

# **Modeling and Forecasting Extended Lactations of Holstein Friesian Dairy Cattle**

by

Linaya Pot

A Thesis

presented to

The University of Guelph

In partial fulfilment of requirements  
for the degree of

Masters of Science

in

Animal Biosciences

Guelph, Ontario, Canada

© Linaya Pot, May 2020

## ABSTRACT

### MODELING AND FORECASTING EXTENDED LACTATIONS OF HOLSTEIN FRIESIAN DAIRY CATTLE

Linaya Pot  
University of Guelph, 2020

Advisor:  
Dr. John P. Cant

Lactation models were used to fit and forecast extended lactations to determine if cows suitable for extended lactations could be identified within their voluntary waiting period (VWP). Models were fitted to a dataset of daily milk yields (MY) ( $n = 651$ ) from extended lactations  $> 305$  days in milk (DIM). Models predicted 305-d MY after fits to data from only the first  $i$  DIM, where  $i$  was incremented from 30 to 300 DIM in weekly intervals. At least 90 DIM of data was needed by the model for a satisfactory fit. However, using 120 DIM of data was recommended for more accurate forecasting. Beyond 120-d there was a lessening impact of additional data on the improvement of fit statistics. Since a 120-d VWP is not optimal for all animals, it was suggested that model forecasts should be used as a tool to optimize lactation length only in suitable animals.

## ACKNOWLEDGEMENTS

I would like to start this off by thanking my parents – and for a number of reasons. First – thank you for being farmers. When asked to describe myself, one of the first things I proudly say is that I’m a farmer’s daughter. My years running wild on the farm, surrounded by animals everywhere, have left me with very cherished childhood memories. Being a farmer’s daughter taught me the value of hard work and kickstarted my desire to work with animals. Secondly, I also wanted to thank my parents for fostering a strong sense of independence and a thirst for knowledge during my childhood. It is these two things that have sustained me throughout my career.

I would also like to express gratitude for all the John’s in my life – first and foremost; Dr. John Cant. Thank you for walking beside me as I entered the world of modeling and being nothing but encouraging and uplifting throughout my entire project. Without you, none of this would be possible.

Dr. John Doelman, thanks for hiring me as a research assistant during my undergraduate degree. You showed me that research in the dairy industry can be both rewarding and fun. Dr. Jan Dijkstra, thank you for facilitating my semester abroad (which was the best semester of my Masters) and for creating that excellent lactation model which served me so well in my project.

A special shout-out goes to my dear friends, James Templeman and Dave Seymour. These two made my years in grad school as fun as it was educational. The development of our friendship throughout the years is one of most treasured experiences from my masters.

Thank you to my boyfriend, Connor Godfrey, who was with me through all the highs and lows of this project. Despite not knowing what I was talking about half of the time, you always

patiently listened to my rants and vents. Thank you for being a breath of fresh air at the end of the day, and always challenging me to think in new and different ways.

Last but not least, I am so grateful for all the opportunities and friendships I was given along the way. Thank you Carolina Reyes, Patty Kedzierski and Matthew Wells for being the best lab mates one could ask for. Thank you, Julie Kim and Kelly Nichols, for the invaluable and fundamental lab skills you taught me. Thanks John (again) for sending me to conferences and across the world to experience new things. I will also treasure the memories from these years.

## TABLE OF CONTENTS

ABSTRACT.....	ii
Acknowledgements.....	iii
Table of Contents.....	v
List of Tables.....	vii
List of Figures.....	ix
List of Abbreviations.....	xi
Chapter 1: Introduction & Literature Review.....	1
Background: The Historical Dairy Cow.....	1
Introduction: Modelling Extended Lactations.....	1
Literature Review.....	3
The Basic Shape of Lactation Curves.....	3
Biological Basis behind the Shape of the Curve.....	4
Factors Influencing the Shape of Lactation Curves.....	7
Voluntary Waiting Period.....	9
Advantages and Disadvantages of Longer Voluntary Waiting Period.....	10
Selective Suitability of Cows Highlights Needs to Individualize VWP.....	13
Modelling the Lactation Curve.....	15
Importance of Lactation Models.....	15
Classification of Lactation Models.....	16
The Development of Lactation Models.....	17
Research Rationale and Study Objectives.....	22

