

PREDICTING MILK YIELD TO MONITOR THE TRANSITION PERIOD IN HIGH YIELDING DAIRY CATTLE

Matthieu SALAMONE

Supervisors:

Prof. Ben Aernouts

Prof. Geert Opsomer

Prof. Miel Hostens

dr. Ines Adriaens

Thesis presented in partial fulfilment of the requirements for the joint degree of Doctor in Veterinary Science at the university of Ghent and Doctor of Engineering Technology at the KU Leuven

Members of the Examination Committee:

Prof. Catharina De Schauwer - Chair

Prof. Jeroen Degroote - Secretary

Prof. Mirjam Nielen

Prof. Bart De Ketelaere

Prof. Dirk Van den Poel

Dr. Jenne De Koster

March 2025

© 2025 MATTHIEU SALAMONE

Uitgegeven in eigen beheer, MATTHIEU SALAMONE, GENT

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaandelijke schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic, or any other means without written permission from the publisher.

Table of Contents

LIST OF ABBREVIATIONS.....	III
CHAPTER 1. GENERAL INTRODUCTION.....	1
1.1 A GLOBAL VIEW OF THE DAIRY SECTOR	2
1.2 CRITICAL URGENCY OF SUSTAINABLE DAIRY PRODUCTION	4
1.3 HEALTH IN THE DAIRY SECTOR	6
CHAPTER 2. STATE OF THE ART.....	8
2.1 THE TRANSITION PERIOD.....	9
2.2 PRECISION LIVESTOCK FARMING.....	16
CHAPTER 3. RATIONALE.....	29
CHAPTER 4. PREDICTION OF FIRST TEST-DAY MILK YIELD USING HISTORICAL RECORDS IN DAIRY COWS	33
4.1 INTRODUCTION	34
4.2 MATERIAL AND METHODS.....	36
4.3 RESULTS	44
4.4 DISCUSSION	54
CHAPTER 5. MILK YIELD RESIDUALS AND THEIR LINK WITH THE METABOLIC STATUS OF DAIRY COWS IN THE TRANSITION PERIOD	58
5.1 INTRODUCTION	59
5.2 MATERIALS AND METHODS	60
5.3 RESULTS	65
5.4 DISCUSSION	72
5.5 CONCLUSION.....	78
CHAPTER 6. ASSOCIATION OF MILK YIELD RESIDUALS TO BEHAVIOURAL PATTERNS AND TRANSITION SUCCESS IN MULTIPAROUS DAIRY COWS.....	79
6.1 INTRODUCTION	80
6.2 MATERIALS AND METHODS	82
6.3 RESULTS	86
6.4 DISCUSSION	94
6.5 CONCLUSION.....	97
CHAPTER 7. THE ASSOCIATION BETWEEN TRANSITION MANAGEMENT AND MODELLED MILK YIELD IN DAIRY CATTLE	99
7.1 INTRODUCTION	100
7.2 MATERIALS AND METHODS	102
7.3 RESULTS AND DISCUSSION.....	104
7.4 CONCLUSION.....	112
CHAPTER 8. VALORISATION.....	113
8.1 FROM MILK RESIDUALS IN THE TRANSITION PERIOD TO LACTATION ONSET VALUE	114
8.2 INFRASTRUCTURE	114
8.3 CUSTOMERS	122
8.4 OFFERING AND FINANCIAL ASPECT	127
8.5 FROM RESEARCH TO REALITY: BRINGING THE LOV TO THE DAIRY FARM	128
CHAPTER 9. GENERAL CONCLUSION AND FUTURE PERSPECTIVES.....	129
9.1 THESIS OVERVIEW.....	130
9.2 GENERAL CONCLUSION	131
9.3 FUTURE PERSPECTIVES.....	135
REFERENCES.....	139

SUMMARY	153
SAMENVATTING.....	156
APPENDIX.....	159

List Of abbreviations

AI	Artificial Intelligence
AIC	Akaike Information Criterion
BCS	Body Condition Score
BHB	β -hydroxybutyrate
CI	Confidence Interval
CRV	Dutch/Flemish Breeding Organisation
DHI	Dairy Herd Improvement
DIM	Days in Milk
DIS	Clinically Diseased
DMI	Dry Matter Intake
FAIR	Findability, Accessibility, Interoperability, and Reuse of digital assets
FAO	Food and Agriculture Organization
FULL	Production, Herd, and Reproduction Features
ICAR	International Committee for Animal Recording
IGF1	Insulin-like growth factor 1
ILVO	Flanders Research Institute for Agriculture and Fisheries
INAGRO	West Flemish R&D Department for Agriculture
kgTD1_{x+1}	Milk Yield on First Test Day Record of Next Lactation
KPI	Key Performance Indicator
LOV	Lactation Onset Value
LSM	Least Square Means
LV	Latent Variables
M21	Cumulative Milk Yield at 21 DIM
M305	Cumulative Milk Yield at 305 DIM
M75	Cumulative Milk Yield at 75 DIM
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
meanM305	Mean 305-Day Milk Yield of Cows that Calved During the Period of Interest
MM	Milk Meter
MPI	Model Performance Indicators
MRT	Milk Yield Residuals in the Transition Period
NDIS	Not Clinically Diseased
NEB	Negative Energy Balance
NEFA	Non-Esterified Fatty Acids
P	Production Features
PH	Production and Herd Features
PLF	Precision Livestock Farming

PLSR	Partial Least Squares Regression
POI	Period of Interest
R²	Coefficient of Determination
RMSE	Root Mean Square Error
RMSECV	RMSE in Cross-Validation
SCC	Somatic Cell Count
SD	Standard Deviation
sdMRT	Standard Deviation of the Individual MRT During the Period of Interest
sMC	Significance Multivariate Correlation
SPP	Standard Peak Production
SST	Serum Separation Tube
TCI	Transition Cow Index
TD	Test Day
TD1_{X+1}	First Test Day Record of Next Lactation
TDR	Test Day Records
ULP	Large Animal Practice at Utrecht University in the Netherlands
USDA	United States Department of Agriculture
VLAIO	Flanders Innovation & Entrepreneurship
VLAIO-LA	Flanders Innovation & Entrepreneurship Agricultural Grant
X	Current Lactation
X+1	Next Lactation

Chapter 1.

General Introduction

1.1 A global view of the dairy sector

Worldwide, it is estimated that 757 million tons of cow milk were produced in 2021 (FAO, 2022a). With an average global consumption of dairy products of 54.4 kg per capita in 2021, these products represent 5.5% of the daily caloric intake and 8.4% of the daily protein intake (FAO, 2022a). Impacted by high regional variations, the preponderant role of milk as a source of nutrients for humans is mostly visible when these global numbers are broken down per continent. Europe and North America are the highest consumers worldwide, with a respective consumption of 117.8 and 116.6 kg per capita in 2021. Relative to the other food sources, dairy products represent 17-18 % of the food consumption in those regions. Compared to global nutrition figures, the importance of dairy in supplying calories and protein in these regions is roughly double the world averages. The evolution of these figures over the last 10 years is plotted in Figure 1.1.

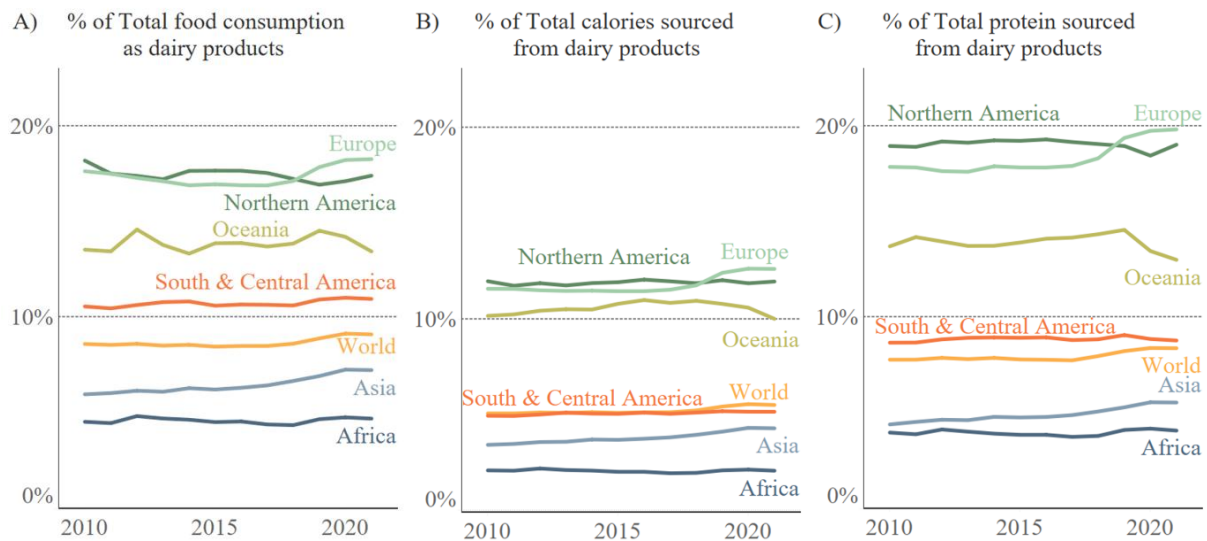


Figure 1.1 Representations of the nutritional importance of dairy products plotted over the last 10 years using data reported by the Food and Agriculture Organization (FAO, 2022a). Over all three panels, the world data is plotted combined with a continental breakdown of the different variables. In panel A, the relative proportion of dairy products is plotted in relationship to the total food consumption and the historical evolutions since 2010. Similarly, panels B and C show the evolution of the proportional importance of dairy products as a source of calories and protein.

While at first sight, these metrics seem to follow a quite stable trajectory, it is important to note that by fixing these figures per capita, the underlying global increase in consumption and production is overshadowed. Using 2021 as a reference year, the increase in production globally has been estimated to be 140% since 1961, or using a more recent time window 26% since 2010 (FAO, 2022a). The continental decomposition of these production figures combined with the

dairy cow population dynamics delineates two different processes at play, plotted in Figure 1.2 and Figure 1.3. On the one hand, Europe and North America’s milk production increased from 1961 to 2021 with respective 19.9% and 72%, the dairy cow population decreased by 48.5% and 60 % in the same period. On the other hand, the exponential increase of milk production over the same period in Asia (1127 %), Africa (424%) and South and Central America (376%) coincides with an increased dairy cow population of 260% in Asia, 305.7% in Africa and 100% in South and Central America. This phenomenon is more apparent when looking at the top-producing countries per continent. For example, over the last 60 years, China increased their milk production by sixtyfold and its dairy population by twentyfold. In other words, they went 600.000 tons of milk (2.3 % of the continent’s total production) to 35 million tons of milk produced in 2021 (13.8% of Asia’s milk production), becoming the second highest producer in Asia. A summarising table of these different figures is available in Appendix A.

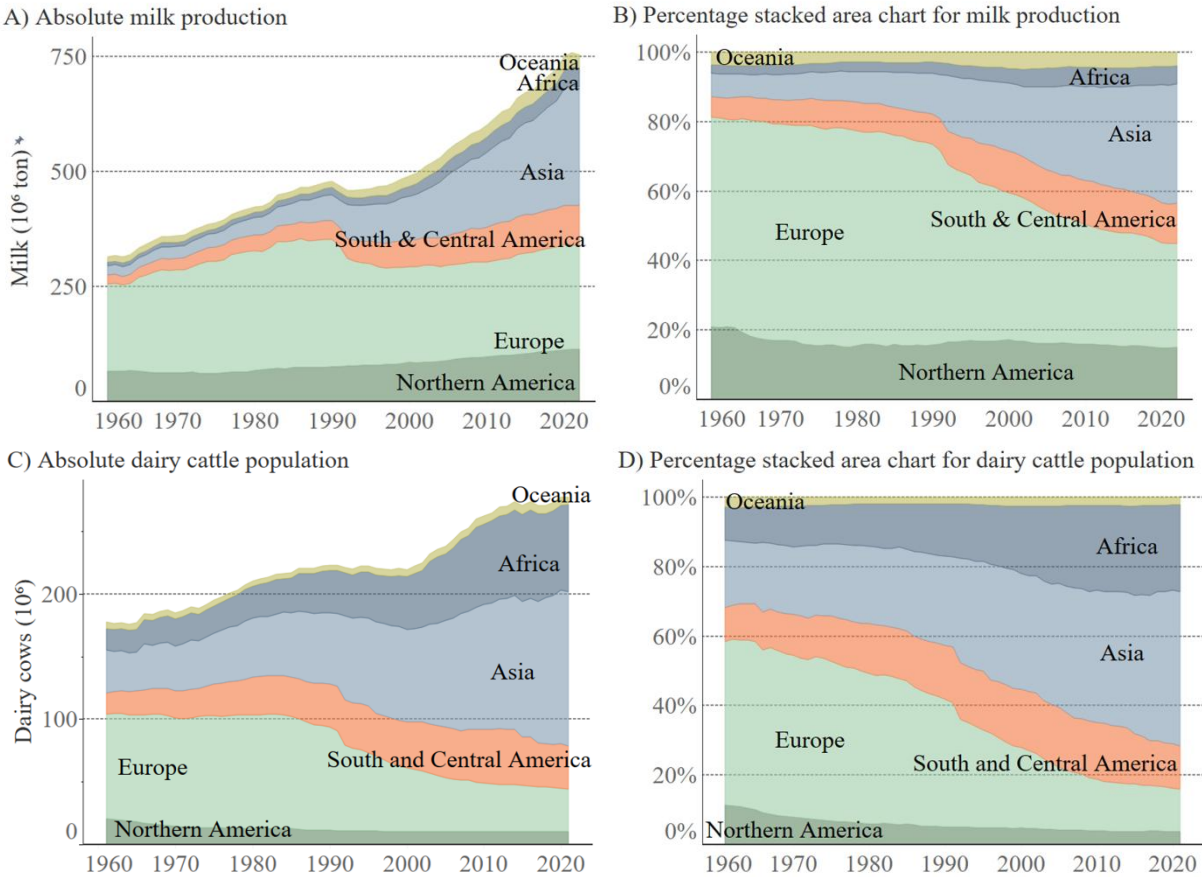


Figure 1.2 These panels plot the evolution of milk production and the dairy cattle population over the last 60 years. The data used originates from FAO reports that compiled estimates and official figures. Within each panel, the data is grouped by continent. The figures concerning milk production and the dairy cattle population are plotted in absolute values in panels A and C. Panels B & D represent the relative proportion of the continental figures over time

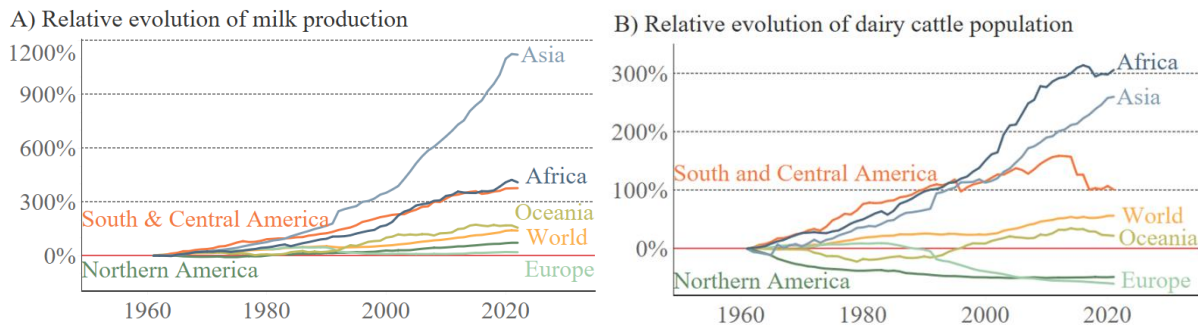


Figure 1.3 These two panels plot the in- and decreasing dynamics of milk production (A) and dairy cattle population (B). The data utilised comes from FAO estimates and official figures. Each panel organises the data by continent. 1961 forms the reference year in both panels.

1.2 Critical urgency of sustainable dairy production

Despite the influence of various complex factors (geopolitical, geographical, environmental, demographic) on the figures shown in the previous section, the general trajectory of these figures emphasises the demand for milk as a source of nutrients for the global population.

Nevertheless, in a context of high scrutiny concerning the sustainability of our food production, this rise in production does not come without challenges. For example, the emissions of global greenhouse gases associated with milk production, represent one of many the flip sides of the coin. The Food and Agriculture Organization of the United Nations (FAO) reported that in 2021, estimated methane and nitrous dioxide emissions of 20.7 and 0.5 megatons worldwide were associated with the dairy industry. Calculating the CO₂ equivalent of these emissions equals 667 megatons of CO₂ emissions; this is roughly equivalent to the total emissions produced by an industrialised country such as Germany in 2021. The evolution of these estimated emissions is plotted in Figure 1.4. This plot indicates that the emissions at the turn of the 21st century globally decreased to 582 tons of CO₂, nearing the levels of the 1960s. Linking these dynamics with the aforementioned evolutions in milk production shows that in 2021, 44.8 % of the world's milk was produced in Europe or North America, with 15.8% of the world's dairy cattle population, only accounting for 27.8 % of the global greenhouse gas emissions attributed to milk production. This highlights the possible major pitfalls for the industry regarding sustainability. The increase in production in certain parts of the world seems to have counteracted the realised emission improvements in other parts of the world.

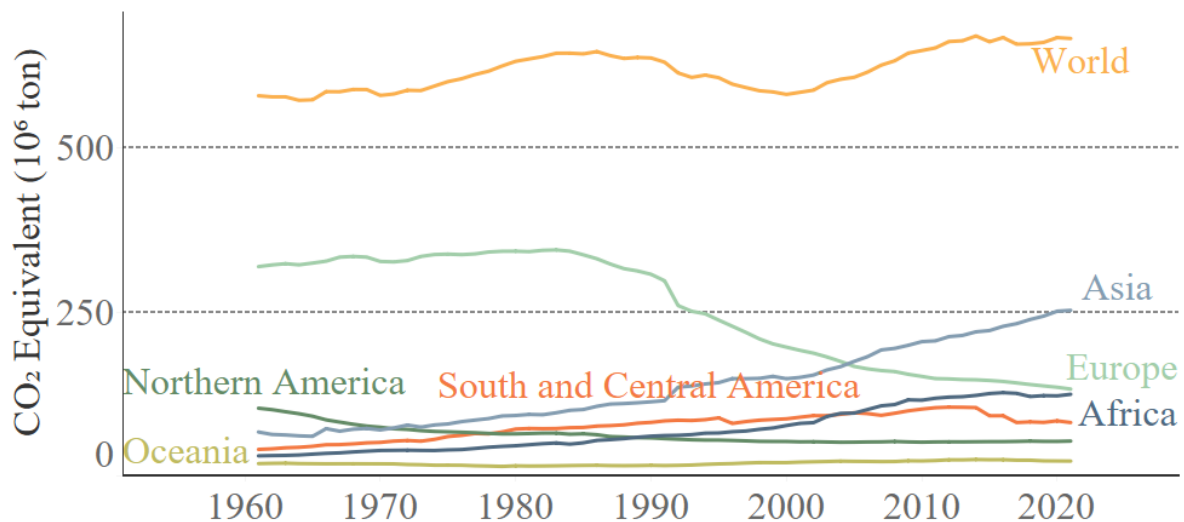


Figure 1.4 Using the data reported by the FAO, this plot illustrates the estimated emission of greenhouse gasses linked with the production of milk over the last 60 years. Methane and nitrous dioxide emissions were converted to their CO₂ equivalent (1kg CH₄ : 25kg CO₂, 1kg N₂O : 298kg CO₂). Data was grouped by continent, highlighting regional differences.

Furthermore, greenhouse gas emissions only represent one of the many challenges emanating from dairy production. Milk production challenges biodiversity, water and energy use and soil fertility (Van Calker et al., 2005). While dairy is not the only sector contributing to these disruptions, disruptions in all these layers of the environment have been recorded at an alarming rate in the last decade, justifying the urgency for a more sustainable approach across all industries, including dairy (Cardinale et al., 2012; Allen et al., 2018; Salehi, 2022).

Thompson (2007) provides an overview of the conceptual definitions of “sustainability” and what it could mean in the agricultural context. Although this work relates more to the philosophical analysis of this broad buzzword, it also brings new understanding to the different positions taken today in the sustainability debate. One of the definitions of sustainability resonating the most with the mainstream paradigm in the dairy industry is the resource sufficiency concept (von Keyserlingk et al., 2013; Britt et al., 2018). In Thompson (2007), the resource sufficiency concept is defined as follows :

Sustainability as a measure of the duration associated with practices that produce (or detract from) well-being. It leaves questions about whose well-being and the relative measure of different forms of satisfaction open. It is consistent with the general form of the utilitarian maxim, and indeed seems to specify nothing more than the temporal dimension of it.

In a certain sense, this definition formulates the global hurdle of sustainability. The crux lies in the openness of whose well-being is assessed and how much it weighs in the general

appreciation of sustainability. In other words, the rise in milk production in developing countries could be interpreted as a sustainable way for these countries to increase the well-being of their population. This is achieved by increasing food or economic security and attributing relatively lower importance to milk production's environmental or social aspects. It also creates a framework to understand evolutions in consumer perceptions in Europe or North America, where the general welfare of cows is included in the sustainability framework in the absence of perceived threats to human well-being. Seeking to achieve sustainability in this context necessitates exploring potential solutions, such as those offered by technological advancements.

At the same time, automation and digitalisation have revolutionised agriculture as a whole and their potential to achieve sustainable food production has often been discussed (FAO, 2022b). For instance, FAO in their report “State of Food and Agriculture 2022” formulates its potential in the following way:

Agricultural automation can raise productivity, build resilience, improve product quality and resource-use efficiency, reduce human drudgery and labour shortages, enhance environmental sustainability, and facilitate climate change adaptation and mitigation. (FAO, 2022b)

However, the report also states that this potential can only be achieved if the developed technologies are available, inclusive, accessible to all, and tailored to local conditions. In a certain sense, the FAO pleads for all actors in the industry to evaluate their potential elements of progress against these limitations. Only then will the industry be able to harness the inherently revolutionising potential described for numerous technologies (AI, Robotics, cloud solutions...) and lead to global sustainable agriculture. For academics, this request could be seen as adding an assessment component when evaluating automation and digitalisation to achieve more sustainable food production. Not only including the direct micro repercussions in implementing their research but also the possible macro repercussions outside of their initial research goals.

1.3 Health in the Dairy Sector

The cow's health has often been central to evaluating the well-being question posed by the sustainability framework. This can either be expressed in the direct effect of health challenges on production as well as the underlying welfare issues. Whereas conceptually, the link between health challenges and detrimental welfare situations is rational, it is challenging to quantify accurately. Numerous studies have shown the effects of disease on production parameters: decreased reproductive performance (Bertoni et al., 2009; Santos and Ribeiro, 2014), decreased

milk yield (Fourichon et al., 1999; Warnick et al., 2001; Adriaens et al., 2021b). Health as the point of action to improve the dairy industry's sustainability takes different forms academically. It consists of studying and understanding the processes at play ((patho)physiologically, epidemiologically, economically...), transferring this understanding to detect detrimental situations and consequently allowing the prevention or cure of health-disrupting events.

The digitalisation of dairy farms in the context of health played a major role in revolutionising the fundamental understanding of health. This mainly due to the increase in data volume and novel data measuring techniques (behaviour, vision...), combined with an increased accessibility in cutting edge methods for analysis (artificial intelligence, machine learning, increased computing power...) (Hermans et al., 2018). As an example, sensor technologies allowed the associations between rumination behaviour and the occurrence of health diseases to be quantified (Paudyal et al., 2018). In the case of the detection and prevention of certain health statuses through automated digital systems, progress has been more nuanced. Academic efforts researching this topic still encounter low implementations in the outside world (Eckelkamp, 2019). Low confidence in the reported disease definitions, necessitating a learning process, cumbersome data integrations, and economic costs are often cited as reasons explaining the difficulty in reaching higher adoption of these technologies (Barkema et al., 2015; Eckelkamp, 2019; von Keyserlingk et al., 2024).

Chapter 2.

State of the Art

2.1 *The Transition Period*

2.1.1 *From Physiology to Pathology*

The typical cyclicality in a dairy cow's life is determined by the cycles of lactations, which, in terms, define the lifetime productivity of a dairy. Each cycle or lactation starts with a calving event initiating the productive stage of the lactation cycle. As the lactation progresses to peak milk, the animal resumes her reproductive stage of lactation, eventually leading to pregnancy. Milk production typically stops around 220 days of pregnancy as the cow is dried off (Goselink et al., 2024). Even though the optimal length of the lactation is currently debated in the industry, in practice, it typically ranges from 370 to 450 days (van Knegsel et al., 2022). Although each of these stages present their own challenges, the transition between two lactations has been the industry's main focus since the beginning of the 21st century (Drackley, 1999; Overton et al., 2017). The transition period is typically defined as the period of 3 weeks before until 3 weeks after calving (Grummer, 1995). Nevertheless, currently discussions exist in the literature, where some authors argue that the transition period should be defined from dry off till 6 weeks after calving (Lopreato et al., 2020). During that period, animals present an elevated risk of developing metabolic and/or infectious diseases, having possible long-lasting effects during the remainder of the lactation. Different authors emphasised on the concept of homeorhesis to understand and explain the physiological processes during this period (Bauman and Bruce Currie, 1980; Horst et al., 2021; Van Saun, 2023a). Contrary to homeostasis, which represents the physiological process to maintain a physiological status quo, homeorhesis represents the physiological prioritisation of essential nutrients by different tissues to support a specific physiological state (e.g. pregnancy). At calving, the physiological status of the dairy cow needs to transition from a non-lactating pregnant state to a lactating non-pregnant state in a matter of hours. In a homeorhetic process, the prioritisation of milk production characterises this transition of states. Concretely, 97% of the net energy during the transition period is utilised by the mammary gland (Drackley, 1999). This, at first sight, could be seen as beneficial, though this unconditional focus of the metabolism seems to have detrimental effects on deprioritised key physiological processes such as immunity and reproduction. Epidemiological studies have shown that 30 to 50% of dairy cows will experience at least one health challenge during the transition period (LeBlanc, 2010; Hostens et al., 2012). These health challenges have been qualified in the past as a man-made production disease caused by the intensification of milk production, straining the animals' natural metabolism (Mulligan and Doherty, 2008). To grasp

how these physiological processes can evolve to vulnerable situations, an outline will be given on the drastic alterations in both the calcium and energy metabolism.

Within the 24 hours after calving, calcium requirements increase by 6-fold (Pascottini et al., 2022). This increase in calcium requirements is the concrete manifestation of the homeorhetic priority towards supporting milk and colostrum synthesis (Couto Serrenho et al., 2021). This prioritisation can potentially jeopardise calcium's other important roles in multiple physiological processes including bone mineralisation, coagulation, cardiac action potentials, cell signalling as a second messenger, muscle contractility and immune response (Wilkens et al., 2020; Horst et al., 2021). Although calcium reserves are abundant within the bones, mobilising the bone-stored calcium needs to be complemented with gastrointestinal and renal absorption of calcium. The organisation of the different bone/intestinal tissues in the function of this calcium demand is modulated by hormones such as parathormone and calcitriol and utilises other minerals such as magnesium and phosphorus (Wilkens et al., 2020). The specific pathways regulating calcium concentrations however fall outside the scope of this manuscript. Nevertheless, the described prioritisation has been associated with detrimental effects on health, fertility, and production (McArt and Oetzel, 2023).

Concurrently, the initiation of the lactation is also characterised by a deficit in energy caused by the disbalance between the nutritional energy input and energy output in the form of basal metabolism and milk. A reduction in dry matter intake (DMI) 2 to 3 weeks before parturition combined with a lagged recovery to the maximal DMI at the start of the lactation causes the restriction in energy input (Bell, 1995; Grummer et al., 2004; Drackley et al., 2005). Oppositely, the output increase at the start of the lactation is mainly due to the production of milk (Drackley, 1999). The intrinsic alterations in the energy metabolism needed to maintain milk synthesis under negative energy balance (NEB) have been an industry-wide research focus since the initial definition of the transition period (Grummer, 1995; Friendly, 2010; De Koster and Opsomer, 2013; McArt et al., 2013b; VandeHaar et al., 2016; Overton et al., 2017). In their review, Horst et al. (2021) expand on the physiological adaptation of the glucose metabolism given its central role in lactogenesis and the subsequent osmoregulatory force driving the milk volume. To compensate for the dietary shortage of gluconeogenic precursors occurring at the start of the lactation, hormonal modulation allows catabolism and mobilisation of respective muscle and adipose tissue (Bell, 1995; Bell and Bauman, 1997). The lipolytic mobilisation manifests itself as the export of non-esterified fatty acids (NEFA) from the adipose tissue under modulation of somatotropin, insulin and catecholamines. Although this influx of NEFA forms

a substantial source of energy for the peripheral tissues and mammary gland, some of it is converted into ketone bodies. The exact mechanism regulating ketogenesis occurring in the liver is not fully understood. Yet, it is likely that a limited supply of oxalacetate combined with the buildup of acetyl-CoA through the β -oxidation of NEFA, pushes the acetyl-CoA towards ketogenesis. Additionally, reduced glucose uptake in the skeletal muscle also occurs because of the oxidation of fatty acids and ketones (De Koster and Opsomer, 2013). Similarly to calcium, Horst et al. (2021) describe a homeorhetic process regarding glucose, where a unidirectional flow of glucose/energy is organised towards the udder to support the drive to produce milk.

In summary, the transition period is characterised by a shift in production status, turning the psychological priorities upside down. Despite the seemingly standard nature of this prioritisation towards milk production, pathological evolutions in other deprioritised processes can challenge the sustainability of milk production.

2.1.2 From Clinical to Subclinical

It is generally assumed that maladaptations caused by the calcium and energy focus induce a state of immunosuppression, facilitating the development of infectious diseases (Wilkens et al., 2020; Horst et al., 2021; Pascottini et al., 2022). The general dogma regarding the transition period has been that clinical disease only represents the ‘tip of the iceberg’, while under the surface a larger proportion of animals face subclinical diseases, insidiously altering the health and welfare of the animal (LeBlanc, 2010). Initially focusing on the clinical diseases, the term ‘production diseases’ is typically used to describe clinical hypocalcemia, retained placenta, metritis, clinical ketosis, displaced abomasum (Drackley, 1999; Probo et al., 2018) as well as, lameness and mastitis occurring early in lactation (Mulligan et al., 2006). Within the production diseases, metritis, mastitis, and certain forms of lameness represent infectious diseases of the uterus, udder and parts of the locomotive system, respectively. Clinical hypocalcemia and clinical ketosis characterise metabolic disruptions in calcium and ketone bodies in the blood. The remaining disorders represent a more eclectic appearance of pathological situations, though etiologically still seen as a consequence of the processes occurring in the transition period, Table 2.1 provides a global view of these production diseases clinical signs and aetiology. Many of the aforementioned production diseases result in reduced milk yield (Seegers et al., 2003; Hostens et al., 2012; Adriaens et al., 2021b), lower reproduction performance (Giuliodori et al.,

2013; Pascottini et al., 2020), higher culling rates (Probo et al., 2018) and altered animal welfare (Proudfoot, 2023).

Table 2.1 General overview of the production diseases. clinical signs and etiology was extracted from literature.

Disease	Typical clinical signs	Etiology
Clinical mastitis (Ruegg, 2017)	Inflammation signs (swelling, heat, hardness, pain, redness) clots in milk	bacterial infection
Clinical Metritis (Sheldon et al., 2006)	Purulent uterine discharge, sometimes with systemic clinical signs such as fever, milk production decrease, anorexia, depression	Polymicrobial infection
Clinical hypocalcemia (Lorenz et al., 2011)	Paresis, subnormal body temperature, cold extremities, tachycardia	Increase in calcium requirements near parturition combined with a dysfunctioning calcium regulation
Clinical ketosis (Fox, 1971)	Anorexia, milk yield decrease, loss of body condition, dry firm faeces, sometimes neurologic signs	Excessive fat mobilisation & accumulation of ketone bodies in blood
Displaced abomasum (LeBlanc et al., 2005)	Anorexia, milk yield decrease	Abomasal hypomotility and/or dysfunction of the intrinsic nervous system
Retained placenta (Sheldon et al., 2006)	Failure to expel fetal membranes within 24h after parturition	Precise etiology unclear, though probably caused by a disturbed immunology
Lameness (Roche et al., 2024)	Abnormal movement and/or posture	Pain in the locomotor system, mainly hoof due to infection (digital dermatitis, foot rot, heel horn erosion, interdigital dermatitis) or non-infectious (sole hemorrhages, sole ulcers, or white line disease)

The prevalence of these diseases has been described by Drackley et al. (1999) and are presented in Table 2.2. This table also combines these prevalences as reported in 1999 with those described in Van Saun et al. (2023) as obtained from more recent American Animal Health Monitoring Data. Van Saun (2023) argues that the trends in disease prevalence demonstrate that little to no progress has been made in reducing disease occurrences over the past 25 years. A more concrete demonstration of this trend can be found in Lean et al. (2023), where associations between parity and production disease were studied by pooling data from different studies from 1992 to 2020. In their logistic regression models, the effect of the study year showed some reduction in the odds ratio in the case of clinical hypocalcemia, displaced abomasum, and retained placenta. An increase in odds ratio over the years was found for clinical ketosis, mastitis, and lameness.

Table 2.2 Compilation of reported production disease incidence found in Drackley (1999) and Van Saun (2023).

Disease	Drackley (1999)	Van Saun (2023)			
	Mean Prevalence (%) [min – max]	Reported prevalence \pm standard error			
		1996	2002	2008	2014
Milk fever	7.2 [0 – 44.1]	5.2 \pm 0.1	5.2 \pm 0.1	4.9 \pm 0.1	2.8 \pm 0.2
Retained Placenta	9.0 [0 – 22.6]	7.8 \pm 0.2	7.8 \pm 0.2	7.8 \pm 0.2	4.5 \pm 0.4
Metritis	12.8 [0 – 66]	NR	NR	NR	6.9 \pm 0.8
Ketosis	3.7 [0 – 20]	NR	NR	NR	4.2 \pm 0.5
Mastitis¹	NR	13.1 \pm 0.3	14.7 \pm 0.3	16.5 \pm 0.5	24.8 \pm 2.4
Lameness	NR	10.5 \pm 0.3	11.6 \pm 0.3	12.9 \pm 0.3	8.2 \pm 0.5
Displaced abomasum	3.3 [0 – 14]	2.8 \pm 0.2	0.8 \pm 0.1	0.6 \pm 0.1	0.6 \pm 0.3

¹ Limited to the postpartum occurring mastitis.

While the precise etiopathogenesis of these diseases was intensively researched (Horst et al., 2021; Van Saun, 2023a), others have attempted to create intervention methods to limit their negative consequences (LeBlanc, 2010). Returning to the iceberg analogy, mitigation strategies have focused on what's happening under the water level, associating subclinical diseases as a gateway to the clinical disease or detrimental on their own. In contrast to infectious subclinical diseases defined by a pathogen, metabolic subclinical diseases are not defined by the presence or absence of a metabolite. They are rather defined by associations for different metabolites at different concentrations emanating from metabolic surveys. In other words, metabolites become biomarkers for their own subclinical disease, with estimated detrimental effects in the real world. Different metabolites/biomarkers have been studied in recent years to estimate the proportion of animals with subclinical diseases. Subsequently, the effects on production, health and reproduction were used to discriminate metabolically favourable from unfavourable states. This section will give a short overview of the studied biomarkers and their associations with an animal's health and metabolic status.

Circulating NEFA and Beta-hydroxy butyrate (BHB) formed the starting point to discriminate favourable from unfavourable metabolic status with regard to the NEB. Overton (2017) links this starting point back to the first rigorous epidemiological study by LeBlanc et al. (2005) linking different biomarkers (NEFA, BHB, glucose, phosphate, calcium, cholesterol and urea) with displaced abomasum. They found that elevated prepartum NEFA and elevated post-partum

BHB were associated with an increased risk of developing displaced abomasum. These increased risks kickstarted a series of studies investigating further associations under the general idea that both biomarkers were indirect measures of the magnitude of NEB and the success of the cow's adaptation to it (Herdt, 2000). McArt et al. (2013a) provide an overview of the different studied cut-off values for BHB and NEFA in blood pre- and post-partum and their association with diseases or diminished production in the transition. Their overview identified the increased BHB post-partum as having consistently been associated with a higher risk of displaced abomasum and culling. Post-partum increased BHB showed inconsistent effects on reproduction and production performances. For NEFA prepartum, it was found that displaced abomasum was more likely, reproduction and production were negatively impacted, and culling risk was higher. Post-partum NEFA concentrations were associated with a higher risk for diseases (e.g. clinical ketosis, metritis, displaced abomasum). However, no associations with reproductive performance and inconsistent associations with milk yield were found.

Other biomarkers have been proposed such as insulin-like growth factor-1 (Wathes et al., 2021), fructosamine (Sorondo and Cirio, 2009; Caré et al., 2018) and haptoglobin (Huzzey et al., 2009, 2011, 2015). These alternate biomarkers presented similar associations as BHB and NEFA with some form of metabolic status. A short overview of these associations: higher haptoglobin concentrations were associated with lower milk yield (Huzzey et al., 2015) and lower reproductive performances (Huzzey et al., 2011). Fructosamine was associated with the glucose concentration over the previous 1-3 weeks (Caré et al., 2018). Lower IGF1 concentrations were found to be associated with low DMI and lower milk production.

2.1.3 From Biomarker to Monitoring

With the aim of creating accurate on-farm possibilities to detect metabolic unfavourable situations and prevent detrimental effects on the animal, in-field monitoring strategies were set up parallel to academic research identifying biomarkers. For the daily management of the transitioning cows the initial strategy formulated by LeBlanc (2010) consisted of measurement of NEFA one week prepartum or BHB measurements in the first week after calving. In practice, NEFA measurements were rapidly abandoned, due to the lack of an accurate cow-side test (Overton et al., 2017). Other biomarkers faced similar limitation to their implementation. In contrast, BHB measurements could be made in blood using a handheld device, urine or milk using colour strips (Iwersen et al., 2009). These tests measurements presented high sensitivity [87-93%] and specificity [93-100%] at identifying elevated BHB in their respective media and

at different cut-offs when validated with lab tests (LeBlanc, 2010). The availability of these accurate cow-side measurement options has formed the basis for the industry-wide focus on BHB as the key biomarker, while research has indicated multiple times its nuanced and inconsistent role (McArt et al., 2013a; Overton et al., 2017; Horst et al., 2021). Under the general idea that “you can’t manage what isn’t measured” for the last 15 years, farmers have been monitoring BHB, albeit with little or no effect on disease prevention (Van Saun, 2023a).

Alternative methods have been researched to circumvent this BHB-centred and cut-off based focus. For instance, researchers used machine learning clustering as unsupervised dimension reduction techniques to combine multiple biomarkers into automated metabolic status classifier models (De Koster et al., 2019; Grelet et al., 2019; Xu et al., 2019; Vosseveld et al., 2022; Heirbaut et al., 2023; Zhang et al., 2023). These models were defined in a similar way in these studies. Study data was clustered based on measured biomarkers or calculated energy balance. The discrimination between imbalanced/unfavourable clusters and balanced/favourable clusters was made using different descriptive statistics such as biomarkers (BHB, NEFA, IGF1), DMI, milk production and disease diagnosis within the clusters. For example, Heirbaut et al. (2023) classified the clusters with high BHB and NEFA combined with low glucose, insulin and IGF1 as imbalanced. The imbalanced cluster presented a 2.97 higher relative risk of developing a production disease. Both De Koster et al. (2019) and Heirbaut et al. (2023) proposed models that are able to predict the cow’s cluster based on readily available milk mid infrared spectra and dairy herd improvement (DHI) data. They state that their model could be used as a way to routinely screen for the metabolic status of animals, though to our current knowledge, this has not been implemented in the field. More recently, behaviour monitoring has also been investigated as a possible alternative to the predominance of biomarkers (Overton et al., 2017). By focusing mostly on the rumination behaviour and the time spent ruminating, associations were found with low rumination times and clinical diseases and elevated key biomarkers (Kaufman et al., 2016; Schirmann et al., 2016). Stangaferro et al. (2016a; b; c) by developing an automated health-monitoring collar system that combined rumination with physical activity studied were able in a large observational experiment to detect disease events (displaced abomasum, hyperketonemia, indigestion, mastitis, and metritis) with a sensitivity of 59% and a specificity of 98%. Despite the limitations of a single-farm setting, those promising results highlight the possible usage of behaviour in short-term monitoring of the transition period in individual animals. However, again the implementation of these seemingly promising

technologies in the real-world farm seemed to have been hampered by external factors outside of the effectivity of these technologies.

On the long term, the advent of standard genomic testing and refined selective breeding solutions open the ability to adapt breeding programs to account for the health and welfare challenges posed by the transition period (Aleri et al., 2016). Genetic parameters have been studied for a plethora of phenotypes such as metabolic disease cases (Pryce et al., 2016), the general immune response (Thompson-Crispi et al., 2014), BHB concentrations in milk and blood (Leal Yepes et al., 2019; Nayeri et al., 2019; Pralle et al., 2020), dry matter intake (Tetens et al., 2014) and metabolic clusters (Atashi et al., 2020). Some of these studies were implemented into commercial products, e.g. CLARIFIED[®] plus (Zoetis, Parsippany-Troy Hills, US; Fessenden et al., 2020).

On the one hand, the long-term breeding solutions that seem to have been developed in recent years allow us to stay positive about tackling the transition period in the future. On the other hand, the short-term options to improve health seem to have failed in the last 15 years or are not able to penetrate the dairy industry. Although the general discussions in (inter)governmental organisations are focusing on the automation of farming to solve challenges like the transition period monitoring. The following section will frame this automation revolution in a historical context and expand on the possible leverage points this phenomenon presents in order to challenge the current status quo in the individual short-term transition period monitoring.

2.2 Precision livestock farming

2.2.1 What is precision livestock farming?

The FAO defines precision livestock farming (PLF) as follows:

“A data-based livestock management strategy that monitors and controls individual animal or group productivity, environment, health and welfare in a continuous, real-time and automated manner. It focuses on improving resource-use efficiency, productivity, quality, profitability and sustainability of livestock production.” (FAO, 2022b)

This definition provides a vague outline of how automation and digitalisation can be used concretely towards sustainability. Concrete examples of how PLF manifests itself as a tool, arise from the early-day publications that conceptualised PLF (Wathes et al., 2008; Berckmans, 2017). Wathes et al. (2008) defined PLF as applying process engineering principles and technologies to livestock production. Furthermore, to apply PLF for biological processes, they state it requires:

1. Continuous sensing of the output with information fed back to the process controller → growth, body condition scores, activity
2. A mathematical model able to predict dynamic responses of the output for variations in each input. → growth curves, nutrient modelling, heat detection from activity change
3. A target for each process output. → accepted deviance from growth curves, accepted fluctuations in body condition scores, threshold for heat event trigger
4. Actuators and a model-based predictive controller for the process inputs. → Alert systems for deviations

Recent advancements in sensor development and miniaturisation, communication technologies, modelling techniques, and modelling technologies, combined with the global digitalisation of agriculture, have shifted PLF-based approaches into higher gears (Aquilani et al., 2022). This resulted in the development of a plethora of new measurement/sensor technologies, sometimes combined with the development of models associating the produced data with real-world events (García et al., 2020; Aquilani et al., 2022). Those recent studies have focused on behaviour identification e.g. grazing behaviour in cattle with accelerometers collars (Aquilani et al., 2022) or on the identification of the health status of cows through different physiological measurements such as body temperature sensors or body weight (Stygar et al., 2021). Some studies going a step further by linking behaviour with health status (Stangaferro et al., 2016a; b; c). Despite the increasing array of technologies focusing on cow physiology/behaviour, the transfer towards real decision-making to improve health management has lacked some key components such as information integration (Rutten et al., 2013). Historically, automation has played a major role in improving the dairy industry. Almost a century before the conceptualisation of PLF, the dairy industry already focused on data collection and how it could improve milk production. Retrospectively, these approaches could be seen as early forms of PLF with all the principles of process engineering present, only lacking in the continuous sensing of the process. The next section will explain how DHI were established and how they fit in a PLF framework.

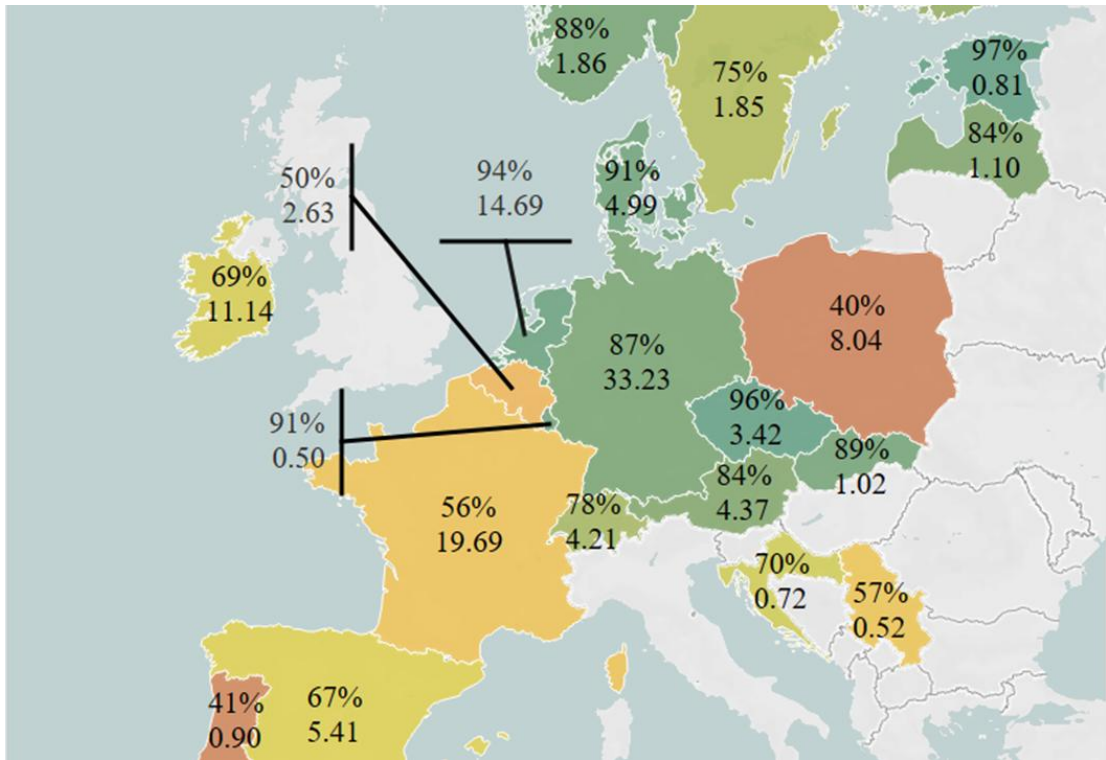
2.2.2 Dairy industry: precision livestock farming before it was cool.

The dairy industry has historically been a pioneer in the standardised recording of production data. In the US, this started as early as 1905 (Hodgson, 1964). In Belgium, the creation of a

generalised milk recording system was instituted in 1919 as a tool to improve dairy breeds (AWE). From those early days, the drive to collect data as a way to improve general production has always remained central in the dairy industry. Digitalisation and automation over the course of the 20th century have allowed the diversification of the data that could be recorded. This digital revolution also allowed technologies to assist the farmer in daily repetitive tasks that require physical or mental labour (Hodgson, 1964; Aernouts and Adriaens, 2023).

The best in-field example of this digital revolution has been the incremental changes to DHI programs. A program originally built as formative tool, which consisted of individual cow recording on farm where production and milk fat were recorded by itinerant specialists (Hodgson, 1964). This program incrementally grew from its initial goal of educating the farmers, to assist farmers in their breeding decisions, to incorporate economic evaluation of production, and to provide approximate herd health evaluations. Although the DHI has not been the only groundbreaking implementation of technological tools, the improvements of the modern dairy industry stand on the shoulders of the DHI programs (Voelker, 1981; Weigel et al., 2017). The focus on production data and derived tools over the last 100 years allowed breeding organisations to calculate the genetic value of high-producing animals and select the best within the population (Wiggans, 1991; Weigel et al., 2017). This potential has been widely understood in the industry, the International Committee for Animal Recording (ICAR) reports compiled the self-reported participation rates from the different DHI organisations all over the world. Figure 2.1 provides an overview of those participation rates in Europe in 2022. In Belgium, only 42.2% of the dairy farms participated in a DHI program in 2022. The DHI programs in Belgium are regionally organised, either by the Dutch/Flemish breeding organisation (CRV; Arnhem, The Netherlands) or the Walloon breeders association AWE (Ciney, Belgium). Reported participation rates in 2022 were 54.3% in Flemish herds and 25.8% in Walloon herds.

A)



B)

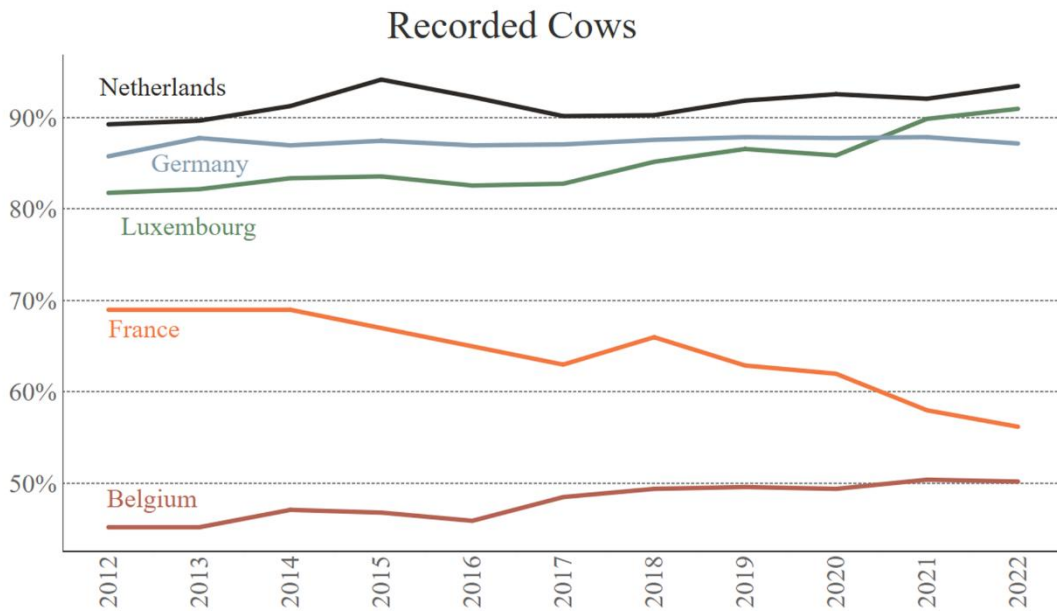


Figure 2.1 Self-reported participation rates in DHI programs across Europe. A) Map of European DHI participation rates. Countries are labelled with the percentage of recorded cows from the global population. Below that is the number of animals recorded (in 100.000 animals). B) Overview of participation in a 10-year window (2012-2022). On the y axis, the percentage of recorded animals is plotted for Belgium and its neighbouring countries.

Figure 2.2 visualises the modern analysis of how DHI programs could be interpreted as a PLF approach to milk production as a biological process. The inputs and output of the PLF framework are clear in the case of milk production: nutrition, management, health and genetics lead to an output of milk. Sensing of the output would be associated with the 4 to 6 weeks recording of milk production and a detailed overview of what is measured today: such as milk solids (fat %, protein %...), somatic cell count, urea in milk. These data are directly transmitted to the farmer serving as a non-model base controller, which will use farm-specific targets to assess the inputs of their milk production. Alternately, the role of a model-based controller could also be assigned to the breeding value models. These are used to estimate the genetic merit of animals, which in term can be used to influence the genetic input of milk production.

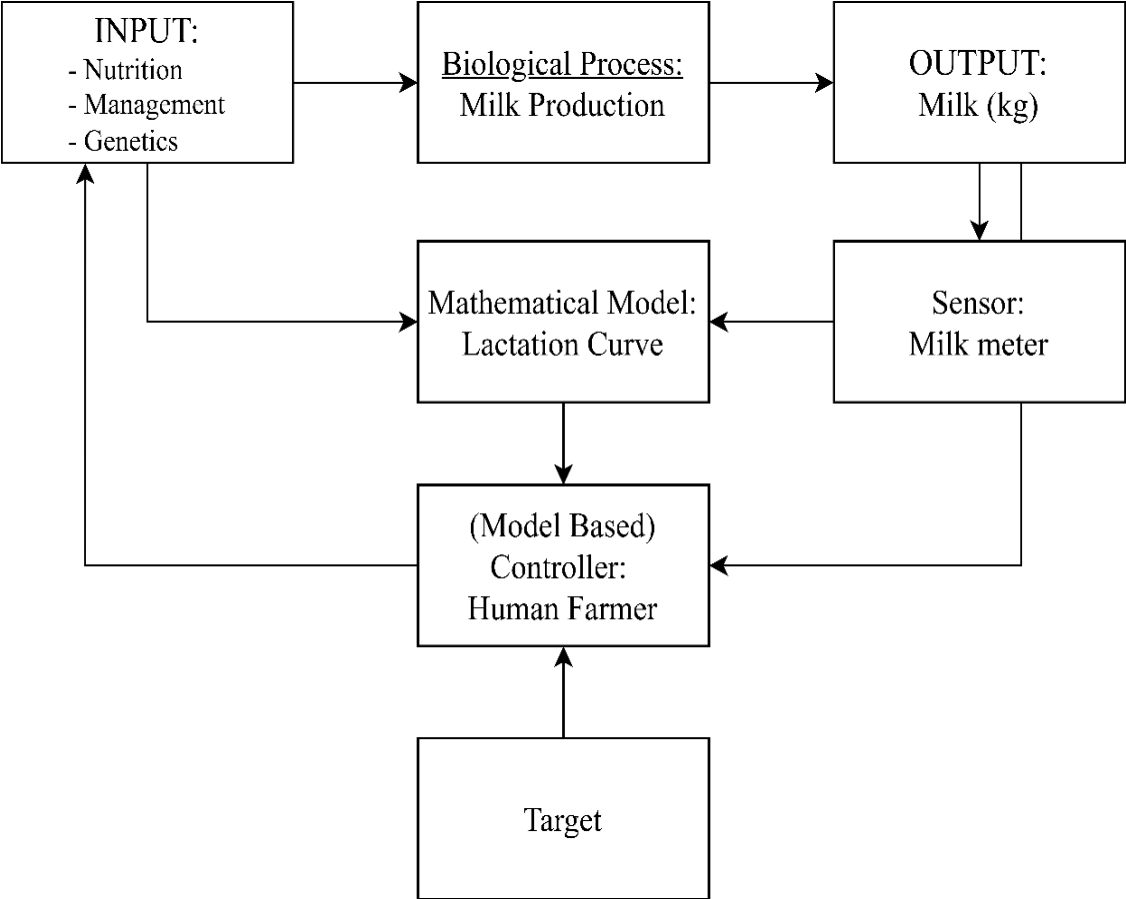


Figure 2.2 Graphical representation of how milk recording in DHI could be interpreted within a PLF framework. Adapted from Wathes et al. (2008)

Ultimately, mathematical lactation curve models perform the milk yield (output) predictions, serving as the PLF framework's core component. An overview of these models is presented in the following section

2.2.3 Lactation curve modelling

Modelling techniques

With the aim to process and evaluate the collected milk production records, mathematical modelling of the lactation process started in the 1920s with Brody et al. (1923). Following on their initial paper describing only the descending phase of the lactation curve, Brody et al. (1924) updated their model to include the ascending phase of the lactation curve. Despite 100 years of research and the numerous models developed over the years, the initial shape of the lactation curve remained unchanged, shown in Figure 2.3. Modelling the lactation curve has been characterised by a high variability of approaches: from empiric to mechanistic, from parametric to non-parametric, from linear to non-linear. Overall, the preponderant approach to lactation curve modelling has been empiric, parametric and non-linear. The subsequent segment will clarify the definitions of these terms, in the context of lactation curve modelling.

