **Jiaoyue Zhu**, **Tianci Zhao**, **Xiaodan Guo**, **Tongxiang Chen**, and **Zao Cheng** from Groningen, and **Bo Fang** from Jülich. These individuals are truly lovable, and we've supported each other during these years. Thank you for giving me strength every time we get together (often while sharing delicious food). Many of you are even willing to physically attend my defense, despite the entire travel time taking up to 5 hours in one day. I am truly touched. Thanks a lot.

To my boyfriend, **Dehuang Kong**, I want to express my heartfelt thanks. Thank you for still being with me and loving me, even after 9 years have passed. We are incredibly fortunate to have each other in the Netherlands and to accompany each other every day. I can't imagine how homesick I would be without you. You and Mimi give me a sense of family that helps alleviate my homesickness. Fortunately, we can continue working and living close to each other after completing our PhD. The commitment to a lifetime together is the deepest confession of love. We still have a long story to write, and I look forward to every further chapter.

To my beloved cat, **Mimi**, I love you so much. I consider myself incredibly lucky that you chose to be friends with me and that you enjoy spending your days with me. I often joke that you are my home office colleague, even though you're usually sleeping beside me or on my lap while I work. Before I met you, I never thought I could love a cat this much. I was initially anxious and worried, wondering if you were a homeless cat. Now, I know where you live, and you have your own family. Saying goodbye to you is incredibly difficult, and I understand that we may never meet again once I return to China. Please stay happy and healthy, living your cat life to the fullest. I've drawn a picture of you on your favorite grey blanket on the last page of my cover. You're standing on top of a hill, watching and knowing everything. Even though we may be apart, I believe you'll always be around me. Kong and I will cherish the memories of us forever.

To my beloved family: my mom and dad, grandparents, aunts, uncles, and cousins. It's been a long time! Thank you for always being there and supporting me throughout. I'm coming back now, and I've grown more mature over the past four years. **Mami** and **baba**, I will take care of our family as you did before for the rest of my life. I love you deeply, just as I've told you so many times. **Grandparents**, please stay healthy and wait for me to come back. Hope you'll recover soon, and the entire family sends you blessings. To my cousin, **Liru**, you're an adorable sister to me. Thanks for reaching out almost every day to share the joys of our lives. Also, to the four funny and sweet girls in my chat group "Sunshine," you are my best friends forever: **Yunyi Luo**, **Jiaqi Zeng**, **Xixin Tan** and **Hailin Ou**. You've all brought lots of sunshine into my life since we met.

Last but not least, I want to express my gratitude for the financial support from the Overseas Study Program of the **Guangzhou Elite Project**. My journey in the Netherlands wouldn't have begun without this crucial support. The staff members were always helpful and took excellent care of us sponsored students. Thank you!

Thank you for taking the time to read my extensive acknowledgment. I extend my best wishes to you, and I look forward to crossing paths with you again in the near future.

(👉 From left to right: Miel, Mirjam, Yongyan, and Wilma, December 14th, 2023)

(👆 Mimi, Yongyan and Kong 👉)

About the author

Yongyan Chen was born in Guangzhou, Guangdong, China, on February 15th, 1995. She and her lovely family lived a happy life near a dairy farm where her father, Jian, worked. Growing up in close proximity to the dairy farm, Yongyan had the opportunity to taste fresh milk and know about dairy farming from a young age. This early exposure sparked her interest in the field. Jian often shared with his daughter the greatness of cows—how they could turn simple grass into high-quality animal protein and therefore, we should always be grateful to these farm animals. Inspired by her childhood experiences, Yongyan desired to learn more about the field of animal science, hoping to understand more about these remarkable animals and, perhaps, contribute to the field herself one day.