Chapter 2: Text and investigation .....	24
Introduction.....	24
Materials and Methods.....	26
Data Source .....	26
Statistical Analysis.....	27
Forecasting MY at 305 DIM.....	30
Results.....	33
Effect of Parity on Lactation Curves and Model Fits .....	33
Effect of Lactation Length on Lactation Curves and Model Fits .....	40
Lactation Curve and Parameters Estimates.....	42
Parameter Comparison.....	45
Analysis of Residuals.....	46
Relationship between early lactation information and 305-d MY Predictions.....	50
Probability of 305-D MY Predictions Exceeding Targets .....	57
Chapter 3: Discussion and Conclusion .....	61
Discussion .....	61
Conclusions.....	65
References.....	67

## LIST OF TABLES

Table 2.1 General description of the dataset.....	26
Table 2.2. Lactation curve models fitted to extended lactations.....	27
Table 2.3 General description of the of the data subset used for incremental fits .....	30
Table 2.3. Characteristics of lactations grouped into primiparous or multiparous groups <sup>1</sup> .....	35
Table 2.4. Comparison of mean model parameter estimates yielded from three lactation models fitted to individual lactations grouped by parity <sup>1</sup> .....	36
Table 2.5. Comparison of mean predicted lactation curve characteristics for primiparous and multiparous cattle <sup>1,2</sup> .....	36
Table 2.6. Mean goodness of fit criteria calculated from individual lactation curves grouped by lactation lengths fitted to Dijkstra, Wood and WL models .....	39
2.7. Lactation characteristics of lactations grouped together by lactation length <sup>1</sup> .....	41
Table 2.8. Comparison of mean lactation model parameters estimates yielded from three lactation models fitted to individual lactations grouped by lactation length <sup>1,2</sup> .....	42
Table 2.9. Comparison of mean predicted lactation curve characteristics from extended lactations grouped by length <sup>1,2,3</sup> .....	43
Table 2.10. Mean goodness of fit criteria calculated from individual lactation curves grouped by lactation lengths fitted to Dijkstra, Wood and WL models <sup>1</sup> .....	45
Table 2.11. Root mean square prediction error (RMSPE), log RSMPE, the decomposition of RSMPE into the mean bias, slope bias and random error components, mean absolute error (MAE) and the concordance correlation coefficient (CCC) calculated from individual 305-DIM (n=548) early lactation MY records increasing by 30-day increments.....	54

Table 2.12 The mean standard error of predictions ( $SE_p$ ) and mean prediction interval (P.I) size of 305-D milk yield predictions (n=548) resulting from fits $i$ days in milk (DIM) data to the Dijkstra et al. model calculated using individual animal and total herd standard deviations.....	55
Table 2.13 The mean probability of predicted MY (n=548) at 305 DIM (n=548) meeting or exceeding various target levels according to fits of the Dijkstra et al. model to MY data up to $i$ DIM.....	60
Table 2.14 Percent of cows (n=548) with observed 305-D MY over a certain target value (kg) compared to the percent of individual predictions over the target value based on probabilities of either 50, 75 or 95 % from fits of $i$ DIM data to the Dijkstra et al. model .....	60