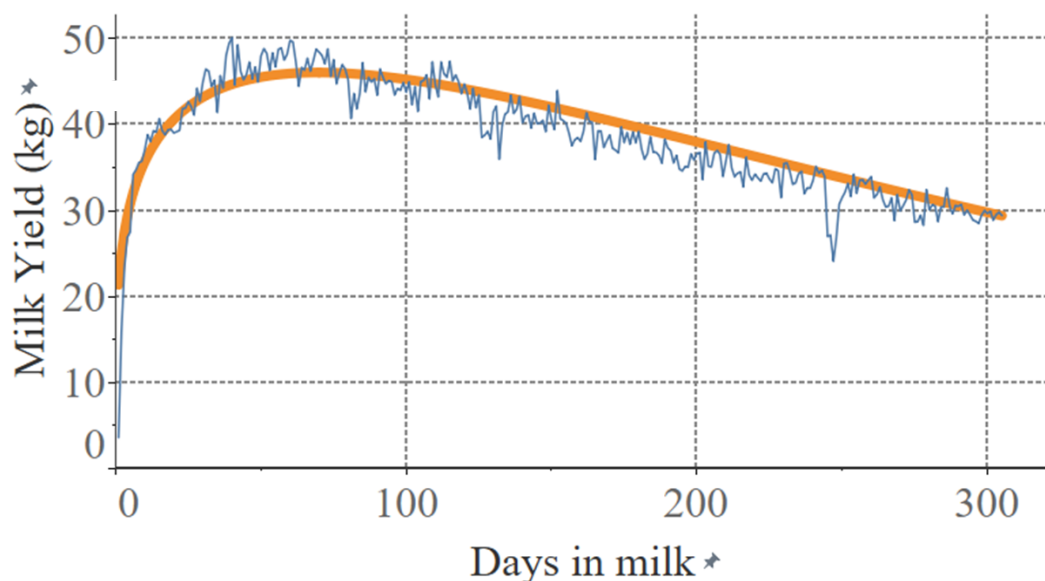


Figure 2.3 General shape of the lactation curve (orange, -) extracted from the fitted value of a wood model, the values used for the fit are the daily milk yields of a single cow (blue, -)

In essence, empiric means that those models are based on the descriptive components of lactation. This is in contrast to mechanistic models, which are based on underlying theoretical assumptions (Macciotta et al., 2011). As an example of this binary distinction, a mechanistic

model could base its milk yield prediction on the underlying energy compartmentation theories, whereas empiric models would only look at the course of milk yield over time. This distinction is often described as choosing between a practical value (empiric) or generating new knowledge (mechanistic; Nestorov et al., 1999). However, nuance is necessary when models are differentiated within this binary view. Assuming theoretical assumptions are only reserved for mechanistic models would overshadow the complexity of the lactation curve modelling (Macciotta et al., 2011). Concretely, even the simplest empirical model contains theoretical assumptions, such as the expectation of a peak yield, ascending phase, and descending phase. In a pharmacological context, Nestorov et al. (1999) underline the important relativity of this distinction, suggesting to focus on the model's predominant element and not underestimating the extrapolation capabilities based on this distinction.

Parametric refers to the model's use of a finite set of parameters to describe the response of milk production in time in a linear or non-linear function. The parametric approach allows the algebraic calculation of lactation curve characteristics, such as yield at peak production and time to peak using the first derivation of the lactation curve function or cumulative milk yields by integrating the function over a specific time interval (Bouallegue and M'Hamdi, 2020). Additionally, the non-linear functions used in most of the parametric models were exponential or included an exponential element in the function. This inherent exponential characteristic, while providing a good fit on the physiological process of the lactation, required some manipulation to reduce the computational power needed to use these models (e.g. linearisation of gamma functions) (Bouallegue and M'Hamdi, 2020). More recently, non-parametric models have also been studied. Those models differ from the parametric approach not by a complete absence of parameters, but by the fact that the model structure is not specified in advance (Salvador, 2017). The structure is extracted from the data itself in a so-called training process, where the number and essence of parameters can vary depending on the training data (e.g. neural networks) (Salvador, 2017). Non-parametric approaches are often referred to as black boxes due to the difficulty of interpreting the inner workings of specific models. The trade-off between interpretability (parametric models) and flexibility (non-parametric models), is not binary and exists on a spectrum dependent on the specific modelling techniques.

Overview Models

The emergence of these different approaches can be analysed as a stepwise search for improvement in lactation modelling. Table 2.3 enumerates the principal lactation curve models and their mathematical properties (Macciotta et al., 2011; Bouallegue and M'Hamdi, 2020). The first viable function was formulated by Wood (1967) in the early days of the computerised era. Wood's function is an incomplete gamma function (exponential), presented a good fit and has a computational less demanding linearised fitting option. Globally, up to the present day, the Wood (1967) model is the reference to model the lactation curve (Dijkstra et al., 2010). The underlying reasons for the central role of the Wood model could lay in the simplicity of the model (3 parameters) or the linearisation possibility. Despite the shortcomings of the Wood model and an array of remedial models (Cobby and Le Du, 1978; Congleton and Everett, 1980; Dhanoa, 1981; Dematawewa et al., 2007), the practicality and interpretability of alternate models seemed to hamper further implementations (Dijkstra et al., 2010).

The addition of a mechanistic layer to lactation curve models was studied at the end of the 1980s. Starting with the publication of Wilmink (1987), the correction of predictions for environmental factors was investigated. While these corrections were computationally challenging, a viable model was presented. This was possible by using clever model engineering, such as dimension reduction by splitting the factors into classes (e.g. 10 age classes) and separating the effect estimation of environmental factors (linear mixed model) from the fitting to the lactation curve (exponential curve). Similarly, Stanton et al. (1992), during the development of the test day model, integrated corrections for environmental factors, though these corrections were not applied to a lactation curve but to group/herd averages. Ptak and Schaeffer (1993), hypothesised that these corrected daily predictions would allow more accurate breeding values, but would require a drastic increase in computational power. By the late 1990s, multiple national breeding organisations had adopted the test day model for evaluating milk production traits (Weller et al., 2017). In Flanders and the Netherlands, the Wilmink model was used.

Fully mechanistic models developed by Dijkstra et al. (1997) and Pollot (2000) approached the milking process by exploiting the close relation between milk production and the number of cells in the mammary gland. These models describe the proliferation and death of mammary gland cells during lactation to model milk yield. An alternate approach within a mechanistic framework was formulated by Martin and Sauvant (2010), where the compartmentation and

prioritisation of energy towards different physiological statuses (growth, maintenance, gestation and lactation), using an array of sub-models. Their approach described changes in dry matter intake, body weight, foetal growth, milk yield and composition. Although these mechanistic models showed similar fits to the standard empiric models (Dijkstra et al., 2010; Martin and Sauvant, 2010), the direct link with biological processes would seemingly increase the interpretability of the model. Nevertheless, Pollot (2000) indicates the limitations formed by their model's overparametrisation and the difficulties it would have to fit in a monthly-based recording system.

The pursuit of greater flexibility, particularly in addressing individual lactations deviating from the average pattern, has driven various studies away from traditional parametric models (Macciotta et al., 2011). Initial studies looked at different non-parametric methods to predict milk yield such as cubic splines regressions (White et al., 1999; Silvestre et al., 2006), neural networks (Lacroix et al., 1995) and kernel density estimations (Weihrather, 1993). In recent years, the emergence and wide accessibility of non-parametric models (specifically machine learning techniques), combined with an ever-increasing variety and volume of data, has provoked an increase in interest in these techniques. The initial benefit of these models was an increased model accuracy when compared to traditional lactation curve models (Grzesiak et al., 2006; Murphy et al., 2014). In the neural network model developed by Liseune et al. (2020), predictions are adapted to account for health events such as mastitis, by making use of the inherent capability of neural networks to capture complex non-linear relations. Although these studies often indicated potential in a non-parametric flexible modelling of milk production, they remained mostly theoretical so far.

Since the first viable lactation curve models, scientific efforts have been made to improve lactation curve modelling. Notably, while improvements were found in the modelling performance, conversion of these into actual implementations within the field are rare. A hypothetical explanation of the underlying reason for this conservative pose could be found within the application field of those lactation models, i.e. the DHI and the PLF framework described above. Within this framework incentives for changing the mathematical model are low if the model can only reasonably accurately predict the output. Reasonability in this context would be dictated by the trade-off between (1) the ability of the controller-based model to impact the inputs based on the prediction, (2) the shortcomings of the model and (3) practicality (e.g. needed computational power, inputs). In other words, early-day milk yield predictions were accurate enough to predict impactful breeding values motivating a status quo.

Table 2.3 Overview of mathematical models representing the lactation curve of dairy cattle.

Publication	Type	Function	Parameters
Brody et al.(1924)	Empiric, parametric, non-linear → exponential	$Y_t = ae^{-k_1t} - ae^{-k_2t}$	a = Milk yield at calving k ₁ = ascending constant k ₂ = descending constant t = time
Wood (1967)	Empiric, parametric, non-linear → exponential	$Y_t = at^b e^{-ct}$ $\ln Y_t = \ln a + b \ln t - ct$	a = scaling factor ~ average milk yield b = pre-peak curvature c = post-peak curvature t = time
Wilmlink (1987)	Empiric, parametric, non-linear → exponential	$Y_t = a + be^{-kt} + ct$	a = scaling factor ~ average milk yield b = production increase towards peak yield c = production decrease after peak yield t = time k = correction factor for the moment of peak
Ali and Schaeffer (1987)	Empiric, parametric, non-linear → quadratic polynomial	$Y_t = a + b\frac{t}{305} + c\left(\frac{t}{305}\right)^2 + d\frac{305}{t} + k\left(\ln\frac{305}{t}\right)^2$	a = regression coefficient ~ milk yield b = regression coefficient ~ increasing slope c = regression coefficient ~ increasing slope d = regression coefficient ~ decreasing slope k = regression coefficient ~ decreasing slope t = time
Dijkstra et al. (1997)	Mechanistic, parametric, non-linear → exponential	$Y_t = ae^{\frac{b(1-e^{-ct})}{c}-dt}$	a = theoretical initial milk production b = cell proliferation rate at parturition c = decay of cell proliferation constant d = cell death constant t = time
Ehrlich (2011)	Mechanistic, parametric, non-linear → exponential	$Y_t = a\left(1 - \frac{e^{-\frac{c-t}{b}}}{2}\right)e^{-dt}$	a = scale, level of milk production b = ramp, increasing slope c = offset, correction between calving and start of lactation d = decay, decreasing slope t = time

2.2.4 (Model-based) Controller

In a historical context, as mentioned above, the standardised recording of production was set up as a tool to educate and provide concrete action points to farmers to manage their farms. In the PLF framework, the farmer could be seen as the ultimate controller step combining the measured and predicted milk yields with real-world constraints (feed, temperature, breeding decisions...) to steer the milk production on the farm. Alternately, the central model-based controller step for DHI programs has primarily been breeding value models. By linking the modelled milk production performance as a hereditary feature, sires and cows with high-producing descendants could be selected for further breeding. An overview of used techniques within breeding evaluation would fall outside of the scope of this manuscript, though the central role of these evaluations as a major force in the increase of milk production was already established in the 1960s (Hodgson, 1964). In Figure 2.4, this increase in milk production is visualised using data from the USDA and FAO (USDA NASS; FAO, 2022a). Unilaterally attributing the progression of milk production to genetic selection would overshadow the inherent interactions between genotype and environment in determining a phenotype such as milk production. Concurrent with the improvements in breeding evaluations, improvements in feeding, breeding and disease management created an environment where genetic merit could express itself. Nevertheless, the average national production follows the same trend as the historical evolution in breeding values reported by different studies and breeding organisations (Weigel et al., 2017; CRV, 2018). Breeding for milk production impacted the whole industry, whether a farmer actively participated in DHI programs or solely used curated semen provided by breeding organisations.

This success story is mostly built on the usage of 1 specific type of trait, the cumulative yields (kg milk, fat, protein) within 305 days in milk (Cole et al., 2023). Since the mid-1990s, breeding organisations have attempted to create more depth in the variety of genetically evaluated phenotypes (such as fertility, health, calving, workability and longevity), in an attempt to balance the possible adverse effects created by mono-trait selection towards high production (Miglior et al., 2017; Cole et al., 2023). The prerogative for this diversification in genetic traits was the diversification in available phenotypes (Miglior et al., 2017). More specifically, this manifested itself in the rollout of new measurements on the collected milk, with some of these phenotypes being presented back to the farmer as monitoring tools for identifying health and management issues. For instance, urea measurements in milk to identify nitrogen imbalances in the ration or individualised somatic cell count numbers for udder health monitoring.

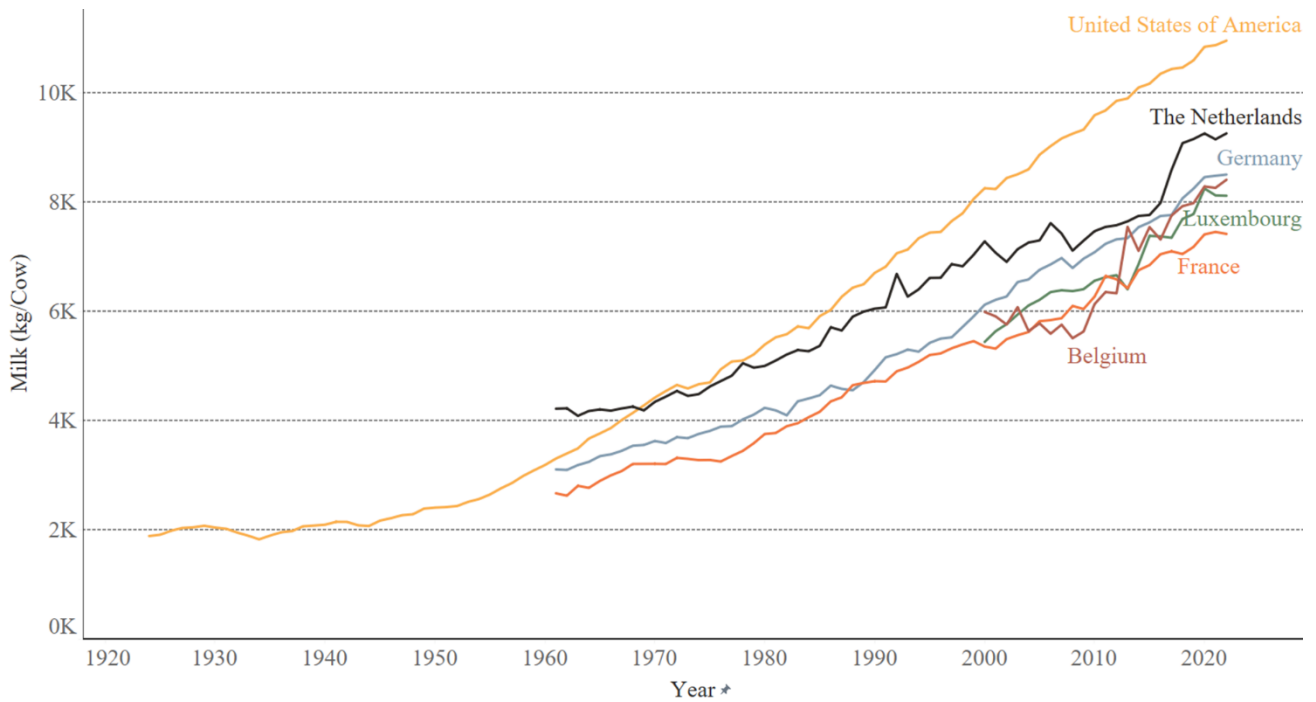


Figure 2.4 Evolution of milk production per cow on average in Belgium, neighbouring countries and the USA. Data was extracted from the USDA (1920-1960) and FAO (1961-2022).

Fundamentally, the usage of DHI as an effective day-to-day health monitoring tool is hampered by the sampling frequency (4 to 6 weeks), but as a monitoring tool at herd-level on the longer term it has been effective in different fields of dairy farm management such as udder health monitoring through somatic cell count determination (Ruegg, 2017) or feed optimisation through the determination of urea concentration in the milk (Jonker et al., 2002).

Currently, the industry's sustainability relies on a more tailored approach to tackle key hurdles such as health monitoring. The answer to this need was recently embodied by the emergence of a data-driven approach identifying health perturbations. Concretely, different studies focused on a similar approach, i.e. an individualised lactation curve is constructed representing the healthy status of the animal and deviations from that lactation curve were seen as health challenges (Adriaens et al., 2018, 2020, 2021a; b; Poppe et al., 2020; Abdelkrim et al., 2021a). These studies moved away from low-frequency data collection inherent to the DHI and harnessed the capabilities of high-frequency automated data collection through automated milking systems and milk meters in conventional milking systems. In doing so, insights on milk perturbation dynamics were unravelled, allowing objective quantifying of milk losses per perturbation (Adriaens et al., 2021a; b). Although promising methodologies were studied, implementations in the field have only been reported by Poppe et al. (2020) and again confined

to the field of genetic improvement. Specifically, The empirical characteristics of the perturbation (recovery, stability) have been regarded as a new phenotype for resilience and have been integrated into a breeding evaluation by CRV (CRV, 2024). While the dairy industry is in dire need of health-focused controller models, transferring the knowledge from literature to real-world implementations seems to be the final frontier.

Chapter 3.

Rationale

This thesis aims to address sustainability challenges within the dairy industry by prioritising improvements in dairy cow health, specifically by developing a tool designed to mitigate factors impacting both production and animal welfare. In the state-of-the-art, we focused on the transition period, clarifying that it has been and still is one of the main challenges on dairy farms. The drastic changes that cows experience during this period can potentially negatively impact animal health and, consequently, jeopardise the sustainability of the dairy industry. We presented an overview of the research efforts over the last 25 years, aiming to understand the physiological changes occurring during the transition period. The clinical signs of physiological maladaptation were categorised into different production diseases, and in addition, the general dogma was introduced that a clinical disease is generally preceded or accompanied by a subclinical state of maladaptation. While the previous chapter highlighted the direct link between this principle and the predominant use of biomarkers for monitoring the herd and individual cows, it also emphasised the potential shortcomings of this focus. To this end, an overview of alternatives listed the long- and short-term monitoring of the transition period on the farm. The second part of the state-of-the-art introduced the concept of PLF and its current recognition by global policymakers as a crucial path toward improving sustainability. Despite the 21st-century origin of the PLF framework, it provides a relevant lens for analysing early 20th-century DHI programs. By mapping DHI elements onto the PLF framework, we reveal how these programs leveraged industry aspects to improve dairy farming.

The state of the art highlights that dairy farms need powerful and validated health monitoring tools that can specifically address the transition period. Given the global urgency for more sustainable production practices across all industries, focusing on broadly applicable action points with a clear emphasis on practical implementations in the field is crucial. In our view, the process that yielded the current widespread focus on biomarkers for such tools, can be traced back to the lack of a clear understanding of etiopathology in the transition period. Resulting from the research studies addressing these processes, monitoring possibilities were formulated. While this approach resulted in some monitoring possibilities, it had limited effects in controlling the challenges associated with the transition period. Novel technologies have been mentioned in the state of the art, yet most of them require a more exhaustive validation or a financial investment from the farmers before they can be reliably used for disease monitoring. By approaching the transition period as a controllable process, the PLF framework offers concrete objectives for developing and understanding a data-based health monitoring tool.

Figure 3.1 provides a schematic overview of this proposed approach as further addressed in this thesis.

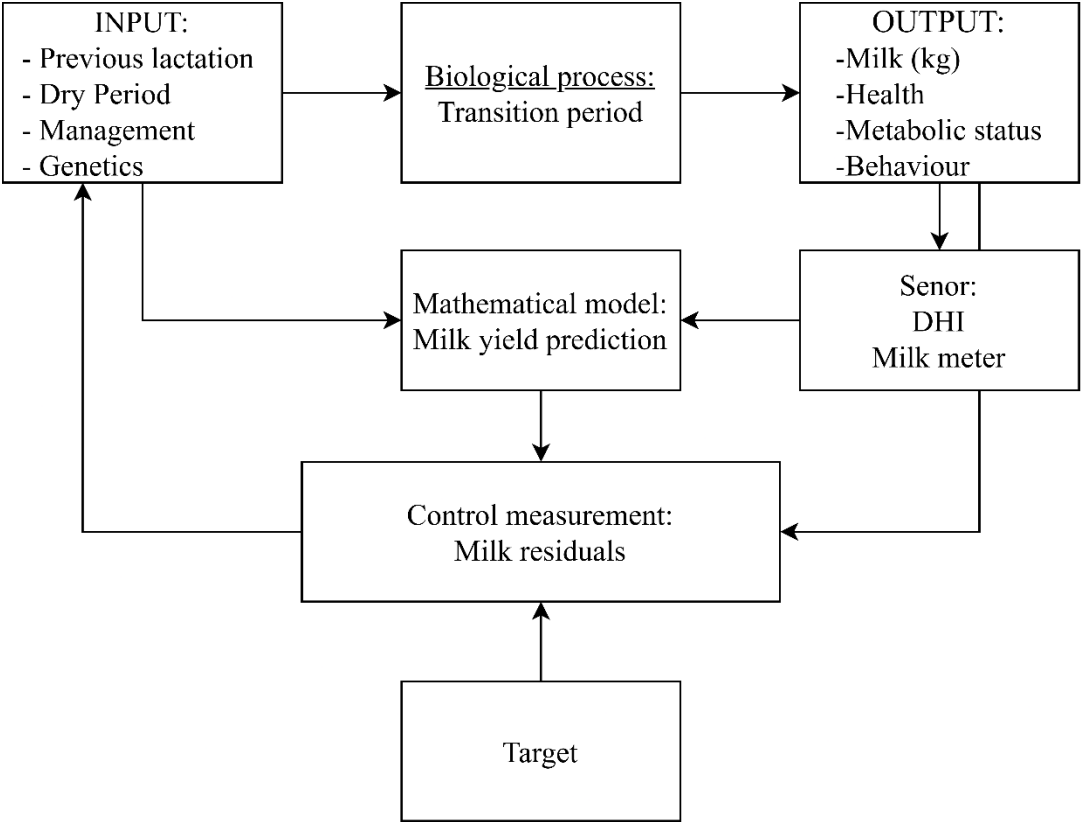


Figure 3.1 The transition period fitted within the PLF framework. The green arrows and the according annotations provide an overview of the following chapters and highlight how they tie in the transition period framework. Adapted from Wathes et al. (2008).

This thesis deliberately focused on sensing milk yield as the basis for evaluating the transition process, within the PLF framework. This approach aligns with health challenge identification methods developed in multiple studies (Adriaens et al., 2018, 2020, 2021a; b; Poppe et al., 2020; Abdelkrim et al., 2021a) and inspired by previous work like the transition cow index developed by Nordlund (2006), within this PLF framework, this thesis deliberately focused on the sensing of milk yield as measure to base the evaluation of the transition process. Nevertheless, using milk yield as the output of interest in the specific case of the transition period also highlighted one of the major gaps in knowledge. Since the transition period can disrupt milk production from the onset of lactation, the traditional way of fitting lactation models might be affected by these disruptions. Moreover, these models generally require milk yield data from the same lactation to predict the remaining lactation curve. Additionally, the

deliberate focus of our approach towards the sensing of milk yield by using DHI programs and milk provides dual advantages. Firstly, it facilitates the exploitation of historical, high-volume datasets which have been routinely collected. Secondly, it enables future implementation on a majority of farms with few extra investments. Investigating novel insights in well-established technologies like milk meters and DHI programs reduces the barrier to finding added value.

The scientific aims of this thesis focused on two key aspects of the proposed framework (Figure 3.1): the mathematical modelling and the development and validation of a novel control measurement. The mathematical model would involve a prediction model, which is able to accurately predict the unperturbed milk yield at the start of the lactation; this formed the primary goal of [Chapter 4]. Based on the predictions of this model, a measure would be derived. This metric needs to provide sufficient insight into the processes at play during the transition period to allow effective monitoring and control of the transition period. For this an extensive validation is necessary where this metric is evaluated with regard to its associations with key elements of the inputs and outputs of the transition period. This constituted the major goal in [Chapter 5],[Chapter 6] and [Chapter 7].

Chapter 4.

Prediction of first test-day milk yield using historical records in dairy cows

M. Salamone ^{a,b}, I. Adriaens ^b, A. Vervaeet ^a, G. Opsomer ^a, H. Atashi ^c, V. Fievez ^d, B. Aernouts ^{b,1} & M. Hostens ^{d, e, 1}

^a Department Internal Medicine, Reproduction and Population Medicine,
Faculty of Veterinary Medicine, Ghent University, Salisburylaan 133, 9820 Merelbeke, Belgium

^b Department of Biosystems, Division of Animal and Human Health Engineering, KU Leuven, Campus Geel,
Kleinhoefstraat 4, 2440 Geel, Belgium

^c Department of Animal Science, Shiraz University, Shiraz, Iran.

^d Department of Animal Sciences and Aquatic Ecology, Faculty of Bioscience Engineering, Ghent University,
Coupure Links 653, 9000 Ghent, Belgium

^e Department of Population Health Sciences, Division of Farm Animal Health, Faculty of Veterinary Medicine,
Utrecht University, Yalelaan 7, 3584 CL Utrecht, the Netherlands

¹ Authors contributed equally to this work

Adapted from:

Salamone, M., Adriaens, I., Vervaeet, A., Opsomer, G., Atashi, H., Fievez, V., Aernouts, B., & Hostens, M. (2022). Prediction of first test day milk yield using historical records in dairy cows. *Animal*, 16(11), 100658. <https://doi.org/10.1016/J.ANIMAL.2022.100658>

4.1 Introduction

The increase in milk yield of dairy cattle coincides with multiple challenges imposed to the cows, especially during the transition period in the six weeks around calving (Probo et al., 2018). In literature, the duration of transition period has been argued in recent years, certain authors define the transition period from dry off till 6 weeks after calving (Lopreiato et al., 2020). In this period, 30 to 50 % of the cows develop metabolic or infectious diseases such as mastitis, metritis, ketosis, lameness or, displaced abomasum (LeBlanc, 2010; Hostens et al., 2012; Pascottini et al., 2020b). To better support animals at risk of developing transition problems, identification methods to date have mainly focused on the lab analysis of metabolic markers as discriminating factor, either requiring milk or blood samples (Saun and Robert, 2006; De Koster et al., 2019; Grelet et al., 2019). Though, these techniques require action from the farmer, which may lead to poor identification performance. The development of an automatic alert system to point out animals at risk, could drastically improve the early identification of sick animals.

The advent of routine data collection on dairy farms has led to the opportunity to develop data-driven and automated health monitoring tools. For example, Adriaens et al. (2021) proposed to monitor udder health through perturbations in milk yield. A bottleneck to this approach is that it requires an accurate estimation of the expected milk yield in an unperturbed state, typically derived from a theoretical lactation shape and the production data available from a certain lactation (Poppe et al., 2020; Abdelkrim et al., 2021b; Adriaens et al., 2021b). In this context, multiple data-based models using high-frequency milk meter data have been developed to predict the unperturbed milk yield within the same lactation (Macciotta et al., 2011; Adriaens et al., 2018). Up till today, these models have had a wide range of application. They have, for instance, been used to estimate the expected production at herd or cow level, as tools to design suitable breeding strategies or genetic selection criteria, as individual health monitoring algorithms and as tools that estimate the response to management and environmental changes (Dematawewa et al., 2007; Ehrlich, 2011; Macciotta et al., 2011). Predicting the unperturbed milk production in the first weeks of lactation, however, is challenging, because at that stage there is a limited amount of milk production data to base the model on. More specifically, estimates of the unperturbed state based on only the first few days of milk production are deemed unreliable because health problems might already have influenced the lactation performance at or even before the start of the lactation. Therefore, differences between predicted milk yield and actual milk yield could be low while health issues are remain unnoticed

in the transition period. Nordlund (2006) identified a potential solution which could overcome the aforementioned issues, by using information from the previous lactation to make predictions on the milk production of the next lactation.

The application of advanced machine learning techniques to the increasing amount of data in dairy farming has led to the development of new insights in animal welfare and to real-time monitoring possibilities (Hermans et al., 2018). For example, recently Liseune et al. (2021) presented a deep learning model to predict the unperturbed daily milk yield of the first 305 days in milk (DIM), using daily milk meter data of the previous lactation in combination with cow and herd key performance indicators. This model demonstrated the application of machine learning to predict milk production of the next lactation. Relying on daily sensor data, this model excludes a wide range of farms that do not have milk meters installed. An alternative approach is to use test-day records (TDR) collected through national dairy herd improvement programs. TDR are recorded with a frequency of four to eight weeks, but present the advantage of having been collected for many years on a majority of dairy farms, thus having plenty of historical data readily available.

In the past, models based on TDR have been proposed for genetic evaluation of dairy cattle as a replacement for the traditional 305-d lactation yield. These models typically have the ability to account for environmental effects occurring on the day of milk recording (Mayeres et al., 2004). Additionally, random regression test-day models have been developed to predict future performance of animals in the same parity (for example, milk yield in the ninth month of the lactation based on the performance of animals in the first eight months of the current lactation). To our current knowledge, no study has evaluated the potential of using TDR of the current lactation to predict the performance of an animal in a next lactation. As TDR are widely available and standardised, they allow training of complex machine learning models. In their turn, these models can create added value to data that is already routinely collected on dairy farms.

In the present study, we aimed to combine powerful machine learning techniques with the idea of using historical data to predict milk production in the next lactation. To this end, a set of random forest regression models was developed to predict the expected milk yield at the first test-day of a lactation, based on features derived from historical TDR of the previous lactation and additional cow and herd information.

4.2 *Material and methods*

4.2.1 *Raw Data*

The raw dataset was accessed via the MmmooOgle™ platform (Bovicom, Puurs, Belgium) and originated from 102 herds located in six countries (BE: $n = 74$, the NL: $n = 16$ DE: $n = 5$, USA: $n = 4$, FR: $n = 2$, IT: $n = 1$), spanning a period of 20 years between 2000 and 2020. In total, data from 83 406 animals with on average 2.6 lactations per animal were available. These data included TDR and the corresponding cow information such as cow identifiers, calving dates, breeding dates and dry-off dates. At herd level, recording of data started between 2000 and 2012 and ended between 2007 and 2020, with a median timespan of 18.1 years (Q1: 15.8 years, Q3: 20.1 years). The TDR contained the following information: daily milk yield, test date, parity and DIM. Additionally, the data contained lactation curve parameters of the MilkBot model fitted on the test-day milk yield (Ehrlich, 2011). These MilkBot parameters (scale, offset, ramp and decay), summarize the shape of the lactation curve in a standardized way. A general flowchart displaying the different steps presented in this methodology can be found in Figure 4.1

4.2.1 *Data Selection*

All data editing, data processing and the model development were done using Apache Spark version 3.0.0 (The Apache Software Foundation, Wakefield, USA) running on the high-performance computing infrastructure of Ghent University, Belgium.

In order to obtain data suitable for the analysis, e.g. in terms of completeness, several selection steps were implemented. First, lactations with missing MilkBot parameters were removed. These parameters were essential to calculate cumulative milk yield at animal and herd level. Second, lactations followed by a subsequent lactation with at least one TDR were selected, referred to as lactation X and lactation $X+1$ respectively. The milk yield (in kg) in that first TDR of lactation $X+1$ (kgTD1_{X+1}). was defined as the dependent variable to be predicted.

The second selection step aimed at identifying lactations for which data quality was sufficient for the analysis. Moreover, lactations $X+1$ where the first test day (TD1_{X+1}) was measured after 60 DIM were excluded. Accordingly, the range for the first test day of the next lactation for which the model was trained and validated encompasses the period in which high producing dairy cows face major transition challenges (LeBlanc, 2010; Probo et al., 2018; Lopreato et al., 2020)..

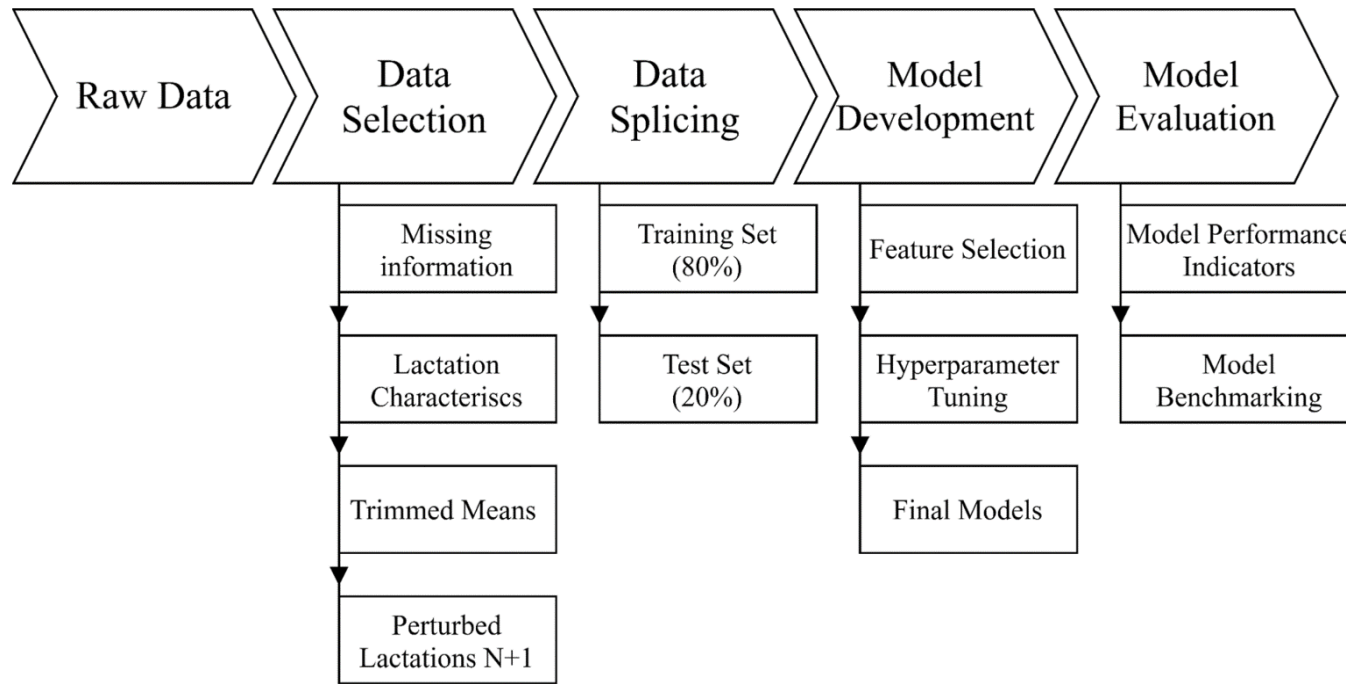


Figure 4.1 Flowchart displaying the different steps applied in this study to develop the nextMILK models for dairy cows.

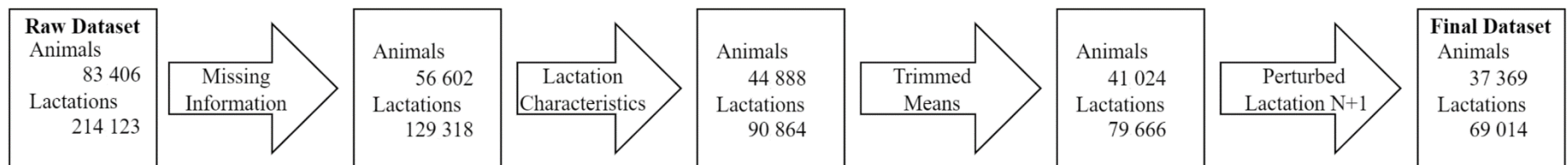


Figure 4.2 Graphical representation of the effects of each filtration step on the number of lactation and the number of distinct cows. This in order to construct the final dataset.

Furthermore, the calving interval in lactation X was at least 300 days and at most 530 days. The minimal and maximal age at first calving had to be between respectively 20 months and 36 months. Additionally, lactations with less than 8 TDR in lactation X were excluded to ensure having sufficient data to make the predictions. When more than 8 TDR were available, the TDR after the 8th were excluded from the predictors. The 8 selected TDR had to be taken regularly throughout the lactation, for example every 4 to 8 weeks. Lactations X for which no regularity between TDR was found were removed from the analysis.

The third and final selection step was applied at lactation level and aimed at identifying lactations from which the first test-day yield potentially belonged to a milk yield perturbation using characteristics of the dependent variable kgTD1_{X+1} . Because the aim of the model is to predict the expected milk yield of healthy cows in the beginning of lactation $X+1$, all 'unhealthy' lactations $X+1$ were removed from the dataset. Because detailed health information of the cows is not available for this dataset, a lactation $X+1$ is 'unhealthy' if it is perturbed at the first test day relative to the MilkBot lactation model fitted to the milk yield data of the entire lactation $X+1$. Moreover, if the actual daily milk production at the first test day of lactation $X+1$ was more than 6 kg below the expected milk yield according to the MilkBot model, then this lactation was identified as being 'unhealthy' and thus rejected. The 6 kg threshold corresponds with the reported RMSE performance of the MilkBot model in second and greater parities (Cole et al., 2012). In Figure 4.2, a flowchart shows the effect of all these data selection steps on the number of lactations and animals present in the final dataset.

4.2.2 Feature description

Several features at different levels (test-day, lactation, cow or herd) were available or could be calculated from the available data to enter in the prediction models. In this section, an overview of these features is given. These features included the DIM and corresponding milk yields at each test-day of lactation X and the DIM of TD1_{X+1} . Features at lactation level were parity number, the lactation length, days open for lactation X , the calving interval and days dry between lactation X and $X+1$. Parity number was considered as an ordinal categorical variable, parity number 5 and above were grouped in the same category. The age at first calving was considered as a cow-specific feature.

A set of additional features were derived from the features presented above. At the lactation level, six seasons of calving classes were defined by the month of calving to take into account

the seasonality of a lactation. The first class was attributed to December and January, the next classes were attributed to each consecutive pair of months. The season of calving was defined for both lactation X and lactation $X+1$. The cumulative milk yield at 21 DIM (M21), 75 DIM (M75) and 305 DIM (M305) of each lactation X was calculated by summing the daily milk yields for the different timespans using the MilkBot equation and parameters.

Some additional features describing the herd performance were calculated, including the average cumulative milk yields (M21, M75, M305), average age at first calving, average calving interval, days dry, lactation length, days open and the maximal and minimal daily milk production during the lactation. To account for different production levels in different lactations, these herd averages were stratified for each parity. Furthermore, to account for evolutions in production in a herd over time, the averages were also calculated in function of time using a sliding window of two years. The resulting herd average features were used to compute absolute differences of a cow in comparison with the herd accounting for the respective herd, year and parity number. Year was defined as the year in which lactation X started.

4.2.3 Data splicing

The aforementioned final dataset was randomly spliced into a test set (20%) and a training set (80%). Data were split before any model development was done to create an independent test dataset to evaluate the final model. The data splicing was performed randomly at animal level, preventing the inclusion data of the same cow in both test and training set. The usage of absolute differences with regard to the herd performances described in the previous section, preserves the independence between both sets.

4.2.4 Model development

In the present study, we developed a random forest regression model to predict the first test-day milk yield in lactation $X+1$ from the above-mentioned features. The following paragraphs describe the steps taken to train and evaluate the model.

Model Description

The random forest regression models were trained on the training set using the native MLlib library function of Apache Spark (The Apache Software Foundation, Wilmington DE, USA). Random forest regression is a supervised ensemble learning method proposed by Breiman

(Breiman, 2001). These models consist of multiple decision trees, where the prediction of each constituting decision tree is combined to create a final prediction. Random forest models can be used in classification or regression applications. When random forest is used for regression, the predictions of the individual decision trees are averaged to obtain the final estimation. Individual decision trees have an inherent tendency to overfit training data, random forest models mitigate this by combining the prediction of a multitude of individual decision trees reducing the overfitting problem (James et al., 2013). Additionally, when compared to black box models such as artificial neural networks, where gathering insights on the model's functioning is rather difficult, random forest provides the possibility to extract feature importance metrics. These metrics represent the relative importance of a feature normalized to sum to 1 calculated by the method presented in Hastie et al. (2009). In dairy research, random forest models have been mainly used as a classifier (Walsh et al., 2007; Shahinfar et al., 2014; Parker Gaddis et al., 2016; Borchers et al., 2017; De Koster et al., 2019). However, in recent years the usage of random forest models in regression applications for livestock data has increased (Dallago et al., 2019; van der Heide et al., 2019). The development of the random forest regression model consisted of the following steps: feature selection, defining the optimal hyper parameters, training of the model to obtain model parameters and evaluation of the final model.

Feature Set Selection

The full feature set comprised of all 40 features described above. We defined three feature categories within this final feature set: (1) the individual production features, (2) herd level production features and (3) reproduction-derived features (age at first calving, calving interval, days dry, days open, lactation length and the absolute difference of those features with the herd average). From this point on, three sets of features were used: only the production features (P),

production and herd features (PH) and finally the full feature set (FULL). A graphical representation of those three sets and their composing features is provided Figure 4.3.

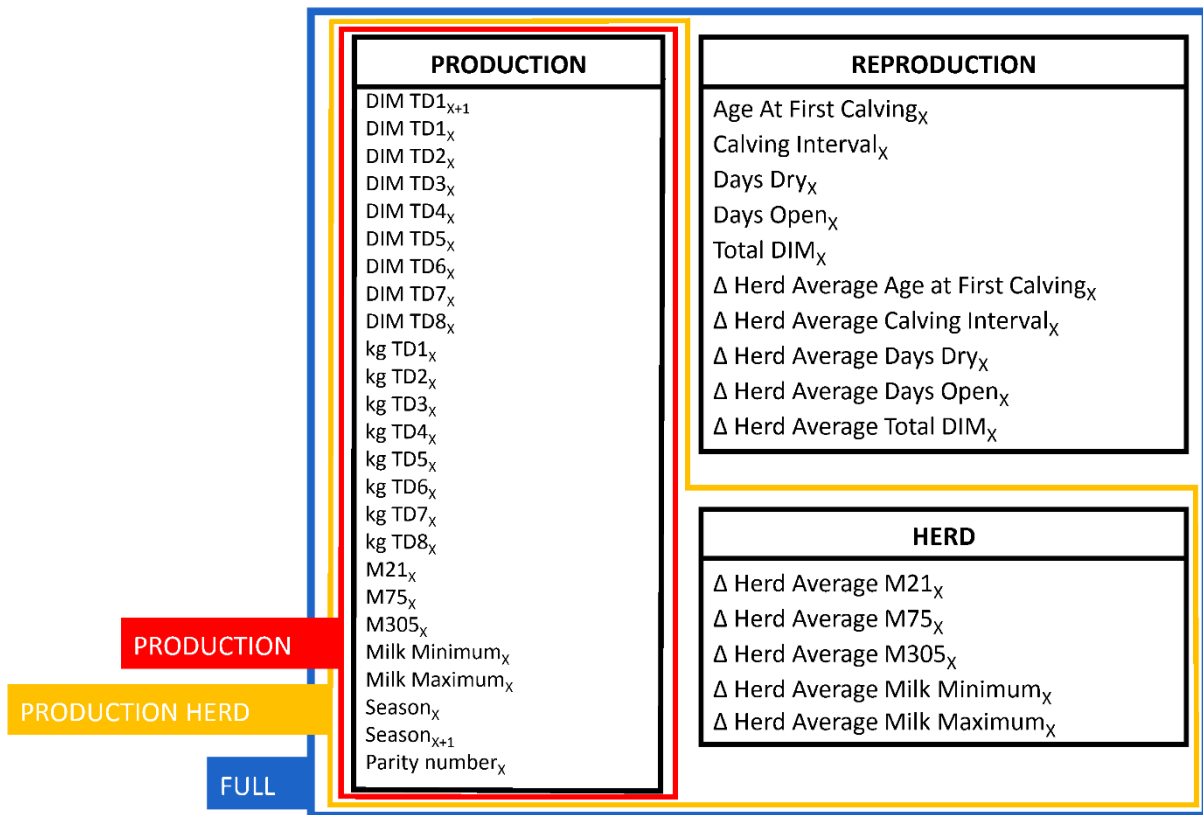


Figure 4.3 Graphical representation of the cow features within their classes and the three feature sets created with each of those classes. Abbreviations: DIM = days in milk, TD = test day, M = cumulative milk yield, X = lactation X, X + 1 = lactation X + 1

Hyperparameter Tuning

In this step, for each feature set the optimal pair of hyperparameters was established, which consists of the number of composing trees (5, 25, 125, 250 and 500 trees) and the maximal depth of those trees (5, 10, 15, 20, 25 and 30 levels deep). A subset of the random forest models was trained by using a random 5-fold cross-validation on the training set

The model performance was evaluated using the error with which the model is able to predict kgTD1_{X+1}. A suitable metric for the purpose of identifying the lowest error RMSE. This performance metric is calculated by taking the square root of the average squared difference between the actual kgTD1_{X+1} (y_i) and the predicted kgTD1_{X+1} (\hat{y}_i). The RMSE is expressed as seen in Eq. 1, with N the total number of predicted values:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (1)$$

The RMSE represents an absolute error with a clear unit (kg of milk) and the interpretation is straightforward: RMSE is a positive value, where the closer to 0, the better the model is able to predict the dependent variable. During this step, the best hyperparameter set model was defined by the lowest RMSE in the cross-validation. Statistically significant differences between the best performing hyperparameter set and the other sets were calculated on the squared value of the residuals by applying one-sided paired t-tests using SparkR. If multiple models had squared residuals that were not statistically higher than the best performing model, the least complex structure of these models was chosen.

Final Models

This optimal set of hyperparameters was then used to train the final random forest models, referred further as nextMILK_{FULL}, nextMILK_{PH}, nextMILK_P models using the respective feature set. They were trained with the full training set (80% of the lactation pairs) and evaluated by predicting the kgTD1_{X+1} for the test set. The feature importance in function of the predicted kgTD1_{X+1} was also analysed, to identify any biases or inconsistencies in the model performances. Ultimately, a set of model performance indicators were computed, to complete the evaluation of the nextMILK models.

4.2.5 Model Evaluation

Model Performance Indicators

The performance of the final nextMILK models was evaluated on the test set using four Model Performance Indicators (MPI) commonly used in similar studies, including the RMSE. In addition, mean absolute error (MAE), mean percentage error (MAPE) and R² were also used to evaluate the final model.

The MAE represents the average absolute difference between the actual kgTD1_{X+1} (y_i) and the predicted kgTD1_{X+1} (\hat{y}_i), represented in Eq. 2:

$$MAE = \frac{1}{N} \sum_{i=1}^{N-1} |y_i - \hat{y}_i| \quad (2)$$

Even though the definition and interpretation of MAE is similar to RMSE, MAE is less influenced by outliers in the residuals. More concretely, in addition to identifying the model

with lower errors overall, RMSE provides a better view on which model is less sensitive to extreme values in the prediction errors. For this reason, MAE will always be lower than RMSE, which could lead to overestimations of model performance in the case of large variation in the residuals.

The MAPE is calculated as the difference between actual kgTD1_{X+1} (y_i) and the predicted kgTD1_{X+1} (\hat{y}_i), divided by the actual kgTD1_{X+1} as shown in Eq. 3:

$$MAPE = \frac{1}{N} \sum_{i=1}^{N-1} \frac{|y_i - \hat{y}_i|}{y_i} \quad (3)$$

This MPI is a relative value between 0 and 1, where lower values indicate better predictions. It has the advantage of displaying errors in function of the actual value. However, MAPE penalizes negative errors more than positive errors, causing it to be a metric biased to favour models which underestimate the dependent variable.

The R^2 is calculated as shown in equation 4. The upper term represents the residual sums of squares, where difference between the actual kgTD1_{X+1} (y_i) and the predicted kgTD1_{X+1} (\hat{y}_i) is squared. The lower term represents the total sum of squares, which is difference between the actual kgTD1_{X+1} (y_i) and the overall mean of the actual kgTD1_{X+1} (\bar{y}_i):

$$R^2 = 1 - \frac{\sum_{i=1}^{N-1} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N-1} (y_i - \bar{y}_i)^2} \quad (4)$$

Model benchmarking

Besides the evaluation of prediction performance, benchmarking a newly developed ML model (i.e., the ones developed by us) against simpler methods can help to assess the added value of this work. To this end, we defined expert-based benchmark models that predict the kgTD1_{X+1} from (1) herd averages, and (2) the lactation curve of lactation X . The herd average milk yield for a certain test day was calculated at year-parity level, using information from the 2 previous years. For example, if lactation X started in 2012, the data from 2010-2011 of that herd was used. More specifically, 2 types of benchmark models were defined, the HERD-HERD and ANIMAL-HERD models.

For the “HERD-HERD” benchmark models, the kgTD1_{X+1} was predicted from the herd milk yield, taking the difference in milk yield between parities during (1) the first 75 days of lactation; HERD-HERD-75 or (2) during 305 days of lactation; HERD-HERD-305 into

account. For example, if the relative milk yield for parity 2 was on average 20% higher compared to parity 1 for a herd-year and the MilkBot model estimated that a cow produced 20kg of milk on DIM of TD_X during the first lactation, the $kgTD_{X+1}$ for that lactation was predicted as $20\text{kg} \times 120\% = 24\text{kg}$.

For the ANIMAL-HERD models, also the performance of the animal compared with the contemporary herd-mates was included. More specifically, when the animal produced e.g. 30% less than the contemporary herd mates in lactation X on TD_X , this factor was taken into account for the ANIMAL-HERD benchmark models. For example, a cow producing 20 kg of milk in lactation X for which the herd produced on average 22 kg on that same test-day, and for which parity 2 of the herd produced on average 25% more compared to the first parity, was predicted to produce $20/22 \times 1.25 \times 20 \text{ kg} = 22.7\text{kg}$. For the herd effect, also here two distinct models were defined, taking either the first 75 DIM of lactation (ANIMAL-HERD-75) and 305 DIM (ANIMAL-HERD-305) into account.

The benchmark models and the nextMILK model were used to predict the $kgTD1_{X+1}$ on the same test set. From these models, the residuals between predictions and actual $kgTD1_{X+1}$ were calculated to evaluate the model performance using the MPI. To compare the nextMILK and benchmark models statistically, the squared residuals were compared using ANOVA followed by a Benjamini & Yekutieli (2001) corrected paired one-sided t-test, significance was defined by $P < .05$.

4.3 Results

4.3.1 Data Descriptive

After the data selection steps, the dataset comprised 102 distinct herds with in total 37 369 unique cow and 69 014 lactation pairs. A general descriptive summary of the dataset after selection can be found in Table 4.1, together with an overview of general production and reproduction characteristics. The average kg of milk produced at $TD1_{X+1}$ was equal to 38.4 ± 8.6 (mean \pm SD). The median DIM at which the $TD1_{X+1}$ took place, was 21 (Q1: 13 days, Q3: 29 days). Figure 4.4 Descriptive plots of $TD1_{X+1}$, in panel A the distribution of the records in function of DIM, where we can see a clear plateau in the number of cow records between 7 and 30 DIM. In panel B an overview of the evolution of milk production during the first test-day is plotted. This average reaches a plateau around days 30 DIM. Abbreviations: DIM = Days in milk, $TD1_{X+1}$ = first test day in lactation $X+1$. A shows the distribution of DIM at $TD1_{X+1}$, demonstrating that the number of $X+1$ lactations with $TD1_{X+1}$ after 30 DIM is drastically reduced. In Figure 4.4 Descriptive plots of

TD1_{X+1}, in panel A the distribution of the records in function of DIM, where we can see a clear plateau in the number of cow records between 7 and 30 DIM. In panel B an overview of the evolution of milk production during the first test-day is plotted. This average reaches a plateau around days 30 DIM. Abbreviations: DIM = Days in milk, TD1 X+1= first test day in lactation X+1B, the average milk production in function of the DIM at TD1 is plotted, from which the low typical production at the start of the lactation can be seen. The split of the final data set yielded a training set and test set counting respectively 57 282 lactations and 11 732 lactations.

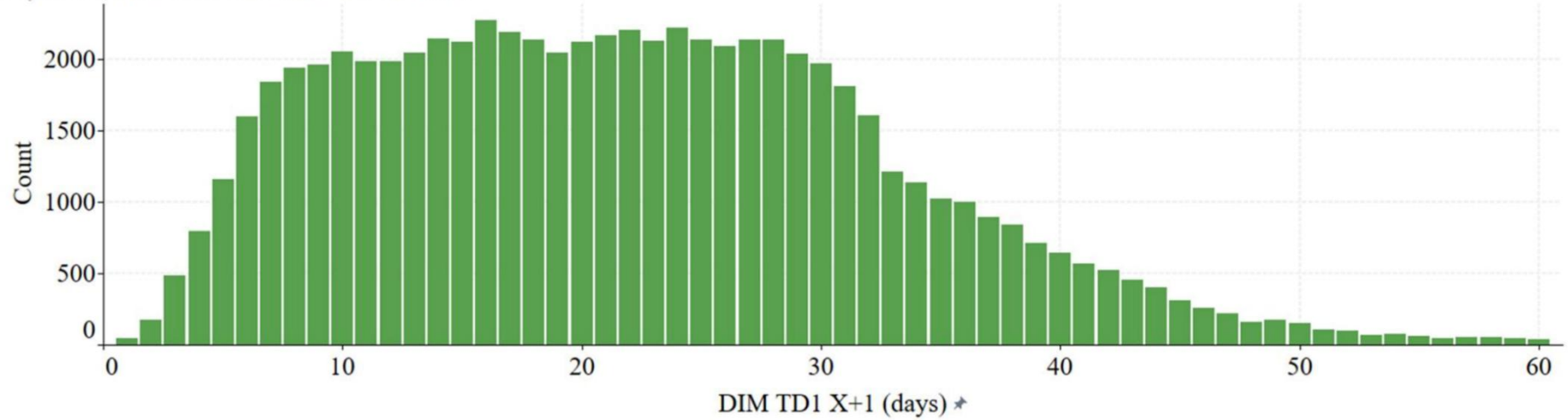
4.3.2 *Hyperparameters Tuning*

A total of 450 (5 composing trees * 6 maximal depth of trees * 5 cross-validation steps * 3 feature sets) subset models were trained for hyperparameter tuning. The performances of each of those models are displayed in Figure 4.5 for each of the three feature sets. When analysing the performances for the P feature set, the hyperparameter combination which yielded the best performance was 500 trees with a maximal depth of 30. The one-sided t-tests pointed out that the optimal feature set as the combination of 125 trees with a maximal depth of 20. This combination showed no significant difference with the best performing hyperparameter set. Similar results were found for the PH feature set and the FULL feature set, where in both cases the best performing model was constructed by 500 trees with a maximal depth of 30. The optimal combination was found by the t-tests to be 125 trees and a maximal depth of 15. These optimal hyperparameters were used in the rest of this study.