After completing high school, Yongyan pursued a bachelor's degree in Animal Nutrition and Feed Science at South China Agricultural University in Guangzhou. Following this, she continued her academic journey at China Agricultural University in Beijing, earning a master's degree in Ruminant Nutrition and Feed Science. In 2019, she received financial support from the Overseas Study Program of the Guangzhou Elite Project, enabling her to embark on a PhD journey at the Department of Population Health Sciences, Faculty of Veterinary Medicine, Utrecht University. Throughout her doctoral studies, Yongyan collaborated with a wonderful and professional team—Mirjam Nielen, Wilma Steeneveld, and Miel Hostens—on the application of lactation curve modeling in dairy production, as detailed in this thesis. During her PhD, she discovered her passion for working with big data, finding excitement in uncovering hidden truths behind seemingly plain numerical data. She aims to further enhance her data analysis skills and apply them in her future career.

As she prepares to graduate, Yongyan plans to return to her hometown of Guangzhou and seek employment there. She is excited about it as she has been away from home for almost four and a half years and has missed many important moments with her lovely family. Although the future is uncertain, Yongyan approaches it with confidence, ready to explore her own path. She can be reached by email at yanz00@foxmail.com.

List of publications

Publications:

Chen, Y., W. Steeneveld, Frankena. K, Leemans. I, Aardema. H, Vos. P, M. Nielen and M. Hostens. 2023b. Association between days post conception and lactation persistency in dairy cattle. Submitted for publication

Chen, Y., W. Steeneveld, M. Nielen, and M. Hostens. 2023. Prediction of persistency for day 305 of lactation at the moment of the insemination decision. Front. Vet. Sci. 10. doi:10.3389/fvets.2023.1264048.

Chen, Y., M. Hostens, M. Nielen, J. Ehrlich, and W. Steeneveld. 2022. Herd level economic comparison between the shape of the lactation curve and 305 d milk production. Front. Vet. Sci. 9. doi:10.3389/fvets.2022.997962.

Chen, Y., M. Hostens, M. Nielen, J. Ehrlich, and W. Steeneveld. 2022. An empirical analysis of economic herd performance in relation to herd lactation curve characteristics. Pages 728–736 in European Conference on Precision Livestock Farming & International Conference on Precision Dairy Farming 2022, Vienna, Austria.

Chen, Y.Y., Y.L. Wang, W.K. Wang, Z.W. Zhang, X.M. Si, Z.J. Cao, S.L. Li, and H.J. Yang. 2020. Beneficial effect of Rhodopseudomonas palustris on in vitro rumen digestion and fermentation. Benef. Microbes 11:91–99. doi:10.3920/BM2019.0044.

Zhang, Z.W., Y.L. Wang, W.K. Wang, **Y.Y. Chen**, X.M. Si, Y.J. Wang, W. Wang, Z.J. Cao, S.L. Li, and H.J. Yang. 2020. The antimethanogenic nitrocompounds can be cleaved into nitrite by rumen microorganisms: A comparison of nitroethane, 2-nitroethanol, and 2-nitro-1-propanol. Metabolites 10:1–14. doi:10.3390/metabo10010015.

Zhang, Z.W., Y.L. Wang, **Y.Y. Chen**, W.K. Wang, L.T. Zhang, H.L. Luo, and H.J. Yang. 2019. Nitroethanol in comparison with monensin exhibits greater feed efficiency through inhibiting rumen methanogenesis more efficiently and persistently in feedlotting lambs. Animals 9:1–14. doi:10.3390/ani9100784.

Oral presentations in conference

Chen, Y., M. Hostens, M. Nielen, J. Ehrlich, and W. Steeneveld. 2022. An empirical analysis of economic herd performance in relation to herd lactation curve characteristics. Dair'Innov congress 2022, Namur, Belgium.

Chen, Y., M. Hostens, M. Nielen, J. Ehrlich, and W. Steeneveld. 2022. An empirical analysis of economic herd performance in relation to herd lactation curve characteristics. European Conference on Precision Livestock Farming & International Conference on Precision Dairy Farming 2022, Vienna, Austria.

Chen, Y., W. Steeneveld, M. Nielen, and M. Hostens. 2023. Can we predict persistency for DIM305? European Buiatrics Congress 2023, Berlin, German.


My beloved mimi 我最爱的小猫