## LIST OF FIGURES

<p>+Figure 1.1 The lactation curve divided into four physiological stages (based on the work of Grossman and Koops, 2003). Phase I representing a steep incline due to a peak in cellular proliferation rates. Phase II representing a more gradual ascending due to a peak in cellular proliferation rates. Phase II representing a more gradual ascending due to increasing synthetic capacity until peak has reached. Phase III representing a gradual declining phase caused by increasing apoptosis rates and Phase IV representing the steepest rate of decline caused by increased cell death rates due to pregnancy .....</p>	6
<p>Figure 2.1. Mean milk yields (observed) over days in milk plotted with lactation curves yielded from the Dijkstra (DJ), Wood (WD) and Wilmink (WL) models using the mean parameter estimates generated by model fits to individual to test day yields grouped by parity .....</p>	37
<p>Figure 2.2 Mean milk yields (observed) over days in milk (DIM) plotted with three lactation curves yielded from the Dijkstra et al, Wood and Wilmink models fitted to test day yields from lactations grouped by length: Group 1 (305-400 DIM; n =448), Group 2 (400-500 DIM, n=157) and Group 3 (&gt;500 DIM; n=37) .....</p>	44
<p>Figure 2.3 Mean daily residuals from individual lactation curves in the primiparous (n=203) or multiparous group (n=445) fitted to the Wood, Wilmink and Dijkstra et al. models.....</p>	48
<p>Figure 2.4 Mean daily residuals from lactation length Group 1 (305-400 DIM; n=448), Group 2 (400-500 DIM; n=157) and Group 3 (&gt; 500 DIM; n=37) fitted to the Wood, Wilmink and Dijkstra et al. models .....</p>	49
<p>Figure 2.5 Relationship between the log residual mean square predicted error (RMSPE) of predicted MY (n=548) at 305 DIM and amount of days in milk (DIM) data available to the DJ model.....</p>	54

Figure 2.6 Lactation curves generated by parameters obtained from the Dijkstra et al model fit to increasing amounts of early lactation data compared to lactation curve from the average observed daily milk production..... 56

**LIST OF ABBREVIATIONS**

CCC	Lin's concordance correlation coefficient
DIM	days in milk
DJ	Dijkstra et al
MAE	mean absolute error
MY	milk yields
NEB	negative energy balance
NMC	National Mastitis Council
PI	prediction interval
PT	time to peak
PY	peak yield
$R^2_{adj}$	adjusted coefficient of variation
RMSE	root mean square error
RMSPE	root mean square prediction error
RSS	residual sum of squares
DW	Durbin Watson coefficient
$s$	standard deviation
$SE_p$	standard error of the predictions
SSE	sum of squares error
TSS	total sum of squares
VWP	voluntary waiting period
WD	Wood
WL	Wilmink

## **CHAPTER 1: INTRODUCTION & LITERATURE REVIEW**

### **Background: The Historical Dairy Cow**

The Neolithic Revolution, in which hunter-gathers changed their lifestyles to that of settled agricultural farmers, was one of the most significant developments in human history. An integral element for the success of this transition was the domestication of agricultural animals, namely cows, sheep, goats and pigs (Scheu et al, 2015). Of these animals, cows were of particular importance due to their multi-use ability, including the provision of milk and meat for human consumption, hides and horns for crafting, manure for fuel and their ability to perform manual labour. As artificial selection progressed, certain breeds became specialized for milk production and became the forefathers for the modern dairy cows today.

These forefathers exhibited seasonal breeding and production patterns. Parturition would occur in early spring, aligning milk production to when food was found in abundance. When food sources declined later in the year and the calf no longer needed milk as its primary sustenance, the cow would cease milk production. Seasonal breeding would induce the cow to calve again the subsequent spring, creating a cycle of calving, lactation, a dry period and gestation that could be contained within a year.

However, dairy cows of today are remarkably different than their ancestors. In the past 50 years alone, milk production of the modern dairy cow has more than doubled. Due to high levels of milk production, problems from fitting high producing modern cows into this traditional lactation cycle are beginning to emerge.

### **Introduction: Modelling Extended Lactations**

Although modern intensive dairy production systems no longer rely on seasonal calving due to availability of feed sources year-round, the 12-month lactation cycle has remained the norm. It had been widely accepted that a lactation of 305 days, followed by

a 60-day dry period, was the most economic beneficial lactation length (Weller and Folman, 1990; Strandberg and Oltenacu, 1989; Inchaisri et al, 2011). To maintain this cycle length, the voluntary waiting period (VWP), which is defined as the time between parturition and when the cow is available for insemination, generally encompasses the first 60-90 days in milk (DIM).