Table 4.1 Overview available dairy cow data in final dataset after the data selection step.
Abbreviations: TD = test day, kgTD1_{X+1}= milk yield of the first test day in lactation X+1

Item	Number for the whole dataset	Mean \pm SD over herds	Range over herds [minimum; maximum]
Number of herds	102		
Number of cows	37 369	366 \pm 365	[3 ; 2 254]
Number of lactation	69 014	677 \pm 654	[4 ; 4 361]
Parity 1	26 097	256 \pm 285.0	[2 ; 1 726]
Parity 2	19 280	189 \pm 189	[2 ; 1 285]
Parity 3	11 778	117 \pm 104	[1 ; 649]
Parity 4	6 268	65 \pm 54	[1 ; 327]
Parity 5+	5 559	56 \pm 55	[1 ; 285]
Age at first calving (years)		2.1 \pm 0.2	[1.7; 3.0]
Interval between TD (days)		33.5 \pm 5.1	[23; 56]
Calving interval (days)		397.5 \pm 49.1	[300; 529]
305d milk yield (kg)			
Parity 1		8 489 \pm 1 732	[3 034 ; 17 916]
Parity 2		9 835.0 \pm 2 034	[2 676 ; 22 030]
Parity 3		10 304 \pm 2 039	[2 628 ; 20 920]
Parity 4		10 294 \pm 2 034	[3 179; 20 326]
Parity 5+		9 971 \pm 1 923	[3 194 22 403]
Dependent variable: kgTD1 _{X+1} (kg)		38.4 \pm 8.6	[3.74; 75.32]

A) Distribution TD1 X+1 records



B) Milk Production TD1 X+1

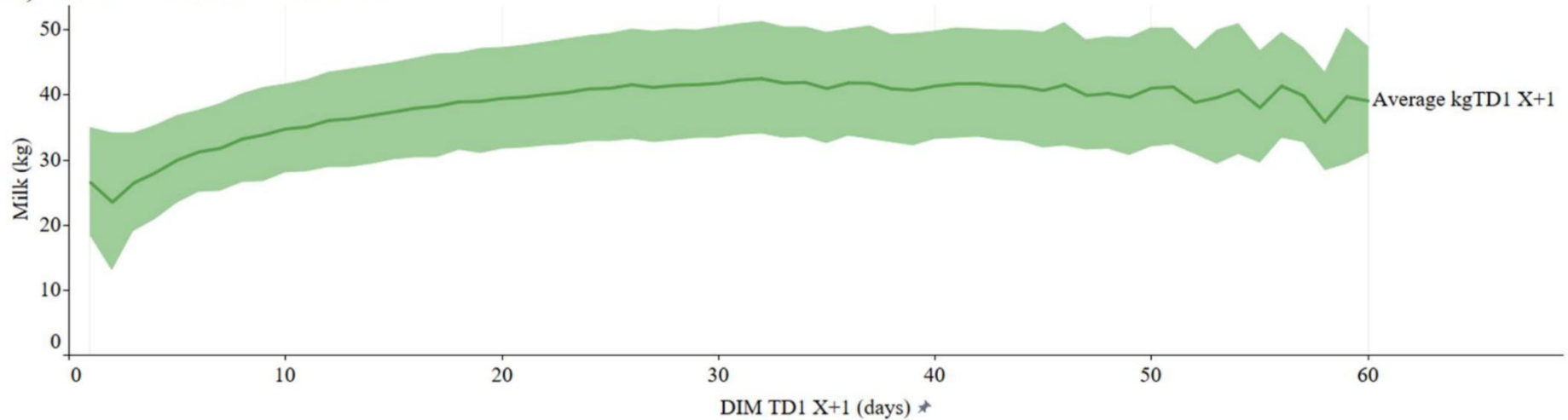


Figure 4.4 Descriptive plots of $TD1_{X+1}$, in panel A the distribution of the records in function of DIM, where we can see a clear plateau in the number of cow records between 7 and 30 DIM. In panel B an overview of the evolution of milk production during the first test-day is plotted. This average reaches a plateau around days 30 DIM. Abbreviations: DIM = Days in milk, TD1 X+1= first test day in lactation X+1

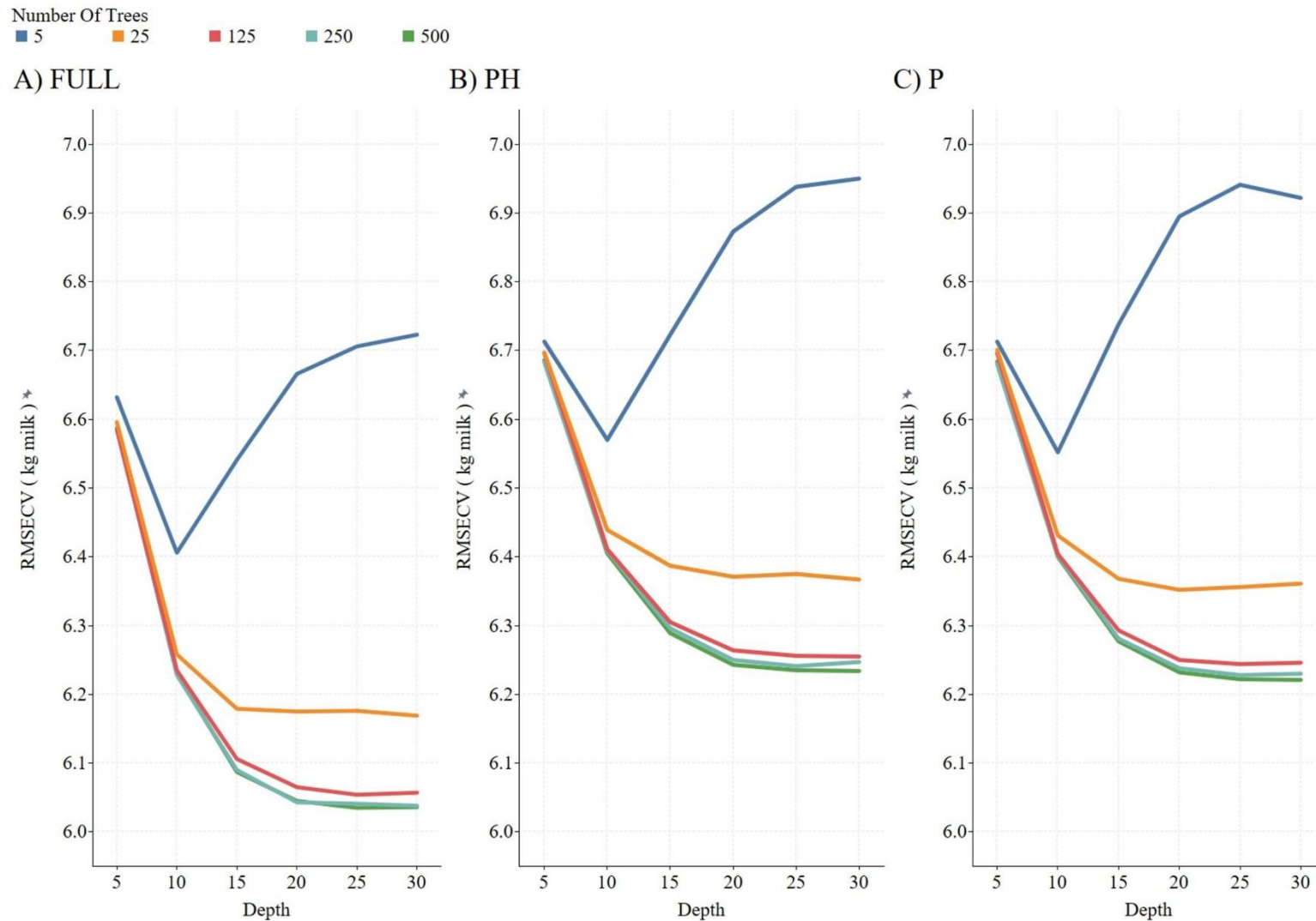


Figure 4.5 The RMSECV plotted for each of the hyperparameter set for each of the 3 feature sets extracted from dairy cows. On the x-axis the max depth of the trees trained in the random forest is showed. Each line representing a number of trees of those random forest models. Abbreviations: RMSECV = RMSE in cross validation, PH = production and herd, P= production

4.3.3 Model performances

The MPI values of the final models computed for the complete test set are summarised in Table 4.2. The full feature set yielded the lowest RMSE, MAE and MAPE and the highest R^2 , though the residuals were not significantly different. In Figure 4.6, the phenomenon of regression to the mean represented by all the nextMILK models is shown. The standard deviation of the dependent variable kgTD1_{X+1} in the test set is equal to 8.79 kg while for the models a standard deviation of 5.97 kg, 5.88 kg and 6.04 kg was found for respectively nextMILK_{FULL}, nextMILK_{PH}, nextMILK_P. In Figure 4.7, the performance of the nextMILK models is plotted in function of DIM TD1_{X+1} . It displays an aspect of the models' performances, where fewer observations of lactations within the range 0 – 5 DIM TD1_{X+1} and 50 – 60 TD1_{X+1} presented in Figure 4.4 Descriptive plots of TD1_{X+1} , in panel A the distribution of the records in function of DIM, where we can see a clear plateau in the number of cow records between 7 and 30 DIM. In panel B an overview of the evolution of milk production during the first test-day is plotted. This average reaches a plateau around days 30 DIM. Abbreviations: DIM = Days in milk, TD1_{X+1} = first test day in lactation $X+1$ A, seem to result in a higher variation in RMSE.

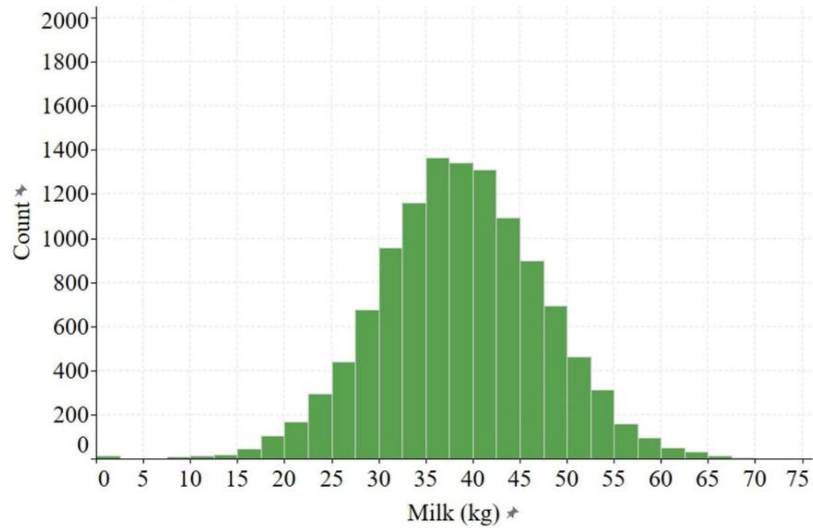
The feature importance has been extracted for each of the final models, the top 10 most important features is displayed in Table 4.3. The consistent presence of the same top 5 features in all the feature importance lists combined with their high importance score, emphasizes the importance of these features. The most important feature is DIM TD1_{X+1} in all the three feature sets, followed by the M305 of the lactation X. The milk production at the 4th and 5th TD is also found to be consistently present in the top 5 of most important features. It can be noted that for those TD_X production features, the corresponding DIM_X represent the tail of the feature importance list in all three final models. Furthermore, the herd and reproduction parameters generally have a relative low importance in the respective nextMILK models.

4.3.4 Model Benchmarking

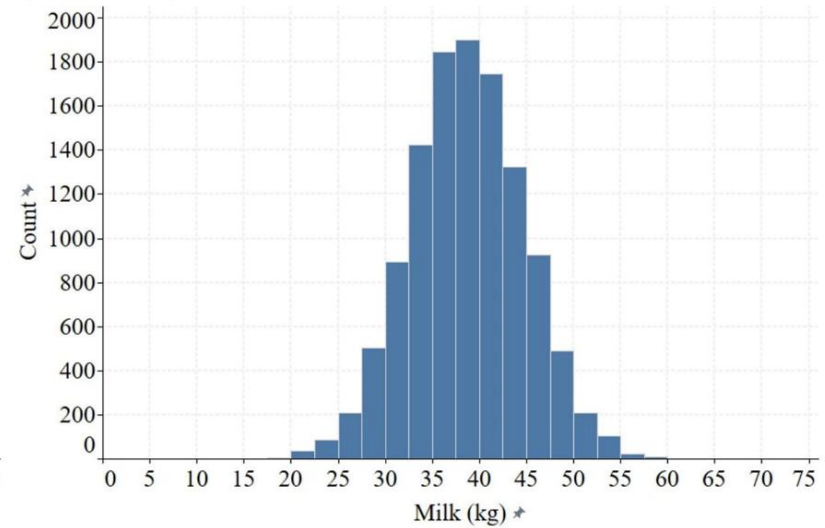
We identified that each of the benchmark models had significantly higher ($P < .05$) residuals (and thus, a lower prediction performance) compared to the three nextMILK models. Additionally, no significant difference in residuals between the nextMILK models was found. These results show the added value of our method compared to less complicated benchmark models. These results are summarised in Table 4.2 presenting also the MPI of both benchmark and nextMILK models. The difference in RMSE between the nextMILK models and the benchmark models ranges from 1.65 to 1.26 kg, in percentage this difference ranges from 23%

to 15%. This difference is not present in MAE where nextMILK_P and nextMILK_{PH} had similar MAE to benchmark HERD-HERD – 75, ANIMAL-HERD – 75 and ANIMAL-HERD – 305.

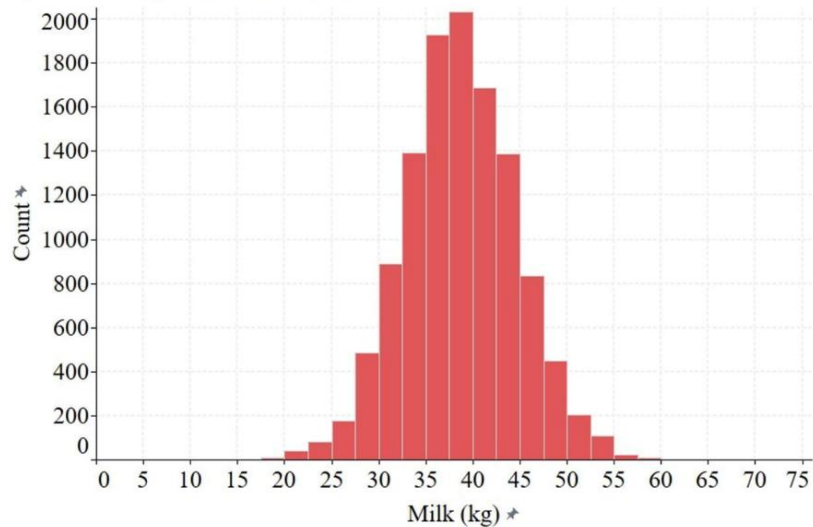
A) Distribution kgTD1 X+1



B) Distribution prediction of nextMILK FULL



C) Distribution prediction of nextMILK PH



D) Distribution prediction of nextMILK P

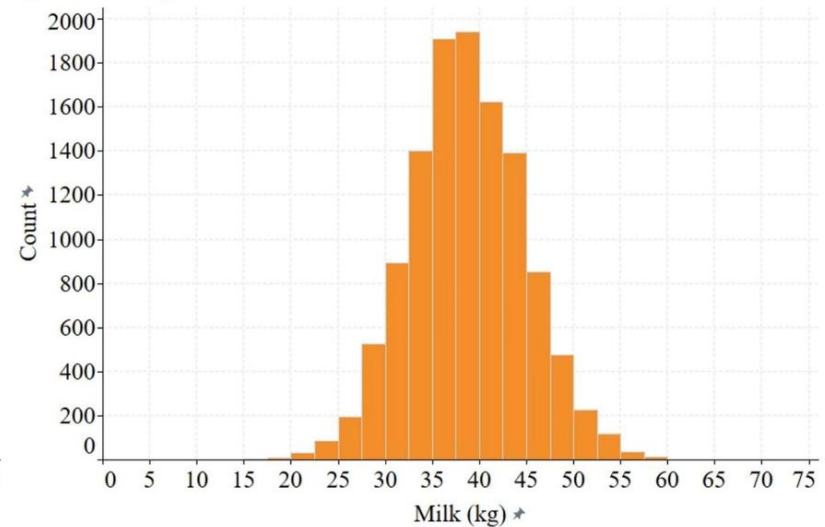


Figure 4.6 Distribution of the dependent variable kg produced by the dairy cow for the complete test set in panel A, in the 3 other panels the predictions of all nextMILK models are plotted for the test set. Abbreviations: TD1 X+1= first test day in lactation X+1, PH = production and herd, P= production

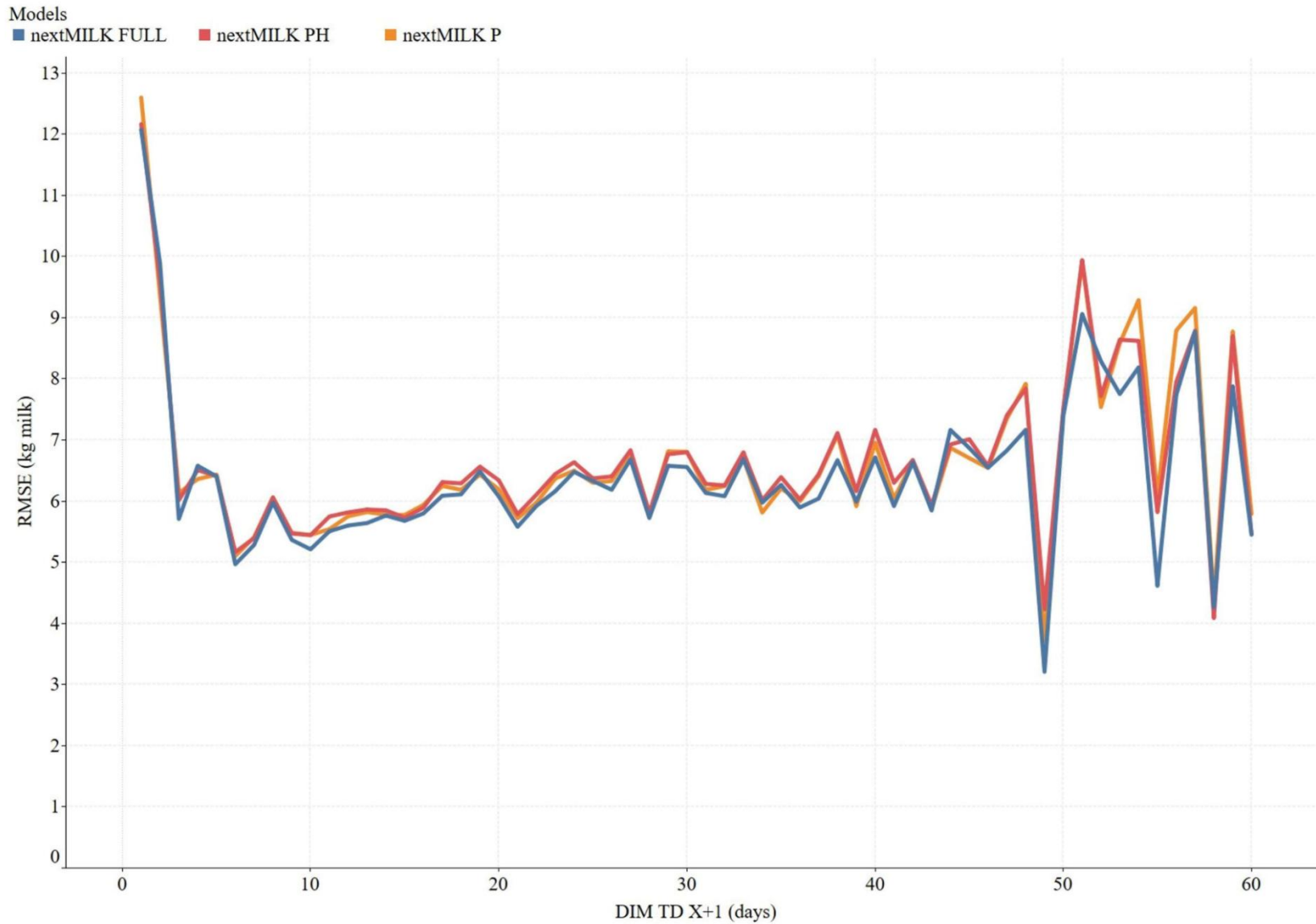


Figure 4.7 In this figure the RMSE is plotted for the test set in function of the DIM of TD1_{X+1} for the nextMILK models for dairy cows. Of note: the lower number of TD1_{X+1} records within the range 0-5 DIM and 40 -60 DIM displayed in figure 4A. Seems to be linked with a high variation of the RMSE in the same region of DIM. Abbreviations: DIM = Days in milk TD1 X+1= first test day in lactation X+1, PH = production and herd, P= production

Table 4.2 Model performance indicators of the nextMILK models & benchmark models in dairy cows. in the last column statistical significance is indicated with three symbols: #, †, ‡ indicate that the squared residuals of these models were significantly higher ($P < 0.05$) when compared with respectively nextMILK_{FULL}, nextMILK_{PH} and nextMILK_P. Abbreviations: MAE = mean absolute error, MAPE = mean absolute percentage error.

Model	RMSE	MAE	R ²	MAPE	Significant difference
nextMILK _{FULL}	6.08	4.56	0.52	0.1327	
nextMILK _{PH}	6.24	4.68	0.49	0.1369	
nextMILK _P	6.18	4.58	0.51	0.1339	
Benchmark I – HERD-HERD – 75	7.48	5.64	0.27	0.1608	#, †, ‡
Benchmark I – HERD-HERD – 305	7.89	6.59	0.09	0.1807	#, †, ‡
Benchmark II – ANIMAL-HERD – 75	7.37	5.50	0.30	0.1585	#, †, ‡
Benchmark II – ANIMAL-HERD – 305	7.40	5.59	0.29	0.1608	#, †, ‡

Table 4.3 Top 10 feature importance extracted from the nextMILK models for dairy cows. The table with the 40 features used in the models can be found in Appendix B. Abbreviations: DIM = days in milk, X = lactation X, X+1 = lactation X+1, TD1 = first test day, TD3 =third test day, TD4 =fourth test day, TD5 =fifth test day, TD6 =sixth test day, M305= cumulative milk yield after 305 days, PH = production and herd, P= production

Feature	Importance		
	FULL	PH	P
DIM TD1 _{X+1}	0.185	0.187	0.193
M305 _X	0.129	0.117	0.128
kgTD5 _X	0.060	0.057	0.065
kgTD4 _X	0.046	0.060	0.065
Milk Maximum _X	0.030	0.038	0.040
kgTD6 _X	0.030	0.038	0.039
Days Dry _X	0.027		
Season _{X+1}	0.026	0.033	0.037
Δ Herd Average Days Dry _X	0.025		
kgTD3 _X	0.023	0.028	0.031

4.4 Discussion

Overall, the calculated performance of the nextMILK models demonstrates the potential of using historical production data to predict milk production in the early stage of the next lactation. The consistent importance of DIM $TD1_{X+1}$ seem to indicate that the models are based on biological process such as the steep incline of the lactation curve at the start of lactation.

In 2012, Cole et al. described a theoretical minimum RMSE of 6 kg for predicting daily milk productions in cows, which is due to the variability resulting from changes in environment and health. The RMSE of the nextMILK models indicate that the nextMILK models are performing in the same order of magnitude as this described theoretical minimum. Furthermore, the comparison of residuals and MPI between the benchmark models and the nextMILK puts these MPIs in perspective of more simple approaches. In all cases, the nextMILK models showed significantly lower residuals and better MPI.

A possible explanation to the small difference in MPI over all the nextMILK models could be found in the feature importance vectors. The low importance of the reproductive and herd features indicates their limited contribution to the prediction, hence the absence of substantial differences between models. In term, comparison of these models with even smaller models should be made to evaluate from which point a real difference can be observed.

The absence of a high-quality disease registration dataset was one of the biggest limitations of the present study. It would have allowed us to select unperturbed lactations in a more objective way. Nevertheless, the applied data selection steps are set up to exclude abnormal lactations (in length, number of records) and possibly perturbed lactations. As a filtering step was applied to obtain high quality data, it should be further investigated how this affects the model performance for an extensive dataset with qualitative disease registration.

The creation of three distinct feature sets was motivated by the variation in quality and ease of collection of the three types of features. The individual and herd level production features are routinely collected with a high-quality standard, whereas the reproduction-derived features are considered poor in quality and challenging to collect on farms, because they often require manual inputs from the farmer.

To our current knowledge, only three studies with a comparable research question have been published. In his initial publication, Nordlund (2006) described the development of the Transition Cow Index (TCI), a tool to evaluate the success of the transition of individual

animals. The TCI uses a mixed model composed of 14 parameters such as DIM on the first test-day, previous 305d milk yield, lactation length of the prior lactation, SCC log score on the last test-day of the prior lactation, days dry and the milking frequency of the current lactation to predict the first test-day production and the 305-day milk production of the next lactation. The publication provides an extensive validation and performance of the TCI, while remaining elusive on the MPI of the model providing the TCI. This makes it impossible to compare the nextMILK models' performance with Nordlund's model.

The study of Dallago et al. (2019) focuses on predicting the production on the first test-day of first lactation animals only. The authors explored three different modelling techniques: multivariate linear regression, random forest and an artificial neural network. The RMSE reported in that study ranges from 5.02 to 5.10 kg of milk, whereas their MAE ranges from 3.9 to 4.0 kg and the R^2 from 0.30 to 0.32 across the three modelling techniques. In this study the author states that the artificial neural network model performed consistently better than the other ones. Using the MPI reported by the authors, compared to the ones of nextMILK, it seems the RMSE, the MAE and R^2 are respectively lower, lower and higher in their study. We believe that the usage of features collected on the 1st test-day of the 1st lactation as inputs for the models such as %fat; % protein, SCC could be the reason for the better performances of the models presented by Dallago et al. (2019) compared to the nextMILK models and even to the theoretical minimum RMSE described by Cole et al. (2012).

In the present study, we chose to not include information of the TD we predict in the feature set, to (1) keep our predictors independent of the outcome variable, and (2) because we aim at predicting the production when no health problems influence the production. If for example composition of the predicted TD would be included, this would affect these estimations.

The the Subsequent Lactation Milk Yield Predictor model presented by Liseune et al. (2021) predicts the 305 production of the next lactation using daily milk meter data. Even though the forecasting horizon of this model and nextMILK models differ, the author provides MPI in function of different forecasting horizons allowing their comparison. They reported an MAE for the 0-60 DIM time window of 5.8 kg, which is higher than the MAE calculated for the nextMILK models (4.56 - 4.68 kg). Moreover, the nextMILK model is designed to predict the daily milk production at the first test-day, between 0 and 60 days after calving, whereas the model of Liseune et al. (2021) aimed at predicting the sequence of individual daily productions

for the entire 305 days, including the transition from an increasing milk production in the first part of the lactation to a decreasing milk production after peak lactation.

The nextMILK models could be used on farms as a data-driven monitoring tool providing information in the short, medium and long term. In the short term, the nextMILK models could be utilized at animal level, providing ad hoc decision support for the farmer around the transition period. In the medium term, the nextMILK models could be applied to identify the generally expected production at the start of a lactation within the farm and assess any differences over different groups (e.g. age, pens) or time. Additionally, aggregating the transition failures over time at herd level could also provide a tangible tool for the farmer with which they could assess their transition management. In the long term, the nextMILK models could be used as part of breeding programs, to evaluate consistency over lactations or general tendency to transition success. Although all these potential applications seem promising, intensive validation is needed to investigate the extent to which the nextMILK models could fulfil these expected goals. In particular, estimating the effect of animals not reaching the first TD due to transition associated diseases on the monitoring capabilities of the nextMILK models. Furthermore the performance increase between the nextMILK model and the benchmark may be significant, but the biological relevance of this reduction in RMSE should be investigated when validating these models.

The usage of TDR in this research on the one hand, provides a wide application basis by allowing these models to be run on all farms that participate in milk recording programs. Additionally, these programs are a familiar resource for the farmer where a benchmark of the herd and individual performance is regularly provided. even though the exact details on the benchmark calculation in these programs remain largely unknown and dependent on the milk recording companies. On the other hand, TDR are intrinsically limited in views of the possible implementation mentioned hereabove. These limitations are due to the interval with which TDR are recorded, which could cause the TDR to be recorded too late. Nevertheless, if during validation the power of predicting transition failure is proven, an altered way of recording production in the early stages of lactation could be envisaged.

In their current state, the nextMILK models do not utilise all the data traditionally being collected on a TDR such as milk fat and milk protein content, as not for all TDR this information was available. Not using the milk constituents allows to, in the future, use a similar modelling approach with data collected automatically by on-farm milk meters. Still, for the development

of future models and when the additional information such as fat%, protein% or SCC is available and reliable, we consider it interesting to consider them as new features for the nextMILK models.

Chapter 5.

Milk yield residuals and their link with the metabolic status of dairy cows in the transition period

M. Salamone ^{a,b}, I. Adriaens ^{b,c}, A. Liseune ^d, S. Heirbaut ^e, X. P. Jing ^e, V. Fievez ^e, L. Vandaele ^f, G. Opsomer ^a, M. Hostens ^{e, g, 1} & B. Aernouts ^{b, 1}

^a Department of Internal Medicine, Reproduction and Population Medicine, Faculty of Veterinary Medicine, Ghent University, Salisburylaan 133, 9820 Merelbeke, Belgium

^b Department of Biosystems, Division of Animal and Human Health Engineering, Campus Geel, KU Leuven, Kleinhoefstraat 4, 2440 Geel, Belgium

^c KERMIT, Department of Data Analysis and Mathematical Modelling, Ghent University, Coupure Links 653, 9000 Ghent, Belgium

^d Faculty of Economics and Business Administration, Ghent University, Tweekerkenstraat 2, B-9000 Ghent, Belgium

^e Department of Animal Sciences and Aquatic Ecology, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, 9000 Ghent, Belgium

^f Institute for Agricultural and Fisheries Research (ILVO), Scheldeweg 68, Melle, 9090 Belgium

^g Department of Population Health Sciences, Division of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University, Yalelaan 7, 3584 CL Utrecht, the Netherlands

¹ Authors contributed equally to this work

Adapted from: Salamone, M., Adriaens, I., Liseune, A., Heirbaut, S., Jing, X. P., Fievez, V., Vandaele, L., Opsomer, G., Hostens, M., & Aernouts, B. (2024). Milk yield residuals and their link with the metabolic status of dairy cows in the transition period. *Journal of Dairy Science*, 107(1), 317–330. <https://doi.org/10.3168/jds.2023-23641>

5.1 Introduction

Monitoring dairy cow health with data-driven tools has been intensively studied in recent years (da Rosa Righi et al., 2020; Adriaens et al., 2021b; Ho et al., 2021). In many cases, these tools are based on the cow's milk yield as a reflection of her general health and welfare status. Dysregulation of the animal's metabolism is oftentimes reflected in her production performance before clinical symptoms arise (Codrea et al., 2011). In essence, most of the data-driven tools start with predicting the expected milk production assuming a healthy cow status. Several data-based models have been developed to predict the unperturbed lactation profile using high-frequency milk meter data or low-frequency test day data (Ehrlich, 2011; Poppe et al., 2020; Abdelkrim et al., 2021a; Adriaens et al., 2021b; Liseune et al., 2021). In these approaches, the unperturbed lactation is calculated as the expected daily milk yield under the assumption that the cow is healthy and not facing abnormal challenges like infections or severe metabolic imbalance. This expected daily milk yield can then be subtracted from the actual yield to obtain milk yield residuals. Severe and consistent negative residuals often referred to as milk yield perturbations can indicate or be linked to health-disturbing events (Poppe et al., 2020; Adriaens et al., 2021b).

The transition period, defined by Grummer (1995) as the period of six weeks around calving, is a critical period for a dairy cow (Horst et al., 2021). Smooth transitioning is associated with good production performance, fewer health problems later in lactation and good fertility. Generally, the cow's energy demand in this transition from gestation to lactation exceeds the energy intake, resulting in a negative energy balance (NEB). In reaction to this NEB, different adaptations in the cow's metabolism of lipids, carbohydrates and minerals occur. It has been estimated that 30 to 50% of the transitioning cows develop metabolic and/or infectious diseases, having long-lasting effects on milk yield (Mulligan and Doherty, 2008; LeBlanc, 2010; Hostens et al., 2012). Also, the metabolic profile of transitioning animals has been studied extensively, often with the aim of identifying (un)favorable status (McArt et al., 2013a; De Koster et al., 2019; Heirbaut et al., 2023). In this regard, the actual milk yield has been varyingly associated with unfavorable or favorable metabolic profiles (Horst et al., 2021). Nevertheless, the actual milk yield is not only reflecting the cow's current metabolic performance, but is also affected by her genetic merit and environment in the past (e.g., rearing, infections, ...). Therefore, the cow's residual milk yield, i.e., the difference between the actual milk yield and the expected milk yield, is hypothesized to better agree with her actual metabolic status and be less influenced by genetic merit or past environmental effects. This relation between the metabolic

profiles of transitioning cows and their residual milk yield has not been studied so far as models to predict the expected milk yield are generally not accurate at the start of lactation. This is because nearly no data on the current lactation is available at this stage to feed these models (Adriaens et al., 2018). Recently, two models have been developed to overcome this limitation in multiparous animals by taking into account data from the previous lactation of the respective cow. One of these two approaches, the Subsequent Lactation Milk Yield Predictor (originally SLMYP, in this thesis nextMILK_{MM}) used high-frequency daily milk yield data from the previous lactation obtained with milk meters (MM) to predict the expected daily milk yield at any time point in a new lactation (Liseune et al., 2021). The second method, further referred to as the nextMILK_{TD} model, used low-frequency daily milk yield data at 8 test days (TD) spanning the previous lactation to estimate the expected daily milk yield at the first TD of the next lactation (Salamone et al., 2022). Both models are fitted on the complete previous lactation and are able to combine information from the whole previous lactation to calculate their prediction. Validation of these models has shown that they both are accurate in predicting the expected ‘healthy’ milk yield at the start of a new lactation (Liseune et al., 2021; Salamone et al., 2022) and can thus be used to calculate the milk yield residuals in the transition period (MRT). The main objective of this observational study was to link these MRT, calculated with MM as well as TD records, to the health, metabolic status and dry matter intake (DMI) of high-producing multiparous Holstein-Friesian dairy cows in the transition period. As such, this study could support the development of a novel method to monitor animals and ultimately detect cows at risk of metabolic dysregulations in the transition period based on routinely collected milk yield data.

5.2 *Materials and methods*

5.2.1 *Animals and management*

Data were collected between October 2018 and October 2020 at the dairy research farm of the Flanders Research Institute for Agriculture and Fisheries (ILVO). This study was approved by the Ethics Committee of ILVO (file no. 2018/329) and it shares the experimental setup, and the analyte and DMI determination protocol with the study published by Heirbaut et al. (2023). Over a time span of 2 years, a total of 120 lactations and corresponding transition periods of Holstein-Friesian dairy cows were monitored. The average daily milk yield of the herd in that period was 31.5 kg/day, with a mean fat and protein content of 4.3 and 3.6%, respectively (milk recording, CRV, Arnhem, the Netherlands).

Dry cows were housed in a free stall barn and moved to a calving pen upon showing signs of imminent calving or, at the latest, 3 days before the expected calving date. After calving, the cows stayed for 3 days in the calving pen before being relocated to a slatted floored cubicle barn with their lactating herd mates. The ration during the 3 weeks before calving consisted of the partially mixed ration of the lactating animals supplemented with a dry cow mineral premix and balanced concentrate (on average 1 kg/cow/day) based on the Belgian-Dutch energy and protein evaluation systems. Lactating cows received a partially mixed ration. This ration is supplemented with balanced concentrate to fulfill the energy and protein needs, the concentrate is ramped up from 21 DIM to 6 kg balanced concentrate and 1 kg CovaSoy (FeedValid, Poederroijen, the Netherlands) and 0.3 kg soy meal. Concentrate intake was measured in the herringbone milking parlor (DeLaval, Tumba, Sweden) and automatic concentrate dispensers (Greenfeed©, C-Lock, Rapid City, US; DeLaval, Tumba, Sweden). Individual roughage intake was recorded by roughage intake control feeding bins (Insentec, Hokofarm Group, Marknesse, the Netherlands) from the start of the dry period until 35 days post-calving. During the period the cow resided in the calving pen, no roughage intake data was collected. A more detailed description of the ration can be found in Heirbaut et al. (2023). During the experiment, fresh lactating cows were milked twice a day in the milking parlor.

5.2.2 Data Collection

Blood samples

Blood samples were collected from the coccygeal or jugular vein in the morning (between 9h00 and 10h00) of days in milk 3, 6, 9 and 21, as described in detail by Heirbaut et al. (2023). Two types of vacuum tubes (BD Vacutainer, Becton Dickinson, Plymouth, UK) were filled using an 18-gauge needle (BD Vacutainer, Precision Glide, Becton Dickinson, Plymouth, UK). In one of the tubes, a clot activation factor (serum separation tube, SST™ II) was present. This sample was used for the analysis of β -hydroxybutyrate (BHB), Non-Esterified Fatty Acids (NEFA), insulin, fructosamine and insulin-like growth factor 1 (IGF1). The other tube contained NaF (2.5 mg/ blood + potassium oxalate 2 mg/mL) and was used for glucose testing. After collection, the blood tubes were gently inverted 10 times. The tubes with SST were kept at room temperature for 30 min, while those with NaF were placed on ice for transportation. Approximately one hour later, the SST and NaF tubes were centrifuged at room temperature at 1500 x g for 15 min and 1000 x g for 10 min, respectively. Serum and plasma were harvested and divided into aliquots of 2 mL each. All samples were stored at -20°C except those for IGF1 analysis, which were held at -80°C. The BHB, NEFAs and glucose concentrations were

analyzed with a Gallery™ Discrete Analyzer (ThermoFisher Scientific, MA, USA) using Randox kits (Randox Laboratories Ltd, Ibach, Switzerland) at the laboratory facility of the Flemish Animal Health Service (DGZ, Torhout, Belgium). Fructosamine was determined with the Nitroblue Tetrazolium method (Johnson et al., 1983; Jensen et al., 1993; Megahed et al., 2018). Insulin was measured with a Mercodia Bovine insulin ELISA kit (Mercodia, Uppsala, Sweden), whereas IGF1 was determined by the Poznań University of Life Sciences (Poland) using the Bovine IGF1 kit (LifeSpan Biosciences, Seattle, USA).

Production data

The following datasets were collected by the milking system, sensor system or the milk recording database at the research farm of ILVO and made available for this study: general cow information (cow identifiers, inseminations and calving dates), the daily milk yield as measured by the on-farm MM, TD milk yields and milk composition, individual disease events and daily dry matter intake at cow level. The TD records were registered with an interval of 5 weeks as part of the dairy herd improvement (DHI) program of the Dutch breeding organisation CRV (Arnhem, The Netherlands). Apart from the lactation in which the transition was monitored, production data (MM and TD) of the previous lactation were available to feed the SLMPY and nextMILK_{TD} models. Disease events were registered after diagnosis by a licensed veterinarian or farm staff, in which diagnoses were based on clinical symptoms only. No outlier removal techniques were used on the received datasets. Strict control for erroneous data had been done at the creation of these datasets, which made further cleaning of the data unnecessary.

5.2.3 Data Processing

Calculation of the milk yield residuals in the transition period.

The MRT were calculated using the two previously mentioned models: the nextMILK_{TD} model which is a random forest model developed by Salamone et al. (2022), and the deep learning framework nextMILK_{MM} model developed by Liseune et al. (2021). In its original publication, 3 nextMILK_{TD} variants were trained using different numbers of features. In the present study the production variant of the nextMILK_{TD} model was used, not relying on herd production performance and fertility traits. Briefly, the nextMILK_{TD} model uses a sequence of 8 TD in the previous lactation to predict the production on the first TD in the subsequent lactation. Our monitoring window ran from DIM 0 to DIM 21, and thus only the milk yield on TD falling within this window are relevant for this study. The test farm of ILVO has a 5-weekly milk recording scheme, only 61 of the 120 monitored transitions had a TD in this monitoring

window. To avoid having to remove nearly half of the monitored transitions, we therefore chose to use a TD milk yield derived from the daily (milk meter) rather than the TD data for all the transitions. To this end, we used a random number generator to sample a DIM between 7 and 21 DIM. Next, the average daily milk yield of the last 7 days prior to the selected DIM was taken as the “virtual” TD, following the standard procedure of the DHI program.

The nextMILK_{MM} model uses the daily milk yields of the entire preceding lactation, as measured with milk meters, to predict the expected daily milk production at any time point in the subsequent lactation. In parts of the previous lactation, cows were milked with a conventional milking system twice a day while being milked by an automated milking system at variable milking intervals in other parts of the lactation. This resulted in an unequal variability in the 24-h daily summed milk yields in the original dataset. To be able to apply the nextMILK_{MM} on the available MM data, an additional pre-processing of the MM data was needed. To correct for the variable milking intervals in the automated milking system, the first milking after midnight was partitioned into a part 'produced' on the previous day and a part on the current day and proportionally assigned to these days.

Additionally, a set of features needed to be calculated from the previous lactation to run the nextMILK_{MM} and nextMILK_{TD} models: lactation number, the cumulative milk yield at DIM 21, 75, 305 and end of the lactation, milk yield min and max, average and standard deviation, age at first calving, calving interval, total days in milk, the season of calving, days open. Herd and within-lactation averages were calculated for these features.

Of the initial 120 transition periods, 6 were excluded due to missing data points (such as historical data and DMI during transition) or the death of an animal during the observation period. Using the nextMILK_{TD} and nextMILK_{MM} models, predictions were made for the remaining 114 transition periods. nextMILK_{TD} variant was used to predict the expected kg of milk produced on the virtual TD (between 7 and 21 DIM) of the lactation under study. The milk yield residuals based on TD (MRT_{TD}) were obtained by subtracting the expected production from the actual production at that first TD. The milk yield residuals based on MM data (MRT_{MM}) were obtained by subtracting the daily prediction of the nextMILK_{MM} from the measured milk production for the first 21 days of lactation and calculating the average for that period.

5.2.4 *Statistical Analysis*

Programming and access to code

All statistical analyses were performed using R (R Core Team, 2020) version 4.1.2 and the following packages: car (3.1-1, Fox and Weisberg, 2019), skimr (2.1.5, Waring et al., 2022), heplots (Friendly, 2007), MASS (7.3-54, Venables and Ripley, 2002), sjPlot (2.8.12, Lüdtke, 2021) and tidyverse (1.3.2, Wickham et al., 2019). The code used for the statistical analyses can be accessed in the following repository: <https://doi.org/10.5281/zenodo.8108763>

Diseased vs. Not Diseased transitions

A distinction was made between monitored transitions during which the cow was clinically diseased (DIS) or not clinically diseased (NDIS), this distinction was solely based on the presence of clinical symptoms, no subclinical state was evaluated. Differences between both MRT were tested using a one-sided paired t-test grouped by clinical state. The differences in plasma analytes, DMI and MRT were compared between the DIS and NDIS groups. Before this comparison, the 6 plasma analyte variables (BHB, NEFA, insulin, IGF1, glucose and fructosamine) were averaged over the 4 sampling days (3, 6, 9 and 21 DIM). Due to the lack of normality of the measured variables, a Mood's median test was performed between the DIS and NDIS transitions for each variable. The significance level was set at $P < 0.05$.

Associations between MRT and blood plasma analytes, lactation number and DMI

Two linear models were developed to link the plasma analytes, lactation number and DMI to the MRT_{MM} and the MRT_{TD}, using data from only NDIS transition periods. The exclusion of DIS transitions was motivated by the fact that these transitions are known to have a different analyte pattern compared to NDIS transitions (LeBlanc, 2010). To avoid confounding, they were excluded from this part of the analysis.

In both linear models, the dependent variable was the respective MRT. The independent variables were the 6 plasma analytes variables averaged over 4 sampling days, the average DMI over 21 days prepartum, the average DMI over 21 days postpartum and the lactation number grouped in 3 categories (2, 3, 4+). All the continuous independent variables were standardized by subtracting the values with their respective means and dividing them by the standard deviation (z-score normalization). Prior to the model building, collinearity between variables was checked by calculating the generalized variation inflation factor with a threshold of $10^{[1/(2*\text{degrees of freedom})]}$.

Variable selection was performed using the stepAIC function from the “MASS” package. It conducts a stepwise selection procedure in both forward and backward directions to include or exclude the independent variables based on the Akaike information criterion. Pairwise interactions between all blood plasma analytes and DMI variables were also evaluated during this step. The significance level of each remaining variable or interaction was assessed, and non-significant effects ($P > 0.1$) were excluded. Ultimately, the remaining variables were used to build the final multivariate model. The sign and effect size of these variables were studied to determine their association with the MRT.

The residuals of the final model were checked for normality using the Shapiro-Wilk normality test, while the Breusch-Pagan Test was used to investigate the homogeneity of variances. The fitting performances of the models were defined as the adjusted R^2 . Additionally, the partial η^2 was calculated to estimate the variable importance within the model. The partial η^2 is a value between 0 and 1, with a higher value indicating more variability being explained by that variable after accounting for the variability already covered by the other model variables. To ensure data independence, correlation between model residuals was evaluated with specific emphasis on animals for which more than one transition was monitored during the study period. Specifically, the correlations of the residuals extracted from the final linear model were tested using Pearson’s product moment correlation coefficients. The significance level in this test was set at $P < 0.05$.

5.3 Results

5.3.1 Descriptive analysis

The historical production levels and reproduction key performance indicator of the experimental farm are presented in Table 5.1. Amongst the monitored transition periods there were 48, 37 and 29 cows in their 2nd, 3rd and 4th or higher lactation numbers, respectively. For 23 lactations, the animal was diagnosed with at least one or multiple disease events in the first 21 days: hypocalcemia ($n = 9$, downer cow responsive to calcium treatment), ketosis ($n = 7$, blood hyperketonemia, anorexia and decreased in milk production), clinical mastitis ($n = 4$, fibrin clots in milk potentially accompanied by fever), left abomasum displacements ($n = 10$, clinical symptoms and laparotomy), uterine infection ($n = 5$, abnormal vaginal discharge and enlarged uterus) and other ($n = 6$, other health issues such as leg problems). On average, these disease events occurred at 5.56 DIM. Accordingly, 23 transitions were categorized as DIS and 91 transitions as NDIS.

Table 5.1 Overview of the general production and reproduction key performance indicators of the research farm.

	Mean \pm SD	Range [min; max]
305d milk yield in previous lactation (kg)	9 974 \pm 1 589	[6 346; 13 455]
Age at first calving (years)	2.14 \pm 0.19	[1.79; 2.81]
Calving Interval (days)	406 \pm 72	[312; 655]

Furthermore, in Figure 5.1A and B, the distributions of both MRT are visualized and grouped for NDIS and DIS transitions. The average MRT_{TD} were -3.05 ± 4.59 kg and -8.57 ± 7.32 kg for NDIS and DIS transitions, respectively, while the average MRT_{MM} were 3.70 ± 5.67 kg and -1.37 ± 5.61 kg, respectively. The MRT_{MM} was higher than the MRT_{TD}, both in NDIS as well as DIS transitions ($P < 0.001$). Figure 5.1C shows the relation between the MRT_{TD} and MRT_{MM}, resulting in a Pearson correlation coefficient of 0.79. The prevalence of a diseased transition in the top 25% (from third quartile to max) of both MRT was 2 out of 28 transitions, corresponding to a prevalence of 7%. In the bottom 25% (from min to first quartile) of both MRT_{TD} and MRT_{MM}, 13 out of 28 transitions were affected by a clinical disease event, corresponding to a prevalence of 46%. In Table 5.2 the median and range of the analyte concentrations, MRT and DMI are presented separately for DIS and NDIS transitions. Mood's median tests revealed both MRT to be lower in DIS transitions (MRT_{TD} = $P < 0.001$, MRT_{MM} = $P < 0.001$), though the absolute ranges of milk residuals between the DIS and NDIS transitions were similar in both. When the MRT were grouped by lactation number, lower values were found for DIS transitions in lactation 2 for MRT_{TD} ($P = 0.02$) and MRT_{MM} ($P = 0.02$) and lactation 4+ for MRT_{TD} ($P = 0.02$). Moreover, the Mood's mean tests showed DIS transitions to be associated with a lower DMI intake postpartum ($P < 0.001$) and plasma IGF1 concentration ($P = 0.04$), while the plasma NEFA concentration was higher ($P < 0.001$). There was a tendency for a higher plasma BHB concentration in the DIS transitions ($P = 0.09$).

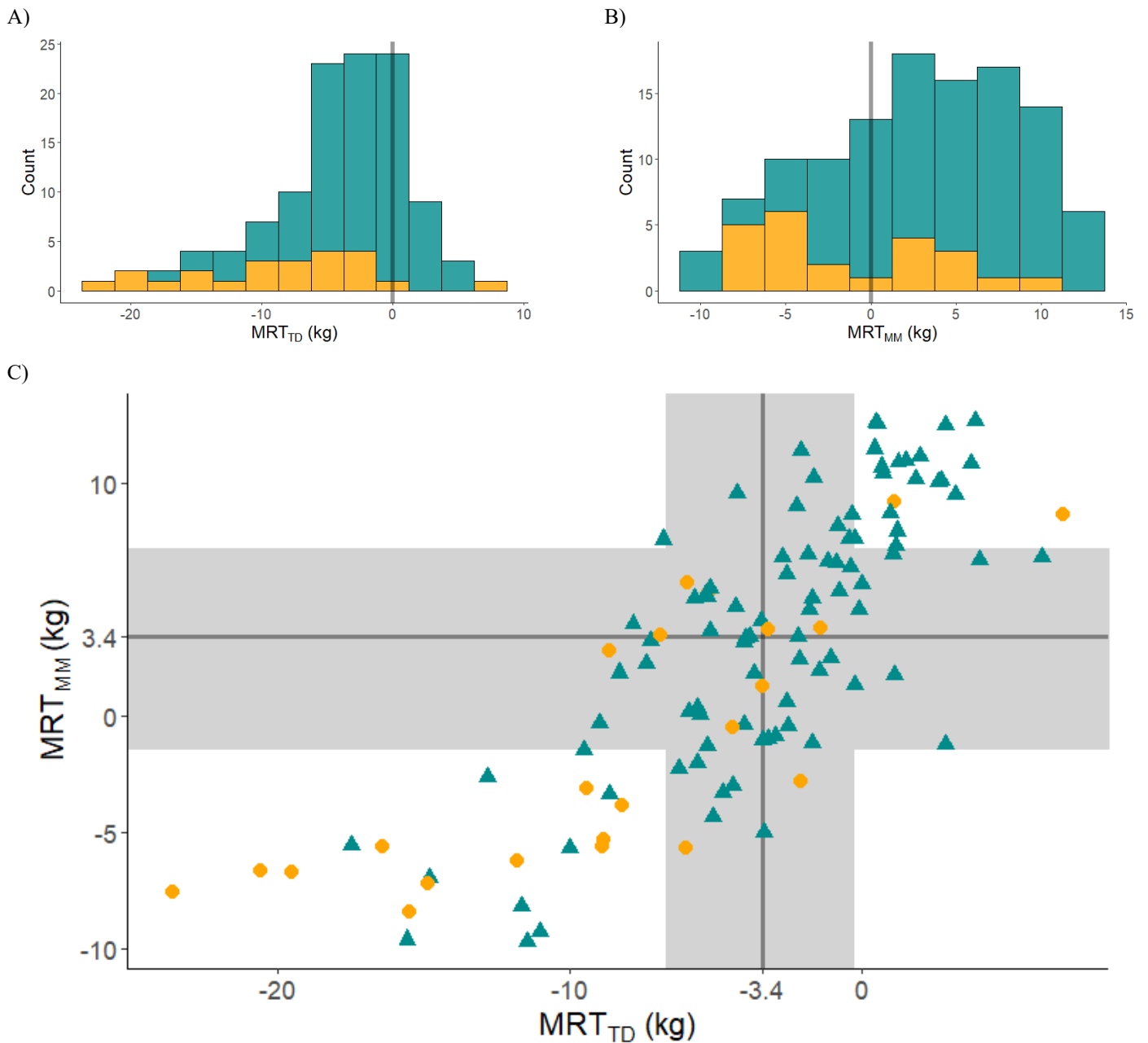


Figure 5.1 Panels A and B show the milk yield residuals in the transition period (MRT) for test days (MRT_{TD}) and milk meter (MRT_{MM}) distribution. The colors distinguish between non-clinically diseased (blue) and diseased (orange). Both colors are stacked on top of each other. The bin width in panels A and B is 2.5. In panel C, the relation between the MRT_{TD} and MRT_{MM} is plotted. A distinction is made between the clinically diseased (●) and non-clinically diseased (▲). The axes are located on the median of each MRT. Distribution bands were also plotted to represent the interquartile ranges.

Table 5.2 Comparison of transitions affected by a disease event (DIS) and not clinically diseased (NDIS) for all measured parameters and milk yield residuals in the transition period (MRT) for test days (TD) and milk meter (MM). Results from the Mood's median test between NDIS and DIS are annotated with *, ° representing $\alpha < 0.05$ and $\alpha < 0.1$, respectively. Abbreviations: BHB = β -hydroxybutyrate, NEFA = Non-Esterified Fatty Acids, IGF1 = insulin-like growth factor 1.

			NDIS (n= 91)		DIS (n= 23)			
			Median	Range [min; max]	Median	Range [min; max]		
MRT								
	TD (kg)		*	-2.52	[-17.5; 6.17]	-8.21	[-23.6; 6.91]	
	MM (kg)		*	4.55	[-9.74; 12.7]	-3.13	[-8.43; 9.19]	
MRT/ lactation number								
	TD (kg)							
	2		*	-0.42 (n = 40)	[-12.8; 3.91]	-3.35 (n = 8)	[-23.6; 1.12]	
	3			-2.74 (n = 27)	[-15.6; 6.17]	-2.89 (n = 10)	[-16.4; 6.91]	
	4+		*	-4.59 (n = 24)	[-17.5; -0.01]	-8.82 (n = 5)	[-20.6; -8.66]	
	MM (kg)							
	2		*	7.64 (n = 40)	[-2.64 ,12.7]	2.39 (n = 8)	[-7.59 ,9.19]	
	3			2.43 (n = 27)	[-9.63 ,10.9]	-4.74 (n = 10)	[-8.43 ,8.64]	
	4+			-0.65 (n = 24)	[-9.74 ,6.63]	-5.29 (n = 5)	[-6.69 ,2.81]	
Dry Matter Intake								
	Average	21	days	14.5	[10.1; 19.3]	13.6	[10.5; 18.9]	
	prepartum (kg/d)							
	Average	21	days	*	21.4	[13.8; 26.5]	17.8	[11.4; 23.0]
	postpartum (kg/d)							
Plasma analytes								
	BHB (mmol/L)		°	0.88	[0.46; 2.22]	1.28	[0.568; 3.66]	
	NEFA (mmol/L)		*	0.56	[0.17; 1.61]	0.76	[0.34; 1.58]	
	IGF1 (ng/mL)		*	87.5	[37.8; 250.0]	58.8	[36.2; 175.0]	
	Fructosamine (μ mol/L)			257	[202; 391]	252	[198; 338]	
	Insulin (ng/mL)			0.224	[0.025; 0.697]	0.158	[0.062; 0.807]	
	Glucose (mmol/L)			3.05	[2.35; 3.98]	3.10	[2.35; 4.41]	

5.3.2 Models

MRT_{TD}

In the case of MRT_{TD}, after the variable selection step, 5 variables and 2 interactions were retained in the final model. None of the variables exceeded the threshold for collinearity.