However, nowadays cows are frequently being dried off while still producing significant quantities of milk, anywhere from 24- 30 kg/d (Chapinal et al, 2014). Since this represents a waste of valuable product, it is becoming a common practice to manage cows for an extended lactation. This allows them to continue milking until production levels have declined to a degree that it is no longer wasteful or uncomfortable for the cow to be dried off. This also has positive implications for improving fertility, feed efficiency, productivity and animal welfare.

Current research has indicated that not all cows are suitable for this management strategy. If an entire herd is managed for extended lactations, there is a decrease in overall herd production (Auld et al, 2007). Conversely, if only suitable cows are selected to be managed for longer lactations, there is an increase in milk production and improved economic returns (Knight, 2001; Mellado et al, 2016, Kolver et al, 2007; Butler et al, 2010). While many cows can potentially be suitable for longer lactations, the length of time that they are able to sustain their milk yields (MY) varies from cow to cow – some are able to milk for up to years, others only suitable milk for a few weeks longer than a traditional lactation length (Auld et al, 2007; Kolver et al, 2007). This indicates the need to tailor lactation lengths for the individual cow.

The successful optimization of individual lactation lengths requires knowledge of each cow's ability to maintain extended lactations. This can be obtained from forecasted MY, which can be generated from mathematical models based on the shape of lactation curves. However, many commonly used lactation models were created and optimized with traditional lactations

and may be inadequate when it comes to forecasting yields in extended lactations. To ensure a thorough understanding of the models, this literature review will discuss the shape of the lactation curve, factors that influence this shape and the benefits to both animal and producer that can come from the implementation of extended lactations. The suitability of modern cows to undergo extended lactations will be discussed, as well as the ability of lactation models to describe them. A short history of the models up to and including the models that will be used in this study will be overviewed to ensure knowledge and understanding of the complexity to model the different parts of a lactation curve.

## **Literature Review**

### **The Basic Shape of Lactation Curves**

To be able to accurately simulate a lactation, models need to be able to predict the evolution of MY throughout time. Despite a wide variation of production levels among dairy breeds, the visualization of milk production throughout a lactation tends to create a repeatable, sigmoidal curve – the ‘lactation curve’. It is created by an initial ascending phase occurring directly after parturition, which lasts until PY is reached. This is followed by a declining phase as the rate of milk production gradually decreases until the cow is dried off. Usually the initial ascending phase has a steeper slope caused by the quickly increasing MY in early lactation, as compared to a more moderate slope in the descending phase from the gradual decline in milk production as the lactation progresses. The repeatability of this shape is the basis for the mathematical modelling of lactation curve.

The two main factors impacting the intensity of the curvature are the peak milk yield (PY) and the persistency. PY is the maximum milk production that is achieved within the lactation cycle, which usually occurs within 35-60 days in milk (DIM). Persistency is a measure

of a cow's ability to maintain the PY throughout the remainder of the lactation. A high level of persistency is associated with a slow rate of declining milk production, and vice versa. A more persistent lactation is characterized by an absence of a high peak with more gradual slope in the declining phase.

There is an inverse relationship between persistency and PY. Cows with high persistency generally have lower PYs which are reached later in lactation (Ferris et al, 1985; Muir et al, 2004). On average, a cow with high persistency has lower milk production in early lactation and higher production in late lactation, when measured against a cow with similar levels of production and an average persistency (Cole and VanRaden, 2006). The opposite is true for cows with a lower persistency. Together, the PY and the persistency are the main determinants of the shape of the lactation curve.

### **Biological Basis behind the Shape of the Curve**

Both the ascending and descending components of the lactation curve are a direct consequence of physiological changes that occur within the cells of the mammary glands. Since the number and secretory capacity of these cells determines the PY and persistency of a lactation, changes in these cells can be used to elucidate the shape of the lactation curve (Capuco et al, 2001; Grossman and Koops, 2003; Vetharanim et al, 2003).

Based on the work of Grossman and Koops (2003), the physiological changes of the mammary gland throughout a lactation can be described in four phases. As seen in Figure 1, Phase I and II both represent the state of increasing MY that occurs in early lactation, with Phase I differentiated by having the steeper increasing slope. Knight et al (1984) found that two-thirds of the increasing yields in the first week of lactation can be explained by increasing cell number, suggesting higher rates of proliferation in Phase I than Phase II. Furthermore, one study found

that lactating mice experience a short-lived peak in cell-cycle activity in the mammary gland directly after parturition, as indicated by DNA content (Hadsell et al, 2007). Similarly, DNA content peaks around day 5 of lactation in goats and in the first week in dairy cows (Knight and Wilde, 1987). Therefore, it can be concluded that cell proliferation rates are highest shortly after parturition, accounting for the rapid increasing milk production in Phase I.

The remaining two phases, Phase III and IV, represent the declining portion of the lactation curve. The cause behind the declining MY is attributed to increased rates of apoptosis (Grossman and Koops, 2003; Vetharaniam et al, 2003; Hadsell et al, 2007). The apoptotic rate follows the opposite trend of proliferation; low at the beginning of lactation but increasing as lactation progresses (Hadsell et al, 2007). As the rate of apoptosis increases and eventually surpasses the rate of proliferation there is a subsequent decline in cell numbers that results in decreasing MY (Hadsell et al, 2007). This is supported by decreasing amounts of total mammary DNA and parenchyma DNA found after peak milk is reached (Capuco et al, 2001). Again, these last two phases are differentiated by the steepness of slope. Phase IV has a steeper decline in MY than Phase III, which can be explained by the growing fetus causing increased rates of mammary cell death (Grossman and Koops, 2003; Boutinaud et al, 2004).