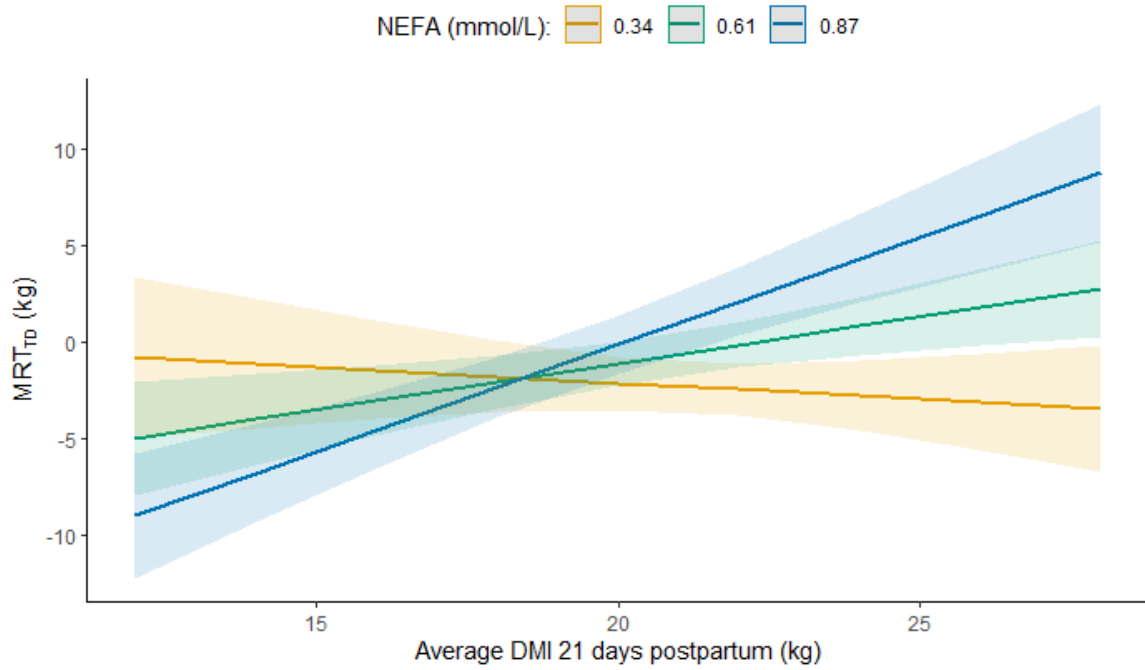
Table 5.3 shows the summary of the final multivariate model. The adjusted R² for this final model was 0.47 and the model's residual errors were found to be normally distributed (P = 0.96) and homoscedastic (P = 0.75), thus respecting the general assumptions of linear regression. Based on the partial η^2 , the most important variable in this model was the interaction between the average DMI 21 days postpartum and the NEFA concentration. This interaction is plotted in Figure 5.2A and suggests that the response of MRT_{TD} to NEFA is highly dependent on the DMI postpartum, a pivot point in the MRT response can be found around 18.4kg average daily DMI post-partum. Transitions where the DMI postpartum was lower (<18.4 kg/day) and blood plasma NEFA was high (0.87 mmol/L) were associated with the highest negative milk residuals, or in other words, the lowest MRT_{TD}. In contrast, transitions for which DMI postpartum and NEFA were high (>18.4 kg/day and 0.87 mmol/L), the response of MRT_{TD} was inversed.

Table 5.3 Final multivariate linear model for the association of the dependent variable milk yield residuals in the transition period for test days (MRT_{TD}) with the blood plasma analyte concentrations, dry matter intake (DMI) pre- and postpartum and the lactation number as independent variables. Only the selected independent variables are presented, ordered by the partial η^2 . The partial η^2 represents each variable's relative contribution to the model's final performance. “*” denotes interaction terms between two independent variables. Abbreviation: NEFA = Non-Esterified Fatty Acids

	Estimate	Std. Error	Pr(> t)	Partial η^2
(Intercept)	-0.69	0.56	2.21E-01	
Average DMI 21 days postpartum * NEFA	1.60	0.34	1.20E-05	0.21
Average DMI 21 days postpartum	1.21	0.40	3.63E-03	0.14
Lactation number ^a				0.13
3	-1.46	0.89	1.04E-01	
4+	-3.14	0.90	7.49E-04	
Glucose	-1.10	0.42	1.01E-02	0.06
Glucose * Insulin	-0.81	0.35	2.33E-02	0.06
NEFA	1.53	0.47	1.73E-03	0.04
Insulin	0.03	0.50	9.53E-01	0.02

^a compared to the reference (lactation number 2).

A)



B)

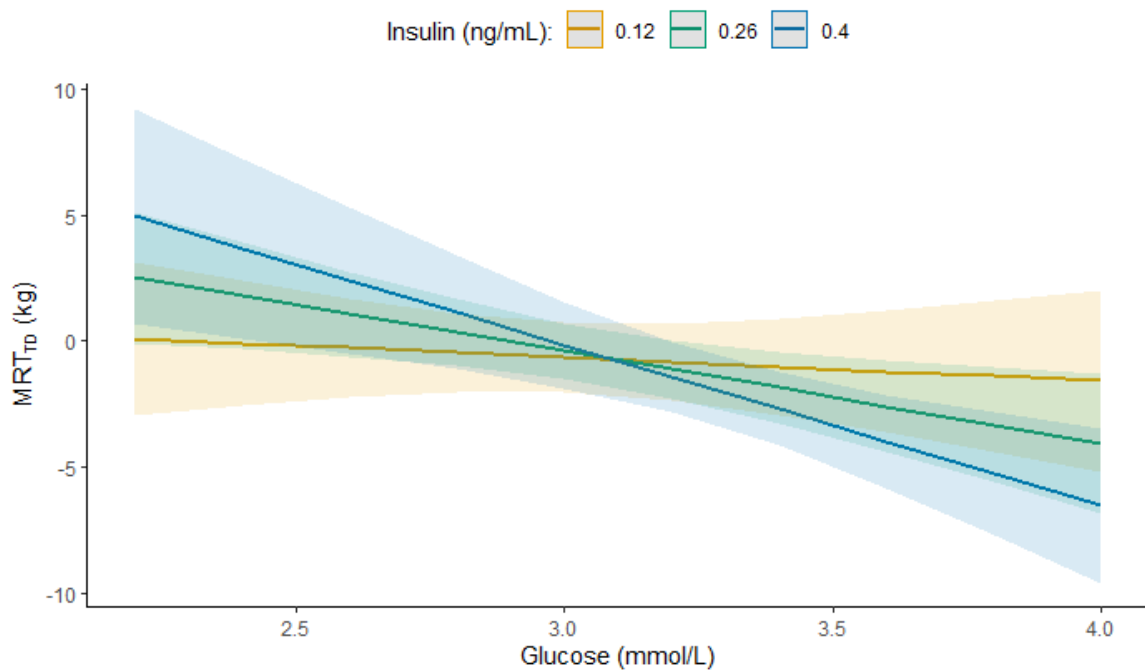


Figure 5.2 Interaction plots showing the 2 significant interactions in the linear regression model with the milk yield residuals in the transition period for test days (MRT_{TD}) as the dependent variable. In panel A, the interaction between dry matter intake (DMI) during 21 days postpartum and plasma non-esterified fatty acids (NEFA) concentrations are shown, while panel B presents the interaction between plasma glucose and plasma insulin concentrations. The marginal effects of plasma NEFA (panel A) and plasma insulin concentrations (panel B) are plotted in 3 categories: average minus one standard deviation (orange), average (green) and average plus one standard deviation (blue). The bands are the 95% confidence interval of these effects.

The lactation number was the third most important variable ranked by partial η^2 . The main effect of these categorical variables indicates that later lactations (3,4+) were negatively associated with MRT_{TD} and had more negative milk yield residuals in the transition period. Glucose, as the fourth important variable, was negatively associated with MRT_{TD}. The fifth in line is the significant interaction between insulin and glucose, plotted in Figure 5.2B. In this interaction the response in MRT pivots at a concentration of glucose 3.1 mmol/L. Where a combination of higher glucose (> 3.1 mmol/L) and high insulin (0.4 ng/mL) concentrations was found to be associated with a lower MRT_{TD}. On the other hand, low glucose (< 3.1 mmol/L) concentrations, in combination with high insulin (0.4 ng/mL) concentrations, resulted in higher values for MRT_{TD}. Ultimately, independency between repeated measures was affirmed by the Pearson's product-moment correlation (P = 0.61)

MRT_{MM}

In the final model for MRT_{MM}, 6 variables and 2 interactions were retained. The adjusted R^2 for the MRT_{MM} model was 0.73. The linear model assumptions were not violated (residual normality, P = 0.20 and homoscedasticity, P = 0.20). The variables kept in the final model are reported in Table 5.4, ordered by partial η^2 .

Table 5.4 Final multivariate linear model for the association of the dependent variable milk yield residuals in the transition period for milk meter data (MRT_{MM}) with the blood plasma analyte concentrations, dry matter intake (DMI) pre- and postpartum and the lactation number as independent variables. Only the selected independent variables are presented, ordered by the partial η^2 . The partial η^2 represents each variable's relative contribution to the model's final performance. A colon denotes interaction terms between two independent variables. NEFA = Non-esterified fatty acids, IGF1 = insulin-like growth factor 1

	Estimate	Std. Error	Pr(> t)	Partial η^2
(Intercept)	8.61	0.49	< 2e-16	
Lactation Number^a				0.65
3	-7.16	0.77	1.60E-14	
4+	-8.43	0.78	< 2e-16	
Average DMI 21 days postpartum	1.91	0.36	1.09E-06	0.26
IGF1	2.01	0.35	1.89E-07	0.23
NEFA	1.80	0.42	4.09E-05	0.15
IGF1 * NEFA	1.08	0.35	2.62E-03	0.11
Insulin	-1.23	0.40	2.90E-03	0.10
Glucose	-0.93	0.36	1.24E-02	0.08
NEFA * Glucose	0.74	0.36	4.63E-02	0.05

^a comparison to the reference (lactation number 2).

The most important variable in this model was the lactation number, for which, similar to the MRT_{TD} model, higher lactations had lower MRT_{MM} . The DMI postpartum was the second most important variable in this linear model. Between DMI postpartum and MRT_{MM} , a positive relation was found. The interaction between NEFA and IGF1 was found to have a significant association with MRT_{MM} . Its effect is visualised in Figure 5.3A. When the plasma IGF1 concentration was low (< 50 ng/mL), the MRT_{MM} was also low, regardless of the plasma NEFA concentration. At higher concentrations of plasma IGF1, an increase in plasma NEFA resulted in an increase in MRT_{MM} . Furthermore, insulin concentrations were found to have a negative effect on the MRT_{MM} . An interaction was found between NEFA and glucose (Figure 5.3B). When NEFA concentrations were low (0.34 nmol/L), there was a negative association between glucose and MRT_{MM} . When NEFA concentrations were high (0.87 nmol), there was no effect of glucose on MRT_{MM} . Also for this model the no significant correlation was found between the residuals of repeated measure ($P = 0.32$).

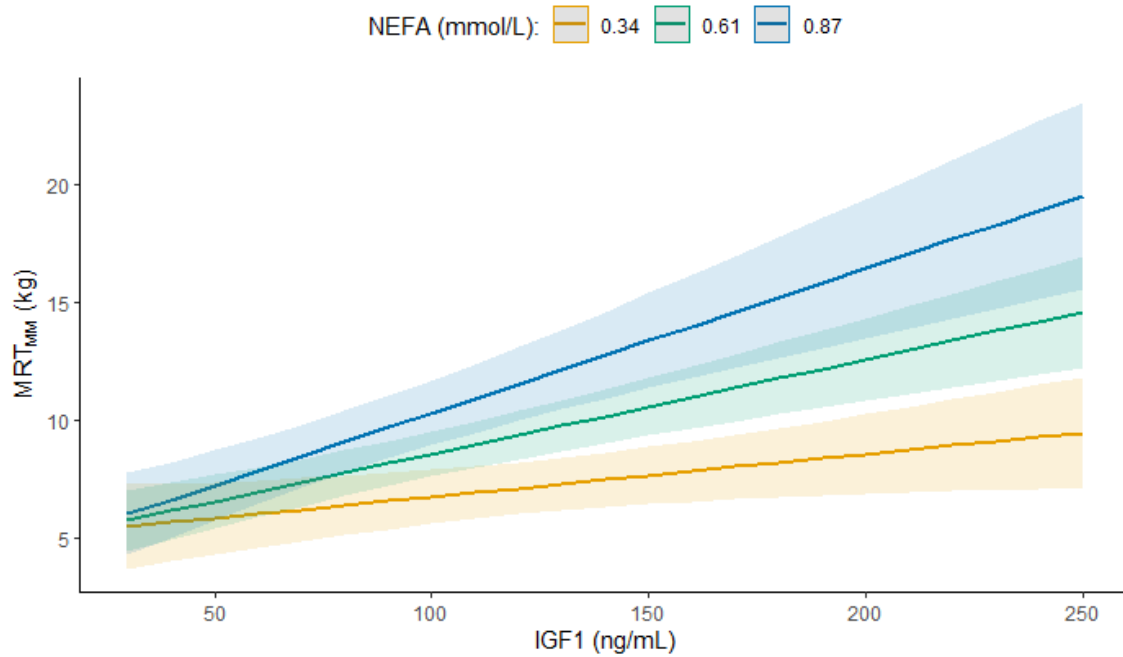
5.4 Discussion

The objective of this study was to evaluate the association between the MRT and indicators of transition success. Even though there is no consensus on the definition of transition success, different authors have described favourable and unfavourable physiological processes or metabolic profiles (LeBlanc, 2010; McArt et al., 2013a; De Koster et al., 2019; Pascottini et al., 2020b; Horst et al., 2021). These definitions of success formed the general framework for the assessment of the MRT. This research evaluated the MRT in two steps. First, compared DIS and NDIS transitions and considered the development of clinical disease during the transition period as a sign of transition failure. Second, the explicit associations between MRT and the metabolic status of NDIS animals were explored. Definitions of transition success, such as being clinically not diseased and having high DMI postpartum, were associated with higher MRT in both approaches. Finally, the potential application of the MRT for early lactation monitoring is being discussed.

5.4.1 DIS vs. NDIS transitions.

The differences in MRT between NDIS and DIS transitions indicated that transitions affected by a disease event result in a lower milk yield during the transition period. This was confirmed by a difference in disease prevalence between the top 25 % and bottom 25 % of both MRT. Clinically diseased animals produce less than their healthy herd mates, which is in line with previous studies (Deluyker et al., 1991; LeBlanc, 2010; Adriaens et al., 2021b).

A)



B)

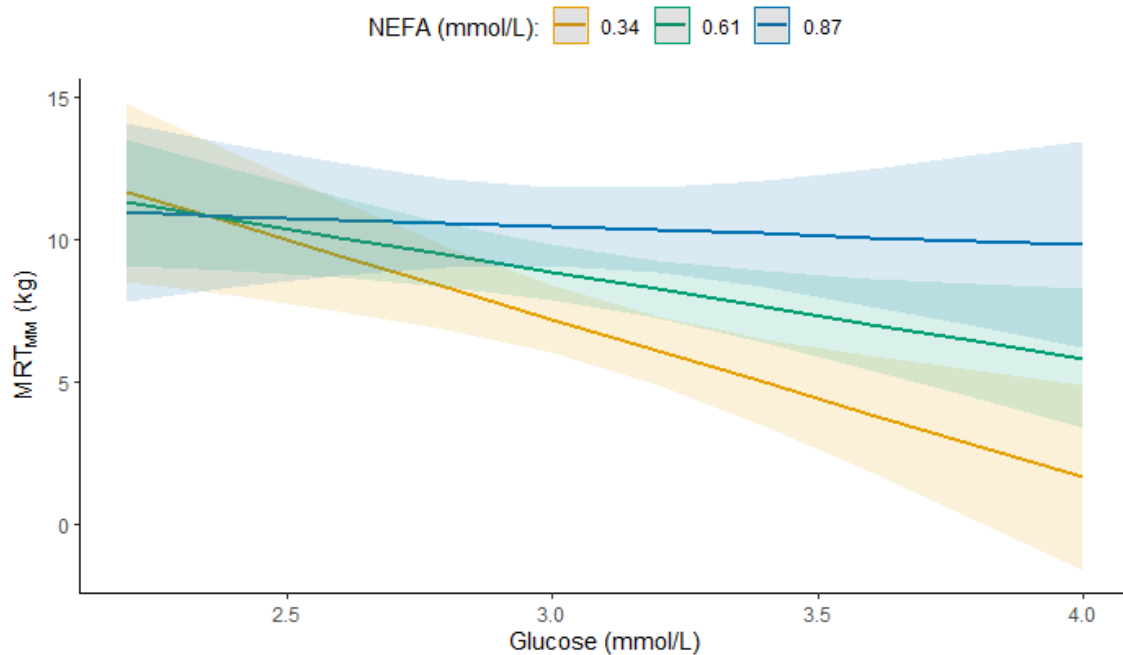


Figure 5.3 Interaction plots showing the 2 significant interactions in the linear regression model with the milk yield residuals in the transition period for milk meter data (MRT_{TD}) as the dependent variable. In panel A, the interaction between plasma insulin-like growth factor 1 (IGF1) and plasma non-esterified fatty acids (NEFA) concentrations is shown, while panel B presents the interaction between plasma glucose and plasma NEFA concentrations. The marginal effects of plasma NEFA concentrations (panels A and B) are plotted in 3 categories: average minus one standard deviation (orange), average (green) and average plus one standard deviation (blue). The bands are the 95% confidence interval of these effects.

In this study, this difference was demonstrated using an individual-animal baseline derived from her previous lactation. Although diseases in the past might have affected the daily milk yields in the previous transition of a cow (Erb and Grohn, 1988; Houben et al., 1993; Rasmussen et al., 1999; Saborío-Montero et al., 2017), the nextMILK_{TD} and nextMILK_{MM} were not affected by that as they seem able to accurately predict the expected healthy production at the start of the next lactation, without any knowledge of disease incidence in the previous lactation.

Differences between the NDIS and DIS transitions with respect to NEFA, DMI postpartum, IGF1 and BHB (Table 5.2) were in line with previously published research. Pascottini et al. (2022) provided an updated overview of how the metabolism adapts during the transition period and described how it can evolve into a pathological state.

It is established that the energy demand for initiating a new lactation is 4 times higher than the maintenance level of a high-yielding dairy cow (VandeHaar et al., 2016). This increase in energy requirement often exceeds the energy intake, making the DMI one of the main determining factors of the cow's energy balance. The NEB resulting from the discrepancy between energy intake and energy demand initiates a cascade of adaptations, reduced insulin sensitivity, lipid mobilisation, gluconeogenesis, ATP production through the β -oxidation of NEFA and production of ketone bodies (BHB, acetone and acetoacetate). Increases of NEFA and BHB in blood plasma are traditionally seen as a key precursor to the development of clinical diseases such as displaced abomasum, mastitis, metritis and retained placenta (Overton et al., 2017; Horst et al., 2021). The differences found between DIS and NDIS transitions are also coherent with the physiological processes explained hereabove. Moreover, clinically diseased animals are known to have characteristically higher plasma BHB and NEFA, while having lower plasma IGF1 and DMI postpartum (LeBlanc, 2010; Wathes et al., 2021). Nevertheless, the central role of BHB and NEFA as the key health and performance indicators during the transition period has been questioned in recent literature (Horst et al., 2021). This controversy is mainly based on the inconsistent effects of these markers over multiple studies and the reporting of numerous different thresholds. Their connection to metabolic imbalance is based on association studies more than controlled and intervening experimentation. In the case of the present study, no conclusion can be made on the intrinsic role of BHB and NEFA during the transition period. It is notable that based on the thresholds reported by McArt et al. (2013), hyperketonemia and elevated NEFA occurred in DIS as well as NDIS transitions.

Using three different thresholds for BHB, hyperketonemia was seen in 14% (>1.4 mmol/L), 20% (>1.2 mmol/L) and 33% (>1.0 mmol/L) of the NDIS transition. When these thresholds are applied to the DIS transitions, 43%, 52% and 56% of the transitions are identified as affected by hyperketonemia. Similarly, when a threshold of 0.7 mmol/L is used to define elevated NEFA, 61% of DIS transitions exceeded this threshold in contrast to 31% of the NDIS transitions. The percentage of animals that could be categorised as elevated in BHB or NEFA are in line with the numbers presented by McArt et al. (2013a).

Based on the results of our study, the fact that only 50% to 60% of DIS transitions present hyperketonemia or elevated NEFA indicates a lack of consistency of BHB and NEFA as a key analyte. Furthermore, as no information was available on subclinical diseases for this study, the presence of hyperketonemia and elevated NEFA in the NDIS transitions could also be linked to subclinically diseased animals.

The significantly lower plasma IGF1 for DIS transitions agreed with the literature. During transition, the circulating IGF1 has been described as an indicator of nutrient availability and is decreased in the case of aggravated nutrient and energy deficiencies or in the presence of an infection (Gross and Bruckmaier, 2019; Wathes et al., 2021).

5.4.2 Associations between MRT, blood plasma analytes and DMI

The linear models revealed associations between the MRT of NDIS transitions and the blood plasma analytes, lactation number and DMI. The preponderant importance of the independent variable, “average DMI postpartum”, based on the partial η^2 , indicated the close relationship between the milk residuals and the DMI. Although the relationship between milk yield and DMI has been described previously (Hristov et al., 2005), the associations between DMI and milk yield residuals have not yet been established. In both linear models (MRT_{MM} & MRT_{TD}), transitions with higher DMI intake resulted in a higher MRT. Furthermore, in the MRT_{TD} model, the interaction between DMI and NEFA concentration was found to be significant, Figure 5.2A. In this interaction, the elevated concentrations of NEFA were beneficial for the transition, yielding more positive MRT when combined with higher DMI. Similarly, in the MRT_{MM} model, positive associations were found between MRT_{MM} and NEFA in interaction with IGF1 concentration, Figure 5.3A. It is generally assumed that the mobilisation of fat tissue is a physiological response to prioritise milk synthesis (Aschenbach et al., 2010; Sordillo and Raphael, 2013; Horst et al., 2021). Based on our results, it appears that the metabolisation of

NEFA towards milk synthesis is dependent upon the availability of sufficient nutrients, as indicated by DMI or IGF1 interactions. Furthermore, these associations are found at NEFA concentrations above what would be considered pathological (McArt et al., 2013a). This hypothesis could also explain the high percentage of elevated NEFA concentration found in the NDIS transitions, these animals were seemingly able to convert their mobilised NEFA without detrimental consequences for their clinical status.

Furthermore, the associations of glucose and insulin in both MRT models seem to confirm the close relationship between the MRT and the physiological priority of milk production at the individual cow level. Glucose and insulin (for MRT_{TD}) were negatively associated with MRT, or in other words, when a higher concentration of glucose/insulin was found, the predicted yield exceeded the produced milk yield. The physiological relation between glucose and insulin can be summarised as follows: the decrease in plasma glucose is accompanied by a decrease in peripheral insulin. This decrease reduces the inhibitory effect of insulin on the lipase enzyme located in the adipose tissues, which causes the release of NEFA from these tissues (Sordillo and Raphael, 2013). In the case that glucose is high combined with high insulin or a low NEFA, the prioritisation of milk production is less pronounced in those animals. Different studies have already described this inverse relationship between plasma glucose concentration and milk yield (Sorondo and Cirio, 2009; Ruoff et al., 2017). The genetic merit of an animal has been linked with plasma glucose concentrations postpartum. Animals with a high genetic merit for milk production generally have lower plasma glucose concentrations compared to animals with a lower genetic merit for production (Snijders et al., 2001). Another low-rank interaction was found between glucose and NEFA in the MRT_{MM} model. This interaction term indicates that a higher grade of fat mobilization seems to negate the negative association of glucose on the milk residuals. Further research is needed to unravel possible pathways for this interaction.

The difference in adjusted R^2 between both MRT is hypothesised to be associated with the difference in partial η^2 for the lactation number. For MRT_{MM} this variable seems to account for the majority of explained variance creating the difference with MRT_{TD} . The central role of lactation number in the MRT_{MM} linear model indicates that the $nextMILK_{MM}$ model is correcting insufficiently for lactation number. This was not reported in Liseune et al. (2021) and should be investigated further when applying this model in the future.

Both linear models found similar associations between MRT and the measured variables. Even though the MRT were extracted from 2 different prediction models based on 2 different

modeling techniques (nextMILK_{TD} model being a random forest model and the nextMILK_{MM} being a deep learning model) running on 2 separate data types. The variables that were found to be associated with the MRT are intrinsically associated with the metabolic status of the cows in the transition period.

5.4.3 Limitations and future perspectives

The primary limitation of this study was the lack of a golden standard for the distinction between the healthy and subclinically diseased animals within the NDIS group. In our view, this does not undermine the general results of this research. Animals in this group are experiencing a varying level of metabolic challenge, which places them on a spectrum between healthy and subclinically diseased. Within the NDIS group, the linear models are still able to differentiate between cows, indicating a clear and logical relation with indicators for metabolic health and feed intake. The MRT can only be calculated for multiparous animals due to the requirement of historic data. This is a clear limitation of the MRT, but it doesn't diminish its significance, given that multiparous animals are the group that is most at risk of disease during the transition period (McArt et al., 2013b; Vanholder et al., 2015; Saborío-Montero et al., 2017; Horst et al., 2021).

The results from this research indicate that the metabolic profile of dairy animals during the transition period shows strong associations with the milk yield residuals as all associations seen within these results are in line with the current literature. While this research is not intended to validate the usage of MRT as a health monitoring tool, the results from the present study indicate the potential value of this indicator towards this goal. This should be further investigated with a longitudinal study including more animals and an experimental design in line with the research question. The impact of diseases in the previous lactation on the prediction of milk was not evaluated in this research. Although, the nextMILK_{TD} and nextMILK_{MM} model were trained with a variety of previous lactation curves (with and without diseases in the previous), the unavailability of a curated dataset of the previous lactation regarding disease registration restricted the capabilities to incorporate this possible effect. It would be highly valuable to understand how health perturbations in the previous lactations affect the MRT and the established metabolic associations. It is important to note that averaging the metabolites over the different collection dates was linked with a loss of information. Yet, this was a necessary step given the high level of collinearity between the different variables if added separately. Further research could investigate the relations between MRT and metabolites

on a daily basis during the transition period, using time series analysis methods. This would help us to better understand how MRT affects daily fluctuations in metabolism.

5.5 Conclusion

This study used two previously developed models (nextMILK_{TD} and nextMILK_{MM}) to predict the expected production of dairy cows during the transition period and subtracted these predictions from the actual milk production to obtain milk yield residuals (MRT). Furthermore, the link between these residuals and different blood plasma analytes, the dry matter intake and the lactation number were studied. We demonstrated that the MRT derived from the nextMILK_{MM} and nextMILK_{TD} were higher for cows without clinical diseases and that there is a strong positive association between these MRT and the DMI postpartum. Moreover, animals producing less than predicted had lower DMI postpartum. Additionally, the MRT seems to be associated with the physiological prioritization of milk production through the effect of NEFA, insulin and glucose. We also found that a certain level of fat mobilization is not directly detrimental to the cows, as long as the DMI is high. Although this study establishes a solid foundation for using the milk yield residuals as a new and valuable metric for health and transition success monitoring, further research and validation is needed.

Chapter 6.

Association of milk yield residuals to behavioural patterns and transition success in multiparous dairy cows

C. Kemel ^{a, 1}, M. Salamone ^{a, b, 1}, B. Aernouts ^b, I. Adriaens ^b, G. Opsomer ^a, P. Hut ^c, M. Hostens ^{c, d}

^a Ghent University, Department of Internal Medicine, Reproduction, and Population Medicine, Faculty of Veterinary Medicine, Salisburylaan 133, 9820 Merelbeke, Belgium

^b KU Leuven, Campus Geel, Department of Biosystems, Division of Animal and Human Health Engineering, Kleinhoefstraat 4, 2440 Geel, Belgium

^c Utrecht University, Department of Population Health Sciences, Division Farm Animal Health, Faculty of Veterinary Medicine, PO Box 80151, 3508 TD Utrecht, the Netherlands

^d Cornell University, College of Agriculture and Life Sciences, 272 Morrison Hall, Ithaca, New York

¹ Authors contributed equally to this work

Adapted from:

Kemel, C., Salamone, M., Aernouts, B, Adriaens, I., Opsomer, G, Hut, P., Hostens, M, (2025). Association of AI-Predicted Milk Yield Residuals to Behavioral Patterns and Transition Success in Multiparous Dairy Cows. *Journal of Dairy Science*, In Review

6.1 Introduction

The continuous increase in milk yield among dairy cattle involves multiple challenges at cow level, especially during the transition period. This period begins roughly 3 weeks before calving and lasts till 3 weeks after parturition (Grummer, 1995; Drackley, 1999; Mezzetti et al., 2021). During this critical phase, dairy cows experience several physiological, nutritional, and social challenges, which make them more vulnerable to infectious and metabolic diseases. It's estimated that 30-50% of the cows develop diseases, such as mastitis, metritis, ketosis, lameness, and displaced abomasum during the postpartum period (LeBlanc, 2010; Hostens et al., 2012; Van Saun, 2023a). This high prevalence of transition diseases negatively affects animal welfare, milk production, and herd profitability.

Therefore, understanding diseases in transitioning dairy cows has been a major point of attention for dairy industry research in the past 25 years, focusing on disease detection, treatment, and prevention (Proudfoot, 2023). Currently, disease detection primarily relies on laboratory analysis or cow-side testing of metabolic markers, requiring milk or blood samples and active farmer involvement, which can lead to poor identification performance (Ospina et al., 2013; Mezzetti and Trevisi, 2023; Van Saun, 2023b). In current practices, beta-hydroxybutyrate (BHB) in blood, milk, and urine receives a lot of attention as a promising biomarker for transition success. At the same time, several recent studies indicate the nuanced and inconsistent role of BHB (Horst et al., 2021). Over the last few years, dairy farming technology has evolved, allowing to collect and analyse data on production, milk quality, behaviour, and environmental conditions. These data provide opportunities to improve disease management through improved detection and monitoring (Adriaens et al., 2021b; Ho et al., 2021). Due to the high complexity and diversity of the collected dairy data, advanced data processing is needed to translate these data into reliable insights and useful decision support. Nevertheless, despite a plethora of on-farm technologies collecting valuable data, and all the technical advancements, we fail to diminish the number of transition problems (Pinedo et al., 2020; Van Saun, 2023b). Thus, several gaps remain in managing the health of transitioning dairy cows.

Although readily available milk yield data can be a solid basis for cow health monitoring, it is underutilized in current on-farm applications. The emergence of data-driven health monitoring tools that use on-farm milk data offers a new opportunity to monitor the transition period in dairy cows. By comparing AI-generated expected milk yields with actual yields, these tools can

help identify health-perturbing events (Adriaens et al., 2021a; Ranzato et al., 2024) and possibly allow for pre-emptive actions in this crucial period (Mezzetti and Trevisi, 2023). Today's models to predict milk yields typically make use of historical milk yield data of the current lactation (Macciotta et al., 2011; Adriaens et al., 2018). However, during the transition period, when lactation just started, the amount of available milk yield data of the current lactation is limited while the lactation curve is highly non-linear. This poses challenges to the model fitting, especially when health issues affect these milk yield data at the onset of lactation, possibly resulting in the transition failure being unnoticed (Nordlund, 2006). For multiparous cows, using information from previous lactations could help to increase the robustness of the estimation of the expected milk yield in the current lactation, assuming the cow is healthy, where it creates opportunities to identify health-perturbing events during the transition period (Adriaens et al., 2018; Liseune et al., 2020; Salamone et al., 2022). Recently, an artificial intelligence (AI) model, nextMILK_{TD}, was developed to predict the expected milk yield on the first test day after calving, between 5 and 45 days in milk, based on milk yield data from the previous lactation (Salamone et al., 2022). Validation of the nextMILK_{TD} has demonstrated its accuracy in predicting the expected "healthy" milk yield at the first test day of a new lactation, resulting in a root mean square error (RMSE) of 6.18 kg (Salamone et al., 2022). The predicted milk yield can be subtracted from the actual production to calculate the milk yield residuals in the transition period (MRT). Salamone et al., (2023) found significantly more negative MRT values in clinically diseased compared to non-clinically diseased cows. Moreover, associations were found between MRT values of non-clinically diseased cows and the dry matter intake and blood plasma metabolites. As such, more negative MRT values were linked with reduced dry matter intake and higher lactation numbers (Salamone et al., 2023).

Over the last 20 years, research has increasingly shown how dairy cow behaviour changes during the transition period (Huzzey et al., 2005; Neave et al., 2017). While many studies leveraging extensive sensor data often focus on isolated behavioural parameters, Hut et al. (2022) integrated feeding, lying, and walking behavioural data with independent variables: parity, milking type, and calving season, to study the complete time budgets of high-producing dairy cows. The latter study highlighted the variability in time budgets from the late dry cow period to the late lactation cycle, illustrating distinct behavioral patterns per parameter that varied between parities but were consistent across farms. Others reported that behavioral changes pre- and or postpartum were associated with transition diseases and could thus be used as an indicator of cow health (Huzzey et al., 2011; Rial et al., 2023; Santos et al., 2024).

Moreover, some of these commercial sensor systems generate health alerts using machine learning algorithms, where AI interprets changes in the measured behavioral patterns. However, in terms of accuracy, low sensitivity and/or specificity levels are typically obtained (Stangaferro et al., 2016a; Melendez and Cook, 2023; Simoni et al., 2024).

Research has shown a link between specific behavioural patterns and transition success or failure in dairy cows. Similarly, milk yield residuals appear to be closely associated with the physiological processes occurring during this critical period. While milk yield data are commonly available on farms, behavioural data are not yet widely adopted. Making more effective use of these milk yield data presents opportunities to enhance the early detection of transition failure through the use of MRT. Therefore, the main objective of this study was to link these MRT to behavioural parameters measured by 3-dimensional neck and leg accelerometers, in high-producing multiparous Holstein-Friesian dairy cows during the transition period.

6.2 *Materials and Methods*

6.2.1 *Farms, animals, and sensor data*

For this study, sensor data of 2,381 lactations at 8 commercial farms in The Netherlands were collected over five years from January 1st, 2016, until December 31st, 2020. All farms had free-stall barns and all dairy cows were fitted with 2 types of sensors. A detailed description of the farms is presented in Table 6.1 and is extensively documented in two previous papers (Hut et al., 2021, 2022). Feeding behaviour (eating time and rumination time) was monitored with the commercially available “Nedap Smarttag Neck” sensors (Nedap, Groenlo, The Netherlands) attached to the neck collar of each cow. The “Nedap Smarttag Leg” sensors were attached to one of the front legs of each cow, to monitor walking (standing time) and lying behaviour (lying time). Previous studies described the validation of both sensors, showing high correlations between observed and reported behavioural parameters (0.88-0.97) (Nielsen et al., 2018; Borchers et al., 2021). The use of these sensors in commercial dairy herds is not considered an animal experiment under the Dutch law, thus no formal ethical approval was needed (Hut et al., 2021, 2022).

Table 6.1 Details of eight commercial farms in the Netherlands used in this retrospective observational study. As cubicle bedding in the far-off or close-up period, “deep litter” is related to cubicle systems with different filling materials, where a “straw yard” means a free-range area. Pasture access is only for lactating cows. Abbreviations: automated milking system = AMS, conventional milking system = CMS, dry period = DP, interquartile range = IQR

Farm	Herd size	Cubicle bedding far-off	Cubicle bedding close-up	Average DP length (25-75% IQR)	Cubicle bedding lactation	Milking system	Pasture access	Production level (kg milk/cow/year)
1	170	Deep litter	Straw yard	41 (31-46)	Deep litter	AMS	No	10, 786
2	130	Deep litter	Straw yard	39 (30-41)	Deep litter	AMS	No	11, 177
3	110	Mattress	Mattress	45 (40-51)	Mattress	AMS	No	9, 341
4	110	Mattress	Straw yard	39 (33-43)	Mattress	CMS	Yes	9, 314
5	140	Deep litter	Deep litter	35 (30-40)	Deep litter	CMS	Yes	9, 256
6	170	Mattress	Mattress	37 (32-42)	Mattress	CMS	Yes	9, 243
7	175	Deep litter	Straw yard	42 (32-48)	Deep litter	CMS	Yes	9, 109
8	120	Mattress	Mattress	45 (37-49)	Mattress	CMS	Yes	9, 197

All data was provided by Nedap Livestock Management (Nedap, Groenlo, The Netherlands). Behavioural parameter data were collected in minutes per 24 h (min/24h) per cow. Because of major effects of calving on behavioural patterns, the day of calving was eliminated from the behavioural data.

6.2.2 Production data and calculation of the milk yield residuals

The available milk yield data consisted of test day (TD) records registered in the context of a dairy herd improvement (DHI) program with an interval of 4 to 6 weeks (CRV, Arnhem, The Netherlands). In addition to the milk yield data from the lactation in which the transition was monitored, TD records of the previous lactation were available as well. As no historical lactation data is available for primiparous cows, this study only considered multiparous cows to link their behavioural data to the calculated MRT. The expected milk was predicted using the nextMILK_{TD} model, a random forest model developed by Salamone et al. (2022). In its original publication, three model variants were developed, using different numbers of features. In the present study, the production variant of the nextMILK_{TD} model was used, not relying on

herd production performance or individual fertility traits. More information about the models and their validation is available in the publications of Salamone et al. (2022, 2023).

Additionally, a set of features from the previous lactation of each cow was included to accurately apply the nextMILK_{TD} model: the season of calving, the cumulative milk yield at DIM 21, 75, 305, and at the end of lactation, and the milk yield minimum, maximum, average and standard deviation. The nextMILK_{TD} model predicted the expected milk yield on the DIM corresponding to the first TD recorded in the lactation under study, ranging from 5 to 45 DIM (Salamone et al., 2022).

The milk yield residuals (MRT) were obtained by subtracting the expected production from the actual production at the first TD. Hence, a positive MRT indicates that the cow produced more than was predicted, while a negative MRT reflects that the cow actually produced less than was predicted. In other words, the more negative or the lower the MRT, the less the cow had produced compared to what was predicted for her. The calculated MRT of all multiparous lactations in this study were grouped into three categories (<33% quantile: LOW; 33-66% quantile: MED; >66% quantile: HIGH).

6.2.3 Statistical analysis

The datasets were analysed using R (version 4.4.1), the packages used in this study were: “cAIC4” (Säfken et al., 2021), “car” (Fox and Weisberg, 2019), “carData” (Fox et al., 2022), “club sandwich” (Pustejovsky, 2025), “cowplot” (Wilke, 2024), “dplyr” (Wickham et al., 2023a), “emmeans” (Lenth, 2025), “geomtextpath”(Cameron and van den Brand, 2022), “ggplot2” (Wickham, 2016), “ggpubr” (Kassambara, 2023), “ggrepel” (Slowikowski, 2024), “grid” (R Core Team, 2020), “lme4” (Bates et al., 2015), “lmerTest”(Kuznetsova et al., 2017), “lsmeans” (Lenth, 2016), “multcomp” (Hothorn et al., 2008), “multcomView” (Graves et al., 2024), “mvtnorm” (Genz et al., 2021), “plyr” (Wickham, 2011), “purrr” (Wickham and Henry, 2025), “readr” (Wickham et al., 2022), “sandwich” (Zeileis, 2006), “skimr” (Waring et al., 2022), “stringr”(Wickham, 2023), “TH.data” (Hothorn, 2025), “tibble”(Müller and Wickham, 2022), and “tidyr” (Wickham et al., 2023b). All code scripts of the statistical analyses can be consulted at <https://github.com/Bovi-analytics/Salamone-et-al-2025>.

Behavioural data and MRT categories were used in linear mixed-effect models for statistical analysis, with behavioural parameters (standing, lying, eating, and rumination) as the dependent

variable of interest. The number of animals included in the four models varied slightly due to missing data and data transfer issues, see Figure 6.1.

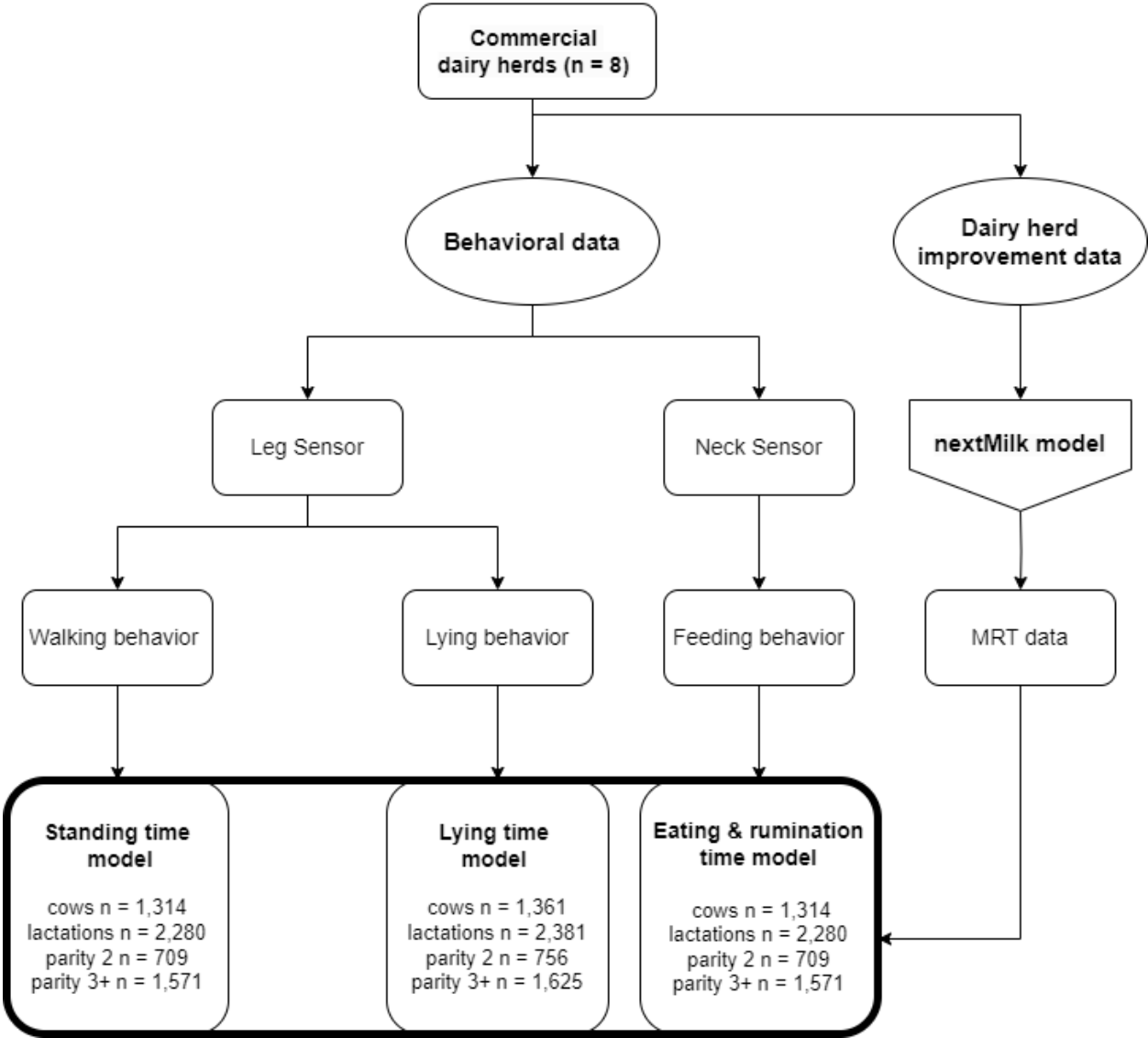


Figure 6.1 Distribution of the cows used in this study equipped with leg and neck sensors for each of the statistical models resulted from the merge data set between behavioral sensor data and dairy herd improvement data

Separate models were built for each behavioral parameter with the unique herd, animal identifier, parity of the animal (2nd or 3+), days in milk (-21 to 21), calving season (spring: April-June; summer: July-September; autumn: October-December; winter: January-March), and the MRT categories (LOW, MED, HIGH) as independent explanatory variables, with random effects for respectively time relative to calving and lactation identifier. In order to perform a variable selection, first a multivariable model was built with all explanatory variables as fixed effects and the DIM nested within each lactation as random effect. A drop 1 function based on Akaike Information Criterion (AIC) was applied to remove any non-informative variables from this initial multivariate model. Next, biologically relevant two-way interactions of DIM with parity, herd, calving season, and MRT category were also evaluated. Iteratively these interactions were added to the initial multivariate model and compared to the model without interaction. The final model incorporated the interactions and their main effect where a significant improvement of AIC was found in the comparison. Final models were reported and plotted as least square means (LSM) with 95% confidence intervals (95% CI). Multiple contrasts were adjusted using the “Sidak” method. Residuals of the final models were checked for normal distribution by q-q plotting. Significant differences were reported ($P < 0.05$) when the 95% CI error bars did not overlap.

6.3 Results

6.3.1 Descriptive analysis

The mean sensor values for each behavioural parameter for the three weeks before and the three weeks after parturition are shown in Table 6.2. Eating, lying and standing times were higher in the pre-calving period relative to after calving. In contrast, rumination time post-calving was higher than in the period before parturition.

Table 6.2 Mean sensor values for eating, rumination, lying, and standing times during the observed period relative to calving in minutes per day (min/d).

	Pre-Calving	Post Calving
Eating time (min/d)	312 ± 99.3	303 ± 97.3
Rumination time (min/d)	498 ± 112	533 ± 109
Lying time (min/d)	777 ± 165	630 ± 161
Standing time (min/d)	608 ± 162	583 ± 160

Furthermore, in Figure 6.2, the milk production on the first TD versus the MRT are visualized and grouped per MRT category (LOW, MED, HIGH). The average milk production and

standard deviation on the first TD within each MRT category were 31.9 ± 6.28 L, 38.9 ± 4.65 L, and 44.8 ± 5.76 L for the LOW, MED, and HIGH MRT categories, respectively. The average MRT and standard deviation calculated on the first TD within each MRT category were $-7, 96 \pm 5.04$, -0.59 ± 2.08 , and 5.53 ± 3.60 for the LOW, MED, and HIGH MRT categories.

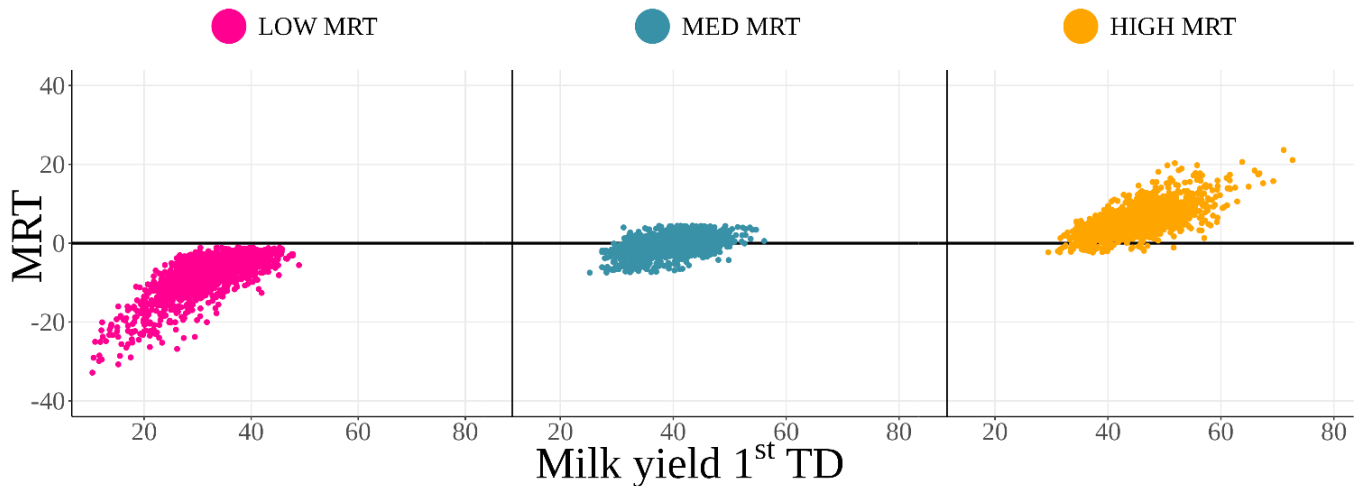


Figure 6.2 The relation between the milk production on the first test day and MRT is plotted over all farms. A distinction is made between the LOW (pink), MED (blue), and HIGH (yellow) MRT category.

6.3.2 Transition time budget models

The complete time budget from 21 days before calving until 21 DIM of all cows in the study is presented in Figures 6.3 to 6.6. These overall estimates show that behavioral parameters followed a specific pattern during the transition period. The final models for all measured behavioral parameters showed significant differences between the MRT categories in some parts of the transition period. The LSM and 95% CI predictions per behavioral parameter for different MRT categories are presented in Figures 6.3 to 6.6, and all exact estimates are available on the open-access repository previously reported.

Eating time during the prepartum period decreased from 334 (95% CI: 327-342) minutes a day on day -21 to 301 (95% CI: 295-308) minutes on the day before calving (-1). The patterns and order of magnitude of eating time in the period before calving were similar for all three MRT categories. Postpartum eating time increased, but patterns differed significantly between the MRT categories over the presented period. Eating time increased to 309 (95% CI: 299-319), 320 (95% CI: 310-330), and 320 (95% CI: 311-330) minutes on day 21 for respectively the LOW, MED, and HIGH MRT category, as seen in Figure 6.3A. Moreover, postpartum eating

time for cows in the LOW MRT category was consistently lower compared to the HIGH MRT category, except on day 21, where differences were not significant. Significant differences were also observed between the MED MRT and LOW MRT categories during the postpartum period, except for a few days. No significant differences were found between the MED MRT and the HIGH MRT categories, as plotted in Figure 6.3B.

Similarly to eating time, rumination time shown in Figure 6.4A decreased in the prepartum period from 521 (95% CI: 513-530) minutes on day -21 to 484 (95% CI: 477-492) minutes on the day before calving. Patterns of rumination time were comparable for all three MRT categories in the period before calving. From day 1 after calving, rumination time increased up to 562 (95% CI: 551-572) and 556 (95% CI: 546-567) minutes on day 21 after calving for respectively the HIGH and the MED MRT categories. These were significantly higher in comparison with the LOW MRT category which had a rumination time of 537 (95% CI: 527-548) minutes. Significant differences between the MED and HIGH MRT categories were observed only on days 4, 8, and 10 to 12 postpartum, these differences are shown in Figure 6.4B.

Lying time slowly increased in the period before calving and reached a maximum on day -8 of 788 (95% CI: 776-799) minutes a day and slowly decreased until the day before calving to 722 (95% CI: 710-733) minutes, as plotted in Figure 6.5A. Patterns of lying time over the three different MRT categories were comparable prepartum, without significant differences between the three MRT categories, except small, but significant effects between the LOW and the HIGH MRT categories on days -16 and -12. On these days, the LOW MRT category had the shortest lying times, shown in Figure 6.5B. Around calving, a large decrease in lying time occurred, reaching its nadir of 606 (95% CI: 594-618) minutes one day postpartum. Figure 6.5A also show that after calving, lying time increased slowly in the first days and then remained relatively stable during the rest of the observed period. Differences in postpartum lying time between the MRT categories were small but significant on days 1, 11, 12, 14, 15, and 17-21, with the LOW MRT category showing higher lying time compared to the HIGH MRT category. No significant differences were noted between the LOW and MED MRT categories, except on day 1 postpartum. Additionally, as shown in Figure 6.5B, no differences were found between the HIGH and MED MRT categories, except on days 11 and 14, where the MED MRT category's lying time was higher than that of the HIGH MRT category. For standing time, all three MRT categories had comparable standing times in the prepartum period, where total standing times per day increased to 636 (95% CI: 626-647) minutes one day before calving. A

further incline up to 644 (95% CI: 633-655) minutes one day after calving was observed and was followed by a decrease to 547 (95% CI: 535-559) minutes on day 19 postpartum after which it remained stable for the rest of the observation period, see Figure 6.6A. In the period after calving, the LOW MRT category had significantly higher standing times compared to both other categories except on day 1 postpartum, plotted in Figure 6.6B . On days 6 and 7 postpartum, significant differences were also found between the MED and HIGH MRT categories, where the HIGH MRT category had the shortest standing times

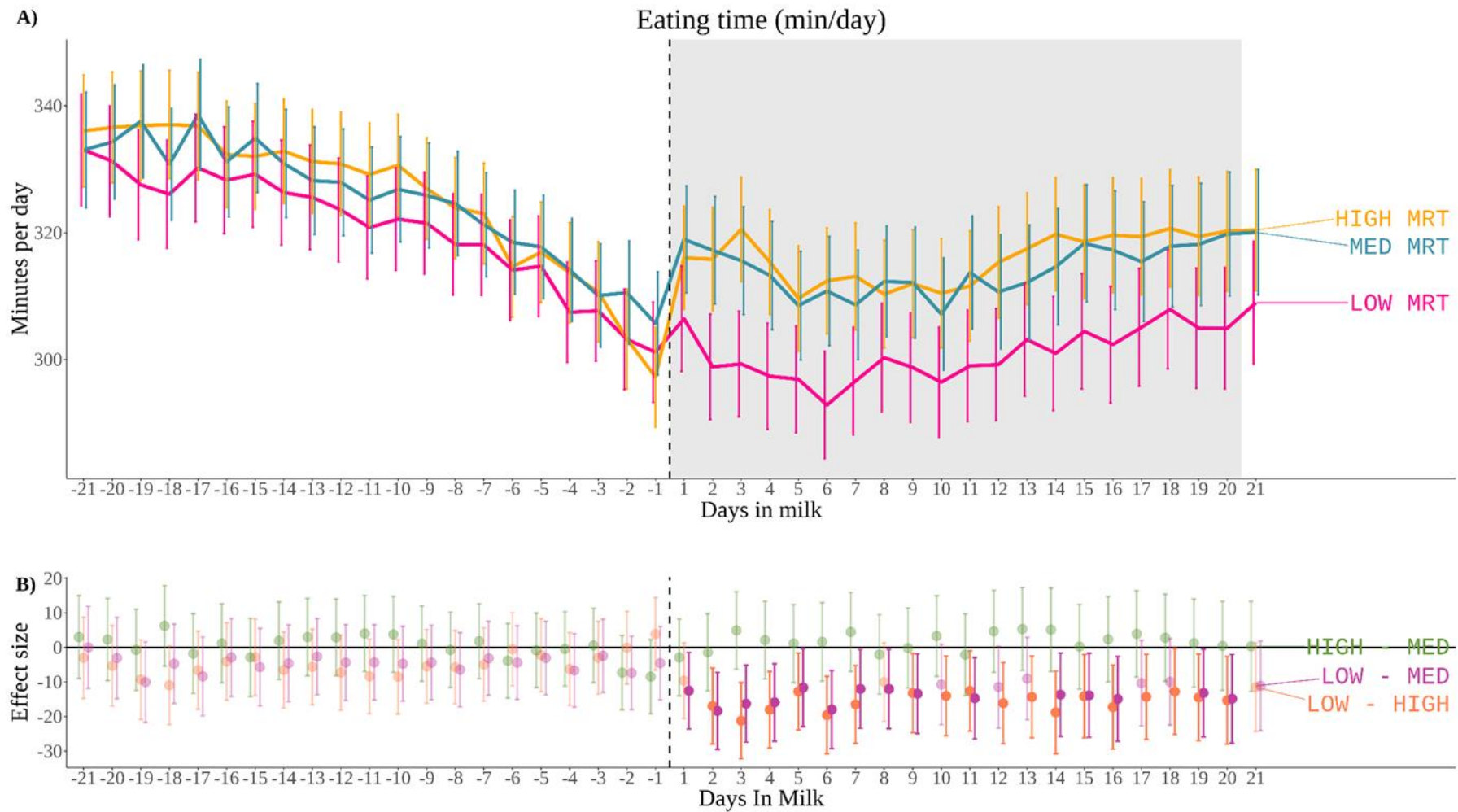


Figure 6.3 A) Eating time in minutes per day (min/day) based on least square means (LSM) with 95% confidence intervals (95% CI) grouped in MRT categories (<33% quantile: LOW, 33-66% quantile: MID, >66% quantile: HIGH) on eight commercial dairy farms in the Netherlands from 21 days before calving until 21 days in milk. Grey zones indicate significant effects. B) Pairwise comparisons between different MRT categories were plotted, where non-significant results were made more translucent

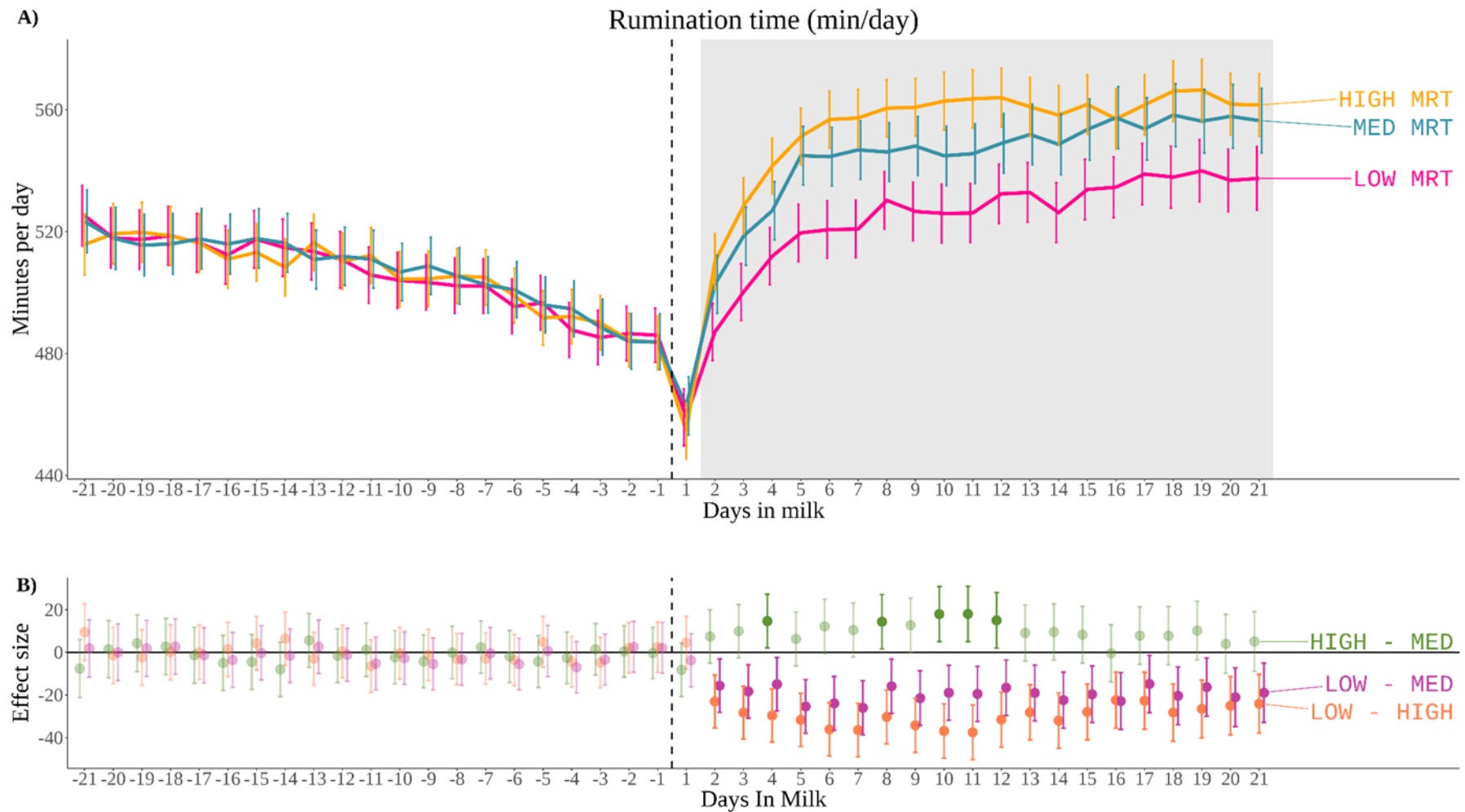


Figure 6.4 A) Ruminant rumination time in minutes per day (min/day) based on least square means (LSM) with 95% confidence intervals (95% CI) grouped in MRT categories (<33% quantile: LOW, 33-66% quantile: MID, >66% quantile: HIGH) on eight commercial dairy farms in the Netherlands from 21 days before calving until 21 days in milk. Grey zones indicate significant effects. B) Pairwise comparisons between different MRT categories were plotted, where non-significant results were made more translucent.

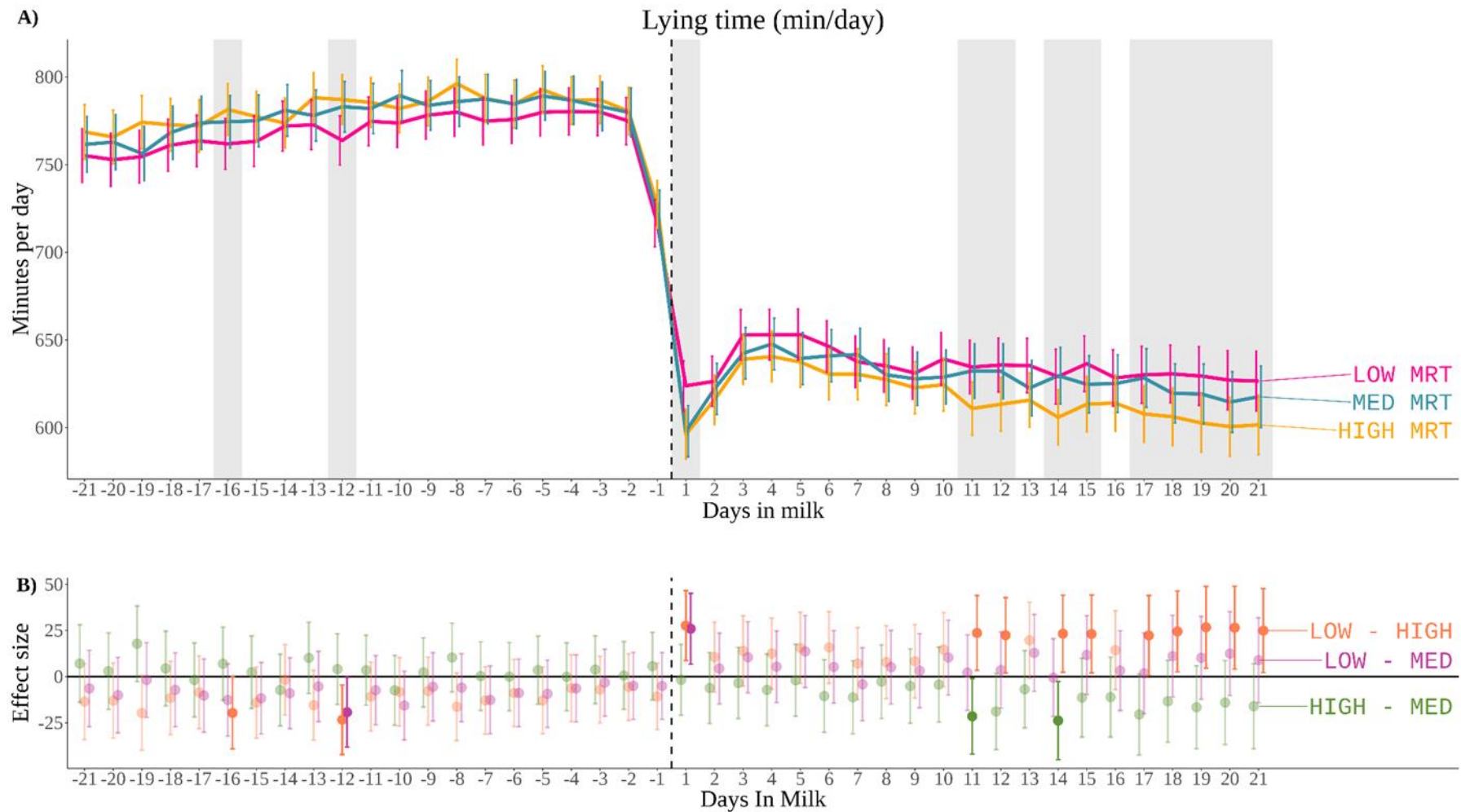


Figure 6.5 A) Lying time in minutes per day (min/day) based on least square means (LSM) with 95% confidence intervals (95% CI) grouped in MRT categories (<33% quantile: LOW, 33-66% quantile: MID, >66% quantile: HIGH) on eight commercial dairy farms in the Netherlands from 21 days before calving until 21 days in milk. Grey zones indicate significant effects. B) Pairwise comparisons between different MRT categories were plotted, where non-significant results were made more translucent.

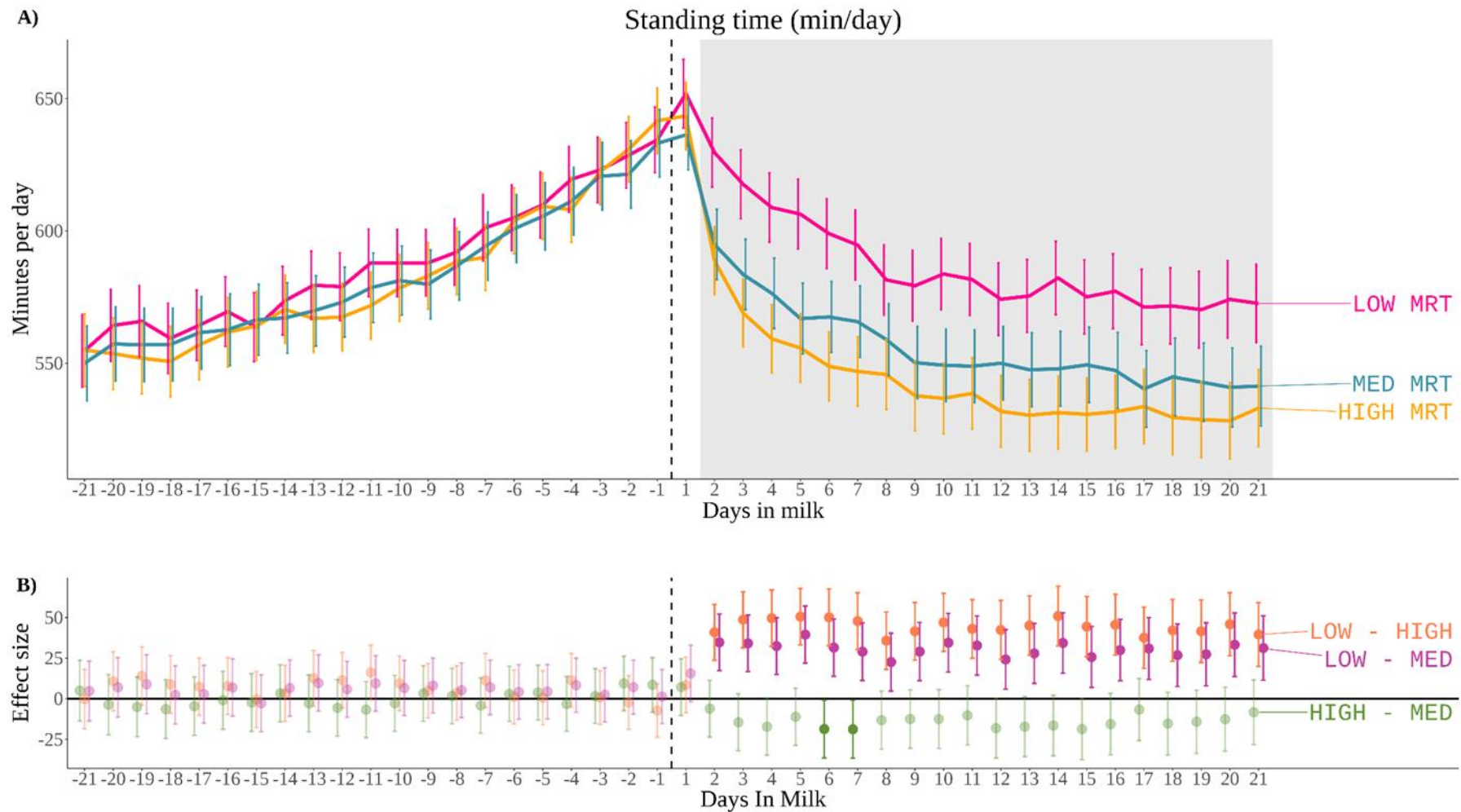


Figure 6.6 A) Standing time in minutes per day (min/day) based on least square means (LSM) with 95% confidence intervals (95% CI) grouped in MRT categories (<33% quantile: LOW, 33-66% quantile: MID, >66% quantile: HIGH) on eight commercial dairy farms in the Netherlands from 21 days before calving until 21 days in milk. Grey zones indicate significant effects. B) Pairwise comparisons between different MRT categories were plotted, where non-significant results were made more translucent..

6.4 Discussion

The objective of this study was to evaluate potential associations between the MRT and behavioural parameters of multiparous dairy cows in the transition period. Over the past 25 years, behavioural-related research in cattle has received increasing attention, where behavioural data pre- and postpartum were typically linked to milk production and transition diseases (Goldhawk et al., 2009; Stangaferro et al., 2016a; Lean et al., 2023). Behavioural patterns in the transition period are subject to change (Huzzey et al., 2005; Neave et al., 2017). Our data corroborated this, presenting differences between 21 days before calving up to 21 days postpartum, with data including healthy as well as (sub)clinically diseased cows. Others reported such behavioural changes in the transition period as useful parameters to relate to transition diseases, where failure of transition in dairy cows is often associated with behavioural changes pre- and or postpartum (Stangaferro et al., 2016a; Rial et al., 2023; Santos et al., 2024). Similarly, in our results, cows within different MRT categories present vastly different time budgets, mainly in the postpartum period. To the best of our knowledge, this is the first study where behavioural patterns of dairy cows over the transition period are linked with milk yield residuals, rather than actual milk production or health. In this regard, actual milk production has been linked to both unfavourable as well as favourable metabolic profiles (Horst et al., 2021), not only reflecting the cow's current health but also being affected by her genetic merit and the historical environment (e.g. rearing, infections). Therefore, using the cow's MRT is hypothesized to better agree with her actual metabolic status. In the period after calving, cows within different MRT categories showed different behavioural patterns. Moreover, from the first day after calving till three weeks postpartum, the LOW MRT category had the shortest rumination time in comparison to both other categories. These results are in line with other studies, where similar differences were observed in early postpartum rumination time between healthy and diseased cows (Calamari et al., 2014; Liboreiro et al., 2015; Kaufman et al., 2018). A reduced postpartum rumination time is associated with clinical diseases in transition cows, where the daily rumination time typically starts to decrease prior to a drop in milk yield (Stangaferro et al., 2016a; Santos et al., 2024). Low postpartum rumination times in the first weeks after calving have also been linked to impaired metabolism and a higher level of inflammation (Soriani et al., 2012; Calamari et al., 2014), while high early postpartum rumination times are indicative of higher milk production at the onset of the lactation and peak yield (Kaufman et al., 2018). Similarly, eating time was consistently lower after calving in the LOW MRT category compared to the HIGH and MED MRT categories. These results align

with previous studies, where healthy cows spent more time eating after calving (Urton et al., 2005; Schirmann et al., 2016; Stevenson et al., 2020). Beauchemin (2018) reviewed factors influencing eating times and noted a strong positive association between eating time and DMI, whilst the positive association between DMI and rumination time is relatively weak. Eating time could be useful in estimating the individual cow's feed intake especially if combined with other information (milk yield, body weight)(Beauchemin, 2018; Johnston and DeVries, 2018). Our findings, with postpartum eating times being longer in the HIGH MRT category, likely relate to the positive association between DMI and MRT found previously (Salamone et al., 2023). Similarly, no association was found between MRT and prepartum DMI either.

Lying times in the postpartum period varied between the different MRT categories, with small, but significant differences on dispersed days, where the lying time of the LOW MRT category was higher than both other categories. Previous studies have shown conflicting results in postpartum lying times for metabolically challenged cows. Longer lying times after calving were found in multiparous cows experiencing only ketosis or ketosis in combination with other metabolic diseases within 4 weeks after calving compared to non-diseased cows (Itle et al., 2015; Kaufman et al., 2016b; Piñeiro et al., 2019). However, it could not be determined whether the increased lying time caused the metabolic problems or vice versa. In contrast, other studies found no differences in postpartum daily lying times among different metabolic status groups (Neave et al., 2018; Cattaneo et al., 2020; Hendriks et al., 2022). The association between MRT and lying behaviour may reflect that healthy cows with higher milk yields have higher energy demands and spend more time eating and ruminating, leaving less time available for lying (Cook, 2020). The same observation was made by Løvendahl and Munksgaard (2016), although they monitored primiparous cows in contrast to our study. They found first-lactation cows with higher yields to have higher eating times and spend less time lying compared to lower-yield herd mates. The potential negative effect of increased lying time on milk production depends on whether the feed intake is reduced because of this increase in lying time (Tucker et al., 2021).

In addition to the increased lying times in the postpartum period in the LOW MRT category, also the highest postpartum standing times were observed for this category in comparison to both categories. This aligns with previous research that showed that primiparous cows with metritis spent more time standing in the barn after parturition, several days before they were diagnosed with the disease (Lomb et al., 2018). The authors hypothesised that increased standing times can be viewed as self-isolating sickness behaviour. Healthy cows spent more of their standing time at the feed bunk eating (higher eating times), whereas cows with metritis

spent this time perching, standing idle in the stall, or in social or abnormal lying-related behaviours (Lomb et al., 2018). Sick cows were more hesitant to lie down in response to visceral pain in the case of metritis. Prolonged standing times have also been reported after liver biopsy (Mølgaard et al., 2012), ruminal acidosis (DeVries et al., 2008), mastitis (Cyples et al., 2012; Fogsgaard et al., 2012), and lameness (Gomez and Cook, 2010). The impact of lameness on standing and lying durations can be bidirectional: some lame cows may struggle to lie down, leading to prolonged standing in the barn, while others might have difficulties standing and consequently spend more time lying down. When comparing herds on sand versus mats, the mat herds exhibited longer standing times (Gomez and Cook, 2010). Another study found no differences in postpartum standing times between different postpartum health categories (Itle et al., 2015).

In contrast to the differences observed between the different MRT categories in the postpartum period, our prepartum models showed no differences for any of the analysed behavioural parameters. Similar to our results, Liboreiro et al. (2015), Stevenson et al. (2020), and Cook (2020) reported no associations between prepartum rumination time and postpartum diseases. Recently, the average daily change in rumination time in the last 3 days before calving had no predictive power for hypocalcemia and hypomagnesemia (Emam et al., 2023). However, other studies found an association between reduced rumination times in the weeks before parturition and a higher level of postpartum diseases, reduced milk production, lower odds of conceiving, and higher culling rates in multiparous cows (Schirmann et al., 2016; Abuelo et al., 2021; Santos et al., 2024). These contradicting findings regarding prepartum rumination times highlight the persisting knowledge gap regarding this behaviour and its associations with postpartum health and performance (Santos et al., 2024). Similarly, no differences in prepartum eating times were found between the different MRT categories, which is in line with a recent study by Stevenson et al. (2020) who found that eating times did not differ by health status before calving. However, these results contrast with a lot of previous studies where eating times were lower during the prepartum period for cows diagnosed with metabolic health disorders after calving (Urton et al., 2005; Huzzey et al., 2007; Goldhawk et al., 2009; Schirmann et al., 2016). Results of lying behaviour analysis reveal comparable lying times in the prepartum period for the different MRT categories. This is comparable with previous studies showing that cows with metritis had similar lying times in the prepartum period compared to healthy cows (Barragan et al., 2018; Cattaneo et al., 2020). Some studies show that metabolically challenged dairy cows lay less, had fewer lying bouts, and were more active for up to three weeks before

calving (Itle et al., 2015; Neave et al., 2018; Hendriks et al., 2022), whereas Piñeiro et al. (2019) found a positive quadratic association between prepartum lying times and blood NEFA concentrations.

In the prepartum period, no significant differences were found in standing times between the different MRT categories, unlike other studies showing that cows with ketosis stood longer than nonketotic cows in the period before calving (Goldhawk et al., 2009; Itle et al., 2015). Rodriguez-Jimenez et al. (2018) found decreased standing duration prepartum in cows diagnosed with ketosis postpartum. For standing behaviour, there seems to be little agreement among the existing studies about the effect of pre- or postpartum standing times on ketosis and other postpartum diseases.

Our results showed an association between MRT and behavioural patterns of transition dairy cows, particularly in the post-calving period. This reinforces previous research that linked MRT to the health status of transition cows based on pre- and postpartum blood metabolites. Although associations were strong, our research was also limited by some factors. One of the major limitations of this study was the lack of a gold standard for the distinction between the healthy, subclinical, or clinically diseased cows within the study period and population, and the different MRT categories. However, dairy cows face different levels of metabolic challenges, with or without clinical symptoms. AI models can distinguish between these levels by analysing the clear and logical relationship between metabolic health indicators and feed intake (Salamone et al., 2023). Another limitation of the MRT is that it can only be used to evaluate the onset of lactation of multiparous cows. Nevertheless, multiparous animals are at the highest risk of developing transition disease (McArt et al., 2013a; Vanholder et al., 2015; Horst et al., 2021), supporting the usage of AI-generated MRT based on historical milk yield data. Still, this also highlights the need to create alternative approaches for primiparous cows. Generating AI-based residuals using our model requires test-day milk data collected through dairy herd improvement programs. This restriction was recently mitigated by developing an AI model using milk meter data (Liseune et al., 2021). Further research should focus on establishing a gold standard method for defining transition success that can be used as a basis for the further validation and optimization of the MRT in relation to behavioural analysis of high-producing dairy cows.

6.5 Conclusion

Based on the present study, we found different behavioural patterns for eating, rumination, lying, and standing times for cows with different milk yield residuals in the transition period

(MRT). After calving, the LOW MRT categories exhibited lower eating and rumination times than the other categories. Additionally, in the first three weeks postpartum, the LOW MRT category had the shortest rumination time. On specific days postpartum, longer lying times were observed in the LOW MRT category. Standing times were longest in the LOW MRT category following calving. The final models indicated no significant differences in any of the analysed behavioural parameters during the prepartum period among the three MRT categories. Our findings strengthen the idea of using MRT as an indicator of transition success in multiparous dairy cows and to predict the cow's adaptive capacity throughout the transition period. As such, this indicator might be a solid basis for the development of new health monitoring tools.

Chapter 7.

The association between transition management and modelled milk yield in dairy cattle

M Salamone ^{a, b}, M. Hostens ^{c, d, e}, E. Canniere ^f, T. Goossens ^a, V.W.M. van Beest ^e, T. van Gasteren ^g, G. Opsomer ^a, B. Aernouts ^{b, 1}, I. Adriaens ^{b, 1}

^a Ghent University, Department of Internal Medicine, Reproduction and Population Medicine, Faculty of Veterinary Medicine, Salisburylaan 133, 9820 Merelbeke, Belgium

^b KU Leuven, Campus Geel, Department of Biosystems, Division of Animal and Human Health Engineering, Kleinhoefstraat 4, 2440 Geel, Belgium

^c Cornell University, College of Agriculture and Life Sciences, 273 Morrison Hall, NY 14853 Ithaca, United States of America

^d Ghent University, Department of Animal Sciences and Aquatic Ecology, Faculty of Bioscience Engineering, Coupure Links 653, 9000 Ghent, Belgium

^e Utrecht University, Department of Population Health Sciences, Division of Farm Animal Health, Faculty of Veterinary Medicine, Yalelaan 7, 3584 CL Utrecht, the Netherlands

^f Inagro vzw, Ieperseweg 87, 8800 Rumbeke-Beitem, Belgium

^g University Farm Animal Practice (ULP), Reijerscopse Overgang 1, 3481 LZ Harmelen, Netherlands

¹ Authors contributed equally to this work

Adapted from : Salamone, M., Hostens, M., Canniere, E., Goossens, T., van Beest, V.W.M., van Gasteren, T., Opsomer, G., Aernouts, B., Adriaens, I., (2025). The association between transition management and modelled milk yield in dairy cattle. Biosystems Engineering, In Review

7.1 Introduction

Making decisions based on appropriate and meaningful information is essential for dairy farmers to maximise production and improve animal welfare (Kristensen et al., 2008). Providing the farmer with timely and insightful information to take (pre-emptive) action has been an industry-wide focus in recent years. Concretely, numerous software and hardware solutions have been developed to assist farmers in their decision process (Costa et al., 2021). Technologies such as milk meter, in-line electrical conductivity sensors, pedometers, body condition scoring cameras and rumination sensors can provide farmers with insights into their cows' production, health and behaviour. To this end, the time series generated by these sensors must be interpreted at individual cow or herd level to detect deviations from the normal condition. For example, pedometer data can be analysed to identify increased activity and subsequently detect heat events (Chanvallon et al., 2014). These technologies that quantify activity and behaviour have proven to be helpful in assisting decision-making processes such as breeding decisions (Hartung et al., 2017).

In the case of transition monitoring, the development and integration of tailored technologies for informing decision-making support are scarce. The transition period encompasses a time window of 6 weeks around calving (Grummer, 1995). It is a physiologically challenging phase during which animals have an increased risk of developing metabolic and infectious disorders. Early detection of at-risk animals has been the primary focus to mitigate the impact of disorders associated with the transition period on animal health, welfare, and herd economic viability (LeBlanc et al., 2005; Esposito et al., 2014; Overton et al., 2017). Detection methods are typically based on biomarkers, which require financial investment, manual labour and regular sampling (Overton et al., 2017). Moreover, broader management guidelines have been added as a way to mitigate the negative effects of transitioning from the dry period to lactation (Mulligan and Doherty, 2008). So far, such guidelines have been defined by extrapolating insights found at cow level (sub-clinical conditions, metabolites, stressors, behaviour) and focused mainly on supporting dry matter intake postpartum through nutritional or operational decisions (Nydham et al., 2017). Within those guidelines, general principles can be identified, such as actions to limit the body condition change during the prepartum period, supporting the calcium metabolism, reducing environmental stress, mitigating the risk of developing diseases, and implementing protocols to identify animals at risk of transition failure (Nydham et al., 2017). Couto Serrenho et al. (2022) provided an overview of strategies applied on dairy farms, highlighting the differences in nutritional, grouping and housing management. These

management differences have been associated with traditional biomarkers of transition success. For instance, the risk of elevated beta-hydroxybutyrate concentration was associated with comfort measures in the fresh period and water access in the close-up period (Couto Serrenho et al., 2023). Nevertheless, the translation of these findings towards decision support has remained limited. Although Mulligan and Doherty (2008) stated that the existing body of knowledge should allow for optimal management and prevention of transition diseases, little progress has been made in reducing the prevalence of these diseases over the last 25 years (Van Saun, 2023a).

Defining "normal conditions" during the transition period presents a significant obstacle to developing early health detection technologies. During the transition period, changes occurring at behavioural, physiological and production levels are so variable that the concrete definition of a successful transition is still a topic of research and debate (Overton et al., 2017; Horst et al., 2021; Van Saun, 2023a). Moreover, the potential influence of transition-associated health disorders on early lactation data further complicates the establishment of a baseline. Despite these challenges, recent advances in data science offer a potential pathway to overcome this limitation. As such, Salamone et al. (2024) used historical data from the previous lactation to model and predict the expected daily milk yield during the transition period. Deviations from the modelled milk yield, calculated by subtracting the realised from the predicted milk yield, were found to be a good proxy for the cows' health status. Moreover, these milk yield residuals in the transition period (MRT) were associated with physiological and metabolic key indicators of the transition success at individual cow level (Salamone et al., 2023). Concretely, cows with a higher dry matter intake early postpartum had a higher MRT and clinically diseased animals had a lower MRT. Associations of MRT with non-esterified fatty acids, insulin, glucose and IGF1 also indicated the positive association of MRT with the physiological prioritisation of milk production during the transition period.

In the present research, we hypothesise that the MRT can be used as a data-driven tool to evaluate the transition management at farm level, as an extension of the individual animal level monitoring. Therefore, we aimed to investigate the association between MRT and favourable or unfavourable farm management practices. To this end, a survey was conducted among Flemish and Dutch farmers to document their transition management practices and associated these with the average MRT at farm level.

7.2 *Materials and methods*

7.2.1 *Farm selection and survey collection*

Farmers who were already involved in established research collaborations with KU Leuven and UGent or that were clients of the Large Animal practice at Utrecht University in the Netherlands (ULP, Harmelen) were invited to participate in this study. All farmers participated in the national Dairy Herd Improvement program. The survey consisted of 50 closed questions and 9 continuous questions divided into five sections, each focusing on a different aspect of the transition period: dry cow nutrition, environmental stressors, infectious diseases, monitoring protocols during the transition period, and prevention of hypocalcaemia. The closed questions consisted of 20 binary categorical, 26 ordinal categorical and 4 nominal questions. The original survey was performed in Dutch, but an English translation of the complete survey is available in the supplementary material (Appendix C). Questions were formulated to distinguish between what is generally accepted as favourable or unfavourable. The survey data were collected via interviews taken by two interviewers, one per country, performed either in person or by video call. The surveys were conducted between September 2020 and March 2021 for the Belgian farms and between September 2021 and November 2021 for the Dutch farms. During the video interviews, the screen was shared so that the questions were visible to both parties. The interviewer collected the answers in both the in-person and video call interviews.

7.2.2 *Calculating milk residual in the transition period*

All production data from the participating farms were extracted from their herd management software through the MmmooOgle streamer (Bovicom, Puurs, Belgium) or provided by the West Flemish R&D Department for Agriculture (INAGRO, Rumbeke-Beitem, Belgium). The data consisted of milk yield data originating from test days (TD), and ancillary cow informations such as birth and calving dates. For all complete lactations, the expected milk yield on the first test day of the next lactation was predicted using the nextMILK_{TD} model (Salamone et al., 2022), which utilises information from the completed lactation. More details on the nextMILK_{TD} model can be found in Salamone et al. (2022). The MRT was obtained by subtracting the realized milk from the predicted milk yield on the first TD.

7.2.3 *Association of meanMRT with management*

The individual MRT were averaged at herd level (meanMRT) for the period from 6 months before until 6 months after the survey date (period of interest: POI). To avoid the inclusion of

farms with too few MRT data points, the meanMRT was only included in the analysis for herds with at least 10 individual MRT within the POI. This means that the meanMRT was calculated based on the individual MRT of at least 10 cows for which the first TD was in the POI. A partial least squares regression (PLSR) model was used to identify specific management practices associated with the meanMRT. The mean 305-day milk yield of cows that calved during the POI (meanM305), the difference between this mean and the mean 305-day milk yield of cows calved in the year prior to the POI (meanM305 Δ), and the standard deviation of the individual MRT during the POI (sdMRT) were calculated to account for potential year and herd effects. Together with the survey answers, they were considered as independent variables to predict the meanMRT with the PLSR. These independent variables are further referred to as features. All continuous features were z-standardised and categorical features were handled as nominal variables through one-hot encoding. PLSR models with up to 10 latent variables (LV) were trained in a group-wise cross-validation with 5 randomly assigned groups. The optimal number of latent variables was found at the lowest RMSECV. Feature selection was performed in subsequent 5-fold cross-validation. In the feature selection step, performance was defined by the significance multivariate correlation (sMC), a measure indicating the association of the feature with the dependent variable meanMRT (Tran et al., 2014). Concretely, the selected features were identified by iteratively excluding the worst-performing features until the lowest RMSECV was obtained. The effect of categorical features in these PLSR models are compared with a reference category. For questions with more than two answers and with at least one of these answers present in the optimal set, the remaining answers were joined together as reference level. This set of aggregated selected features was used to train the final PLSR model. From this final model, the regression coefficients and the individual performance (sMC) weights of each selected feature were evaluated on their contribution to the prediction of the meanMRT. Moreover, the X-loadings of the selected LV of the PLSR model were evaluated to study the role of the selected features in the selected LV. All modelling was conducted in R (R Core Team, 2020) version 4.1.2 using the following packages: *tidyverse* (1.3.2, Wickham et al., 2019), *skimr* (2.1.5, Waring et al., 2022), *pls* (2.8-2 Liland et al., 2023) and *plsVarSel* (0.9.12, Mehmood et al., 2012). The code used for the model building can be accessed in the following repository: <https://doi.org/10.5281/zenodo.14886820>

7.3 *Results and Discussion*

7.3.1 *Survey results*

A total of 45 dairy farmers participated in the survey: 28 were located in Belgium and 17 in the Netherlands. The median farm size was 100 lactating cows, with the smallest and largest farm having 49 and 376 cows respectively. Automatic and conventional milking systems were represented equivalently in the participating farms, 22 and 23, respectively. Descriptive statistics of the answer distribution to each survey question can be found in the supplementary material, Appendix C.

Although best practices for managing transition cows exist for a long time, our findings suggest that these guidelines are not widely implemented. For instance, the use of metabolic biomarkers to monitor the transition period, as defined by LeBlanc et al. (2005), is still being considered the most effective follow-up method. However, our results show this practice was only utilised at 57.8% of the farms participating in the survey. The majority of the farms that did not use such biomarkers reported the usage of sensors or milk measurements to monitor the performance of the cows. Six farmers based their transition monitoring on more subjective measures such as rumen fill, appetite, general impression and absolute milk production. This disjunction between the general guidelines and their application in the field is supported by answers to several questions of our survey. For instance, separating sick cows from the calving pen, a practice to limit disease spread, was implemented at only 48.9 % of the farms. Around 77.8% of the herds fed the same diet to all dry cows, while a small proportion of farms (22.2%) had a separate far-off and close-up diet. The implementation of fans as heat abatement measures for the dry cows was only observed in 28.9% of the farms. This mismatch highlights the continued need for practical tools to monitor and promote good herd-level management practices, especially regarding the transition management (Mulligan and Doherty, 2008).

7.3.2 *Partial least squares regression model*

The MRT could be calculated for 35 out of the 45 participating farms. Ten farms had insufficient data to predict the individual MRT. Additionally, three farms were excluded as they had less than 10 individual MRT during the POI or insufficient data to calculate the meanM305 Δ . The meanMRT within the remaining 35 farms ranged from -9.51 to 5.06 , with a median meanMRT of -1.63 . The sdMRT laid between 4.1 and 10.1, with a median of 6.6 kg/day. The meanM305 was 9800 \pm 975 kg. An overview of these values at herd level is also provided in the supplementary materials section, Appendix D.

Results from the cross-validation to identify the optimal number of latent variables are plotted in Figure 7.1. The lowest RMSECV was found for a PLSR model with two LV. During feature selection, the lowest RMSECV was found after six exclusion steps, as plotted in Figure 7.2. After feature selection, 12 from the initial 91 features (88 survey features, 3 farm yield properties features) were retained. An overview of the selected features and their relationship with best practice transition management guidelines can be found in Table 7.1. In the final PLSR model, the R^2 of the cross-validation was equal to 69.1 %. Latent variable 1 explained 67.8% of the variance in meanMRT, while latent variable 2 explained 1.4%. The regression coefficients, the X-loadings for the two selected LV and the computed sMC for the selected features are presented in Table 7.2.

The regression coefficient associated with the final PLSR model quantifies the modelled effect of each variable on the meanMRT. For instance, the negative regression coefficient of -1.07 for “FollowupRummenFill - Never” compared to the reference levels could correspond to a less involved attitude of the farmer towards the follow-up of the animals, leading to lower meanMRT on those farms. In contrast, the positive effect of 0.79 for “FollowupRummenFill - Only at-risk animals” could indicate that focusing on at-risk animals leads to a positive effect on the meanMRT. This could be due to a more focused approach, where the attention of the farmer is not diluted among all cows. The presence of an unbalanced nature of the answer distribution urges us not to over-interpret these regression coefficients. The ability of PLSR to account for multicollinearity can also lead to the attribution of an effect where a small subset of data points differs in one confounding question. Some features in the model were only found in less than 10% of participating farms, such as “FollowupRummenFill - Never”, “ComfortDiagonallyDRY - Sometimes”, ”TrimmingOutsiedFixed - Min. 3/year”, “ProfylacticCalcium - None”. This phenomenon does not undermine the results of our study. However, the found effects in these unbalanced features are potentially influenced by farm-specific situations than the general effects of these practices. It might explain less obvious regression coefficients attributed to intermediary responses such as “ComfortDiagonallyDRY - Sometimes”, ”TrimmingOutsiedFixed - Min. 3/year”. The relatively high percentage of explained variance found for this PLSR model could also be linked with this phenomenon.

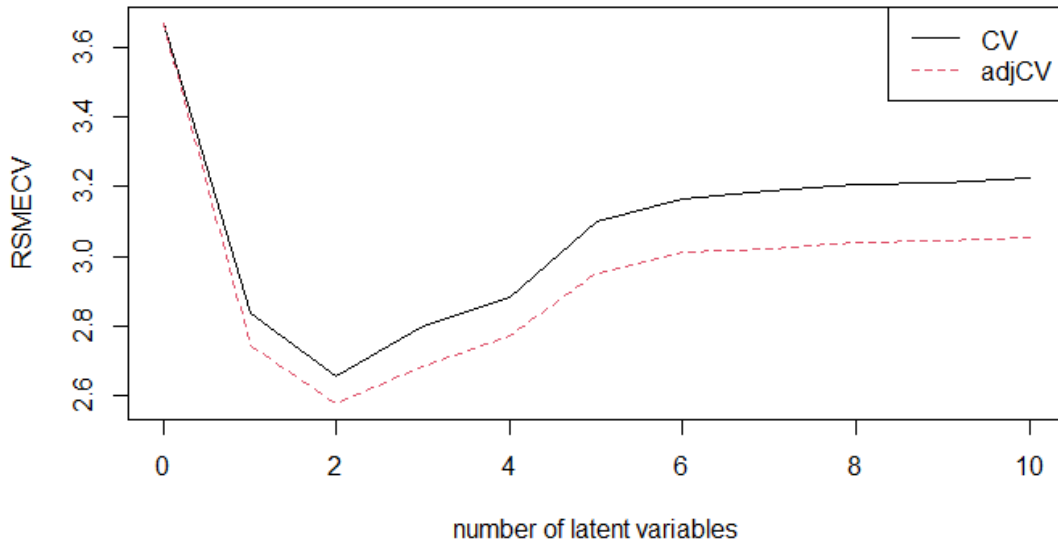


Figure 7.1 Evolution of the RMSECV in function of the number of latent variables of the partial least squares regression model. The CV line represents the RMSECV in the cross-validation. The adjCV line represents the bias-corrected RMSECV. In both cases, the lowest RMSECV is found at 2 latent variables.

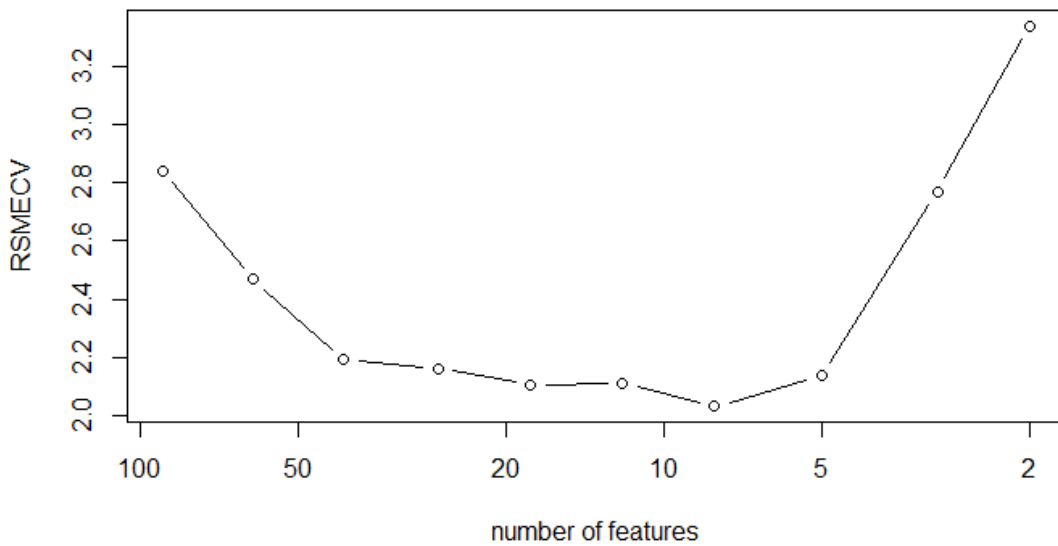


Figure 7.2 Graphical representation of the feature selection step. Each point on the graph represents the exclusion of the lower third of worst-performing features compared to the previous point. The lowest RMSECV was found after 6 reductions.

Table 7.1 An overview of the selected features in the final model. The survey answers and the general interpretation according to the transition period best practice management guidelines are also provided. The following symbols ⁺,^{*},⁰ correspond to optimal, suboptimal and poor practice respectively.

Question	Answer	Guideline
Are cows monitored during the first weeks after calving based on Rumen Fill? (FollowupRummenFill)	All Animals ⁺ Only at-risk animals [*] Never ⁰	A decrease in rumination and eating behaviour have been associated with the wellbeing of cows in the transition period (Nydam et al., 2017; Brown and Bradford, 2021). Establishing protocols based on clinical signs of diseases allows the timely identification of animals requiring intervention.
How often do you see dry cows lying diagonally in the cubicles? (ComfortDiagonallyDRY)	Rare to never ⁺ Sometimes [*] Regularly ⁰	Cow comfort contributes to higher transition success (Couto Serrenho et al., 2022). Inadequate cubicle design could lead to lesions and increased environmental stress.
Do you feed grassland products to the dry cows (fresh grass, silage, or hay)? (GrasslandProductsToDRY)	No ⁺ Yes ⁰	High concentration of potassium in grassland products lead to higher dietary cation-anion difference which is opposed the current strategies to reduce this difference in order to reduce hypocalcaemia cases (Nydam et al., 2017).
In which animals are calcium supplementations (bolus, infusion) used to prevent transition diseases? (ProfylacticCalcium)	All animals ⁺ At-risk animals ⁺ None ⁰	Postpartum calcium supplementation aims to provide cows with readily available calcium until their natural calcium regulation mechanisms are restored (Wilkins et al., 2020).
meanM305	Continuous production characteristics	
Apart from the (possible) fixed times in the production cycle, how many times a year are the claws trimmed? (TrimmingOutsideFixed)	Only when lame ⁰ Min. 1/year [*] Min. 2/year ⁺ Min. 3/year ⁺	2 or 3 hoof trimming during the first lactation has been associated with improved hoof health and survivability of primiparous cows in their second lactation (Åkerström et al., 2024).
Are cows monitored during the first weeks after calving based on appetite (visually)? (FollowupApetite)	All animals ⁺ Only at-risk animals [*] Never ⁰	See FollowupRummenFill.
Do you feed straw to the dry cows? (FeedingStrawDRY)	Yes, chopped ⁺ Yes, not chopped [*] No ⁰	Adequate energy nutrition is crucial to prevent body condition score fluctuations and by extension prevent postpartum metabolic disorders. Straw can be used to dilute the energy content of rations. To reduce sortability straw should be chopped (Nydam et al., 2017).
In which animals is monensin used to prevent transition diseases? (ProfylacticMonensin)	All animals ⁺ At-risk animals [*] None ⁰	Providing monensin to dry animals improved indicators of energy balance significantly (Duffield et al., 2003)
How often do you see lactating cows standing in the cubicles? (ComfortPerchingLACT)	Rare to never ⁺ Sometimes [*] Regularly ⁰	See ComfortDiagonallyDRY.

Table 7.2 Output from the final PLSR model and the coefficient of regression are displayed for each retained question/answer and are ordered by sMC. X-loadings 1 and 2 represent the correlation between the different questions/answers within each of the respective latent variables.

Question	Answer	Percent (n =33)	Coeff	X-loading 1	X-loading 2	sMC
FollowupRummenFill						
	All Animals	69.7	Ref.	-	-	-
	Never	9.1	-1.07	-0.11	-0.31	23.73
	Only at-risk animals	21.2	0.79	0.05	0.27	3.98
ComfortDiagonallyDRY						
	Rare to never, Regularly	90.9	Ref.	-	-	-
	Sometimes	9.1	-1.20	-0.07	-0.42	20.05
GrasslandProductstoDRY						
	No	36.4	Ref.	-	-	-
	Yes	63.6	-1.86	-0.32	-0.36	18.00
ProfylacticCalcium						
	All animals, At-risk Animals	90.9	Ref.	-	-	-
	None	9.1	-0.75	-0.18	-0.08	17.99
meanM305	Continuous		1.82	0.86	-0.39	10.87
TrimmingOutsiedFixed						
	Min. 1/year, Only when lame	75.8	Ref.	-	-	-
	Min. 2/year	21.2	-0.88	-0.19	-0.12	8.98
	Min. 3/year	3.0	-0.53	-0.04	-0.17	13.25
FollowupApetite						
	All Animals, Never	72.7	Ref.	-	-	-
	Only at-risk animals	27.2	-1.08	-0.17	-0.24	10.79
FeedingStrawDRY						
	No, Yes, chopped	75.8	Ref.	-	-	-
	Yes, not chopped	24.2	1.44	0.09	0.50	9.10
ProfylacticMonensin						
	All animals, At-risk Animals	78.8	Ref.	-	-	-
	None	21.2	-0.67	-0.18	-0.04	4.99
ComfortPerchingLACT						
	Rare to never, Sometimes	87.8	Ref.	-	-	-
	Regularly	12.1	-0.32	-0.01	-0.13	1.41

A feature where this unbalance is not present is “GrasslandProductstoDRY - Yes”. The negative effect associated with this feature of -1.86 also seems to reflect current preventive strategies to mitigate hypocalcaemia by reducing dietary potassium present in grassland products (Nydam et al., 2017). Furthermore, the consistent association of MRT with calcium metabolism is furthermore amplified in the feature “ProfylacticCalcium – None”, where an effect of -0.75 was found on the meanMRT if no calcium was routinely supplemented around calving. This relationship between one of the core physiological processes during the transition period and MRT should be investigated further. In a previous study using MRT at cow level (Salamone et al., 2023), 9 out of the 23 animals showed symptoms of milk fever, which was reflected in a significantly lower MRT. The positive effect of 1.85 for the meanM305 on the meanMRT is in line with the current usage of this measurement on dairy farms as the level of milk production has been associated with different degrees of profitability and other herd performance indicators related to reproduction, health and culling risk (Windig et al., 2005; Kristensen et al., 2008; Haine et al., 2017). Furthermore, the meanM305 is widely used as a metric to assess overall performance in genetic improvement, nutrition and animal health (Fleischer et al., 2001; Buckley et al., 2003). The positive effect of meanM305 on meanMRT reveals that higher 305-day milk yield is associated with better transition management and higher MRT (less negative). Similar associations between yield at herd level and traditional biomarkers of transition success were also found by Couto Serrenho et al. (2023), where a greater prevalence of hyperketonaemia was associated with lower absolute milk yield.

For the features “TrimmingOutsiedFixed - Min 2/year”, “TrimmingOutsiedFixed - Min 3/year” and “FollowupApetite - Only at-risk animals” the following effects are found: -0.88, -0.53 and -1.08, respectively. These effects are in opposition with their best practice guidelines, where a closer follow-up with regard to claw health and general health would yield a higher meanMRT. Different potential clarifications to these contradictory effects could be given. Such as, the potential presence of a delayed effect between the implementation of practices and realised improvement in the meanMRT, the unbalanced answer distribution mentioned above or the presence of compensatory practices where a practice is implemented to offset or compensate for known deficiencies in other aspects of the management practices. A longitudinal follow-up of management practices over time could help bridge this gap between reported practice and observed effects. The regression coefficient of 1.44 attributed to “FeedingStrawDRY - Yes, chopped” aligns with the intermediary practice of reducing the energy content of the dry period feed Nydam et al., (2017), the fact that the optimal and poor practices are seen as the reference

level seems counterintuitive. The negative effects observed for “ProfylacticMonensin - None” (-0.67) and “ComfortPerchingLACT - Regularly” (-0.32) reflect, respectively, the protective effects of monensin supplementation prepartum on postpartum metabolic status (Duffield et al., 2003) and the negative effect of inadequate cow comfort on the transition success (Couto Serrenho et al., 2022). Further research on these associations within a controlled environment could help in transforming the findings from our observational studies towards identifying the underlying preventive effects.

The first LV (LV1) of a PLSR model explains as much as possible of the covariance between the features and the dependent variable meanMRT. The second latent variable is orthogonal (uncorrelated) to the first latent variable and explains the next largest portion of the covariance between the features and meanMRT. The second latent variable thus explains the additional variation that is not accounted for by the first latent variable. The X-loadings of each LV help to interpret the magnitude and direction of the relationship between the selected features and the respective LV of the PLSR model. As the features were scaled, X-loadings of these features can be compared and the highest absolute X-loadings thus indicate the most influential features in forming the respective LV. X-loadings in a single latent variable with the same sign indicate that the respective features are positively correlated in how they contribute to that LV. The Y-loading of the LV1 is 2.87, this indicates a positive relation between LV1 and the dependent variable meanMRT, the X-loading of 0.86 for the feature meanM305 indicates a strong positive association between this feature and the LV1. Conversely, the X-loading of -0.32 for the feature “GrasslandProductstoDRY – Yes” suggests that using grassland products for dry cows has a weak negative association with LV1. Smaller negative X-loadings, also indicating a negative association with LV1, were observed for “ProfylacticCalcium – None” (-0.18), “TrimmingOutsiedFixed - Min. 2/year” (-0.19), “FollowupApetite - Only at-risk animals” (-0.17), and “ProfylacticMonensin – None” (-0.18). Moreover, these negative loadings indicates that the corresponding features are negatively correlated with meanM305 in how they contribute to LV1. This correlation is the strongest in “GrasslandProductstoDRY – Yes”. In other words, these loadings suggest that meanM305 is most influential in LV1 and, consequently, influential in meanMRT. Furthermore, farms with higher meanM305 tend not to use grassland products for dry cows. They also tend to provide calcium or monensin prophylactically to some animals, a practice associated with a higher meanMRT. However, the interpretation of the remaining associations is less clear, as contradicting poor and optimal practices show similar associations.

The Y-loading of the second latent variable (LV2) is 2.31, indicating a positive relation between LV2 and the dependent variable meanMRT. Multiple features present high absolute X-loadings in LV2. The features positively associated with LV2 are features “FeedingStrawDRY – yes, but not chopped” and “FollowupRummenFill – Only at-risk animals” with their respective X-loadings of 0.5 and 0.27. Surprisingly, these features do not represent the optimal choice based on literature but rather the intermediary solution. The following features are negatively associated with the second latent variable: 'ComfortDiagonallyDRY– Sometimes' (-0.42), 'GrasslandProductstoDRY – Yes' (-0.39), meanM305 (-0.36), and 'FollowupRummenFill – Never' (-0.31). Similar to the features with a positive X-loading in LV2, the selected feature “ComfortDiagonallyDRY– Sometimes” reflects an intermediary response between “Rare to never” and “Regularly”. The negative X-loading for LV2 of “FollowupRummenFill – Never”, similar to this feature regression coefficient, could advocate towards better follow-up of animals during the transition period. The features meanM305 and "GrasslandProductsToDRY – Yes" exhibit contradictory X-loadings when compared to the first latent variable, but this could be because the major effect of these features is already represented in LV1. The unexpected X-loadings for some of the selected features, giving more relative weight to the intermediate instead of the optimal or the poor solution, combined with a relatively low explained variance, suggests that the second latent variable also reflects unbalanced responses and hence warrants caution in the interpretation.

In summary, results from the PLSR model identified possible meaningful associations. However, these results also highlighted the possible effect of unbalanced distributions in the answers making interpretation more difficult. Moreover, surveys are a snapshot fixed in time where complex interactions in management and individual animal performances are difficult to capture. Combined with limited ability to allow nuances in the response, this possibly resulted in less interpretable effects.

Mills et al. (2020) identified barriers to improving the management in the transition period, categorised into different themes: socio-psychological factors (farmer attitude, veterinarian involvement), individual farmer perspectives (nutrition, overstocking), and the lack of a clear definition of transition success. Our results tie in with the latter theme, where farmers report sources of information such as conferences, websites, and data from herd management software as sources to implement certain management on their farms. Still, they also indicate that the inconsistencies between research findings make it difficult to interpret and implement these findings into their management (Mills et al., 2020). The reality of these conflicting findings and

how they translate into re-evaluations of best practice guidelines can also be found in recent academic work. Couto Serrenho et al. (2022) classified management practices into optimal, suboptimal and poor and motivated this classification to be true at the time of assessment. This caveat also highlights that the interpretation of management guidelines could be subject to change over time.

To tackle the aforementioned limitations of the present study, efforts could be made to increase the number of participating farms. This could be done by using a commercial survey as was done by Couto Serrenho et al., (2023). Nevertheless, more efforts should also be considered to simultaneously collect production data to reduce the loss of farms between survey and MRT calculation (27 % in our study). Alternatively, a focus group of farmers could be set up in order to establish a more longitudinal approach, studying the evolution of MRT over a longer time span and allowing assessment of the effect of change in management within a farm. The incorporation of expert knowledge of routine farm advisors (veterinarians, feed suppliers, ...) in these focus groups could create a broader understanding of farm-specific events. The results of the present study suggest an association between MRT and management practices, highlighting the feasibility in developing a practical and integrated assessment tool tailored to the transition period.

7.4 Conclusion

As an initial step towards the development of an assessment tool for practices related to transition management, this paper reveals associations between the meanMRT and farm management practices. In addition to the relationship with herd production performance (meanM305), the PLSR model identified relevant management practices that could be linked with physiological processes during the transition period. These are the inclusion of grassland products in dry cow rations and the use of calcium or monensin prophylactically to prevent transition-associated diseases. Follow-up protocols for transition cows based on rumen fill and cow comfort were also associated with meanMRT. However, confounding aspects were also identified in this research urging caution in generalising our findings. The inclusion of more farms in the analysis or a longitudinal set-up could alleviate this study's limitations.

Chapter 8.

Valorisation

8.1 From milk residuals in the transition period to lactation onset value

The necessity to better control the transition management and success has been extensively discussed in the introduction. In [Chapter 4], the development of the nextMILK_{TD} model was described in depth after which its key derived metric (MRT) was linked to key transition period in and outputs such as metabolic parameters [Chapter 5], behaviour [Chapter 6], and management [Chapter 7]. Although MRT plays a crucial role in the process-based framework outlined in [Chapter 3], research is needed to validate its use in the field as a tool for providing cow-side decision assistance or long-term monitoring of herd transition success. Nevertheless, in addition to these essential validation steps, an initial outline of how MRT could be applied in practice can be developed to assist farmers in their day-to-day operations. As such, a negative MRT would flag a potential failure to transition in an animal and allow measures to prevent economic losses and negative effects on the welfare of animals. On a more general farm level, the MRT could provide insights into the potential processes at play regarding the transition period for the farmer and his advisors.

Similar to the systematic approach utilized in the PLF framework for developing the MRT, this chapter employs the business model canvas to delineate various aspects of a potential implementation and valorisation strategy. Originally introduced by Osterwalder (2004), this canvas focuses on the infrastructure (key activities, key partners, key resources), offering (value propositions), customers (customer segments, channels, customer relationships) and financial aspects (costs and revenue streams). Figure 8.1 provides a visual representation of this canvas. By analysing the constituting elements of this canvas in the context of transition monitoring and the MRT, the outline of an initial valorisation plan can be drawn. The following subsections are structured around this canvas and highlight the elements of a future business model that are currently already in place and identify parts of the business plan that require further developments.

8.2 Infrastructure

8.2.1 Key Resources

A first step in identifying the key resource for this valorisation plan involves rebranding the MRT to a more evocative name: the lactation onset value (LOV) is proposed. This transformation clarifies the MRT as a numerical value specifically developed to assess the onset of lactation, thus forming the central key resource in this valorisation plan. Its numeric nature provides inherent versatility, allowing for a wide range of applications in the field.