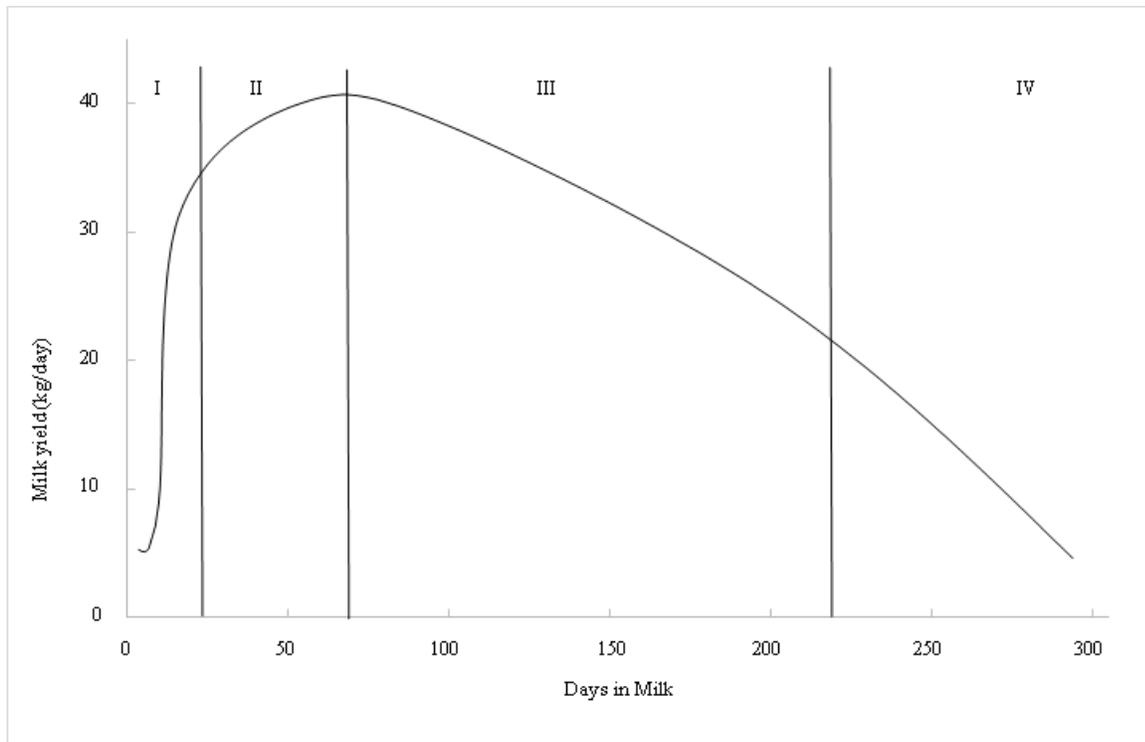


Figure 1.1 The lactation curve divided into four physiological stages (based on the work of Grossman and Koops, 2003). Phase I representing a steep incline due to a peak in cellular proliferation rates. Phase II representing a more gradual ascending due to a peak in cellular proliferation rates. Phase II representing a more gradual ascending due to increasing synthetic capacity until peak has reached. Phase III representing a gradual declining phase caused by increasing apoptosis rates and Phase IV representing the steepest rate of decline caused by increased cell death rates due to pregnancy

However, milk production is not just dependent on the number of cells but also the secretory capacity of each cell. Secretory activity has been found to peak several weeks after parturition and usually coincides with the timing of PY (Capuco et al, 2003; Mellenberger et al, 1973). After the cells have reached their maximum secretory activity, they tend to maintain this high level of activity. This is exhibited in dairy goats, where Wilde et al (1986) found that after PY was reached, MY declined while cellular activity remained unchanged. Consequently, the ascending phase of the lactation curve can be attributed to both increasing cell number and activity, but the decreasing phase is thought to be caused solely by changes in cell numbers.

Therefore, it can also be concluded that the rate of the cell death is a direct causative factor for a cow's persistency. Management strategies that improve mammary cell turnover rates should thereby improve the persistency of lactation.

### **Factors Influencing the Shape of Lactation Curves**

Since many factors can impact cell turnover, thereby influencing persistency or the magnitude of PY, it is important to have a thorough understanding of their impacts on lactation curves. This supports both the accuracy of lactation models as well as management decisions to promote a successful lactation. Two of the more well-known factors that heavily influence the shape of the curve are parity and pregnancy.

A cow's parity has a notable and very distinct effect on the shape of a lactation curve. On average, lactation curves from primiparous cows are markedly flatter, with lower peaks that occur later in lactation (Borman et al, 2004; Niozas et al 2019b). This flatness is caused not only by a moderate ascending phase, but more importantly, a slow and gradual decline in MY in the latter portion of lactation (Borman et al, 2004; Kocak and Ekiz, 2008). This slow decline is caused by a higher degree of persistency that is consistently seen among primiparous cows. Arbel et al (2001) quantified this difference, reporting that primiparous cattle had an average monthly decline in energy-corrected milk (ECM) of 1.5%, compared to a decline of 5.2% in multiparous cows. The excellent persistency in primiparous cattle allows them to retain high production levels reminiscent of their PY; despite their lower PY they tend to have similar milk production at 305 DIM as multiparous cows (Haile-Mariam and Goddard, 2008; Cole et al, 2009). Mellado et al (2016) calculated the average daily MY between 305 and > 450 days and found values of 29.9 and 31.9 kg for primiparous and multiparous cows, respectively. Due to this significant

difference in persistency, Arbel et al (2001) suggested that the criteria for determining the optimal lactation length will differ for primiparous versus multiparous cows.

It is not well understood why parity has such drastic effects the shape of the lactation curve. Since primiparous cattle are still growing at the time of parturition, perhaps conflicting demands between body growth and lactation reduces nutrient availability for the mammary gland (Wathes et al, 2007). Research also shows that multiparous cattle have a higher density of secretory cells, with these cells having higher metabolic activity (Miller et al, 2006). Although this explains higher MY seen in multiparous cattle, the reason for the differences in persistency is less clear. However, Miller et al (2006) also found that primiparous cattle had higher circulating IGF-I concentrations, which promotes cell survival. This is supported by the work of Safayi et al (2010) who found a higher expression of IGF-I genes during late lactation in primiparous goats than compared to multiparous goats. Both researchers theorized that these higher IGF-I levels support the survival of mammary cells throughout a lactation, encouraging higher persistency in primiparous dairy animals.