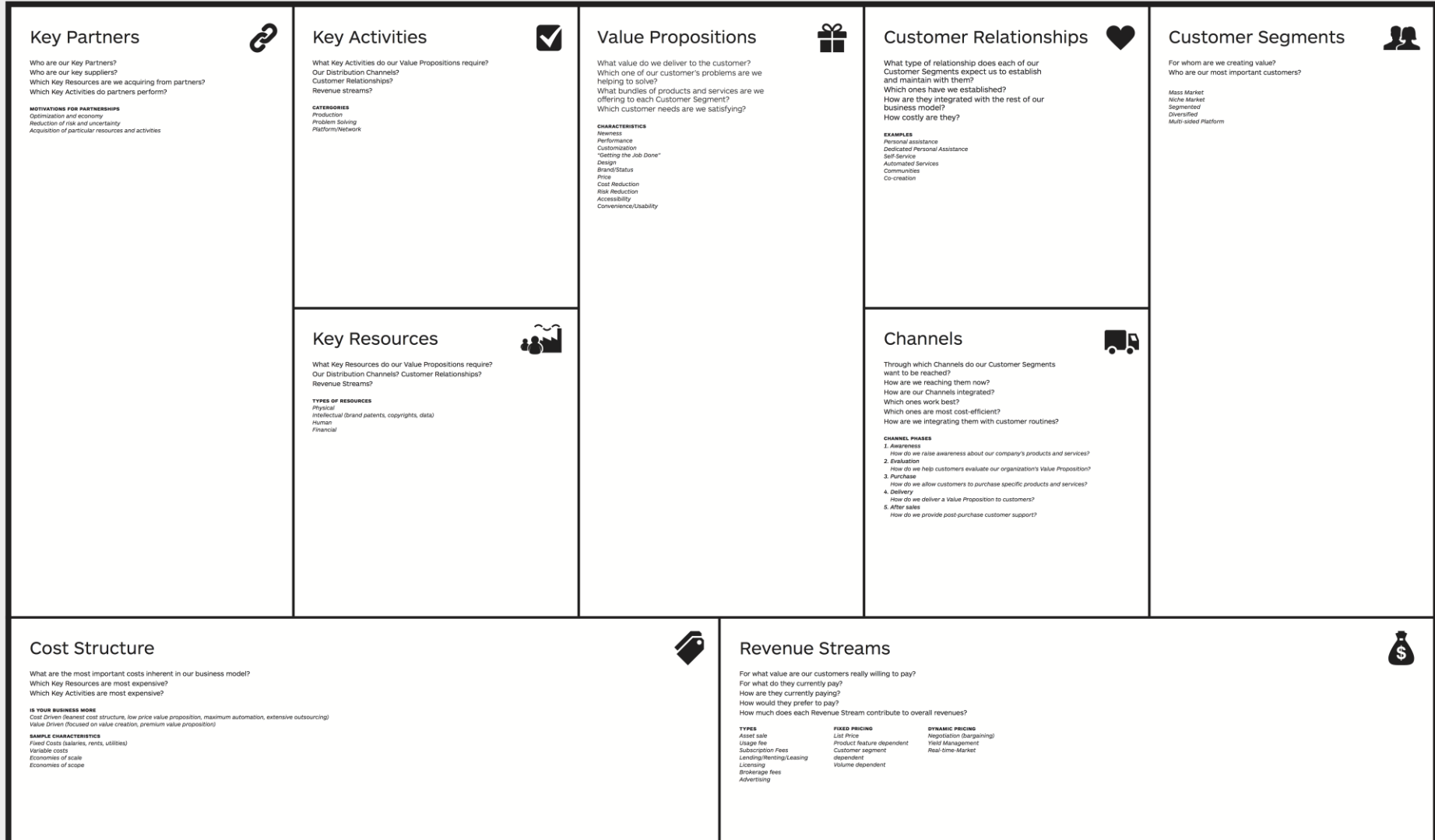
The Business Model Canvas

Designed for:

Designed by:

Date:

Version:



DESIGNED BY: Business Model Foundry AG
The makers of Business Model Generation and Strategyzer

This work is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported License. To view a copy of this license, visit: <http://creativecommons.org/licenses/by-sa/3.0/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.



Figure 8.1 Business model canvas

Next, examples of these applications and their relevance to current practices in the dairy farming industry will be presented. The resources associated with the LOV shown in this section aim to present the insights generated by the LOV in the most understandable form, which would allow for the evaluation of ongoing processes on the farm over time. The LOV could be implemented in the form of a report, where the LOV and all related information (i.e., animal ID, calving date, milk production) would be assembled in a one-pager similar to milk recording reports. In the case the LOV is computed using DHI data (LOV_{TD}), the report could be annexed to the standard DHI report (every four to eight weeks). In contrast, a report on the LOV derived from milk meter data (LOV_{MM}) could be generated with greater flexibility according to customer preference, for example, two-weekly, monthly, quarterly.... An example of such a report for LOV_{TD} is presented in Figure 8.2. This draft includes three main components. The first component contains a text describing the meaning of the LOV. In this report, the descriptive text could be adjusted based on the target audience, depending on the age and education level of the audience. The second component aggregates LOV data from the past year in the mean LOV and provides a top/bottom ranking of the LOV of all individual cows in that period. This section provides a snapshot of the LOV at the moment of generating the report, while the evolution of the mean LOV since the last report provides a broad overview of the evolution of the farm. Additionally, the rankings provide a rapid overview regarding the possible reasons justifying evolutions in the mean LOV. For instance, a series of new highly negative LOV in the bottom ranks could already indicate a downwards shift of the LOV. The final component of this report offers a graphical representation of the different calving events (x-axis) and the computed LOV for those individual events (y-axis). The rolling average line is also plotted, with the LOV for animals entering their second lactations grouped separately from older cows. This section provides the specific individual data points for each calving allowing the identification of subtle drifts over time within well-defined groups. This grouping is currently based on the results of [Chapter 5], where higher parity animals had lower MRT in general. Depending on the customer and future findings regarding the MRT, grouping could be organised differently, for example, based on breeding values for resilience or based on previous LOV (lower LOV in previous lactation – vs others). The temporal resolution of the LOV is an aspect that requires further clarification. Investigating questions such as “How much information is present in a year of calving events?” and “Is it dependent on customer profile or herd size?” will be essential hurdles to define the optimal temporal resolution for this report and, more generally, for the LOV.

Lactation onset value report

A cow that successfully navigates the transition period has a higher chance of a successful lactation and a longer lifespan. To monitor this transition period, we have developed a new metric called the **Lactation Onset Value (LOV_{TD})**. This metric represents the difference between the actual milk production recorded on the first test day (TD) and the predicted healthy milk production. Cows that perform worse than predicted have a negative **LOV_{TD}** and are considered potential risk animals. By examining the average **LOV_{TD}** at the farm level, one can assess the effectiveness of management strategies during the transition period. This report has been prepared for **herd X**

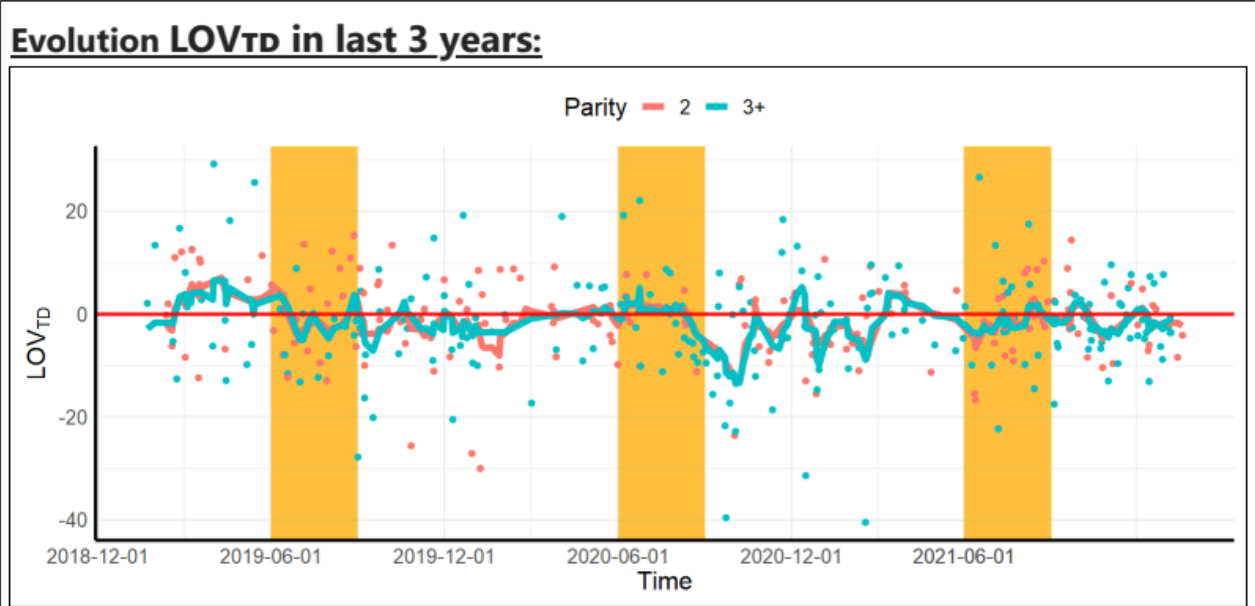
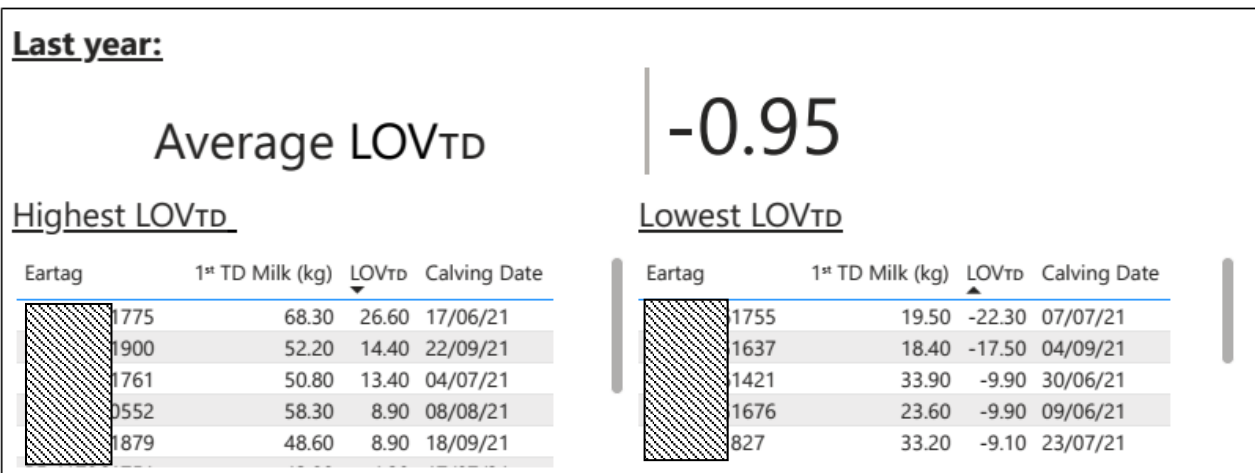


Figure 8.2 Example of a report which could be generated using the LOV_{TD} and ancillary cow and herd information

Additionally, an exhaustive list could also be attached to the report, detailing the numeric LOV and related data for the most recent n number of calving events. This list would offer an exhaustive overview of the different summarising plots in the report and allow for more targeted decision-making at cow level. For example, this list could facilitate the tailoring of the monitoring protocols for soon-to-calf cows based on their previous LOV, or provide additional information with regard to operational decisions regarding culling or breeding. The plots and lists featured in the report could be accessed via a web application or mobile phone app, enabling more interactive engagement from customers with the presented results. Alternatively, the LOV could also be assembled in a digital dashboard illustrated in Figure 8.3. Such a dashboard could incorporate data coming from multiple herds and, with the help of a control algorithm, generate alerts for significant deviations in individual animals or general herd trends.

Dashboard Multiple herds

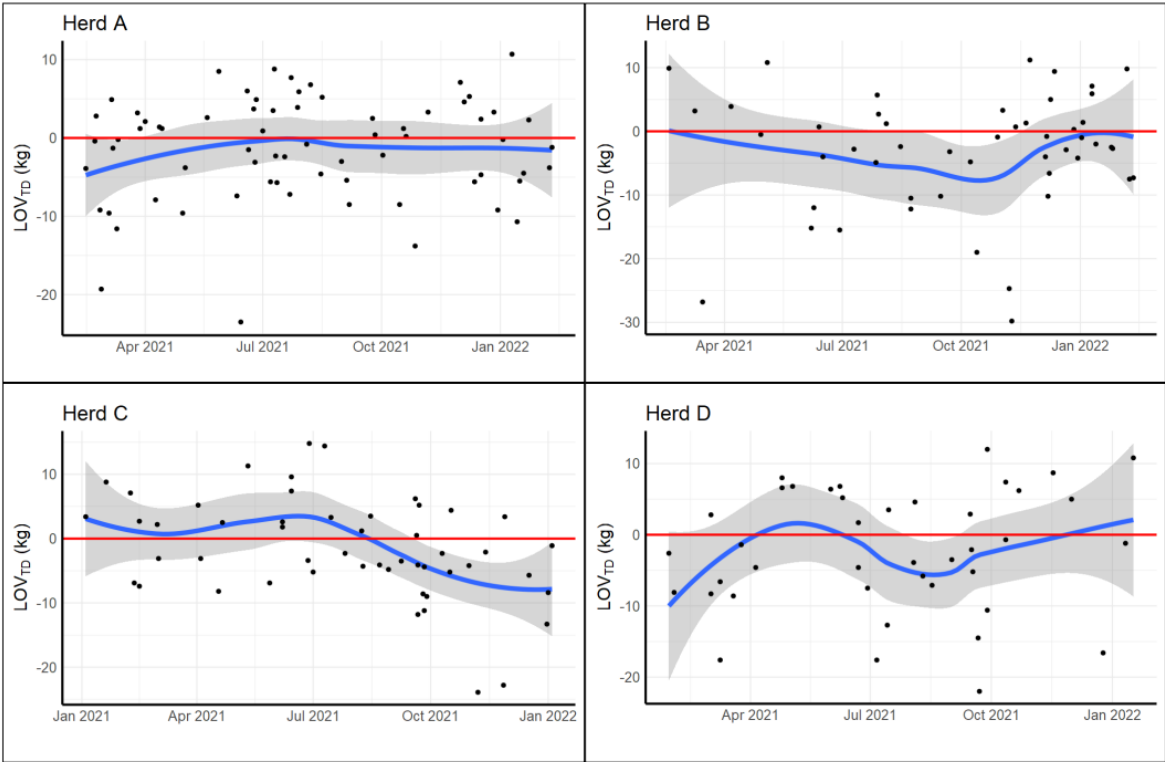


Figure 8.3 Example of a dashboard combining the information of different herds in one visual environment.

8.2.2 Key Partners

The LOV fundamentally depends on the data needed for its computation. The availability of these data, comprising TD records or milk meter records and ancillary data depends on farmers participating in DHI programs, or having implemented technology to measure and record milk

yield daily. Additionally, acquiring these data requires collaboration with external partners capable of providing it. Currently, the potential key partners in the dairy industry fulfilling this role are limited to a few actors in the dairy industry. The primary actors in the current dairy farm “ecosystem” are hardware and technology suppliers such as Lely, DeLaval, GEA. These companies supply machinery with the primary goal of automating and assisting processes on the farm (milking systems conventional and or automated, feed pusher robots, concentrate feeders, pedometers...). To varying degrees, the data collected is utilised within a framework resembling the PLF framework presented in [Chapter 3], either as a primary control metric or as a byproduct of a fully automated process. Furthermore, this generated data can be processed locally or through a cloud processing service presented through a proprietary herd management system.

Historically, milk recording data collected through DHI programs represent a major part of the data collected at a dairy farm. As extensively discussed in the introduction of this thesis, these programs have been existing for over a century and have been identified as key target of the LOV, by fitting a model specifically to work with TD milk yield data. Again, the generated data is presented through proprietary channels, making access to raw data challenging without partnering with these organisations.

While this aspect of the valorisation plan may appear straightforward at first sight, establishing partnerships with the two primary types of data providers has proven to be complex. The reluctance of these actors to collaborate became evident during the development of the LOV. LOV was developed as part of a Flanders Innovation & Entrepreneurship funded project (VLAIO), which involves co-funding by industry partners, allowing them to express interest and participate closely in the development process. Key actors such as Lely (a technology supplier) and CRV (the Dutch/Flemish DHI program and breeding organisation) were part of this consortium and showed interest in the LOV concept. However, no concrete partnerships with these actors were realised by the project’s completion. One of the potential reasons for this outcome, alongside the technical difficulties of standardisation of workstreams and data format, may lie in the closed data environments these companies have established. This position allows them to act as monopolists with regard to their customer base. Collaborating on the development of third-party products such as the LOV could increase competition in their respective markets, which may explain their reluctance to enable the emergence of an open solution. These challenge their closed models on the farm, which currently grant them the ability to selectively integrate with third-party innovation or develop them internally. One could

argue that these companies exploit the innate reluctance of farmers to switch between new/additional digital tools, thereby ensuring they remain the primary partner of the farmers. Despite these challenges, partnerships with these companies should still be encouraged because success in this endeavour would facilitate the adoption of LOV across numerous farms.

An alternative for identifying partnerships to acquire the necessary data and provide the LOV in return could be the different third-party herd management software, such as UNIFORM-Agri (Assen, The Netherlands), Agrovision (Gelderland, The Netherlands) or dairyCOMP 305 (Visalia CA, US). Over the years, these herd management software systems have established themselves as a central data hub for dairy farmers by coupling the data from different sources (on-farm technologies, national registration programs, medication registration) and presenting them to the farmer. Methods used to achieve this coupling are based on formal partnerships with the major players of the industry. Nevertheless, the implementation of those overarching platforms remains fragmented. A potential reason for this could be the individual farmer's choice to adopt this software or not, restricting adoption to individual preferences or cost evaluations. The LOV could benefit from partnerships with these companies to increase adoption, though a thorough evaluation of the cost-benefit ratio should be performed.

The emergence of third-party data management software providers could alleviate the fragmented character of the herd management software mentioned above. Companies such as Bovicom (Puurs, Belgium), Dairy Data Warehouse (Assen, The Netherlands) and iYOTAH (Westminster CO, US) have emerged over the past 15 years, aiming to facilitate access to data beyond the farms walls. The coupling of data does not serve the farmer directly but can be used by stakeholder companies (e.g. feed providers and veterinary practices) to monitor their clients. Partnerships with these companies could yield more valuable access to a broader number of farms. An additional advantage is that the challenger role of these third-party software solutions in the dairy industry positions them to be less reluctant to share data if, in return, it would allow them to further differentiate themselves from the established actors. Notably, Bovicom was part of the VLAIO project consortium and provided essential data needed for the development of the LOV through their MmmooOgle platform.

8.2.3 *Key Activities*

This valorisation chapter explores the potential for a practical use case for the LOV. As stated before, the LOV could be implemented in various contexts. Three concrete workstreams can be identified to substantiate these implementations: calculate, implement, and disseminate. Computing the LOV in all possible forms would form the core activity surrounding this data-based tool. The reliance on third-party partners in this workstream will require strategic decision-making to identify the potential candidate in which collaboration would allow the LOV to be computed but remain aware of its platform dependency. These broad collaborations should be sought after with multiple partners in the industry and avoid getting stuck in exclusivity partnerships. These types of partnerships could be an actual threat to the existence of the LOV, with the loss of interest over time and the obsolescence of the LOV in an evolving dairy industry.

The second workstream's primary goal is to mitigate this potentiality by maximising the number of potential implementations, i.e. software, mobile app, food label requirements... the LOV should acquire more notoriety and a broader customer base. The development of tailored pipelines to incorporate the LOV in existing tools will necessitate close collaboration with the providing partners. In this regard, the LOV presents intrinsic benefits and the models behind it are relatively simple, well-documented and programmed in object-oriented languages (Scala and Python), which makes them straightforward to implement. In essence, if the models' inputs are provided in the form they expect, the computation can happen independently. Nevertheless, complexity will arise in the formulation of the shape and form of the inputs, in addition to tailor the LOV to the different user interfaces. The possibility of relying on specialised software porting companies should be investigated to facilitate this implementation workstream.

Finally, the dissemination workstream can rest on the academic facet of the LOV. The close partnership with key regional actors of dissemination (e.g. koesensor, Inagro) in the development of the LOV provides the network to kickstart access to a broad network of potential clients. Furthermore, the numerous educational moments organised at the university or within the context of the project could also be used as a platform to disseminate the knowledge acquired on the LOV as a way to improve the transition period and present its potential applications.

8.3 Customers

8.3.1 Customers Segment

The LOV was developed to assist decision making during the transition period. Potential customer segments are characterised by their involvement in the decision-making process on the farm. The primary and central actor in this process on most average-sized farms is the farmer. Although consulting advisors on farms could also be seen as potential customers due to their close participation in the decision-making process, this valorisation plan primarily identified these actors as channels to reach the farmer. The following section will focus on this potential customer segment in a Belgian and European context, characterising it, and analysing how the LOV could be valuable for them.

According to Eurostat data, in 2020, 8.710 dairy farmers were active in Belgium (Eurostat, 2024), with a predominantly male (91.58%) profile. These data highlight an ageing population, with 67.76% being between 45 and 64 years old in Belgium, which is representative for the European Union as well. Data about the education levels of farmers across all branches of agriculture illustrate that 48.2 % of farmers only have practical experience acquired through working on the farm, 29.3% followed a basic agricultural training (general college, apprenticeship) and 22.5% had higher level agricultural training (agricultural college, university or other institute of higher education) (Eurostat, 2023). The impact of age on these distributions is notable: 44.6% of farmers aged 40 to 64 reported only practical experience as training, compared to 26.2% of farmers aged <40. Conversely, the proportion with full agricultural training increased from 24.3 % to 37.2 % in this age group (Eurostat, 2023). Within this context, the LOV is particularly relevant and could play an important role in alleviating the challenges the dairy industry is facing. Demographically, as the number of farmers decreases and many approach retirement age, the need to automate will become more important. Nevertheless, the increasing education level combined with the younger farmers increased incline to implement PLF systems (Bianchi et al., 2022) creates a window of opportunities to develop and implement tools such as the LOV to shape the automated farms of the future.

The structure of labour on dairy farms was thoroughly described in the 2021 European Commission dairy report and associated data (European Commission, 2021). In Belgium, 94% of farm labour was family labour, with an average of 2.0 full-time equivalents employed per farm. The average number of dairy cows per farm was 88, nearly double the number of cows per farm in 2008. In contrast to this increase, the number of farms over the same time period

decreased from 7.937 to 5.338. Meanwhile, the forage area per farm increased by 50% (from 40 to 59 ha) and the tons of milk produced per fodder area (crop + grassland area) increased by 65% (from 7.9 to 12 t/ha). This increase in production is also evident from the absolute milk production on the average Belgian farm, which doubled from 318.6 to 711.9 tons over those 12 years. This shift in dairy farm size will require the ability to manage larger herds, manage more land and produce more milk with a diminishing amount of labour. This reality justifies that global policies are pushing for automation and digitalisation to tackle these and future evolutions. In this regard, the LOV and similar tools could become foundational in reducing labour demands during crucial periods, allowing to focus on animals in need of attention and reducing labour in animals actually thriving.

A comprehensive economic characterisation of dairy farming would require an extensive study beyond the scope of this valorisation plan. Nevertheless, some key figures for Belgium can be extracted from the latest data sourced from the European Commission's agri-food data portal (European Commission, 2024). In 2022, milk sales represented 89.4% of the economic output of the average dairy farm. The economic assets of a dairy farm were estimated at approximately 1 447 499 €, with 82% of these assets being fixed (land, buildings, machinery and livestock) (European Commission, 2024). The evolution of milk prices in the national and European markets over the last 10 years is plotted in Figure 8.4, characterised by significant fluctuations. For instance, in Belgium, in the last 2 years, milk lost 35% of its value, dropping from 59.06 €/ton milk in October 2022 to 38.04 €/ton milk in August 2023, before regaining 31% of its value by December 2024. Figure 8.5 presents the cost breakdown per 100 kg of milk produced, alongside the average remuneration for family labour on dairy farms over the past 15 years. These plots suggest that operating costs have followed milk price fluctuations, and the importance of family-imputed costs (labour and capital) have been diminishing over time for each produced unit of milk. From the remuneration for family labour shown in Figure 8.6, it seems that dairy farmers in Belgium, on average, can remunerate themselves above minimum wage levels. However, the trends under challenging years (e.g. 2015, 2016) seem to indicate that farmers reduce their income when the economic situations are less favourable. The potential for the LOV to minimise wastes and reduce costs, through the prevention of milk production losses, reduction of veterinary costs and involuntary culling, allows it to contribute to a more resilient farmers' income.

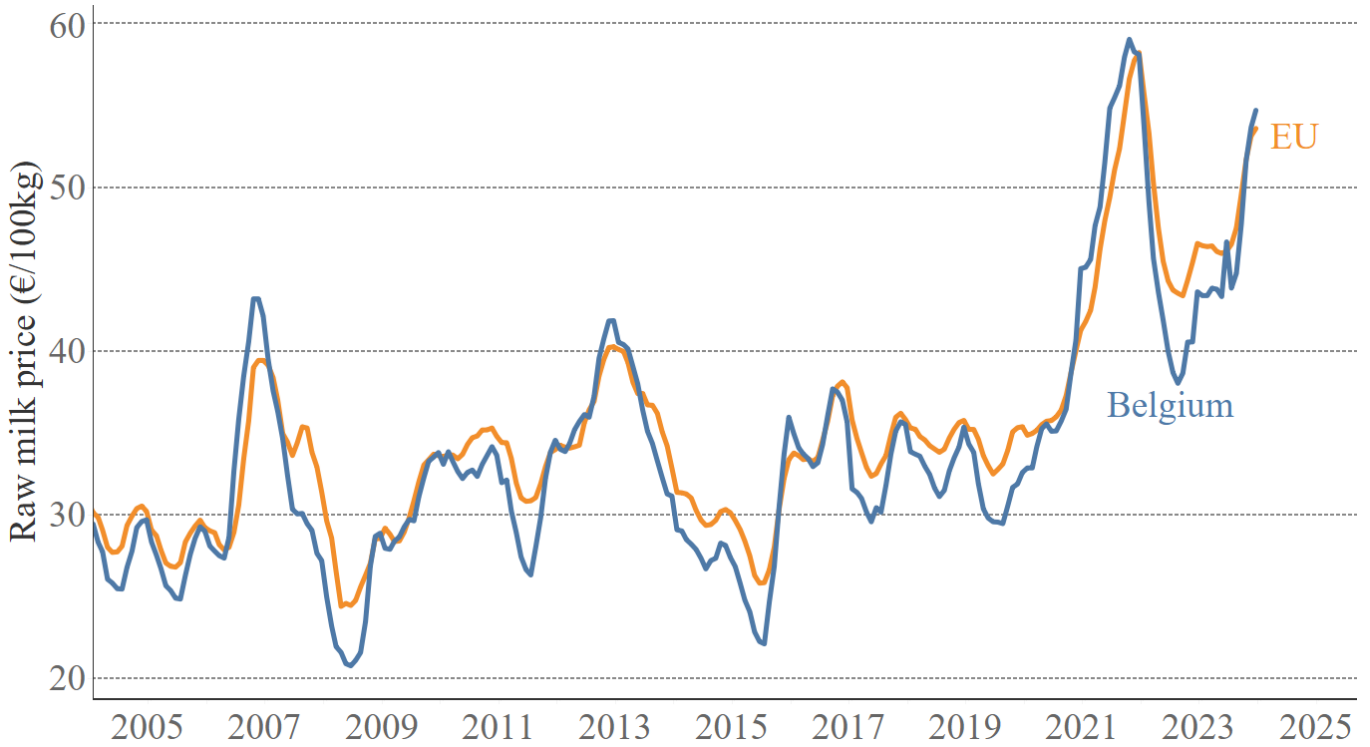


Figure 8.4 Average monthly prices for raw milk prices over the last 10 years on the European and Belgian market. Using data from the European Commission (European Commission)

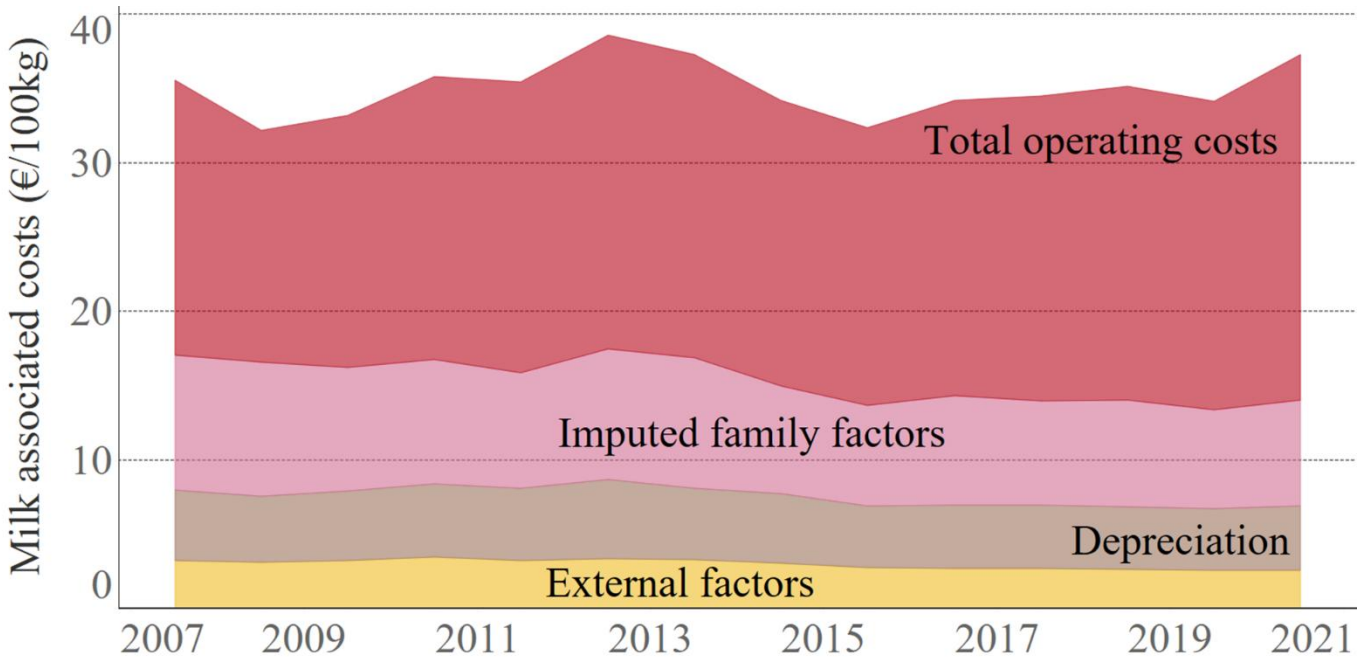


Figure 8.5 The of the cost breakdown of the milk production in Belgium, Operating costs include specific costs of milk production such as feed costs, herd renewal, milk levy costs... It also includes nonspecific costs such as machinery and building upkeep, energy, taxes... External factors include non-family labour wages, rent and interest. Imputed family factors incorporate family labour cost and own capital costs. This plot is based on data from European Commission through their Agri-food Data portal.

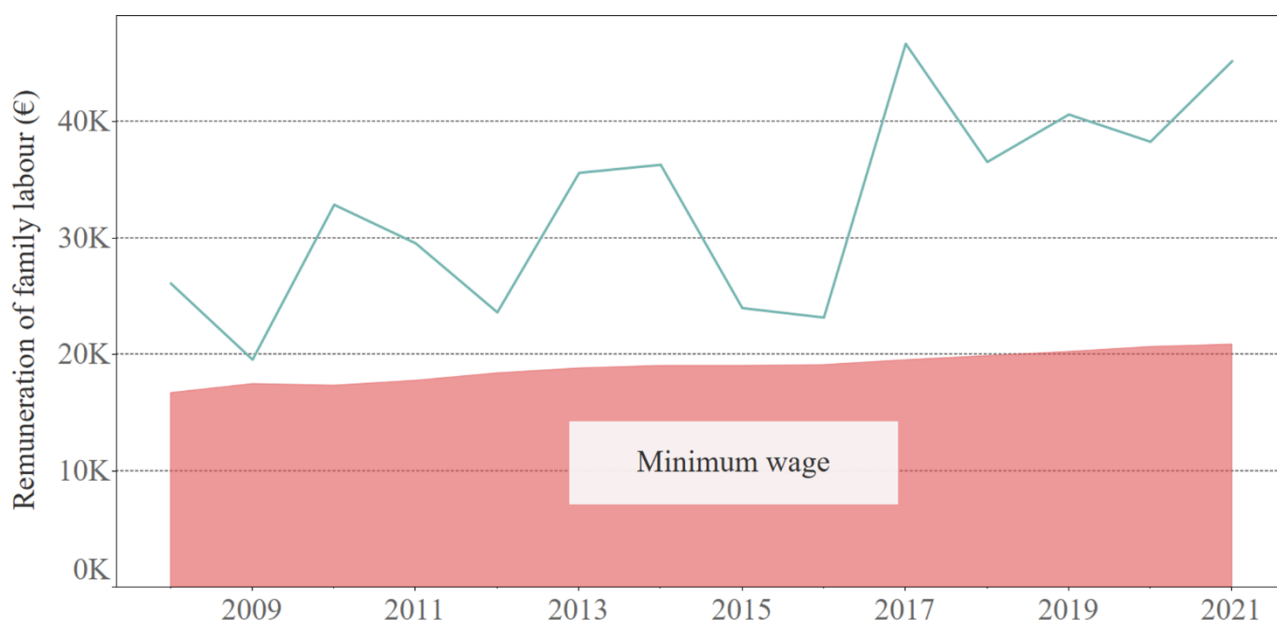


Figure 8.6 The bottom panel translates the family labour into a yearly average wage for each working unit for Belgium. Minimum wage for agricultural labour is also provided as a measure to interpret the remuneration. This plot is based on data from European Commission through their Agri-food Data portal, minimum wage data is based Belgian governmental data (Federale Overheidsdienst Werkgelegenheid)

These economic aspects further influence farmers' considerations regarding the implementation of technology on their farms. Specifically, when deciding to invest, farmers prioritise the benefit-to-cost ratio, total investment cost and user-friendliness (Borchers and Bewley, 2015). Complementary to economic considerations, the socio-psychological influences on technology adoption have been examined in recent studies. For instance, Ahmed et al. (2024) identified that the adoption of PLF tools depends on the interplay between cognitive factors such as the farmer's overall perception of PLF, the farmer's belief in their ability to work with the PLF tool and environmental factors such as the technological networks in which farmers are already embedded. These technological networks relate to the social network in which farmers operate and their opportunities to access peer-to-peer learning or access to problem-solving support. These findings suggest that enhancing PLF implementation could involve the integration of farmers into knowledge-sharing networks, these user groups could thereby minimise perceived risks associated with PLF and boosting farmers' confidence in their ability to utilize new tools. Recognising the importance of economic considerations in adopting technological tools, the LOV has been designed to interoperate with technologies already widely present on dairy farms (DHI and MM), lowering the economic barrier to adoption. Efforts to promote the use of LOV in dissemination strategies, should also focus on socio-psychological factors emphasising this

aspect to achieve sustainable implementations (Ahmed et al., 2024). Identifying effective channels and tailored resources to reach farmers will be key in this process.

8.3.2 *Customer channels & customer relations*

As an academic endeavour, establishing direct customer relationships between our research group and farmers as our customers is not a feasible approach for bringing LOV to the market. Therefore, the LOV should be introduced to the farmers through existing services provided by third-party actors. These third-party actors are involved in the decision-making process, either by counselling farmers in their decision-making (veterinarian, feed advisor...) or by providing the tools to inform farmers' decisions (software/hardware companies). The difficulties in collaborating with these third parties have been discussed in previous sections. The primary channel for reaching farmers is through the plethora of advisors who interact with them daily. Historically, veterinarians have been identified as the most important consultants related to animal health (Pothmann et al., 2014). The changes in the industry seem to have altered that dominant position. The rapid pace of industry evolution, the increased importance of preventive medicine, and the rising importance of automation and data analysis seem to have surpassed the traditional dairy veterinarian. Veterinarians are challenged by different factors such as the diversification of knowledge needed to assist the farmer in a systemic way (automated feeding units, milking robots, sensors...) and the necessity to stay updated, coupled with a low willingness among farmers to pay for quality consulting in dairy farming (Gerber et al., 2020).

In other words, aside from their clinical work, veterinarians primarily offer their expertise in a market where farmers are reluctant to pay for it, while competitors can offset consulting costs through the sale of products (feed additive, machines, milk levy...). Nonetheless, veterinarians are not entirely disadvantaged. Their regular presence on the farm to treat animals enables them to detect early signs of farm-wide issues. Additionally, their lack of pure commercial interest and the advanced education linked with their profession should encourage them to remain critical of other consultants and question commercial claims.

The work presented in this thesis is crucial for elucidating the processes at play and formulating potential strategies to improve the sustainability related to the key processes on dairy farms. Veterinarians can be crucial in fostering an evidence-based ecosystem of knowledge providers to support farm health improvement. The presence of farmers' social networks connected through their common veterinarian practice presents the potential to become the user groups

mentioned above. These groups would enable farmers and veterinarians to collaborate on using, exchanging, and experimenting with new digital tools and data analysis platforms, collectively pushing for broader adoption of revolutionising tools. On a smaller scale, advisory initiatives like those established by Inagro, a regional research institute in Belgium, exemplify this potential. They organise regular gatherings for farmers to disseminate knowledge, share experiences and participate in research studies. In these events, the LOV and its generated reports could become a tool to facilitate knowledge exchange by standardising and objectively measuring the subject of discussion. More generally, numerous veterinary practices host annual gatherings for farmers and clients, but these focus more on relationship-building than knowledge exchange. During the course of this PhD, the LOV was presented at different types of gatherings, including the meetings of the user groups established by INAGRO as well as the clients of ULP veterinary practices. The farmers' interest and the relevant discussions and obtained insights highlighted the huge value of these type of interactions for all involved parties.

However, further efforts are needed to educate veterinarians, enhancing their training by providing more in-depth knowledge on technical aspects of PLF (i.e. data processing, engineering, commercial selling techniques). Additionally, the education of young farmers could also be leveraged by introducing key concepts, fostering a more favourable attitude towards these technologies. The faculty of veterinary medicine at UGent has recognised these necessities and created a new PLF-focused professor chair. The digital education of the veterinarians of tomorrow should not only focus on the technical aspects but also their potential role as evidence-based experts and educate them in the steps towards developing effective user groups around their potential future clients.

8.4 Offering and financial aspect

The value proposition of the LOV aligns with the general proposition of PLF, which aims to automate and analyse processes for improvement. Each research chapter of this thesis aimed at finding evidence for the inherent potential of the LOV to fulfil these goals. Specifically, the extensive validation of the MRT and its relationship with the transition period provides foundational validation for the LOV's value proposition. However, future research should focus on quantifying the LOV's practical value in the field. Key questions remain, such as whether the LOV provides sufficient insights into the transition period to enhance management when used for animal monitoring, whether it can serve as an auxiliary tool for detecting diseased

animals, and what the tangible economic benefits are of responding to insights generated by the LOV.

This gap in practical knowledge complicates the formulation of a financial plan. In the long term, different options are feasible: licensing the LOV through revenue-share licences or usage-based license to multiple parties, or selling the intellectual property to a single software provider. Well-established dairy key performance indicators (KPI) have been negotiated in this fashion among major actors. For example, the standard peak production (SPP) developed by CRV has been integrated into different software. In its current state, the LOV could be proposed to smaller, more innovative dairy data companies to be implemented in the field, gain notoriety and create a basis to investigate the remaining value proposition questions.

8.5 From research to reality: bringing the LOV to the dairy farm

In conclusion, this outline of a valorisation plan is a necessary supplement to the academic work, aimed at analysing the strengths of this research and identifying the weaknesses and opportunities that must be addressed to bring it to the field. Understanding and researching industry challenges has been central to the academic work, though efforts to translate these findings into practical applications have remained limited. The dairy industry is a challenging market, characterised by the consolidated nature of service providers, the cost-conscious nature of dairy farmers and their unwillingness to pay for consulting. Whereas these characteristics are real hurdles in implementing research output, commercial actors can innovate and present new products to farmers. In that regard, the presented reflection on the valorisation of the LOV identified several key factors critical to a successful implementation. By focusing on strong partnerships, effective dissemination strategies, and efforts to demonstrate value, the LOV can become a valuable tool for dairy farmers, contributing to improved health, productivity, and economic sustainability in the dairy industry.

Chapter 9.

General conclusion and future perspectives

9.1 Thesis overview

This thesis aims to leverage the increasing automation and resulting data on dairy farms to tackle one of the major challenges on those farms: managing the transition period effectively. Turning to the image presented in [Chapter 3], an adapted version is shown in Figure 3.1. This adaptation highlights how the backbone of this thesis ties in the proposed PLF framework, with the various chapters collectively contributing towards enhancing control over this critical period. In summary, the first steps in this endeavour were presented in [Chapter 4], where the nextMILK_{TD} model was presented. This model utilises the DHI records and information from the previous lactation (production, herd, reproduction) to predict the milk yield on the first test day of the next lactation. Next, a monitoring metric was introduced in [Chapter 5] using the milk yield residuals in the transition period (MRT) as obtained from the nextMILK models. Furthermore, this newly defined metric was associated with health [Chapter 5], behaviour

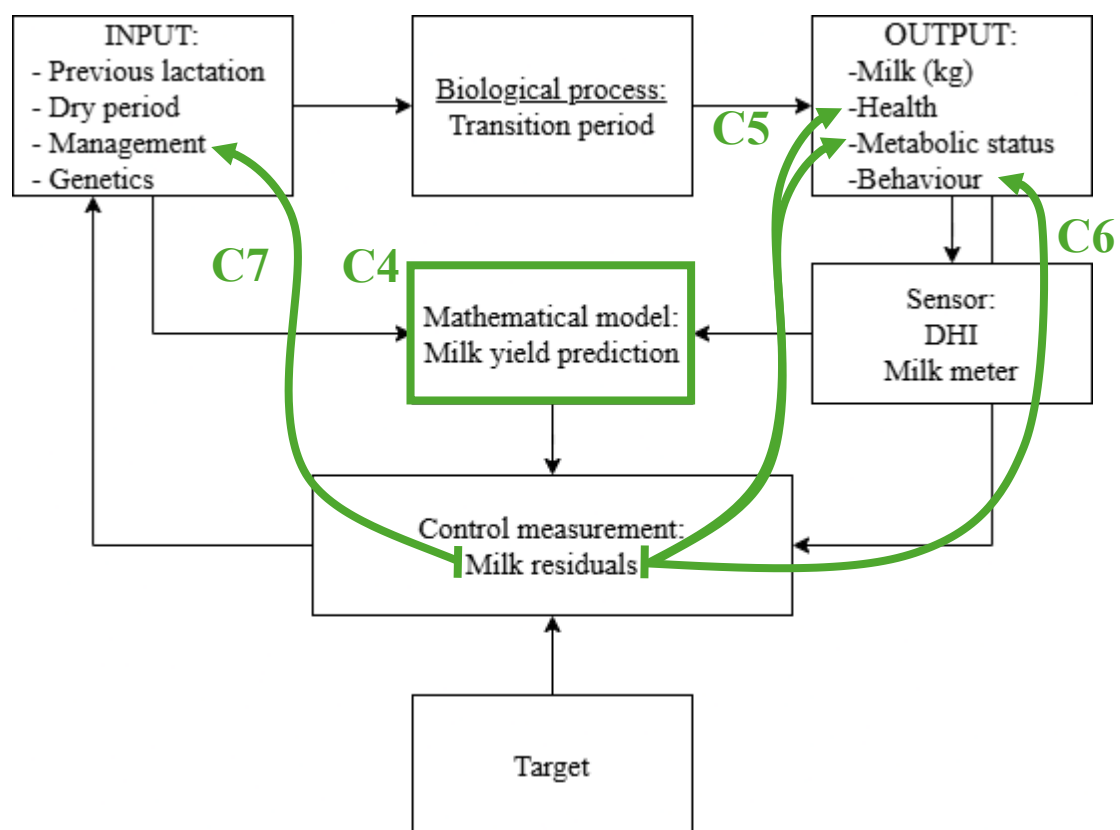


Figure 9.1 The transition period fitted within the PLF framework. This graphically represents the approach used in this thesis. The green arrows and the according annotations provide an overview of the following chapters and highlight how they tie in the transition period framework. Adapted from Wathes et al. (2008).

[Chapter 6] and management [Chapter 7] related to transition success at different levels. Specifically, in [Chapter 5], the MRT was linked to the general health performance of individual dairy cows and to metabolic parameters measured in a field trial. Next, in [Chapter 6] the

associations between behavioural changes in the transition period and the MRT were investigated. As a final validation step, the MRT at herd level was linked to farm management in [Chapter 7]. By exhaustively validating the MRT, a scientific basis for using MRT as a metric was created that ensures and substantiates its use to monitor transition success or failure, with a grasp on how it relates to the current state-of-the-art monitoring tools and scientific knowledge. The introduction of the LOV in [Chapter 8] as the potential field-ready version of the MRT also allowed for the identification of remaining efforts necessary to bring this metric to the field outside of the scientific scope.

9.2 *General conclusion*

9.2.1 *The nextMILK models the milk yield prediction for the transition period*

Central to the PLF framework and this thesis has been the ability to predict milk yield in the transition period, unaffected by possible pathophysiological processes altering milk production at the start of the lactation. The latter aspect was crucial, as it allowed to estimate the deviation when pathophysiological processes affecting milk production are present. This resulted in the nextMILK models based on test-day DHI data presented in [Chapter 4]. In parallel, models based on high-frequency milk meter data (Liseune et al., 2021) were developed by Liseune et al. (2021). Despite employing different modelling methods, these models achieved relatively similar performances in predicting milk yield at the start of lactation using historical data from the previous lactation, with RMSE of approximately 6 kg and 7 kg for nextMILK_{TD} (Salamone et al., 2022) and nextMILK_{MM} (Liseune et al., 2021) respectively. In the absence of similar peer-reviewed models, comparing model performance with the existing corpus of milk yield prediction models is challenging. Although prediction performances of lactation curve models, such as the wood model (Wood, 1967) have been found to be around 4kg RMSE per DIM over the full lactation (Grzesiak et al., 2006; Adriaens et al., 2021a), the inherent difference in the prediction window between the nextMILK models (modelled at the onset of the lactation) and these traditional models (modelled throughout the entire lactation and often using current-lactation data) doesn't allow a fair comparison. Therefore, both models were compared with a benchmark based on herd average lactation curves or individual productions transposed from the previous lactation. These comparisons showed that both nextMILK models performed better in terms of RMSE than the respective benchmark models.

9.2.2 *Addressing the Limitations of Traditional KPIs*

The common practice in milk yield recording inspired the definition of these benchmark models, where standard lactation curves based on population, breed, farm characteristics and physiology are used to interpolate milk productions until the end of lactation (Wilmink, 1987; Schaeffer and Jamrozik, 1996; Vanraden, 1997). These standard lactation curves are used to standardise milk production to a fixed time point in the life cycle of the individual animal. The resulting KPI serves to compare lactation performance, independently of the environment and production stage, e.g., the 305-age average used by the Council on Dairy Cattle Breeding in the US (Bowie MD, US) or standard peak production (SPP) used by CRV in Flanders and the Netherlands. For SPP, the milk yield is multiplied by a correction factor that transposes the production in any lactation at any DIM to the production at 50 DIM of a 3-year-old cow who calved in February/March, similar to the methodology presented by Wilmink (1987). Although these transformations could be applied on the first TD and naively serve as a monitoring metric of the transition period, mathematically, these transformations are the result of the milk production on that TD multiplied with a correction factor. Since the SPP (and similar standardization measures), especially in the transition period, rely on records potentially affected by an altered physiological or pathological state (e.g. udder health problem at the start of lactation), they were deemed unsuitable for addressing the issues raised in this thesis. Despite the perceived benefits that could be found in a comparative investigation of these regional standard KPIs and the MRT, we were unable to access updated documentation, rendering this exercise exceedingly difficult. Even in a collaborative effort with the organisation behind these KPI, undocumented changes in methodologies, such as the one performed in 2020 for the SPP (CRV), would have led to unreliable conclusions.

In the realm of dairy KPI, the Transition Cow Index (TCI) developed by Nordlund (2006) conceptually relates closely to the MRT. Similar to the MRT_{TD} , it attempts to predict the production of the first TD of the next lactation using data from the previous lactation. This prediction is made using a linear mixed model with the regression coefficients published in the associated patent (Nordlund et al., 2011). Nevertheless, the TCI was considered unsuited for benchmarking of the nextMILK models in this thesis. For instance, this KPI was originally developed for the US and fitted on US-based dairy herds only. Including bovine somatotropin in the linear regression, which is forbidden in Europe, for example, highlights this highly regional character of the TCI. This regional aspect combined here again with little to no updated documentation on this methodology accentuated the necessity to develop a new model.

9.2.3 *The importance of openness*

Developing the nextMILK models instead of relying on the existing KPI allowed us to fine-tune these models to perfectly fit our needs, with the added advantage of full ownership over these models. This allowed for complete access to the models, rendering insights into how their predictions are formulated and flexibility in incorporating features, i.e. the different versions of nextMILK_{TD}. It also reflects a conscious choice within all chapters of this thesis to contribute to an open science effort, where openness was deliberately maximised by providing open access to the statistical code. While the data used in the different studies cannot be shared publicly due to privacy and data protection restrictions, the openness to share the data from the different studies of this thesis for research purposes through bilateral agreements is present. Similarly, whereas the models themselves could not be openly distributed due to their technical immaturity (lack of documentation, lack of bug testing...), the possibility of sharing these models with any researcher without needing approval from external partners further underscores the open nature of our research. Outside of this academic context, the situation could prove to be more complex, where the involvement of a whole consortium of researchers from different faculties and universities could be seen as a potential hurdle. Yet, the preliminary contractual work defined within the VLAIO-LA project also provides a backbone to sort forthcoming issues in this regard.

9.2.4 *Data on dairy farms encountered challenges*

Throughout the different studies, the complexity of working with dairy data was a persistent challenge. The lack of standardisation, quality and consistency were issues encountered in the transformation from the raw data towards clear and useful results, as presented in the different chapters. One of the major issues encountered was that, despite the current views on automation and digitalisation as sources of reliable and clear data, data regarding disease registration was missing or inconsistent. This is a known limitation in dairy research, and efforts to quantify and alleviate this limitation are actively investigated (Hermans et al., unpublished data). This lack of quality outside of experimental set-ups also impacts the registration of routinely collected cow-side diagnostics (e.g. BHB, serum calcium tests). This proved to be quite challenging in a period where success or failure is exclusively characterised by these cow-side diagnostics (metabolic status) or human determination of disease (health status). Except for the study presented in [Chapter 5], which by nature presented high-quality records, this hurdle was present across all the other studies. In the development of the nextMILK [Chapter 4], the

absence of good health data could be circumvented by fitting traditional lactation curves and identifying and removing deviating lactations. However, this strategy was unsuitable for [Chapter 6] and [Chapter 7], as associating MRT with potentially invalid interpolated health status would not advance their aims. Nevertheless, the availability of well-documented disease records would have allowed a better understanding of the dynamics at play between MRT and the transition period. Regarding the standardisation and access to dairy data, or the lack thereof, this was primarily present in the association of MRT with management practices ([Chapter 7]). Although weaker associations found in that study are not entirely attributable to a lack of standardisation. 27% of the surveyed farms dropped out because of a lack of TD data. This highlights once more the difficulties encountered regarding data acquisition, even though third-party data management software providers were involved in this research to mitigate the difficulties in accessing data. Nevertheless, the plethora of software, software versions and closed-source digital environments on farms gathering different types of data and their lack of interoperability hampered the efforts presented in that study.

9.2.5 A plea for FAIRness

Although not mainstream at the start of this PhD, the emergence of the FAIR principles in the mid-2010s are crucial in understanding the challenges mentioned above (Wilkinson, 2016). These principles were formulated to improve the Findability, Accessibility, Interoperability, and Reuse of digital assets in research. The key elements have been indirectly elaborated on throughout this thesis, underlining the compliance of our research with these principles but also its limitations. In these last chapters, the contrast with existing practices by industry leaders highlights the conceptual mismatch between the presented scientific approach and many mainstream tools. If dairy data research were not solely reliant on commercial partners, adherence to the FAIRness principle would potentially be less of an issue. Though like a relay race where the slowest element could define the outcome for the whole team, the result of research efforts is determined by the least compliant involved actor. Mitigations in these research efforts are sometimes possible but require additional steps, which bring their own limitations. In conclusion, an industry-wide change is necessary to exploit the full potential of data in the dairy industry. The identified needs regarding data in this thesis focus mostly on the need for conceptual improvements to facilitate the recording of diagnostics, diseases and treatments on the farm, openness and common standards regarding any type of data recording on the farm. A potential driver for these necessary changes could come from scientists working with the major industry partners, by transparently and by name addressing the FAIRness of

their collaborations. This could incentivise companies to profile themselves as innovative and close to research, which could motivate others to adopt more FAIR approaches in their digital assets.

9.2.6 The MRT as a new window in the transition period?

Derived from these nextMILK models, the milk yield residuals during the transition period, or the MRT, were identified as the metric of interest given its association with the transition period. This metric is obtained by subtracting the predicted from the realised value, often called the residual. Moreover, these residuals also present a straightforward interpretation: negative residuals denote that the actual milk yield was lower than predicted, indicating the possible presence of a process that negatively impacts milk production. As there is no consensus on a clear, common definition of transition success, this thesis focused on a combination of indicators defining the success of transition. Moreover, studies at the metabolomic, behavioural and management levels all confirmed MRT is a suitable proxy to identify a failing transition. Throughout the different chapters, the MRT was associated with the characterising elements of the transition period: health, DMI, biomarkers and behaviour. These associations were also consistent over the different approaches investigated in this thesis. MRT was associated with the health of the animals directly in [Chapter 5] and indirectly through their behaviour in [Chapter 6]. A strong link between MRT and DMI was consistently observed across [Chapter 5] and [Chapter 6]. These results highlighted the potential close relationship of MRT with the NEB in the transition period, which is caused by a disbalance between energy demands for milk production and energy supply through DMI. The biomarkers in the inference models of [Chapter 5] showed that the effects of key metabolites (IGF1, NEFA, insulin, glucose) on MRT could be attributed to their role in the energy balance. In summary, these results across different studies all point towards the concrete potential of MRT to be used as a monitoring tool specifically tailored to evaluate the success of the transition period. Nevertheless, the conversion of this potential into a cow-side decision tool or herd monitoring tool will still require further studies, as discussed below.

9.3 Future perspectives

This thesis opens the path towards a novel automated approach to monitoring the transition period. The different steps that were taken to develop this novel methodology based on empirical data provide a robust basis on which future efforts could rest. Further efforts should focus on validation of this methodology, both in expanding the association of the MRT with

new physiological insights into the transition period and concurrently challenging the MRT as a practical decision-making tool.