Pregnancy is also known to effect persistency. Cows that are simultaneously lactating and gestating have a lower degree of persistency than non-pregnant cows. This decrease in persistency contributes to a steep declining phase and low MY in late lactation (Olori et al, 1997; van Amburgh et al, 1997). Calculations to quantify the magnitude of this effect found that pregnant cows experience a culminative loss of 122 kg of milk within the standard 305-day lactation (van Amburgh et al, 1997). Lee et al (1997) found a more extreme response, with the difference in MY in a cow bred 100 days later amounting to 876 kg. Delaying the insemination another 100 days was also found to have a positive but less extreme effect of increasing the yields by another 174 kg (Lee et al, 1997). When analyzing how pregnancy affected the shape of

the lactation curve, Tekerli et al (2000) found that cows with a longer VWP experience a more gradual slope in the declining phase of their lactation curve.

Significant effects of pregnancy on MY is usually seen five months onward (Olori et al, 1997; Roche, 2003). Olori et al (1997) found that the effect of pregnancy was highest in mid lactation opposed to that of late lactation. Interestingly, Auran (1974) found pregnancy's effect on a cow's persistency was less pronounced in high producing cows. Perhaps this is related to the fact that high producing cattle are optimized to divert nutrients to lactogenesis versus other physiological processes.

As previously mentioned, the declining phase in the lactation cycle of dairy cows is thought to be caused by declining cell numbers. Pregnancy is thought to exacerbate this process by diverting nutrients that could be used for lactogenesis to support gestation, thereby increasing the rate of cell death and decreasing the cell's secretory capacity (Capuco et al, 2003; Knight and Peaker, 1984). Pregnancy also causes changes in the hormonal profile of the cow, promoting mammary gland regression (Akers, 2006).

Overall, this awareness of how parity and pregnancy influences lactation curve can support management decisions around extended lactation. Producers can be encouraged to extend the VWP for primiparous cattle, knowing they are more likely to have a high persistency to support a successful extended lactation. This delay in pregnancy will also support an extended lactation by postponing the negative impact of pregnancy on cow's persistency.

### **Voluntary Waiting Period**

The timing of pregnancy, as determined by the length of the VWP, not only impacts lactation length and milk production but also influences animal welfare, fertility, feed efficiency, herd turnover dynamics and economics.

### **Advantages and Disadvantages of Longer Voluntary Waiting Period**

Currently, over 50% of Holsteins in North America have lactations beyond 305 days, with the average increase in lactation length being estimated at 30 days (VanRaden et al, 2006; Tsuruta et al, 2005; Gonzalez-Recio et al, 2004). Although many of these lactations are a result of unsuccessful conception in early lactation, the views of producers are beginning to shift in favour of purposefully managing for extended lactations due to the many purported benefits.

Many high-producing dairy cows experience low conception and fertility rates thought to be caused by genetic differences developed from selective breeding and exacerbated by the timing of insemination. To maintain a traditional 305-lactation length, producers must inseminate cattle within a VWP of 60-90 DIM. This timing unfortunately coincides with a period of high metabolic stress caused by both a negative energy balance and rapidly increasing MY that is commonly seen in early lactation. This creates a high-risk period within the first two months of lactation which is believed to negatively impact conception (Bertilsson et al, 1997; Lucy, 2001). Extended lactations allow insemination to be delayed until after a cow's PY have been reached and the cow has recovered from a state of negative energy balance; this gives the cow more time to recuperate and move past the high-stress transition period. Multiple studies have shown the benefits of extended lactation on fertility measurements, such as higher successful artificial insemination rates, decreased inseminations per pregnancy and decreased need for anestrus treatments (Larsson and Berglund; 2000; Niozas et al 2019a; Schindler et al, 1991). Improved fertility has economic benefit to the farmer, due to decreased veterinary visits and insemination costs (Cole et al, 2009). This is reflected in a study done by De Vries (2006) in which a bioeconomic model determined it was economically beneficial to delay breeding by an extra 60 days for more persistent cows.