9.3.1 Generalising the scope of associations

In its current state, the MRT has been associated with the key elements of the transition period: health, metabolic profiles and behaviour. To confirm these associations, a series of studies could re-iterate these set-ups, which is necessary to generalise our findings to other farms. Across these different re-iterations, the associations found between health, metabolic profile, and MRT should be consistent to confirm the results of this thesis. In an exploratory step, the experimental set-up such as the one used in chapter [Chapter 5] was proven useful, where a low number of animals were enrolled. Still, a broad panel of both health and metabolic parameters was investigated. Yet now that the candidate associations have been identified, the following steps should primarily focus on confirming these candidate associations and possibly broadening them in their different facets, e.g. investigation of potential differences in disease types (infectious, metabolic...) with regard to their association with MRT or markers of NEB in milk fatty acids profiles and their relationship with the MRT.

The associations of health and MRT could be reinvestigated on a small group (e.g. 4) of large farms (> 1000 lactating cows), with an emphasis on qualitative disease registration. These farms present the advantage of having a high number of calvings in shorter time periods, thus a higher number of computed MRT, i.e., farms with 1000 animals have approximately 20 calvings a week. Additionally, these farms heavily rely on hired labour, which is organised around well-defined standard operating procedures, such as caring for calves and young stock, overseeing calvings, managing fresh cows, managing lactating cows, feeding, and breeding. The different people in charge of dry and fresh cows could be consulted to create a uniform and well-defined protocol that would allow a complete and qualitative recording of disease events on their farms. In the best case, these protocols would be straightforward so that data could be generated over longer periods of time. Moreover, an envisaged window of one year would constitute a valuable starting point as it allows to correct for seasonal effects. A potential approach to improve the quality of these disease records could lie in the increased digitalisation of dairy veterinary practices. An increasing number of veterinary practices rely on specialised programs to manage their drug inventories, accounting, and report antibiotic use with dedicated software. Further studies should investigate if these veterinary records could also serve as a data source to confirm or replace farmers' records. While the increased amount of calvings on the larger farms and,

consequently, the higher number of diseased animals would further confirm the associations found in [Chapter 5], it could allow differentiation of the dynamics associated with different disease processes, i.e., clinical hypocalcaemia vs NEB associated diseases. Complementary to this, by selecting large farms of which the animals are equipped with commercial behavioural sensors, the behavioural association study could also be reinvestigated and linked with health data, which were missing in [Chapter 6]. The longitudinal aspect could also be used to reassess the association with management practices on these larger farms over time to circumvent the limitations encountered in [Chapter 7]. A final advantage of focusing on these large herds can be found in their strong reliance on genomic testing to evaluate the genetic potential of their herd. This routinely collected data may open up an entirely new facet of the production process and can be taken into account when evaluating the MRT. These genomic values could also be integrated as a correction factor for the MRT, allowing an evaluation of the transition corrected for the genetic merit of the animals. This could further improve the genetic resilience of dairy herds, allowing the differentiation of animals with high genetic merit for milk, which completely failed to transition from animals with lower genetic drive to produce milk.

9.3.2 *A decision-making tool*

Alongside these future possibilities to broaden the scope of the MRT's associations, future work should also focus on the MRT's capabilities to serve as a monitoring tool and, by extension, a decision-making tool. Although this goal motivated the different steps taken in this thesis, the concrete translation of the results into a tool will require more work. The other variations of MRT open different pathways to investigate this role. Starting with the MRT based on milk meter data, the flexibility inherently present in this MRT_{MM}, allows for investigating various transformations. Currently calculated as the mean over the 21 first days of the lactation, it could also be approached as a daily prediction of the milk yield. A statistical process control step could be applied to these predictions to detect perturbations in the daily milk yields, similar to the approach proposed in Huybrechts et al.(2014) . This, in turn, could allow the early detection of diseased animals or individuals that experience metabolic challenges. Quantifying the sensitivity and specificity of these future techniques will represent an essential step to confirm the strengths of the MRT as a monitoring tool for individual animals.

In the case of the MRT_{TD}, whose calculation is based on DHI data, day-to-day individual monitoring of animals in the current DHI system (measurement every 4 to 8 weeks) will be hampered. A cow calving the day after a TD will have crossed the entire transition period before

her next TD. However, it could still be used to identify broader dynamics occurring on the farm level to detect and monitor feed formulation changes, increased infection risks, etc. The possibility of transforming the MRT_{TD} into a more individual monitoring tool would require a change in how the DHI programs sample farms. In a reality, where the MRT has proven to be an effective metric to assist cow-side decisions, altering the sampling frequency specifically for freshly calved cows could be envisaged. This would require a small revolution to become reality. Nevertheless, the number of farms solely relying on DHI data will diminish in the foreseeable future, increasing the potential of MRT_{MM} instead of relying on the MRT_{TD} primarily.

9.3.3 The MRT as the path forward in the future challenges

In Belgium, the SPP necessitated 30 years to grow from its initial formulation to the mainstream tool it is today, demonstrating the long road to the adoption of new data-based technologies in the dairy industry. This thesis provides a roadmap for the MRT, built on a systemic process engineering approach, thorough validation, and a comprehensive valorisation plan. The central role of the joint nature of the doctoral studies behind this thesis should not be overlooked. The dual nature of this PhD, characterised by the veterinary and engineering disciplines converged and synergised to yield results greater than the sum of their parts. Ultimately, the approach to develop and further improve the MRT could serve as a concrete blueprint for alternate changes and improving sustainability of the dairy industry, particularly within the digital dairy farm environment, the new final frontier.

References

- Abdelkrim, A.B., L. Puillet, P. Gomes, and O. Martin. 2021a. 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.
- Abdelkrim, A.B., T. Tribout, O. Martin, D. Boichard, V. Ducrocq, and N.C. Friggens. 2021b. Exploring simultaneous perturbation profiles in milk yield and body weight reveals a diversity of animal responses and new opportunities to identify resilience proxies. *J. Dairy Sci.* 104:459–470. doi:10.3168/jds.2020-18537.
- Abuelo, A., L. Wisnieski, J.L. Brown, and L.M. Sordillo. 2021. Rumination time around dry-off relative to the development of diseases in early-lactation cows. *J. Dairy Sci.* 104:5909–5920. doi:10.3168/JDS.2020-19782.
- Adriaens, I., I. van den Brulle, L. D’Anvers, J.M.E. Statham, K. Geerinckx, S. De Vliegheer, S. Piepers, and B. Aernouts. 2021a. 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.
- Adriaens, I., I. Van Den Brulle, K. Geerinckx, L. D’Anvers, S. De Vliegheer, and B. Aernouts. 2021b. Milk losses linked to mastitis treatments at dairy farms with automatic milking systems. *Prev. Vet. Med.* 194:105420. doi:10.1016/J.PREVETMED.2021.105420.
- Adriaens, I., N.C. Friggens, W. Ouweltjes, H. Scott, B. Aernouts, and J. Statham. 2020. Productive life span and resilience rank can be predicted from on-farm first-parity sensor time series but not using a common equation across farms. *J. Dairy Sci.* 103:7155–7171.
- Adriaens, I., T. Huybrechts, B. Aernouts, K. Geerinckx, S. Piepers, B. De Ketelaere, W. Saeys, B. De Ketelaere, and W. Saeys. 2018. Method for short-term prediction of milk yield at the quarter level to improve udder health monitoring. *J. Dairy Sci.* 101:10327–10336. doi:10.3168/JDS.2018-14696.
- Aernouts, B., and I. Adriaens. 2023. *Robotic Technologies for Dairy Farming*. Springer International Publishing.
- Ahmed, H., L. Ekman, and N. Lind. 2024. Planned behavior, social networks, and perceived risks: Understanding farmers’ behavior toward precision dairy technologies. *J. Dairy Sci.* 107:2968–2982. doi:10.3168/jds.2023-23861.
- Åkerström, F., C. Bergsten, and A.-K. Nyman. 2024. Association between hoof trimming frequency in primiparous cows and hoof health and survival in second lactation. *J. Dairy Sci.* 107:9676–9693. doi:10.3168/jds.2023-24543.
- Aleri, J.W., B.C. Hine, M.F. Pyman, P.D. Mansell, W.J. Wales, B. Mallard, and A.D. Fisher. 2016. Periparturient immunosuppression and strategies to improve dairy cow health during the periparturient period. doi:10.1016/j.rvsc.2016.07.007.
- Ali, T.E., and L.R. Schaeffer. 1987. Accounting for covariances among test day milk yields in dairy cows. *Can. J. Anim. Sci.* 67:637–644. doi:10.4141/cjas87-067.
- Allen, M.R., O.P. Dube, W. Solecki, F. Aragón-Durand, W. Cramer, S. Humphreys, M. Kainuma, J. Kala, N. Mahowald, Y. Mulugetta, R. Perez, M. Wairiu, and K. Zickfeld. 2018. *Framing and Context*. Cambridge. Cambridge University Press, Cambridge, UK.
- Aquilani, C., A. Confessore, R. Bozzi, F. Sirtori, and C. Pugliese. 2022. Review: Precision Livestock Farming technologies in pasture-based livestock systems. *Animal* 16:100429. doi:10.1016/J.ANIMAL.2021.100429.
- Aschenbach, J.R., H.M. Hammon, G.B. Penner, N.B. Kristensen, S.S. Donkin, H.M. Hammon, and G.B. Penner. 2010. Gluconeogenesis in dairy cows: The secret of making sweet milk from sour dough. *IUBMB Life* 62:869–877. doi:10.1002/IUB.400.
- Atashi, H., M. Salavati, J. De Koster, M.A. Crowe, G. Opsomer, M. Hostens, M. Niamh, F. Alan, C. Fiona, M. Elizabeth, S. Andreia, B. Colin, R. Pauline, O. Roisin, H. Sinead, W. Claire, C. Zhangrui, F. Ali, P. Geoff, W. Dirk, B.S. Beatriz Sanz, F. Conrad, W. Alistair, B. Matt, V. Mieke, H. Kristof, M. Sander, B. Hannes, V. Jan, V. Leila, V. Bonny, I. Klaus, S.T. Martin Tang, H. Johanna, D. Susanne, O. Soren, R. Janne, K. Mogens, M. Else, F. Leslie, G. Charlotte, E. Jehan, R. Tine, L. Torben, de O. de, M. Cinzia, S. Federica, N. Francesco, M. Bianca, C. Alessandra, B. Luca, M. Jennifer, M. Daragh, K. Francis, C. Andrew, M. Matt, Z. Shujun, C. Xing, C. Huanchun, Z. Junlong, Y. Liguohua, H. Guohua, T. Chen, W. Guiqiang, B. Michel, S. Marlène, P. Armin, E. Arnold, K. Linda, F. Anders, A. Thomas, L. Matthew, E. Chris, C. Gavin, T. Jerry, T. Deborah, G. Nicolas,

- G. Michel, C. Frederic, P.R. Marilou Ramos, H. Hedi, B. Catherine, T. Haruko, L. Aurelie, V.L. Anne-Sophie, M. Rodrigo, D.N. Saied Naderi, D. Frederic, G. Clement, V. Amelie, F. Eric, B. Frank, S. Martin, and V.P. Sergio Palma. 2020. Genome-wide association for metabolic clusters in early-lactation Holstein dairy cows. *J. Dairy Sci.* 103:6392–6406. doi:10.3168/JDS.2019-17369.
- AWE. Rapport Activités 10 Ans. Accessed March 11, 2025. https://www.awenet.be/awe/UserFiles/file/RapportActivites/BD-AWE_rapport_activites_10_ans-FR.pdf.
- Barkema, H.W., M.A.G. von Keyserlingk, J.P. Kastelic, T.J.G.M. Lam, C. Luby, J.P. Roy, S.J. LeBlanc, G.P. Keefe, and D.F. Kelton. 2015. Invited review: Changes in the dairy industry affecting dairy cattle health and welfare. *J. Dairy Sci.* 98:7426–7445.
- Barragan, A.A., J.M. Piñeiro, G.M. Schuenemann, P.J. Rajala-Schultz, D.E. Sanders, J. Lakritz, and S. Bas. 2018. Assessment of daily activity patterns and biomarkers of pain, inflammation, and stress in lactating dairy cows diagnosed with clinical metritis. *J. Dairy Sci.* 101:8248–8258. doi:10.3168/jds.2018-14510.
- 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.
- Bauman, D.E., and W. Bruce Currie. 1980. Partitioning of Nutrients During Pregnancy and Lactation: A Review of Mechanisms Involving Homeostasis and Homeorhesis. *J. Dairy Sci.* 63:1514–1529. doi:10.3168/jds.S0022-0302(80)83111-0.
- Beauchemin, K.A. 2018. Invited review: Current perspectives on eating and rumination activity in dairy cows. *J. Dairy Sci.* 101:4762–4784. doi:10.3168/JDS.2017-13706.
- Bell, A.W. 1995. Regulation of organic nutrient metabolism during transition from late pregnancy to early lactation. *J. Anim. Sci.* 73:2804. doi:10.2527/1995.7392804x.
- Bell, A.W., and D.E. Bauman. 1997. Adaptations of Glucose Metabolism During Pregnancy and Lactation.
- Benjamini, Y., and D. Yekutieli. 2001. The Control of the False Discovery Rate in Multiple Testing under Dependency. *Ann. Stat.* 29:1165–1188.
- Berckmans, D. 2017. General introduction to precision livestock farming. *Anim. Front.* 7:6–11. doi:10.2527/AF.2017.0102.
- Bertoni, G., E. Trevisi, and R. Lombardelli. 2009. Some new aspects of nutrition, health conditions and fertility of intensively reared dairy cows Giuseppe. *Ital. J. Anim. Sci.* 8:491–518. doi:10.4081/ijas.2009.491.
- Bianchi, M.C., L. Bava, A. Sandrucci, F.M. Tangorra, A. Tamburini, G. Gislou, and M. Zucali. 2022. Diffusion of precision livestock farming technologies in dairy cattle farms. *Animal* 16. doi:10.1016/j.animal.2022.100650.
- Borchers, M.R., and J.M. Bewley. 2015. An assessment of producer precision dairy farming technology use, prepurchase considerations, and usefulness. *J. Dairy Sci.* 98:4198–4205. doi:10.3168/JDS.2014-8963.
- Borchers, M.R., Y.M. Chang, K.L. Proudfoot, B.A. Wadsworth, A.E. Stone, and J.M. Bewley. 2017. Machine-learning-based calving prediction from activity, lying, and ruminating behaviors in dairy cattle. *J. Dairy Sci.* 100:5664–5674. doi:10.3168/jds.2016-11526.
- Borchers, M.R., S. Gavigan, A. Harbers, and J. Bewley. 2021. An evaluation of a novel device for measuring eating, rumination, and inactive behaviors in lactating Holstein dairy cattle. *Animal* 15:100008. doi:10.1016/J.ANIMAL.2020.100008.
- Bouallegue, M., and N. M’Hamdi. 2020. Mathematical Modeling of Lactation Curves: A Review of Parametric Models. IntechOpen.
- Breiman, L. 2001. *Random Forests*. Springer.
- Britt, J.H., R.A. Cushman, C.D. Dechow, H. Dobson, P. Humblot, M.F. Hutjens, G.A. Jones, P.S. Ruegg, I.M. Sheldon, and J.S. Stevenson. 2018. Invited review: Learning from the future—A vision for dairy farms and cows in 2067. *J. Dairy Sci.* 101:3722–3741. doi:10.3168/jds.2017-14025.
- Brody, S., C.A. Ragsdale, and W.C. Turner. 1923. The rate of decline of milk secretion with the advance of the period of lactation. *J. Gen. Physiol.* 5:441–444.
- 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.
- Brown, W.E., and B.J. Bradford. 2021. Invited review: Mechanisms of hypophagia during disease. *J. Dairy Sci.* 104:9418–9436. doi:10.3168/JDS.2021-20217.

- 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.
- Calamari, L., N. Soriani, G. Panella, F. Petrer, A. Minuti, and E. Trevisi. 2014. Rumination time around calving: An early signal to detect cows at greater risk of disease. *J. Dairy Sci.* 97:3635–3647. doi:10.3168/JDS.2013-7709.
- Van Calker, K.J., P.B.M. Berentsen, C. Romero, G.W.J. Giesen, and R.B.M. Huirne. 2005. Development and application of a multi-attribute sustainability function for Dutch dairy farming systems. doi:10.1016/j.ecolecon.2005.05.016.
- Cameron, A., and T. van den Brand. 2022. geomtextpath: Use Curved Text in Ggplot2.
- Cardinale, B.J., J.E. Duffy, A. Gonzalez, D.U. Hooper, C. Perrings, P. Venail, A. Narwani, G.M. MacE, D. Tilman, D.A. Wardle, A.P. Kinzig, G.C. Daily, M. Loreau, J.B. Grace, A. Larigauderie, D.S. Srivastava, and S. Naeem. 2012. Biodiversity loss and its impact on humanity. *Nat.* 2012 4867401 486:59–67. doi:10.1038/nature11148.
- Caré, S., E. Trevisi, A. Minuti, A. Ferrari, J.J. Looor, and L. Calamari. 2018. Plasma fructosamine during the transition period and its relationship with energy metabolism and in fl ammation biomarkers in dairy cows. *Livest. Sci.* 216:138–147. doi:10.1016/j.livsci.2018.08.003.
- Cattaneo, L., V. Lopreiato, E. Trevisi, and A. Minuti. 2020. Association of postpartum uterine diseases with lying time and metabolic profiles of multiparous Holstein dairy cows in the transition period. *Vet. J.* 263:105533. doi:10.1016/J.TVJL.2020.105533.
- Chanvallon, A., S. Coyral-Castel, J. Gatien, J.M. Lamy, D. Ribaud, C. Allain, P. Clément, and P. Salvetti. 2014. Comparison of three devices for the automated detection of estrus in dairy cows. *Theriogenology* 82:734–741. doi:10.1016/J.THERIOGENOLOGY.2014.06.010.
- 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.
- Codrea, M.C., S. Højsgaard, and N.C. Friggens. 2011. Differential smoothing of time-series measurements to identify disturbances in performance and quantify animal response characteristics: An example using milk yield profiles in dairy cows. *J. Anim. Sci.* 89:3089–3098. doi:10.2527/JAS.2010-3753.
- Cole, J.B., J.L.L. Ehrlich, and D.J.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., B.O. Mankanjuola, C.M. Rochus, N. van Staaveren, and C. Baes. 2023. The effects of breeding and selection on lactation in dairy cattle. *Anim. Front. Rev. Mag. Anim. Agric.* 13:62. doi:10.1093/AF/VFAD044.
- Congleton, W.R., and R.W. Everett. 1980. Error and Bias in Using the Incomplete Gamma Function to Describe Lactation Curves. *J. Dairy Sci.* 63:101–108. doi:10.3168/jds.S0022-0302(80)82894-3.
- Cook, N.B. 2020. Symposium review: The impact of management and facilities on cow culling rates. *J. Dairy Sci.* 103:3846–3855. doi:10.3168/JDS.2019-17140.
- Costa, J.H.C., M.C. Cantor, and H.W. Neave. 2021. Symposium review: Precision technologies for dairy calves and management applications *. *J. Dairy Sci.* 104:1203–1219. doi:10.3168/JDS.2019-17885.
- Couto Serrenho, R., C. Church, D. McGee, and T.F. Duffield. 2022. Environment, nutrition, and management practices for far-off, close-up, and fresh cows on Canadian dairy farms—A retrospective descriptive study. *J. Dairy Sci.* 105:1797–1814. doi:10.3168/jds.2021-20919.
- Couto Serrenho, R., C. Church, D. McGee, and T.F. Duffield. 2023. Association of herd hyperketolactia prevalence with transition management practices and herd productivity on Canadian dairy farms— A retrospective cross-sectional study. *J. Dairy Sci.* 106:2819–2829. doi:10.3168/jds.2022-22377.
- Couto Serrenho, R., T.J. DeVries, T.F. Duffield, and S.J. LeBlanc. 2021. Graduate Student Literature Review: What do we know about the effects of clinical and subclinical hypocalcemia on health and performance of dairy cows?. *J. Dairy Sci.* 104:6304–6326. doi:10.3168/JDS.2020-19371.
- CRV. CRV Actualiseert Berekening Kerngetallen NO, ISK, LM En BSK - CRV. Accessed March 11, 2025. <https://crv4all.nl/nl/news/crv-actualiseert-berekening-kerngetallen>.
- CRV. 2018. Genetische trends van koeien in Nederland 199.
- CRV. 2024. Breeding value Resilience.
- Cyples, J.A., C.E. Fitzpatrick, K.E. Leslie, T.J. DeVries, D.B. Haley, and N. Chapinal. 2012. Short communication: The effects of experimentally induced *Escherichia coli* clinical mastitis on lying behavior of dairy cows. *J. Dairy Sci.* 95:2571–2575. doi:10.3168/jds.2011-5135.

- Dallago, G.M., D.M. de Figueiredo, P.C. de Resende Andrade, R.A. dos Santos, R. Lacroix, D.E. Santschi, D.M. Lefebvre, D.M. de Figueiredo, P.C. de R. Andrade, R.A. dos Santos, R. Lacroix, D.E. Santschi, and D.M. Lefebvre. 2019. Predicting first test day milk yield of dairy heifers. *Comput. Electron. Agric.* 166:105032. doi:10.1016/j.compag.2019.105032.
- Deluyker, H.A., J.M. Gay, L.D. Weaver, and A.S. Azari. 1991. Change of Milk Yield with Clinical Diseases for a High Producing Dairy Herd. *J. Dairy Sci.* 74:436–445. doi:10.3168/JDS.S0022-0302(91)78189-7.
- Dematawewa, C.M.B.M.B., R.E.E. Pearson, and P.M.M. Vanraden. 2007. Modeling extended lactations of Holsteins.. *J. Dairy Sci.* 90:3924–36. doi:10.3168/jds.2006-790.
- DeVries, T.J., F. Dohme, and K.A. Beauchemin. 2008. Repeated Ruminant Acidosis Challenges in Lactating Dairy Cows at High and Low Risk for Developing Acidosis: Feed Sorting. *J. Dairy Sci.* 91:3958–3967. doi:10.3168/JDS.2008-1347.
- Dhanao, M.S. 1981. A note on an alternative form of the lactation model of Wood. *Anim. Prod.* 32:349–351. doi:10.1017/S0003356100027276.
- Dijkstra, J., J. France, M.S. Dhanoa, J.A. Maas, M.D. Hanigan, A.J. Rook, and D.E. Beever. 1997. A Model to Describe Growth Patterns of the Mammary Gland during Pregnancy and Lactation. *J. Dairy Sci.* 80:2340–2354. doi:10.3168/jds.S0022-0302(97)76185-X.
- Dijkstra, J., S. Lopez, A. Bannink, M.S. Dhanoa, E. Kebreab, N.E. Odongo, M.H. Fathi Nasri, U.K. Behera, D. Hernandez-Ferrer, and J. France. 2010. Evaluation of a mechanistic lactation model using cow, goat and sheep data. *J. Agric. Sci.* 148:249–262. doi:10.1017/S0021859609990578.
- Drackley, J.K. 1999. Biology of Dairy Cows During the Transition Period: the Final Frontier?. *J. Dairy Sci.* 82:2259–2273. doi:10.3168/JDS.S0022-0302(99)75474-3.
- Drackley, J.K., H.M. Dann, G.N. Douglas, N.A. Janovick Guretzky, N.B. Litherland, J.P. Underwood, and J.J. Looor. 2005. Physiological and pathological adaptations in dairy cows that may increase susceptibility to periparturient diseases and disorders. *Ital. J. Anim. Sci.* 4:323–344. doi:10.4081/IJAS.2005.323.
- Duffield, T.F., S. LeBlanc, R. Bagg, K. Leslie, J. Ten Hag, and P. Dick. 2003. Effect of a Monensin Controlled Release Capsule on Metabolic Parameters in Transition Dairy Cows. *J. Dairy Sci.* 86:1171–1176. doi:10.3168/JDS.S0022-0302(03)73700-X.
- Eckelkamp, E.A. 2019. INVITED REVIEW: Current state of wearable precision dairy technologies in disease detection. *Appl. Anim. Sci.* 35:209–220. doi:10.15232/aas.2018-01801.
- Ehrlich, J.L. 2011. Quantifying shape of lactation curves , and benchmark curves for common dairy breeds and parities. *Bov. Pract.* 45:88–96.
- Emam, M.H., E. Shepley, M.M. Mahmoud, M. Ruch, S. Elmaghawry, W. Abdelrazik, A.M. Abdelaal, B.A. Crooker, and L.S. Caixeta. 2023. The Association between Prepartum Rumination Time, Activity and Dry Matter Intake and Subclinical Hypocalcemia and Hypomagnesemia in the First 3 Days Postpartum in Holstein Dairy Cows. *Animals* 13:1621. doi:10.3390/ANI13101621/S1.
- Erb, H.N., and Y.T. Grohn. 1988. Epidemiology of Metabolic Disorders in the Periparturient Dairy Cow. *J. Dairy Sci.* 71:2557–2571. doi:10.3168/JDS.S0022-0302(88)79845-8.
- Esposito, G., P.C. Irons, E.C. Webb, and A. Chapwanya. 2014. Interactions between negative energy balance, metabolic diseases, uterine health and immune response in transition dairy cows. *Anim. Reprod. Sci.* 144:60–71. doi:10.1016/J.ANIREPROSCI.2013.11.007.
- European Commission. DG AGRI - Monthly Market Prices. Accessed January 20, 2025. <https://agridata.ec.europa.eu/extensions/DashboardPrice/DashboardMarketPrices.html>.
- European Commission. 2024. Agri-Food Data Portal. Accessed January 20, 2025. <https://agridata.ec.europa.eu/extensions/DataPortal/home.html>.
- Eurostat. 2024. Main Agricultural Labour Force by Age and Sex of the Farm Manager, Farm Type of the Agricultural Holding and NUTS 2 Regions. Accessed January 14, 2025. https://ec.europa.eu/eurostat/databrowser/view/ef_lf_main/default/table?lang=en&category=agr.ef.lf.
- FAO. 2022a. World Food and Agriculture - Statistical Pocketbook 2022. FAO, Rome.
- FAO. 2022b. The State of Food and Agriculture 2022. Leveraging automation in agriculture for transforming agrifood systems. FAO, Rome.
- Federale Overheidsdienst Werkgelegenheid, A. en S.O. Minimumlonen per Paritair (Sub)Comité. Accessed January 28, 2025. <https://www.minimumlonen.be/index.html>.
- Fessenden, B., D.J. Weigel, J. Osterstock, D.T. Galligan, and F. Di Croce. 2020. Validation of genomic predictions for a lifetime merit selection index for the US dairy industry. *J. Dairy Sci.* 103:10414–10428. doi:10.3168/JDS.2020-18502.

- 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.* 84:2025–35. doi:10.3168/jds.S0022-0302(01)74646-2.
- Fogsgaard, K.K., C.M. Røntved, P. Sørensen, and M.S. Herskin. 2012. Sickness behavior in dairy cows during *Escherichia coli* mastitis. *J. Dairy Sci.* 95:630–638. doi:10.3168/jds.2011-4350.
- Fourichon, C., H. Seegers, N. Bareille, and F. Beaudeau. 1999. Effects of disease on milk production in the dairy cow: A review. *Prev. Vet. Med.* 41:1–35. doi:10.1016/S0167-5877(99)00035-5.
- Fox, F.H. 1971. Clinical Diagnosis and Treatment of Ketosis. *J. Dairy Sci.* 54:974–978. doi:10.3168/JDS.S0022-0302(71)85953-2.
- Fox, J., and S. Weisberg. 2019. An $\{R\}$ Companion to Applied Regression. Third. Sage, Thousand Oaks $\{CA\}$.
- Fox, J., S. Weisberg, and B. Price. 2022. carData: Companion to Applied Regression Data Sets.
- Friendly, M. 2007. HE plots for Multivariate General Linear Models. *J. Comput. Graph. Stat.* 16:421–444.
- Friendly, M. 2010. HE Plots for Repeated Measures Designs. *J. Stat. Softw.* 37:1–40. doi:10.18637/jss.v037.i04.
- García, R., J. Aguilar, M. Toro, A. Pinto, and P. Rodríguez. 2020. A systematic literature review on the use of machine learning in precision livestock farming. *Comput. Electron. Agric.* 179:105826. doi:10.1016/j.compag.2020.105826.
- Genz, A., F. Bretz, T. Miwa, X. Mi, F. Leisch, F. Scheipl, and T. Hothorn. 2021. $\{mvtnorm\}$: Multivariate Normal and t Distributions.
- Gerber, M., S. Dürr, and M. Bodmer. 2020. Decision-Making of Swiss Farmers and the Role of the Veterinarian in Reducing Antimicrobial Use on Dairy Farms. *Front. Vet. Sci.* 7. doi:10.3389/fvets.2020.00565.
- Giuliodori, M.J., R.P. Magnasco, D. Becu-Villalobos, I.M. Lacau-Mengido, C.A. Risco, and R.L. De la Sota. 2013. Metritis in dairy cows: Risk factors and reproductive performance. *J. Dairy Sci.* 96:3621–3631. doi:10.3168/JDS.2012-5922.
- Goldhawk, C., N. Chapinal, D.M. Veira, D.M. Weary, and M.A.G. von Keyserlingk. 2009. Parturition feeding behavior is an early indicator of subclinical ketosis. *J. Dairy Sci.* 92:4971–4977. doi:10.3168/JDS.2009-2242.
- Gomez, A., and N.B. Cook. 2010. Time budgets of lactating dairy cattle in commercial freestall herds. *J. Dairy Sci.* 93:5772–5781. doi:10.3168/jds.2010-3436.
- Goselink, R.M.A., A.T.M. van Kneegsel, A. Bannink, R.M. Bruckmaier, J. Dijkstra, G. van Duinkerken, J.T. Schonewille, and W.H. Hendriks. 2024. Dry period length affects rumen adaptation in dairy cattle precalving and during the first weeks after calving. *J. Dairy Sci.* 107:8569–8581. doi:10.3168/JDS.2023-24090.
- Graves, S., H.-P. Piepho, and L.S. with help from Sundar Dorai-Raj. 2024. multcompView: Visualizations of Paired Comparisons.
- Grelet, C., A. Vanlierde, M. Hostens, L. Foldager, M. Salavati, K.L. Ingvarsten, M. Crowe, M.T. Sorensen, E. Froidmont, C.P. Ferris, C. Marchitelli, F. Becker, T. Larsen, F. Carter, F. Dehareng, and F. Dehareng. 2019. Potential of milk mid-IR spectra to predict metabolic status of cows through blood components and an innovative clustering approach. *animal* 13:649–658. doi:10.1017/S1751731118001751.
- Gross, J.J., and R.M. Bruckmaier. 2019. Invited review: Metabolic challenges and adaptation during different functional stages of the mammary gland in dairy cows: Perspectives for sustainable milk production. *J. Dairy Sci.* 102:2828–2843. doi:10.3168/JDS.2018-15713.
- Grummer, R.R. 1995. Impact of changes in organic nutrient metabolism on feeding the transition dairy cow. *J. Anim. Sci.* 73:2820–2833. doi:10.2527/1995.7392820x.
- Grummer, R.R., D.G. Mashek, and A. Hayirli. 2004. Dry matter intake and energy balance in the transition period 20:447–470.
- Grzesiak, W., P. Błaszczuk, and R. Lacroix. 2006. Methods of predicting milk yield in dairy cows—Predictive capabilities of Wood’s lactation curve and artificial neural networks (ANNs). *Comput. Electron. Agric.* 54:69–83. doi:10.1016/J.COMPAG.2006.08.004.
- Haine, D., H. Delgado, R. Cue, A. Sewalem, K. Wade, R. Lacroix, D. Lefebvre, J. Arsenault, É. Bouchard, and J. Dubuc. 2017. Contextual herd factors associated with cow culling risk in Québec dairy herds: A multilevel analysis. *Prev. Vet. Med.* 144:7–12. doi:10.1016/J.PREVETMED.2017.05.014.
- Hartung, J., T. Banhazi, E. Vranken, and M. Guarino. 2017. European farmers’ experiences with

- precision livestock farming systems. *Anim. Front.* 7:38–44. doi:10.2527/af.2017.0107.
- Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. S.S. in Statistics, ed. Springer, New York.
- van der Heide, E.M.M., R.F. Veerkamp, M.L. van Pelt, C. Kamphuis, I. Athanasiadis, and B.J. Ducro. 2019. Comparing regression, naive Bayes, and random forest methods in the prediction of individual survival to second lactation in Holstein cattle. *J. Dairy Sci.* 102:9409–9421. doi:10.3168/jds.2019-16295.
- Heirbaut, S., X.P.P. Jing, B. Stefańska, E. Pruszyńska-Oszmałek, L. Buysse, P. Lutakome, M.Q.Q. Zhang, M. Thys, L. Vandaele, and V. Fievez. 2023. Diagnostic milk biomarkers for predicting the metabolic health status of dairy cattle during early lactation. *J. Dairy Sci.* 106:690–702. doi:10.3168/jds.2022-22217.
- Hendriks, S.J., C.V.C. Phyn, S.A. Turner, K.R. Mueller, B. Kuhn-Sherlock, D.J. Donaghy, J.M. Huzzey, and J.R. Roche. 2022. Associations between peripartum lying and activity behaviour and blood non-esterified fatty acids and β -hydroxybutyrate in grazing dairy cows. *animal* 16:100470. doi:10.1016/J.ANIMAL.2022.100470.
- Herdt, T.H. 2000. Ruminant adaptation to negative energy balance. Influences on the etiology of ketosis and fatty liver.. *Vet. Clin. North Am. Food Anim. Pract.* 16. doi:10.1016/S0749-0720(15)30102-X.
- Hermans, K., B. Van Ranst, G. Opsomers, and M. Hostens. 2018. Promises and Challenges of Big Data Associated With Automated Dairy Cow Welfare Assessment. A. Butterworth, ed. CABI, Bristol.
- Ho, P.N., T.D.W.W. Luke, and J.E. Pryce. 2021. Validation of milk mid-infrared spectroscopy for predicting the metabolic status of lactating dairy cows in Australia. *J. Dairy Sci.* 104:4467–4477. doi:10.3168/JDS.2020-19603.
- Hodgson, R.E. 1964. Past, Present, and Future of the Dairy Herd Improvement Section. *J. Dairy Sci.* 47:310–315. doi:10.3168/jds.S0022-0302(64)88651-3.
- Horst, E.A., S.K. Kvidera, and L.H. Baumgard. 2021. Invited review: The influence of immune activation on transition cow health and performance—A critical evaluation of traditional dogmas. *J. Dairy Sci.* 104:8380–8410. doi:10.3168/JDS.2021-20330.
- 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.
- Hothorn, T. 2025. TH.data: TH's Data Archive.
- Hothorn, T., F. Bretz, and P. Westfall. 2008. Simultaneous inference in general parametric models. *Biom. J.* 50:346–363. doi:10.1002/BIMJ.200810425.
- Houben, E.H.P., A.A. Dijkhuizen, J.A.M. Van Arendonk, and R.B.M. Huirne. 1993. Short- and Long-Term Production Losses and Repeatability of Clinical Mastitis in Dairy Cattle. *J. Dairy Sci.* 76:2561–2578. doi:10.3168/JDS.S0022-0302(93)77591-8.
- Hristov, A.N., W.J. Price, and B. Shafii. 2005. A Meta-Analysis on the Relationship Between Intake of Nutrients and Body Weight with Milk Volume and Milk Protein Yield in Dairy Cows. *J. Dairy Sci.* 88:2860–2869. doi:10.3168/JDS.S0022-0302(05)72967-2.
- 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.
- Hut, P.R., S.E.M. Kuiper, M. Nielen, J.H.J.L. Hulsen, E.N. Stassen, and M.M. Hostens. 2022. Sensor based time budgets in commercial Dutch dairy herds vary over lactation cycles and within 24 hours. *PLoS One* 17:e0264392. doi:10.1371/JOURNAL.PONE.0264392.
- Huybrechts, T., K. Mertens, J. De Baerdemaeker, B. De Ketelaere, and W. Saeys. 2014. Early warnings from automatic milk yield monitoring with online synergistic control. *J. Dairy Sci.* 97:3371–3381.
- Huzzey, J.M., T.F. Duffield, S.J. LeBlanc, D.M. Veira, D.M. Weary, and M.A.G. Von Keyserlingk. 2009. Short communication: Haptoglobin as an early indicator of metritis. *J. Dairy Sci.* 92:621–625. doi:10.3168/jds.2008-1526.
- Huzzey, J.M., M.A.G. Von Keyserlingk, and D.M. Weary. 2005. Changes in Feeding, Drinking, and Standing Behavior of Dairy Cows During the Transition Period. *J. Dairy Sci.* 88:2454–2461. doi:10.3168/JDS.S0022-0302(05)72923-4.
- Huzzey, J.M., S. Mann, D. V. Nydam, R.J. Grant, and T.R. Overton. 2015. Associations of peripartum markers of stress and inflammation with milk yield and reproductive performance in Holstein dairy cows. *Prev. Vet. Med.* 120:291–297. doi:10.1016/J.PREVETMED.2015.04.011.

- Huzzey, J.M., D. V. Nydam, R.J. Grant, and T.R. Overton. 2011. Associations of prepartum plasma cortisol, haptoglobin, fecal cortisol metabolites, and nonesterified fatty acids with postpartum health status in Holstein dairy cows. *J. Dairy Sci.* 94:5878–5889. doi:10.3168/jds.2010-3391.
- Huzzey, J.M., D.M. Veira, D.M. Weary, and M.A.G. Von Keyserlingk. 2007. Prepartum Behavior and Dry Matter Intake Identify Dairy Cows at Risk for Metritis. *J. Dairy Sci.* 90:3220–3233. doi:10.3168/JDS.2006-807.
- ICAR. . Accessed April 28, 2024. <https://my.icar.org/stats/list>.
- Itle, A.J., J.M. Huzzey, D.M. Weary, and M.A.G. von Keyserlingk. 2015. Clinical ketosis and standing behavior in transition cows. *J. Dairy Sci.* 98:128–134. doi:10.3168/JDS.2014-7932.
- Iwersen, M., U. Falkenberg, R. Voigtsberger, D. Forderung, and W. Heuwieser. 2009. Evaluation of an electronic cowside test to detect subclinical ketosis in dairy cows. *J. Dairy Sci.* 92:2618–2624. doi:10.3168/JDS.2008-1795.
- James, G., D. Witten, T. Hastie, and R. Tibshirani. 2013. *An Introduction to Statistical Learning - with Applications in R*. G. Casella, S. Fienberg, and I. Olkin, ed. Springer, New York, New York, USA.
- Jensen, A.L., M.B. Petersen, and H. Houe. 1993. Determination of the fructosamine concentration in bovine serum samples. *Zentralbl. Veterinarmed. A* 40:111–117. doi:10.1111/J.1439-0442.1993.TB00607.X.
- Johnson, R.N., P.A. Metcalf, and J.R. Baker. 1983. Fructosamine: a new approach to the estimation of serum glycosylprotein. An index of diabetic control. *Clin. Chim. Acta.* 127:87–95. doi:10.1016/0009-8981(83)90078-5.
- Johnston, C., and T.J. DeVries. 2018. Short communication: Associations of feeding behavior and milk production in dairy cows. *J. Dairy Sci.* 101:3367–3373. doi:10.3168/JDS.2017-13743.
- Jonker, J.S., R.A. Kohn, and J. High. 2002. Use of Milk Urea Nitrogen to Improve Dairy Cow Diets. *J. Dairy Sci.* 85:939–946. doi:10.3168/JDS.S0022-0302(02)74152-0.
- Kassambara, A. 2023. ggpubr: “ggplot2” Based Publication Ready Plots.
- Kaufman, E.I., V.H. Asselstine, S.J. LeBlanc, T.F. Duffield, and T.J. DeVries. 2018. Association of rumination time and health status with milk yield and composition in early-lactation dairy cows. *J. Dairy Sci.* 101:462–471. doi:10.3168/JDS.2017-12909.
- Kaufman, E.I., S.J. LeBlanc, B.W. McBride, T.F. Duffield, and T.J. DeVries. 2016a. Association of rumination time with subclinical ketosis in transition dairy cows. *J. Dairy Sci.* 99:5604–5618. doi:10.3168/jds.2015-10509.
- Kaufman, E.I., S.J. LeBlanc, B.W. McBride, T.F. Duffield, and T.J. DeVries. 2016b. Short communication: Association of lying behavior and subclinical ketosis in transition dairy cows. *J. Dairy Sci.* 99:7473–7480. doi:10.3168/JDS.2016-11185.
- von Keyserlingk, M.A.G., N.P. Martin, E. Kebreab, K.F. Knowlton, R.J. Grant, M. Stephenson, C.J. Sniffen, J.P. Harner, A.D. Wright, and S.I. Smith. 2013. Invited review: Sustainability of the US dairy industry. *J. Dairy Sci.* 96:5405–5425. doi:10.3168/JDS.2012-6354.
- von Keyserlingk, M.A.G., K.E. Mills, and D.M. Weary. 2024. Attitudes of western Canadian dairy farmers toward technology. *J. Dairy Sci.* 107:933–943. doi:10.3168/jds.2023-23279.
- 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.* 100:1–10. doi:10.1093/JAS/SKAC220.
- De Koster, J., M. Salavati, C. Grelet, M.A.A. Crowe, E. Matthews, R. O’Flaherty, G. Opsomer, L. Foldager, M. Hostens, M. Niamh, F. Alan, C. Fiona, M. Elizabeth, S. Andreia, B. Colin, R. Pauline, O. Roisin, H. Sinead, W. Claire, C. Zhangrui, F. Ali, P. Geoff, W. Dirk, B.S. Beatriz Sanz, F. Conrad, W. Alistair, B. Matt, V. Mieke, H. Kristof, M. Sander, B. Hannes, V. Jan, V. Leila, V. Bonny, I. Klaus, S.T. Martin Tang, H. Johanna, D. Susanne, O. Soren, R. Janne, K. Mogens, M. Else, F. Leslie, G. Charlotte, E. Jehan, R. Tine, L. Torben, de O. de, M. Cinzia, S. Federica, N. Francesco, M. Bianca, A. Crisà, B. Luca, M. Jennifer, M. Daragh, K. Francis, C. Andrew, M. Matt, Z. Shujun, C. Xing, C. Huanchun, Z. Junlong, Y. Liguu, H. Guohua, T. Chen, W. Guiqiang, B. Michel, S. Marlène, P. Armin, E. Arnold, K. Linda, F. Anders, A. Thomas, L. Matthew, E. Chris, C. Gavin, T. Jerry, T. Deborah, G. Nicolas, G. Michel, C. Frederic, P.R. Marilou Ramos, H. Hedi, B. Catherine, T. Haruko, L. Aurelie, V.L. Anne-Sophie, M. Rodrigo, D.N. Saied Naderi, D. Frederic, G. Clement, V. Amelie, F. Eric, B. Frank, S. Martin, and V.P. Sergio Palma. 2019. Prediction of metabolic clusters in early-lactation dairy cows using models based on milk biomarkers. *J. Dairy Sci.* 102:2631–2644. doi:10.3168/jds.2018-15533.
- De Koster, J.D., and G. Opsomer. 2013. Insulin resistance in dairy cows. *Vet. Clin. North Am. - Food Anim. Pract.* 29:299–322. doi:10.1016/j.cvfa.2013.04.002.

- 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.
- Kuznetsova, A., P.B. Brockhoff, and R.H.B. Christensen. 2017. ImerTest Package: Tests in Linear Mixed Effects Models. *J. Stat. Softw.* 82:1–26. doi:10.18637/JSS.V082.I13.
- Lacroix, R., K. M. Wade, R. Kok, and J. F. Hayes. 1995. Prediction of Cow Performance with a Connectionist Model. *Trans. ASAE* 38:1573–1579. doi:https://doi.org/10.13031/2013.27984.
- Leal Yepes, F.A., D. V. Nydam, S. Mann, L. Caixeta, J.A.A. McArt, T.R. Overton, J.J. Wakshlag, and H.J. Huson. 2019. Longitudinal Phenotypes Improve Genotype Association for Hyperketonemia in Dairy Cattle. *Anim.* 2019, Vol. 9, Page 1059 9:1059. doi:10.3390/ANI9121059.
- Lean, I.J., S.J. LeBlanc, D.B. Sheedy, T. Duffield, J.E.P. Santos, and H.M. Golder. 2023. Associations of parity with health disorders and blood metabolite concentrations in Holstein cows in different production systems. *J. Dairy Sci.* 106:500–518. doi:10.3168/jds.2021-21673.
- LeBlanc, S.J. 2010. Monitoring Metabolic Health of Dairy Cattle in the Transition Period. *J. Reprod. Dev.* 56:S29–S35. doi:10.1262/jrd.1056S29.
- LeBlanc, S.J., K.E. Leslie, and T.F. Duffield. 2005. Metabolic predictors of displaced abomasum in dairy cattle. *J. Dairy Sci.* 88:159–170. doi:10.3168/jds.S0022-0302(05)72674-6.
- Lenth, R. V. 2016. Least-Squares Means: The {R} Package {lsmeans}. *J. Stat. Softw.* 69:1–33. doi:10.18637/jss.v069.i01.
- Lenth, R. V. 2025. emmeans: Estimated Marginal Means, aka Least-Squares Means.
- Liboreiro, D.N., K.S. Machado, P.R.B. Silva, M.M. Maturana, T.K. Nishimura, A.P. Brandão, M.I. Endres, and R.C. Chebel. 2015. Characterization of peripartum rumination and activity of cows diagnosed with metabolic and uterine diseases. *J. Dairy Sci.* 98:6812–6827. doi:10.3168/JDS.2014-8947.
- Liland, K.H., B.-H. Mevik, and R. Wehrens. 2023. pls: Partial Least Squares and Principal Component Regression.
- Liseune, A., M. Salamone, D. Van den Poel, B. van Ranst, and M. Hostens. 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.
- Lomb, J., D.M. Weary, K.E. Mills, and M.A.G. von Keyserlingk. 2018. Effects of metritis on stall use and social behavior at the lying stall. *J. Dairy Sci.* 101:7471–7479. doi:10.3168/JDS.2017-14149.
- Lopreato, V., M. Mezzetti, L. Cattaneo, G. Ferronato, A. Minuti, and E. Trevisi. 2020. Role of nutraceuticals during the transition period of dairy cows: a review. *J. Anim. Sci. Biotechnol.* 11:96. doi:10.1186/S40104-020-00501-X.
- Lorenz, M.D., J.R. Coates, and M. Kent. 2011. Systemic or Multifocal Signs. *Handb. Vet. Neurol.* 432–487. doi:10.1016/B978-1-4377-0651-2.10015-3.
- Løvendahl, P., and L. Munksgaard. 2016. An investigation into genetic and phenotypic variation in time budgets and yield of dairy cows. *J. Dairy Sci.* 99:408–417. doi:10.3168/JDS.2015-9838.
- Lüdecke, D. 2021. SjPlot: Data Visualization for Statistics in Social Science. Accessed. <https://cran.r-project.org/package=sjPlot>.
- Macciotta, N.P.P., C. Dimauro, S.P.G. Rasso, R. Steri, and G. Pulina. 2011. The mathematical description of lactation curves in dairy cattle. *Ital. J. Anim. Sci.* 10:e51. doi:10.4081/ijas.2011.e51.
- Martin, O., and D. Sauviant. 2010. A teleonomic model describing performance (body, milk and intake) during growth and over repeated reproductive cycles throughout the lifespan of dairy cattle. 1. Trajectories of life function priorities and genetic scaling. *Anim. Consort.* 2030–2047. doi:10.1017/S1751731110001357.
- Mayeres, P., J. Stoll, J. Bormann, R. Reents, and N. Gengler. 2004. Prediction of Daily Milk, Fat, and Protein Production by a Random Regression Test-Day Model. *J. Dairy Sci.* 87:1925–1933. doi:10.3168/JDS.S0022-0302(04)73351-2.
- McArt, J.A.A., D. V. Nydam, G.R. Oetzel, T.R. Overton, and P.A. Ospina. 2013a. Elevated non-esterified fatty acids and β -hydroxybutyrate and their association with transition dairy cow performance. *Vet. J.* 198:560–570. doi:10.1016/J.TVJL.2013.08.011.
- McArt, J.A.A., D.V. Nydam, and G.R. Oetzel. 2013b. Dry period and parturient predictors of early lactation hyperketonemia in dairy cattle. *J. Dairy Sci.* 96:198–209. doi:10.3168/JDS.2012-5681.
- McArt, J.A.A., and G.R. Oetzel. 2023. Considerations in the Diagnosis and Treatment of Early Lactation Calcium Disturbances. *Vet. Clin. North Am. - Food Anim. Pract.* 39:241–259.

- doi:10.1016/j.cvfa.2023.02.009.
- Megahed, A.A., M.W.H. Hiew, and P.D. Constable. 2018. Clinical Utility of Plasma Fructosamine Concentration as a Hypoglycemic Biomarker during Early Lactation in Dairy Cattle. *J. Vet. Intern. Med.* 32:846–852. doi:10.1111/JVIM.15049.
- Mehmood, T., K.H. Liland, L. Snipen, and S. Sæbø. 2012. A review of variable selection methods in Partial Least Squares Regression. *Chemom. Intell. Lab. Syst.* 118:62–69.
- Melendez, P.G., and J. Cook. 2023. Association between Prepartum Alerts Generated Using a Commercial Monitoring System and Health and Production Outcomes in Multiparous Dairy Cows in Five UK Herds. *Anim.* 2023, Vol. 13, Page 3235 13:3235. doi:10.3390/ANI13203235.
- Mezzetti, M., L. Cattaneo, M.M. Passamonti, V. Lopreato, A. Minuti, and E. Trevisi. 2021. The Transition Period Updated: A Review of the New Insights into the Adaptation of Dairy Cows to the New Lactation. *Dairy 2021*, Vol. 2, Pages 617-636 2:617–636. doi:10.3390/DAIRY2040048.
- Mezzetti, M., and E. Trevisi. 2023. Methods of Evaluating the Potential Success or Failure of Transition Dairy Cows. *Vet. Clin. North Am. Food Anim. Pract.* 39:219–239. doi:10.1016/J.CVFA.2023.02.008.
- Miglior, F., A. Fleming, F. Malchiodi, L.F. Brito, P. Martin, and C.F. Baes. 2017. A 100-Year Review: Identification and genetic selection of economically important traits in dairy cattle. *J. Dairy Sci.* 100:10251–10271. doi:10.3168/JDS.2017-12968.
- Mills, K.E., D.M. Weary, and M.A.G. von Keyserlingk. 2020. Identifying barriers to successful dairy cow transition management. *J. Dairy Sci.* 103:1749–1758. doi:10.3168/jds.2018-16231.
- Mølgaard, L., B.M. Damgaard, V. Bjerre-Harpøth, and M.S. Herskin. 2012. Effects of percutaneous needle liver biopsy on dairy cow behaviour. *Res. Vet. Sci.* 93:1248–1254. doi:10.1016/J.RVSC.2012.04.001.
- Müller, K., and H. Wickham. 2022. *tibble: Simple Data Frames.*
- Mulligan, F.J., and M.L. Doherty. 2008. Production diseases of the transition cow. *Vet. J.* 176:3–9. doi:10.1016/j.tvjl.2007.12.018.
- Mulligan, F.J., L. O’Grady, D.A. Rice, and M.L. Doherty. 2006. A herd health approach to dairy cow nutrition and production diseases of the transition cow. *Anim. Reprod. Sci.* 96:331–353. doi:10.1016/j.anireprosci.2006.08.011.
- Murphy, M.D.D., M.J. O’Mahony, L. Shalloo, P. French, J. Upton, M.J. O’Mahony, L. Shalloo, P. French, and J. Upton. 2014. Comparison of modelling techniques for milk-production forecasting. *J. Dairy Sci.* 97:3352–3363. doi:10.3168/JDS.2013-7451.
- Nayeri, S., F. Schenkel, A. Fleming, V. Kroezen, M. Sargolzaei, C. Baes, A. Cánovas, J. Squires, and F. Miglior. 2019. Genome-wide association analysis for β -hydroxybutyrate concentration in Milk in Holstein dairy cattle. *BMC Genet.* 20:1–17. doi:10.1186/S12863-019-0761-9/FIGURES/5.
- Neave, H.W., J. Lomb, M.A.G.G. von Keyserlingk, A. Behnam-Shabahang, and D.M. Weary. 2017. Parity differences in the behavior of transition dairy cows. *J. Dairy Sci.* 100:548–561. doi:10.3168/JDS.2016-10987.
- Neave, H.W., J. Lomb, D.M. Weary, S.J. LeBlanc, J.M. Huzzey, and M.A.G. von Keyserlingk. 2018. Behavioral changes before metritis diagnosis in dairy cows. *J. Dairy Sci.* 101:4388–4399. doi:10.3168/JDS.2017-13078.
- Nestorov, I., M. Rowland, S.T. Hadjitodorov, and I. Petrov. 1999. Empirical versus mechanistic modelling: Comparison of an artificial neural network to a mechanistically based model for quantitative structure pharmacokinetic relationships of a homologous series of barbiturates. *AAPS PharmSci* 1:5–13. doi:10.1208/ps010417.
- Nielsen, P.P., I. Fontana, K.H. Sloth, M. Guarino, and H. Blokhuis. 2018. Technical note: Validation and comparison of 2 commercially available activity loggers. *J. Dairy Sci.* 101:5449–5453. doi:10.3168/JDS.2017-13784.
- Nordlund, K. 2006. Transition Cow Index. Pages 139–143 in 39th Proceedings American Association Bovine Practitioners, , Sept. 20-24. Frontier Printers, Inc, St. Paul, MN, USA.
- Nordlund, K., T. Bennett, G. Oetzel, M. Clayton, and N. Cook. 2011. Method for optimizing health and productivity of milk producing animals. US Pat. No. US7886691B2.
- Nydam, D. V, T.R. Overton, J.A.A. McArt, M.M. McCarthy, B. Leno, and S. Mann. 2017. Management of transition cows to optimize health and production. FASS Inc. C.A. Risco, ed. American Dairy Science Association, Champaign, IL 61820.
- Ospina, P.A., J.A. McArt, T.R. Overton, T. Stokol, and D. V. Nydam. 2013. Using nonesterified fatty acids and β -hydroxybutyrate concentrations during the transition period for herd-level monitoring of increased risk of disease and decreased reproductive and milking performance. *Vet. Clin. North*

- Am. - Food Anim. Pract. 29:387–412. doi:10.1016/j.cvfa.2013.04.003.
- Osterwalder, A. 2004. The business model ontology a proposition in a design science approach. [s.n.], Lausanne TA - TT -.
- Overton, T.R., J.A.A. McArt, and D. V. Nydam. 2017. A 100-Year Review: Metabolic health indicators and management of dairy cattle. *J. Dairy Sci.* 100:10398–10417. doi:10.3168/JDS.2017-13054.
- Parker Gaddis, K.L., J.B. Cole, J.S. Clay, and C. Maltecca. 2016. Benchmarking dairy herd health status using routinely recorded herd summary data. *J. Dairy Sci.* 99:1298–1314. doi:10.3168/jds.2015-9840.
- Pascottini, O., M. Probo, S.J. LeBlanc, G. Opsomer, and M. Hostens. 2020a. Assessment of associations between transition diseases and reproductive performance of dairy cows using survival analysis and decision tree algorithms. *Prev. Vet. Med.* 176:104908. doi:10.1016/j.prevetmed.2020.104908.
- Pascottini, O.B., J.L.M.R. Leroy, and G. Opsomer. 2020b. Metabolic Stress in the Transition Period of Dairy Cows: Focusing on the Prepartum Period. *Animals* 10:1419. doi:10.3390/ani10081419.
- Pascottini, O.B., J.L.M.R. Leroy, and G. Opsomer. 2022. Maladaptation to the transition period and consequences on fertility of dairy cows. *Reprod. Domest. Anim.* 57:21–32. doi:10.1111/RDA.14176.
- Paudyal, S., F.P. Maunsell, J.T. Richeson, C.A. Risco, D.A. Donovan, and P.J. Pinedo. 2018. Rumination time and monitoring of health disorders during early lactation. *animal* 12:1484–1492. doi:10.1017/S1751731117002932.
- Pinedo, P., J.E.P. Santos, R.C. Chebel, K.N. Galvão, G.M. Schuenemann, R.C. Bicalho, R.O. Gilbert, S.R. Zas, C.M. Seabury, G. Rosa, and W.W. Thatcher. 2020. Early-lactation diseases and fertility in 2 seasons of calving across US dairy herds. *J. Dairy Sci.* 103:10560–10576. doi:10.3168/JDS.2019-17951.
- Piñeiro, J.M., B.T. Menichetti, A.A. Barragan, A.E. Relling, W.P. Weiss, S. Bas, and G.M. Schuenemann. 2019. Associations of postpartum lying time with culling, milk yield, cyclicity, and reproductive performance of lactating dairy cows. *J. Dairy Sci.* 102:3362–3375. doi:10.3168/JDS.2018-15387.
- Pollott, G.E. 2000. A Biological Approach to Lactation Curve Analysis for Milk Yield. *J Dairy Sci* 83:2448–2458. doi:10.3168/jds.S0022-0302(00)75136-8.
- Poppe, M., R.F. Veerkamp, M.L. van Pelt, and H.A. Mulder. 2020. Exploration of variance, autocorrelation, and skewness of deviations from lactation curves as resilience indicators for breeding. *J. Dairy Sci.* 103:1667–1684. doi:10.3168/jds.2019-17290.
- Pothmann, H., K. Nechanitzky, F. Sturmlechner, and M. Drillich. 2014. Consultancy to dairy farmers relating to animal health and herd health management on small- and medium-sized farms. *J. Dairy Sci.* 97. doi:10.3168/jds.2013-7364.
- Pralle, R.S., N.E. Schultz, H.M. White, and K.A. Weigel. 2020. Hyperketonemia gwas and parity-dependent SNP associations in holstein dairy cows intensively sampled for blood β -hydroxybutyrate concentration. *Physiol. Genomics* 52:347–357. doi:10.1152/PHYSIOLGENOMICS.00016.2020/ASSET/IMAGES/LARGE/ZH70072044240003.JPEG.
- Probo, M., O.B. Pascottini, S. LeBlanc, G. Opsomer, and M. Hostens. 2018. Association between metabolic diseases and the culling risk of high-yielding dairy cows in a transition management facility using survival and decision tree analysis. *J. Dairy Sci.* 101:9419–9429. doi:10.3168/jds.2018-14422.
- Proudfoot, K.L. 2023. ADSA Foundation Scholar Award: What makes for a good life for transition dairy cows? Current research and future directions. *J. Dairy Sci.* 106:5896–5907. doi:10.3168/JDS.2022-23194.
- Pryce, J.E., K.L. Parker Gaddis, A. Koeck, C. Bastin, M. Abdelsayed, N. Gengler, F. Miglior, B. Heringstad, C. Egger-Danner, K.F. Stock, A.J. Bradley, and J.B. Cole. 2016. Invited review: Opportunities for genetic improvement of metabolic diseases. *J. Dairy Sci.* 99:6855–6873. doi:10.3168/JDS.2016-10854.
- Ptak, E., and L.R. Schaeffer. 1993. Use of test day yields for genetic evaluation of dairy sires and cows. *Livest. Prod. Sci.* 34:23–34. doi:10.1016/0301-6226(93)90033-E.
- Pustejovsky, J. 2025. clubSandwich: Cluster-Robust (Sandwich) Variance Estimators with Small-Sample Corrections.
- R Core Team. 2020. R: A Language and Environment for Statistical Computing. Accessed February 16, 2023. <https://www.r-project.org/>.
- Ranzato, G., B. Aernouts, I. Lora, I. Adriaens, A. Ben Abdelkrim, M.J. Gote, and G. Cozzi. 2024.

- Comparison of 3 mathematical models to estimate lactation performance in dairy cows. *J. Dairy Sci.* 107:6888–6901. doi:10.3168/JDS.2023-24224.
- Rasmussen, L.K., B.L. Nielsen, J.E. Pryce, T.T. Mottram, and R.F. Veerkamp. 1999. Risk factors associated with the incidence of ketosis in dairy cows. *Anim. Sci.* 68:379–386. doi:10.1017/S1357729800050372.
- Rial, C., A. Laplacette, L. Caixeta, C. Florentino, F. Peña-Mosca, and J.O. Giordano. 2023. Metabolic-digestive clinical disorders of lactating dairy cows were associated with alterations of rumination, physical activity, and lying behavior monitored by an ear-attached sensor. *J. Dairy Sci.* 106:9323–9344. doi:10.3168/JDS.2022-23156.
- Roche, S.M., D.L. Renaud, J. Saraceni, D.F. Kelton, and T.J. DeVries. 2024. Invited review: Prevalence, risk factors, treatment, and barriers to best practice adoption for lameness and injuries in dairy cattle—A narrative review. *J. Dairy Sci.* 107:3347–3366. doi:10.3168/JDS.2023-23870.
- Rodriguez-Jimenez, S., K.J.J. Haerr, E. Trevisi, J.J.J. Loor, F.C.C. Cardoso, and J.S.S. Osorio. 2018. Prepartal standing behavior as a parameter for early detection of postpartal subclinical ketosis associated with inflammation and liver function biomarkers in periparturient dairy cows. *J. Dairy Sci.* 101:8224–8235. doi:10.3168/JDS.2017-14254.
- da Rosa Righi, R., G. Goldschmidt, R. Kunst, C. Deon, and C. André da Costa. 2020. Towards combining data prediction and internet of things to manage milk production on dairy cows. *Comput. Electron. Agric.* 169:105156. doi:10.1016/J.COMPAG.2019.105156.
- Ruegg, P.L. 2017. A 100-Year Review: Mastitis detection, management, and prevention. *J. Dairy Sci.* 100:10381–10397. doi:10.3168/JDS.2017-13023.
- Ruoff, J., S. Borchardt, and W. Heuwieser. 2017. Short communication: Associations between blood glucose concentration, onset of hyperketonemia, and milk production in early lactation dairy cows. *J. Dairy Sci.* 100:5462–5467. doi:10.3168/JDS.2016-12237.
- 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.
- Saborío-Montero, A., B. Vargas-Leitón, J.J. Romero-Zúñiga, and J.M. Sánchez. 2017. Risk factors associated with milk fever occurrence in grazing dairy cattle. *J. Dairy Sci.* 100:9715–9722. doi:10.3168/JDS.2017-13065.
- Säfken, B., D. Rügamer, T. Kneib, and S. Greven. 2021. Conditional model selection in mixed-effects models with caic4. *J. Stat. Softw.* 99:1–30. doi:10.18637/JSS.V099.I08.
- Salamone, M., I. Adriaens, A. Liseune, S. Heirbaut, X.P. Jing, V. Fievez, L. Vandaele, G. Opsomer, M. Hostens, and B. Aernouts. 2023. Milk yield residuals and their link with the metabolic status of dairy cows in the transition period. *J. Dairy Sci.* 0:317–330. doi:10.3168/jds.2023-23641.
- Salamone, M., I. Adriaens, A. Vervaeke, G. Opsomer, H. Atashi, V. Fievez, B. Aernouts, and M. Hostens. 2022. Prediction of first test day milk yield using historical records in dairy cows. *animal* 16:100658. doi:10.1016/j.animal.2022.100658.
- Salehi, M. 2022. Global water shortage and potable water safety; Today's concern and tomorrow's crisis. *Environ. Int.* 158:106936. doi:10.1016/J.ENVINT.2021.106936.
- Salvador, J. 2017. A Taxonomy of Example-Based Super Resolution. *Example-Based Super Resolut.* 15–29. doi:10.1016/B978-0-12-809703-8.00002-2.
- Santos, J.E.P., and E.S. Ribeiro. 2014. Impact of animal health on reproduction of dairy cows. *Anim. Reprod.* 11:254–269.
- Santos, M.G.S., N. Antonacci, C. Van Dorp, B. Mion, D. Tulpan, and E.S. Ribeiro. 2024. Use of rumination time in health risk assessment of preparturient dairy cows. *J. Dairy Sci.* doi:10.3168/JDS.2023-24610.
- Van Saun, R.J. 2023a. Ruminant Metabolic Diseases: Perturbed Homeostasis. *Vet. Clin. North Am. - Food Anim. Pract.* 39:185–201. doi:10.1016/j.cvfa.2023.02.001.
- Van Saun, R.J. 2023b. Metabolic Profiling in Ruminant Diagnostics. *Vet. Clin. North Am. Food Anim. Pract.* 39:49–71. doi:10.1016/J.CVFA.2022.10.004.
- Saun, V., and J. Robert. 2006. Metabolic profiles for evaluation of the transition period. *Proc. Am. Assoc. Bov. Pract.* 130–138.
- Schaeffer, L.R., and J. Jamrozik. 1996. Multiple-Trait Prediction of Lactation Yields for Dairy Cows. *J. Dairy Sci.* 79:2044–2055. doi:10.3168/JDS.S0022-0302(96)76578-5.
- Schirmann, K., D.M. Weary, W. Heuwieser, N. Chapinal, R.L.A.A. Cerri, and M.A.G.G. von Keyserlingk. 2016. Short communication: Rumination and feeding behaviors differ between healthy and sick dairy cows during the transition period. *J. Dairy Sci.* 99:9917–9924. doi:10.3168/JDS.2015-10548.

- Seegers, H., C. Fourichon, and F. Beaudeau. 2003. Production effects related to mastitis and mastitis economics in dairy cattle herds. *Vet. Res.* 34:475–491. doi:10.1051/vetres:2003027.
- Shahinfar, S., D. Page, J. Guenther, V. Cabrera, P. Fricke, and K. Weigel. 2014. Prediction of insemination outcomes in Holstein dairy cattle using alternative machine learning algorithms. *J. Dairy Sci.* 97:731–742. doi:10.3168/jds.2013-6693.
- Sheldon, I.M., G.S. Lewis, S. LeBlanc, and R.O. Gilbert. 2006. Defining postpartum uterine disease in cattle. *Theriogenology* 65:1516–1530. doi:10.1016/J.THERIOGENOLOGY.2005.08.021.
- 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.
- Simoni, A., F. König, K. Weimar, A. Hancock, C. Wunderlich, M. Klawitter, T. Breuer, M. Drillich, M. Iwersen, M.R. Borchers, J.M. Bewley, M.C. Codrea, S. Højsgaard, N.C. Friggens, I. Adriaens, I. van den Brulle, K. Geerinckx, L. D’Anvers, S. De Vlieghe, B. Aernouts, R.R. Grummer, K. Nordlund, H. Wickham, M. Averick, J. Bryan, W. Chang, L.D. McGowan, R. François, G. Grolemond, A. Hayes, L. Henry, J. Hester, M. Kuhn, T.L. Pedersen, E. Miller, S.M. Bache, K. Müller, J. Ooms, D. Robinson, D.P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, A. Liseune, M. Salamone, D. Van den Poel, B. Van Ranst, M. Hostens, I. Adriaens, I. van den Brulle, L. D’Anvers, J.M.E. Statham, K. Geerinckx, S. De Vlieghe, S. Piepers, B. Aernouts, O. Pascottini, M. Probo, S.J. LeBlanc, G. Opsomer, M. Hostens, J.M. Piñeiro, B.T. Menichetti, A.A. Barragan, A.E. Relling, W.P. Weiss, S. Bas, G.M. Schuenemann, P.N. Ho, T.D.W. Luke, J.E. Pryce, G. Ranzato, B. Aernouts, I. Lora, I. Adriaens, A. Ben Abdelkrim, M.J. Gote, G. Cozzi, G.L. Hadley, S.B. Harsh, C.A. Wolf, T. Hothorn, F. Bretz, P. Westfall, A. Genz, F. Bretz, T. Miwa, X. Mi, F. Leisch, F. Scheipl, B. Bornkamp, M. Maechler, T. Hothorn, F. Bretz, P. Westfall, A. Kuznetsova, P.B. Brockhoff, R.H.B. Christensen, B. Säfken, D. Rügamer, T. Kneib, S. Greven, A. Gomez, N.B. Cook, R.J. Van Saun, M.L. Stangaferro, R. Wijma, L.S. Caixeta, M.A. Al-Abri, J.O. Giordano, A. Simoni, F. König, K. Weimar, A. Hancock, C. Wunderlich, M. Klawitter, T. Breuer, M. Drillich, M. Iwersen, P.G. Melendez, J. Cook, M. Mezzetti, and E. Trevisi. 2024. Evaluation of sensor-based health monitoring in dairy cows: Exploiting rumination times for health alerts around parturition. *J. Dairy Sci.* 107:6052–6064. doi:10.3168/JDS.2023-24313.
- Slowikowski, K. 2024. ggrepel: Automatically Position Non-Overlapping Text Labels with “ggplot2.”
- Snijders, S.E.M., P.G. Dillon, K.J. O’Farrell, M. Diskin, A.R.G. Wylie, D. O’Callaghan, M. Rath, and M.P. Boland. 2001. Genetic merit for milk production and reproductive success in dairy cows. *Anim. Reprod. Sci.* 65:17–31. doi:10.1016/S0378-4320(00)00217-7.
- Sordillo, L.M., and W. Raphael. 2013. Significance of metabolic stress, lipid mobilization, and inflammation on transition cow disorders. *Vet. Clin. North Am. - Food Anim. Pract.* 29:267–278. doi:10.1016/j.cvfa.2013.03.002.
- Soriani, N., E. Trevisi, and L. Calamari. 2012. Relationships between rumination time, metabolic conditions, and health status in dairy cows during the transition period. *J. Anim. Sci.* 90:4544–4554. doi:10.2527/JAS.2011-5064.
- Sorondo, M.L., and A. Cirio. 2009. Evaluation of the serum fructosamine test to monitor plasma glucose concentration in the transition dairy cow. *J. Dairy Res.* 76:173–178. doi:10.1017/S0022029908003750.
- Stangaferro, M.L., R. Wijma, L.S. Caixeta, M.A. Al-Abri, and J.O. Giordano. 2016a. Use of rumination and activity monitoring for the identification of dairy cows with health disorders: Part III. Metritis. *J. Dairy Sci.* 99:7422–7433. doi:10.3168/jds.2016-11352.
- Stangaferro, M.L., R. Wijma, L.S. Caixeta, M.A. Al-Abri, and J.O. Giordano. 2016b. Use of rumination and activity monitoring for the identification of dairy cows with health disorders: Part II. Mastitis. *J. Dairy Sci.* 99:7411–7421. doi:10.3168/JDS.2016-10908.
- Stangaferro, M.L., R. Wijma, L.S. Caixeta, M.A. Al-Abri, and J.O. Giordano. 2016c. Use of rumination and activity monitoring for the identification of dairy cows with health disorders: Part I. Metabolic and digestive disorders. *J. Dairy Sci.* 99:7395–7410. doi:10.3168/JDS.2016-10907.
- Stanton, T.L., L.R. Jones, R.W. Everett, and S.D. Kachman. 1992. Estimating Milk, Fat, and Protein Lactation Curves with a Test Day Model. *J. Dairy Sci.* 75:1691–1700. doi:10.3168/jds.S0022-0302(92)77926-0.
- Stevenson, J.S., S. Banuelos, and L.G.D. Mendonça. 2020. Transition dairy cow health is associated with first postpartum ovulation risk, metabolic status, milk production, rumination, and physical activity. *J. Dairy Sci.* 103:9573–9586. doi:10.3168/JDS.2020-18636.

- Stygar, A.H., Y. Gómez, G. V Berteselli, E. Dalla Costa, E. Canali, J.K. Niemi, P. Llonch, and M. Pastell. 2021. A Systematic Review on Commercially Available and Validated Sensor Technologies for Welfare Assessment of Dairy Cattle. *Front. Vet. Sci.* 8:634338. doi:10.3389/fvets.2021.634338.
- Tetens, J., G. Thaller, and N. Krattenmacher. 2014. Genetic and genomic dissection of dry matter intake at different lactation stages in primiparous Holstein cows. *J. Dairy Sci.* 97:520–531. doi:10.3168/jds.2013-7301.
- Thompson-Crispi, K.A., M. Sargolzaei, R. Ventura, M. Abo-Ismael, F. Miglior, F. Schenkel, and B.A. Mallard. 2014. A genome-wide association study of immune response traits in Canadian Holstein cattle. doi:10.1186/1471-2164-15-559.
- Thompson, P.B. 2007. Agricultural sustainability: What it is and what it is not. *Int. J. Agric. Sustain.* 5:5–16. doi:10.1080/14735903.2007.9684809.
- Tran, T.N., N.L. Afanador, L.M.C. Buydens, and L. Blanchet. 2014. Interpretation of variable importance in Partial Least Squares with Significance Multivariate Correlation (sMC). *Chemom. Intell. Lab. Syst.* 138:153–160. doi:10.1016/J.CHEMOLAB.2014.08.005.
- Tucker, C.B., M.B. Jensen, A.M. de Passillé, L. Hänninen, and J. Rushen. 2021. Invited review: Lying time and the welfare of dairy cows. *J. Dairy Sci.* 104:20–46. doi:10.3168/JDS.2019-18074.
- Urton, G., M.A.G. Von Keyserlingk, and D.M. Weary. 2005. Feeding Behavior Identifies Dairy Cows at Risk for Metritis. *J. Dairy Sci.* 88:2843–2849. doi:10.3168/JDS.S0022-0302(05)72965-9.
- USDA NASS. . Accessed May 17, 2024. <https://quickstats.nass.usda.gov/>.
- VandeHaar, M.J., L.E. Armentano, K. Weigel, D.M. Spurlock, R.J. Tempelman, and R. Veerkamp. 2016. Harnessing the genetics of the modern dairy cow to continue improvements in feed efficiency. *J. Dairy Sci.* 99:4941–4954. doi:10.3168/jds.2015-10352.
- Vanholder, T., J. Papen, R. Bemers, G. Vertenten, and A.C.B. Berge. 2015. Risk factors for subclinical and clinical ketosis and association with production parameters in dairy cows in the Netherlands. *J. Dairy Sci.* 98:880–888. doi:10.3168/jds.2014-8362.
- Vanraden, P.M. 1997. Lactation Yields and Accuracies Computed from Test Day Yields and (Co)Variances by Best Prediction. *J. Dairy Sci.* 80:3015–3022. doi:10.3168/jds.S0022-0302(97)76268-4.
- Venables, W.N., and B.D. Ripley. 2002. *Modern Applied Statistics with S*. Fourth. Springer, New York.
- Voelker, D.E. 1981. Dairy Herd Improvement Associations. *J. Dairy Sci.* 64:1269–1277. doi:10.3168/JDS.S0022-0302(81)82700-2.
- Vossebeld, F., A. van Knegsel, and E. Saccenti. 2022. Phenotyping metabolic status of dairy cows using clustering of time profiles of energy balance peripartum. doi:10.3168/jds.2021-21518.
- Walsh, R.B., J.S. Walton, D.F. Kelton, S.J. LeBlanc, K.E. Leslie, and T.F. Duffield. 2007. The Effect of Subclinical Ketosis in Early Lactation on Reproductive Performance of Postpartum Dairy Cows. *J. Dairy Sci.* 90:2788–2796. doi:10.3168/jds.2006-560.
- Waring, E., M. Quinn, A. McNamara, E. Arino de la Rubia, H. Zhu, and S. Ellis. 2022. Skimr: Compact and Flexible Summaries of Data. Accessed February 16, 2023. <https://cran.r-project.org/package=skimr>.
- Warnick, L.D., D. Janssen, C.L. Guard, and Y.T. Gröhn. 2001. The effect of lameness on milk production in dairy cows. *J. Dairy Sci.* 84:1988–1997. doi:10.3168/jds.S0022-0302(01)74642-5.
- Wathes, C.M., H.H. Kristensen, J.-M. Aerts, and D. Berckmans. 2008. Is precision livestock farming an engineer’s daydream or nightmare, an animal’s friend or foe, and a farmer’s panacea or pitfall?. *Comput. Electron. Agric.* 64:2–10. doi:10.1016/j.compag.2008.05.005.
- Wathes, D.C., F. Becker, L. Buggiotti, M.A. Crowe, C. Ferris, L. Foldager, C. Grelet, M. Hostens, K.L. Ingvarsen, C. Marchitelli, M. Salavati, F. Signorelli, H. Takeda, L. Tang, Z. Cheng, and the GplusE Consortium. 2021. Associations between Circulating IGF-1 Concentrations, Disease Status and the Leukocyte Transcriptome in Early Lactation Dairy Cows. *Ruminants 2021*, Vol. 1, Pages 147-177 1:147–177. doi:10.3390/RUMINANTS1020012.
- Weigel, K.A., P.M. VanRaden, H.D. Norman, and H. Grosu. 2017. A 100-Year Review: Methods and impact of genetic selection in dairy cattle—From daughter–dam comparisons to deep learning algorithms. *J. Dairy Sci.* 100:10234–10250. doi:10.3168/JDS.2017-12954.
- Wehrather, G. 1993. Testing a Linear Regression Model Against Nonparametric Alternatives.
- Weller, J.I., E. Ezra, and M. Ron. 2017. Invited review: A perspective on the future of genomic selection in dairy cattle. *J. Dairy Sci.* 100:8633–8644. doi:10.3168/JDS.2017-12879.
- White, I.M.S., R. Thompson, and S. Brotherstone. 1999. Genetic and environmental smoothing of lactation curves with cubic splines. *J. Dairy Sci.* 82:632–638. doi:10.3168/jds.S0022-

0302(99)75277-X.

- Wickham, H. 2011. The Split-Apply-Combine Strategy for Data Analysis. *J. Stat. Softw.* 40:1–29.
- Wickham, H. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Wickham, H. 2023. *stringr: Simple, Consistent Wrappers for Common String Operations*.
- Wickham, H., M. Averick, J. Bryan, W. Chang, L.D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T.L. Pedersen, E. Miller, S.M. Bache, K. Müller, J. Ooms, D. Robinson, D.P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani. 2019. Welcome to the {tidyverse}. *J. Open Source Softw.* 4:1686. doi:10.21105/joss.01686.
- Wickham, H., R. François, L. Henry, K. Müller, and D. Vaughan. 2023a. *dplyr: A Grammar of Data Manipulation*.
- Wickham, H., and L. Henry. 2025. *purrr: Functional Programming Tools*.
- Wickham, H., J. Hester, and J. Bryan. 2022. *readr: Read Rectangular Text Data*.
- Wickham, H., D. Vaughan, and M. Girlich. 2023b. *tidyr: Tidy Messy Data*.
- Wiggans, G.R. 1991. National genetic improvement programs for dairy cattle in the United States.. *J. Anim. Sci.* 69:3853–3860. doi:10.2527/1991.6993853X.
- Wilke, C.O. 2024. *cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2.”*
- Wilkens, M.R., C.D. Nelson, L.L. Hernandez, and J.A.A. McArt. 2020. Symposium review: Transition cow calcium homeostasis—Health effects of hypocalcemia and strategies for prevention. *J. Dairy Sci.* 103:2909–2927. doi:10.3168/JDS.2019-17268.
- Wilkinson, M.D. 2016. Comment: The FAIR Guiding Principles for scientific data management and stewardship. doi:10.1038/sdata.2016.18.
- 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.
- Windig, J.J., M.P.L. Calus, and R.F. Veerkamp. 2005. Influence of Herd Environment on Health and Fertility and Their Relationship with Milk Production. *J. Dairy Sci.* 88:335–347. doi:10.3168/JDS.S0022-0302(05)72693-X.
- Wood, P.D.P. 1967. Algebraic model of the lactation curve in cattle. *Nature* 216:164–165. doi:10.1038/216164a0.
- Xu, W., A.T.M. van Knegsel, J.J.M. Vervoort, R.M. Bruckmaier, R.J. van Hoeij, B. Kemp, and E. Saccenti. 2019. Prediction of metabolic status of dairy cows in early lactation with on-farm cow data and machine learning algorithms. *J. Dairy Sci.* 102:10186–10201. doi:10.3168/JDS.2018-15791/ATTACHMENT/3CA1D2FE-F34F-4E81-9C31-8A476014D379/MMC1.PDF.
- Zeileis, A. 2006. Object-Oriented Computation of Sandwich Estimators. *J. Stat. Softw.* 16:1–16. doi:10.18637/jss.v016.i09.
- Zhang, M.Q., S. Heirbaut, X.P. Jing, B. Stefańska, L. Vandaele, N. De Neve, and V. Fievez. 2023. Transition cow clusters with distinctive antioxidant ability and their relation to performance and metabolic status in early lactation. doi:10.3168/jds.2022-22865.

Summary

Analysing the global evolution of the dairy industry over recent decades reveals a consistent increase in milk production, accompanied by a growing awareness of sustainability challenges in this sector. While pioneering regions have made significant contributions in reducing green gas emissions, these efforts are often offset by the expansion of the dairy industry in other areas. This confirms the crucial leading role of Europe and the critical importance of continuing to organize research and innovation to further drive existing efforts towards sustainability while facilitating the introduction of sustainable milk production practices in developing industries. Technological advancements, such as automation and digitalization, present promising solutions for enhancing efficiency and minimizing environmental impact in dairy production. Critically, animal health and welfare are central to the sustainability discussion, as diseases can lead to significant resource losses. Digitalisation and novel data analysis methods have provided useful insights to understanding and monitoring animal health, confirming the potential for developing data-driven tools to improve health management and foster more sustainable dairy practices.

Central to this thesis is the observation that the transition period remains one of the most crucial periods in the productive life cycle of dairy cows. This period, typically defined as the six weeks surrounding calving, involves processes where animals successfully transition from one lactation to the other or struggle to adapt to the physiological changes required to support milk production in early lactation. Failure to cope with these changes can lead to diseases with enduring impacts throughout the lactation, including reduced milk production, impaired reproduction and involuntary culling. Despite the development of different monitoring strategies, effective control of transition-associated diseases on dairy farms has yet to be fully realized.

Since the early 20th century, the dairy industry has undergone significant transformation through data-driven improvements, with dairy improvement programs (DHI) playing a pivotal role in this. This thesis builds on these established foundations and leverages modern technological advancements to explore the potential of a data-driven methodology for monitoring the transition period. The research was conceptualised within a framework based on precision livestock farming (PLF) principles, which guided the different steps needed to be taken.

The first step involved developing a model capable of predicting healthy milk yields in the next lactation based on data from the previous lactation. This led to the creation of the nextMILK_{TD} model, a random forest regression model developed in 3 variants. These variants based their predictions on an increasing number of parameters extracted from DHI data, herd data and reproduction data. Each variant could predict milk yields on the first test day of the next lactation with an RMSE of approximately 6 kg. Although there was no significant difference in performance among the variants, they all performed significantly better than benchmark models.

The nextMILK_{TD} was then complemented by the nextMILK_{MM} model which, unlike the DHI-based model, utilises data from milk meters to predict a sequence of milk yields for the next lactation. For both models, a numeric value was derived from the milk residuals in the transition period (MRT). These values represent the differences between the predicted milk yield and the actual milk yield on the first test day for the MRT_{TD} or averaged over the 21 DIM for MRT_{MM}.

The subsequent step in this thesis involved investigating the MRT as a potential monitoring tool in the transition period. This monitoring capability of the MRT is based on the premise that animals failing to transition would be identified by negative MRT. The presence of these negative values would indicate the presence of a perturbations that prevent milk production from reaching it's the expected level. The initial steps towards investigating these assumptions involved examining the relationship between MRT and an animal's health status. In an observational experimental setup, clinically diseased animals had significantly lower MRT. Additionally, in animals without clinical diseases, MRT was negatively associated with indicators of unfavourable metabolic status such as dry matter intake (DMI) post-partum, insulin-like growth factor 1, and nonesterified fatty acids. Despite being based on different data sources of data, MRT_{MM} and MRT_{TD} showed similar associations.

In the next steps, MRT_{TD} was associated with more indirect measures of transition success, such as behaviour and herd management practices. In terms of behaviour, lower MRT was linked to lower rumination and eating times post-partum, further associating the MRT with the DMI post-partum. Similarly, lower MRT was also associated with higher standing times, which have been linked to the occurrence of diseases (metritis, rumen acidosis, mastitis, lameness) in various studies. Finally, in the association study of MRT and herd management practices, MRT was strongly associated with herd milk yield performances. While management practices

showed weaker associations, the use of grassland products during the dry period had a negative effect on MRT.

The insights derived from various direct and indirect associations underscore the potential of MRT as a transition monitoring metric. To realise this potential, a broad valorisation plan was developed, addressing the challenges and opportunities within the increasingly digitalised dairy industry. This potential monitoring metric was rebranded from MRT to lactation onset value (LOV), and from there, an outline of the key resources, partners, activities, and customer segments was presented. The valorisation plan particularly explored the envisioned methods for presenting the LOV to future customers. Moreover, it highlighted the challenges in establishing industry partnerships and the need for effective dissemination strategies. This valorisation plan concludes by emphasising the importance of demonstrating the LOV's value to farmers and other stakeholders in the dairy industry.

In conclusion, this thesis provided a comprehensive examination of the development of a data-based key performance indicator. Diverging from industry standards, the MRT was developed through an open initiative aimed at creating a tangible metric to facilitate sustainable practices during the transition period and to advance the views on digital assets in a dairy sector. The importance of the FAIRness principle in this process is highlighted, and potential improvements to current industry methods are explored. Still, this thesis represents only the first step in this direction, and further research is needed to concretely transform the MRT into a robust evidence-based tool that can serve as a cow-side decision-making assistant. Nonetheless, the work presented in this thesis provides a solid foundation for future endeavours in this area

Samenvatting

Terugkijkend op de wereldwijde ontwikkelingen in de zuivelindustrie van de afgelopen decennia, zijn de volgende trends zichtbaar, een consistente toename van de melkproductie, die gepaard gaat met een groeiend bewustzijn rond duurzaamheid van deze sector. Hoewel vooruitstrijvende regio's vooruitgang hebben geboekt in het verminderen van broeikasgasemissies, worden deze inspanningen vaak tenietgedaan door de uitbreiding van de zuivelindustrie in andere gebieden. Dit bevestigt de cruciale rol van Europa als koploper en het onmisbaar belang om onderzoek en innovatie hier te blijven organiseren om eenerzijds de bestaande inspanningen richting duurzaamheid verder te drijven en anderzijds de invoering van duurzame melkproductiepraktijken in ontwikkelende industrieën te introduceren en stimuleren. Technologische vooruitgang zoals automatisering en digitalisering bieden veelbelovende oplossingen voor het verbeteren van de efficiëntie en het verminderen van milieu-impact door de zuivelproductie. Centraal in het duurzaamheidsgesprek staan de gezondheid en het welzijn van de dieren, waar ziekten kunnen leiden tot aanzienlijke verliezen. Digitalisering en nieuwe methoden voor data analyse hebben nuttige inzichten opgeleverd in het begrijpen en monitoren van de diergezondheid en bevestigen het potentieel voor de ontwikkeling van data gedreven hulpmiddelen om de diergezondheid te verbeteren en duurzamere zuivelpraktijken te bevorderen.

Centraal in dit proefschrift staat de transitieperiode als een van de meest cruciale periodes in de productiecyclus van melkvee. Deze periode, doorgaans gedefinieerd als de zes weken rondom afkalven, omvat processen waarbij dieren ofwel succesvol van de ene lactatie overgaan naar de andere, of moeite hebben om zich aan te passen aan de vele fysiologische veranderingen die nodig zijn om de melkproductie aan het begin van de lactatie te ondersteunen. Het niet kunnen omgaan met deze veranderingen, kan leiden tot ziekten met blijvende gevolgen gedurende de ganse lactatie, zoals een verminderde melkproductie, verminderde vruchtbaarheid en het vroegtijdig moeten vervangen van dieren. Ondanks de ontwikkeling van verschillende monitoringsstrategieën is een effectieve controle van transitieziekten op melkveebedrijven nog niet volledig gerealiseerd.

Sinds het begin van de 20^{ste} eeuw heeft de zuivelindustrie een aanzienlijke transformatie ondergaan, geleid door data gedreven verbeteringen, waarbij melkproductieregistratie (MPR) een centrale rol heeft gespeeld. Dit proefschrift bouwt voort op deze gevestigde fundamenten en benut moderne technologische ontwikkelingen om het potentieel van data gedreven

methodologieën voor het monitoren van de transitieperiode te onderzoeken. Het onderzoek werd geconceptualiseerd binnen de principes van de precisieveehouderij (PLF), welke tevens als leidraad diende doorheen de verschillende stappen die genomen moesten worden.

De eerste stap bestond uit het ontwikkelen van een model dat in staat is om gezonde melkproducties in de volgende lactatie te voorspellen op basis van gegevens uit de vorige lactatie. Dit leidde tot de ontwikkeling van het nextMILK_{TD}-model, een random forest regressiemodel dat in 3 varianten is ontwikkeld. De voorspellingen van deze drie varianten zijn gebaseerd op een toenemend aantal parameters uit de MPR-gegevens, kuddegegevens en voortplantingsgegevens. Elke variant kon de melkopbrengst op de eerste testdag van de volgende lactatie voorspellen met een RMSE van ongeveer 6 kg. Hoewel er geen significante verschillen zijn tussen de varianten, werkten ze allemaal significant beter dan de eenvoudige benchmarkmodellen.

Het nextMILK_{TD}-model werd vervolgens aangevuld met het nextMILK_{MM}-model dat, in tegenstelling tot het op MPR gebaseerde model, gegevens van melkmeters gebruikt om een reeks melkgiften in de volgende lactatie te voorspellen. Voor beide modellen werd de melk residuen berekend in de transitie (MRT). Deze waarden vertegenwoordigen de verschillen tussen de voorspelde melkgift en de werkelijke melkgift op de eerste test dag voor de MRT_{TD} en gemiddeld over de eerst 21 dagen in melk voor MRT_{MM}.

De volgende stap in dit proefschrift omvatte het onderzoeken van de MRT als een potentiële monitoringtool in de transitieperiode. De monitoringcapaciteit van de MRT is gebaseerd op de veronderstelling dat dieren die falen in deze transitieperiode gekenmerkt worden door een negatieve MRT. De aanwezigheid van negatieve waarden zou duiden op de aanwezigheid van een verstoring die het bereiken van de verwachte melkproductie belemmert. De eerste stap om deze veronderstellingen te onderzoeken was het analyseren van de relatie tussen de MRT en de gezondheidsstatus van een dier. In een observationele, experimentele opstelling hadden klinisch zieke dieren een significant lagere MRT. Bijkomend, in de groep met dieren zonder klinische symptomen, was de MRT negatief geassocieerd met indicatoren van een ongunstige metabole status zoals droge stof opname (DSO) post-partum, insuline-like groeifactor 1 en niet-veresterde vetzuren. Ondanks het feit dat MRT_{MM} en MRT_{TD} gebaseerd zijn op verschillende gegevensbronnen, vertoonden ze vergelijkbare associaties.

In de volgende stappen werd MRT_{TD} geassocieerd met meer indirecte indicators van transitie-succes, zoals diergedrag en bedrijfsmanagement. Op het gebied van gedrag werd een lagere MRT gelinkt met lagere herkauw- en eettijden na het afkalven, waardoor de MRT verder in verband werd gebracht met de DSO. Bijkomend werd een lagere MRT ook geassocieerd met hogere statijden, die in verschillende voorgaande onderzoeken al in verband werden gebracht met het optreden van verschillende ziekten (metritis, pensverzuuring, mastitis, kreupelheid). In de associatiestudie van MRT en bedrijfsmanagement was MRT bovendien sterk geassocieerd met de gemiddelde melkproductie van de veestapel. De managementpraktijken als zodanig vertoonden echter zwakkere associaties, het voornaamste negatief effect op MRT was gelinkt met het gebruik van graslandproducten in de droogstand .

De inzichten die voortkomen uit verschillende directe en indirecte associaties ondersteunen het potentieel van MRT als een monitoringstool in de transitie periode. Om dit potentieel uit te bouwen, werd een breed valorisatieplan ontwikkeld dat de uitdagingen en kansen binnen een steeds meer gedigitaliseerde zuivelindustrie bestudeert. Deze potentiële monitoringsmetriek werd vertaald van MRT tot lactatie opstart waarde (LOV), en van daaruit werd een overzicht van de belangrijkste middelen, partners, activiteiten en klantsegmenten gepresenteerd. Het valorisatieplan onderzocht met name de beoogde methoden om de LOV aan toekomstige klanten te presenteren. Bovendien werd gewezen op de uitdagingen bij het opzetten van samenwerking met de industrie en de behoefte aan effectieve verspreidingsstrategieën. Dit valorisatieplan eindigt met het benadrukken van het belang van het aantonen van de waarde van de LOV aan boeren en andere belanghebbenden in de zuivelindustrie.

Als conclusie biedt dit proefschrift een uitgebreide discussie omtrent de ontwikkeling van een op data gebaseerde prestatie-indicator. Verschillend van de industriestandaarden, was de MRT ontwikkeld via een open initiatief met als doel een tastbare waarde te creëren om duurzame praktijken tijdens de transitieperiode te identificeren en de visie op digitale middelen in een zuivelsector te bevorderen. Het belang van het FAIRness-principe in dit proces werd benadrukt en mogelijke verbeteringen van de huidige methoden in de sector werden onderzocht. Echter, dit proefschrift vertegenwoordigt slechts de eerste stap in deze richting waarbij verder onderzoek nodig is om de MRT concreet om te zetten in een robuust, op feiten gebaseerd hulpmiddel dat kan dienen als beslissingstool naast de koe. Desalniettemin biedt het werk in dit proefschrift al een stevige basis voor toekomstige inspanningen op dit gebied

Appendix

Appendix A.

Overview of absolute milk production per continent and country since the first reports to the FAO till 2021. For most country this reference year is 1961, the following symbol † and * indicate 1992 and 1993 as reference year, respectively.

Country	First reported Milk Production (tons)	Milk Production 2021 (tons)	% of Total continental production in 2021	Relative difference production (first record - 2021)	Dairy Cattle Population 2021	% of Total continental population in 2021	Relative difference in Dairy population (first record - 2021)
Africa							
Egypt	391 138	4 794 000	12.43	1125.65	1 735 327	7.18	199.19
Ethiopia *	750 000	3 866 052	10.02	415.47	8 796 506	36.41	131.49
Kenya	669 000	4 640 860	12.03	593.70	5 017 991	20.77	234.53
South Africa	2 557 162	3 825 000	9.92	49.58	1 001 481	4.15	0.15
United Republic of Tanzania	232 200	3 101 384	8.04	1235.65	7 605 976	31.49	424.01
Asia							
India	8 753 000	108 339 000	41.78	1137.74	57 600 000	60.37	178.71
China, mainland	600 000	36 827 000	14.20	6037.83	12 166 506	12.75	2 333.30
Pakistan	1 686 000	22 189 150	8.56	1216.08	15 192 000	15.92	699.58
Türkiye	4 830 000	21 370 116	8.24	342.45	4 022 907	4.22	89.76
Uzbekistan †	3 531 100	11 242 745	4.34	218.39	6 422 632	6.73	72.87
South & Central America							
Brazil	5 227 380	36 238 558	41.82	593.25	15 944 584	65.66	115.58
Mexico	2 327 000	13 237 209	15.28	468.85	2 642 246	10.88	164.22
Argentina	4 150 806	11 899 905	13.73	186.69	1 527 097	6.29	-33.31
Colombia	1 762 000	6 992 670	8.07	296.86	3 565 527	14.68	85.70
Chile	764 229	2 336 542	2.70	205.74	604 826	2.49	41.98
Europe							
Germany	24 796 400	32 506 910	14.32	31.10	3 832 720	21.97	-51.95
Russian Federation †	47 015 000	32 078 587	14.13	-31.77	6 394 372	36.66	-68.33
France	19 069 296	24 760 620	10.91	29.85	3 322 030	19.04	-53.47
United Kingdom of Great Britain and Northern Ireland	12 004 927	15 670 420	6.90	30.53	1 859 000	10.66	-55.90
Poland	12 770 997	14 881 110	6.56	16.52	2 035 200	11.67	-65.59
Northern America							
United States of America	57 019 696	102 644 900	91.28	80.02	9 442 400	90.59	-45.24
Canada	8 325 517	9 807 284	8.72	17.80	981 300	9.41	-67.15
Oceania							
New Zealand	5 217 000	21 886 376	70.66	319.52	4 804 582	77.01	149.10
Australia	6 277 000	9 067 000	29.27	44.45	1 384 000	22.18	-56.23


Appendix B.

Complete importance list for the 3 nextMILK models, Abbreviations: DIM = days in milk, TD = test day, X = lactation X, X+1 = lactation X+1, M = cumulative milk yield , PH = production and herd, P= production

Feature	Importance		
	FULL	Feature	FULL
DIM TD1 _{X+1}	0.1851	0.1867	0.1926
M305 _X	0.1285	0.1166	0.1283
kgTD5 _X	0.0597	0.0566	0.0654
kgTD4 _X	0.0459	0.0599	0.0651
Milk Maximum _X	0.0304	0.0384	0.0395
kgTD6 _X	0.0303	0.0380	0.0394
Days Dry _X	0.0266		
Season _{X+1}	0.0261	0.0331	0.0365
Δ Herd Average Days Dry _X	0.0245		
kgTD3 _X	0.0229	0.0282	0.0313
kgTD7 _X	0.0221	0.0281	0.0324
Season _X	0.0210	0.0255	0.0259
M75 _X	0.0200	0.0278	0.0319
Milk Minimum _X	0.0187		
Age At First Calving _X	0.0176	0.0225	0.0295
kgTD2 _X	0.0170	0.0250	
kgTD8 _X	0.0160		
Δ Herd Average Milk Minimum _X	0.0160	0.0231	0.0285
kgTD1 _X	0.0159	0.0224	
Δ Herd Average M305 _X	0.0156	0.0223	0.0270
Δ Herd Average Age at First Calving _X	0.0153	0.0236	0.0255
Δ Herd Average Milk Maximum _X	0.0148	0.0212	
Δ Herd Average Calving Interval _X	0.0145		
Δ Herd Average M21 _X	0.0141	0.0209	0.0261
Δ Herd Average Lactation Length _X	0.0138		
DIM TD1 _X	0.0135	0.0192	
Δ Herd Average M75 _X	0.0130	0.0189	
M21 _X	0.0130		
Δ Herd Average Days Open _X	0.0125		
Parity number _X	0.0123		
Days In Milk _X	0.0121	0.0182	0.0226
Days Open _X	0.0113	0.0170	0.0204
DIM TD2 _X	0.0108	0.0164	0.0203
DIM TD8 _X	0.0107	0.0163	0.0203
DIM TD6 _X	0.0102	0.0127	0.0158
DIM TD3 _X	0.0102	0.0157	0.0193
DIM TD4 _X	0.0102	0.0154	0.0189
DIM TD7 _X	0.0100	0.0152	0.0187
DIM TD5 _X	0.0099	0.0152	0.0189
Calving Interval _X	0.0076		

Appendix C.

Appendix C.1 Descriptive overview of the answers given by 45 farms in Belgium and the Netherlands in a survey relating to management practices implemented to prevent changes in body condition scores (BCS) during the dry period.

Question	Answer	Percent (n = 45)
Does the dry period consist of a far-off and a close-up group?		
	No	53.3
	Yes	46.7
How many days before the expected calving date are your cows dried off?		
	Median [min, max]	42 [30, 55]
Are there different basic diets during the dry period?		
	No (the same basic diet throughout the entire dry period, possibly with a different concentrate feed)	77.8
	Yes (far-off and close-up diet)	22.2
Do you feed straw to the dry cows?		
	No	37.8
	Yes, Chopped	40.0
	Yes, but not chopped	22.2
Is concentrate feed provided to the dry cows?		
	No	26.7
	Yes	73.3
Are leftovers from the lactating cows fed to the dry cows?		
	No	88.9
	Yes	11.1
Is the diet of the dry cows based on feed analyses?		
	No	28.9
	Yes	71.1
Which picture most closely matches the amount of residual feed present in the dry cow group at the time fresh feed is provided?		
	1) No, surplus	28.9
	2) Some surplus	60
	3) Too Much, surplus	11.1
		
How often is fresh feed provided to dry cows at lower temperatures (like during the winter)?		
	Multiple times per day	11.1
	Daily	40

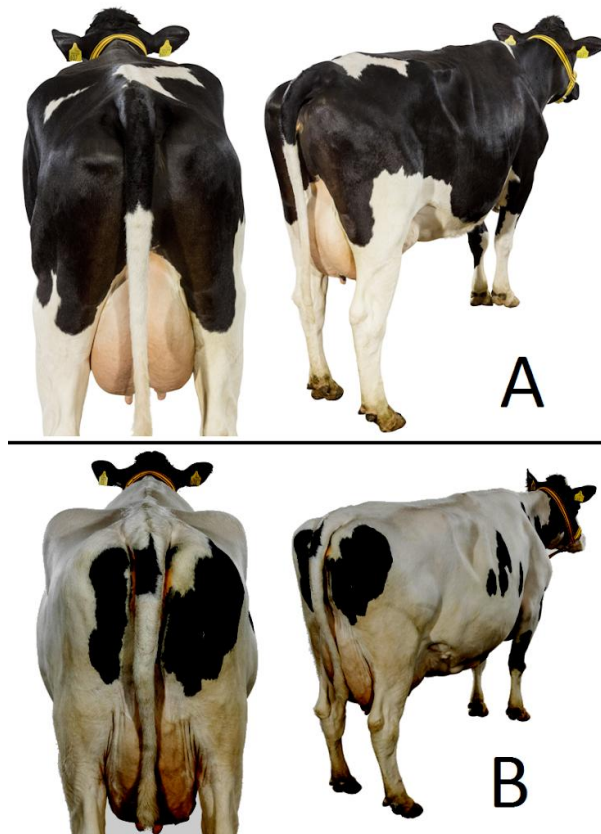
Every 2 Days	44.4
More than every 2 Days	4.4

How often is fresh feed provided to dry cows?

Multiple times per day	15.6
Daily	55.6
Every 2 Days	24.4
More than every 2 Days	4.4

How many cows on your farm are dried off with a similar body condition score (BCS)?

Even distribution of BCS A and B	26.7
Less BCS A than B	42.2
More BCS A than B	31.1



What percentage of all cows are dried off with a body condition score as shown in the picture?

High (> 15%)	15.6
Intermediary (10 -15 %)	53.3
Low (< 10%)	31.1



Appendix C.2 Descriptive overview of the answers given by 45 farms in Belgium and the Netherlands in a survey relating to management practices implemented to prevent environmental stress around the transition period

Question	Answer	Percent (n = 45)
At what point in the production cycle are the claws of your cows trimmed?		
	The claws are not trimmed at a fixed time in the production cycle.	31.1
	Around the time of drying off & During lactation (> 3 weeks after calving)	22.2
	Other Time points such as: during the dry period, around the time of calving, beginning of lactation (< 3 weeks after calving)	46.7
Apart from the (possible) fixed times in the production cycle, how many times a year are the claws trimmed?		
	min.1/year	17.8
	min.2/year	28.9
	min.3/year	4.4
	Only when lame	48.9
Do you work with a fixed time to dry off cows?		
	No (mainly based on expected calving date)	73.3
	Yes (e.g. weekly)	26.7
Presence of cubicles or free stalls in Dry group		
	Deep litter	15.6
	Deep litter, free stall	15.6
	Free stall	68.9
How often do you see cows standing in the cubicles as in the picture below?		
	Rare to never	60.0
	Regularly	6.7
	Sometimes	33.3
	Often	0



How often do you see dry cows lying diagonally in the cubicles as in the picture below?

Rare to never	82.2
Sometimes	15.6
Regularly	2.2
Often	0



How is the calving pattern on your farm?

Concentrated during a certain period of the year	24.4
Uniform throughout the year	75.6

What is the maximum number of cows that calved in the same month during the past year?

Median [min, max]	16 [6, 47]
-------------------	------------

Where do the cows calve?

In the dry group	24.4
In the lactating group	0
Separated calving pen	75.6

Where do the heifers calve?

In the dry group	15.6
In the lactating group	4.4
Separated calving pen	80.0

When are the heifers grouped with the cows?

Before calving (in dry group)	86.7
-------------------------------	------

Around calving	11.1
During their first lactation	2.2
After their first lactation	0

Is there a separate, secluded group for freshly calved cows (calving area not included)?

No	75.6
Yes	24.4

Do the lactating cows have access to outdoor areas?

No	24.4
Yes	75.6

How often do you see lactating cows standing in the cubicles as in the picture below?

Rare to never	35.6
Sometimes	46.7
Regularly	15.6
Often	2.2



How often do you see lactating cows lying diagonally in the cubicles as in the picture below?

Rare to never	57.8
Sometimes	28.9
Regularly	13.3
Often	0



What measures are taken to prevent heat stress? Fans In lactating group	
No	46.7
Yes	53.3
What measures are taken to prevent heat stress? Roof Insulation lactating group	
No	68.9
Yes	31.1
What measures are taken to prevent heat stress? Misting lactating group	
No	86.7
Yes	13.3
What measures are taken to prevent heat stress? Fans In Dry group	
No	71.1
Yes	28.9
What measures are taken to prevent heat stress? Roof Insulation dry group	
No	68.9
Yes	31.1
What measures are taken to prevent heat stress? Misting dry group	
No	100.0
Yes	0.0

Appendix C.3 Descriptive overview of the answers given by 45 farms in Belgium and the Netherlands in a survey relating to management practices implemented to mitigate the risk of infectious diseases.

Question	Answer	Percent (n = 45)
Do you use antibiotics when drying off?		
	No	4.4
	Yes, in all animals.	48.9
	Yes, but not in all animals (selective dry-off)	46.7
Do you take certain measures to reduce milk production before drying off ? (e.g. fewer milkings, feeding a different diet, providing less concentrate...)		
	In all animals	68.9
	Never	8.9
	In some animals	22.2
What is the maximum daily production (kg milk/day) at which cows are still dried off?		
	Median [min, max]	18 [3, 35]
Do you use teat sealers when drying off?		
	No	11.1
	Yes	88.9
Do you use a barrier dip when drying off?		
	No	51.1

Yes	48.9
Is the calving area sometimes also used as a sick bay?	
No	48.9
Yes	51.1
In which animals is monensin used to prevent transition diseases?	
All animals	6.7
Only at-risk animals (e.g. fat cows, cows from 3rd lactation...)	66.7
None	26.7
In which animals are glucose precursors used to prevent transition diseases?	
All animals	20.0
Only at-risk animals (e.g. fat cows, cows from 3rd lactation...)	55.6
None	24.4

Appendix C.4 Descriptive overview of the answers given by 45 farms in Belgium and the Netherlands in a survey relating to management practices implemented to monitor the transition period combined with the reported prevalences of the most common transition diseases.

Question	Answer (Abbreviation within model if present)	Percent (n = 45)
Are cows monitored during the first weeks after calving based on general impression?		
	All animals	86.7
	Only at-risk animals (e.g. fat cows, cows from 3rd lactation...)	13.3
	Never	0
Are cows monitored during the first weeks after calving based on Rumen Fill?		
	All animals	75.6
	Only at-risk animals (e.g. fat cows, cows from 3rd lactation...)	15.6
	Never	8.9
Are cows monitored during the first weeks after calving based on appetite (visually)?		
	All animals	64.4
	Only at-risk animals (e.g. fat cows, cows from 3rd lactation...)	24.4
	Never	11.1
Are cows monitored during the first weeks after calving based on temperature?		
	All animals	8.9
	Only at-risk animals (e.g. fat cows, cows from 3rd lactation...)	60.0
	Never	31.1
Are cows monitored during the first weeks after calving based on manual ketone test?		

All animals	24.4
Only at-risk animals (e.g. fat cows, cows from 3rd lactation...)	33.3
Never	42.2
Are cows monitored during the first weeks after calving based on cow sensors (e.g. activity meters, rumination sensor)?	
All animals	46.7
Only at-risk animals (e.g. fat cows, cows from 3rd lactation...)	0
Never	53.3
Are cows monitored during the first weeks after calving based on measurements on milk (e.g. fat content, conductivity)?	
All animals	55.6
Only at-risk animals (e.g. fat cows, cows from 3rd lactation...)	4.4
Never	40.0
What percentage of cows experienced any form of ketosis on your farm in the past year?	
Median [min, max]	6.3 [0, 40]
What percentage of cows experienced any form of milk fever on your farm in the past year?	
Median [min, max]	3.4 [0, 22]
What percentage of cows experienced any form of clinical mastitis (in the transition period) on your farm in the past year?	
Median [min, max]	3.5 [0, 13.4]
What percentage of cows experienced any form of retained membranes on your farm in the past year?	
Median [min, max]	5.6 [0.9, 13.4]
What percentage of cows experienced any form of metritis on your farm in the past year?	
Median [min, max]	3.5 [0, 15.2]
What percentage of cows experienced any form of abomasal displacement on your farm in the past year?	
Median [min, max]	1.5 [0, 7.4]

Appendix C.5 Descriptive overview of the answers given by 45 farms in Belgium and the Netherlands in a survey relating to management practices relating to the calcium metabolism and practices implemented to prevent milk fever.

Question	Answer (Abbreviation within model if present)	Percent (n = 45)
Do the dry cows have access to outdoor areas?		
	No	66.7
	Yes , low risk pasture (not grassland, unfertilized grassland)	22.2
	Yes, fertilized grassland	11.1
Do you feed grassland products to the dry cows (fresh grass, silage, or hay)?		
	No	33.3
	Yes	66.7

In which animals are calcium supplementations (bolus, infusion) used to prevent transition diseases?	
All animals	8.9
Only at-risk animals (e.g. fat cows, cows from 3rd lactation...)	80.0
None	11.1
In which animals is vitamin D injection used to prevent transition diseases?	
All animals	0
Only at-risk animals (e.g. fat cows, cows from 3rd lactation...)	24.4
None	75.6
Which animals are drenched to prevent transition diseases?	
All animals	0
Only at-risk animals (e.g. fat cows, cows from 3rd lactation...)	22.2
None	77.8
Which of the products below do you feed to the dry cows?	
Anionic salts	2.2
Anionic salts, dry cow minerals	26.7
Anionic salts, dry cow minerals, calcium binders	6.7
Dry cow minerals	64.4

Appendix D.

Table providing a descriptive overview of the meanM305 and MRT

HerdID	meanMRT	sdMRT	meanM305	ΔmeanM305
1	-8.46	5.11	8809	44.2
2	1.55	6.00	10609	-186.9
4	0.03	6.08	9445	5.7
6	-6.79	7.77	9362	-173.9
7	-3.32	5.42	9320	-162.3
9	-2.60	5.50	9699	55.4
10	-2.16	6.96	9306	-151.6
14	-8.61	9.45	8175	-1005.4
15	-1.20	5.54	10442	4267.9
17	-3.28	8.90	9938	78.9
18	1.31	6.14	9958	-524.3
19	-8.06	8.36	8804	-99.2
20	-1.63	5.71	9063	-126.0
21	-0.83	8.18	9990	283.5
22	3.71	4.08	10796	190.0
23	-5.30	8.90	8806	-71.0
25	-1.59	8.68	9410	-211.4
26	2.04	8.50	10721	197.4
27	-1.58	8.16	11300	-659.3
28	-1.92	6.38	10113	694.3
29	1.34	10.06	11671	-119.5
30	-9.51	8.00	6940	-2113.4
31	5.06	4.85	10612	-231.5
32	1.15	8.13	10381	-75.6
33	-4.01	6.64	8344	759.8
34	-0.28	6.41	10412	-52.1
35	0.84	5.83	10119	305.1
36	-0.31	6.43	10390	-301.7
37	-2.86	9.03	10069	-204.2
38	-6.70	6.59	9030	192.2
39	-2.94	9.30	10498	591.6
44	0.24	4.86	10876	803.3
45	-4.50	7.74	9988	-423.9