Other benefits are incurred by reducing the proportion of time a cow spends in this high-risk period of early lactation. When a cow is managed for longer lactations, she spends a larger percentage of her life lactating and decreased amount of time in the transition period. Since 50% of health disorders and 30% involuntary culling occurs within first 30-40 days of lactation (Roemer, 2011; Lehmann et al, 2017), decreasing the amount of calvings in a cow's lifetime limits the amount of time the cow spends at risk within the transition period. This improves overall herd health and animal longevity.

Another welfare benefit of extended lactations would be the reduction in the number of animals that are being dried off while still producing high quantities of milk. Not only is this uncomfortable for the cow, but it also increases her risk of contracting mastitis in the following lactation (Bates and Dohoo, 2016; Dossing, 1994). Rajala-Schultz et al (2005) found that for every 5 kg increase of milk production at dry off above 12.5 kg, risk of contracting mastitis in the subsequent parturition increased by 77%. Therefore, continuing milking a cow until her production levels have declined to a lower degree has direct improvements for cow welfare by improving cow comfort and reducing risk of mastitis.

From an economic basis, an extended lactation could also reduce feed costs due to a decreased dependency on expensive concentrates during peak lactation and increased utilization of forages during the extended descending phase. Using a bioeconomic model that evaluated feed efficiency for different lactation lengths, Dekkers et al (1998) found that cost of feed per unit of milk decreased for extended lactations of multiparous cows. This was attributed to dry matter intake over time more closely following milk production over time, allowing most of the energy requirements to be met by less costly home-grown roughages.

Another economic and potential environmental benefit of longer lactations is related to changes in herd dynamics. Longer lactations lead to fewer calvings per year and improved longevity decreases the need for replacement heifers. Less calves and replacement animals leads to decreased herd sizes, reducing both the amount of feed needed for non-lactating animals and greenhouse gas emissions.

To quantify these benefits, one study using an economic analysis showed extending a lactation by 4-6 months resulted in \$0.19 and \$0.12 in net return/cow per day of the extended lactation, for primiparous and multiparous cows, respectively (Arbel et al, 2001). If a farm decides to change its management practices to allow for extended lactations, it can expect to save costs from to the benefits previously mentioned; namely reduced fertility costs, reduced animal health care costs in early lactation, decreased antibiotics with lower mastitis rates, increased animal longevity and decreased feed per kg of milk costs (Borman et al, 2004; Knight, 1997; De Vries 2006).

However extended lactations also have risks associated with longer DIM. Cows in late lactation are at higher risk of high milk somatic cell count (SCC) which negatively impacts milk production (Hagnestam-Nielsen et al, 2009). High SCC can also lead to penalties for the producer if it is not remedied. Extended lactations are also associated with risks of obesity, as cows start gaining body weight later in lactations to replenish body storages lost in early lactation. Cows managed for these long lactations are at risk of gaining too much condition, which has negative downstream effects on subsequent parturitions and lactations (Marett et al, 2011).

### **Selective Suitability of Cows Highlights Needs to Individualize VWP**

The main risk of extended lactations would be a high incidence of involuntary dry-offs if cows could not maintain their milk production into late lactation, leading to an extended dry period rather than an extended lactation. Despite advantages in animal welfare, fertility and feed costs, the main determinant to ensure extended lactations can be a viable management strategy on commercial farms will depend on cows being able to produce similar or higher MY in an extended lactation as compared to a conventional one.

Research has shown variable responses in total MY when either groups or herds of cows are managed for extended lactations. Some studies found that increased lactations lengths lead to decreased annualized MY and economic losses; Inchaisri et al (2011) found with a simulation study that extending the VWP beyond 6 weeks was associated with economic losses and Auldist et al (2007) found a negative relationship between lactation length and annual production of milk and milk solids. Niozas et al (2019b) found a numeric decrease in kg of milk per DIM in the extended lactation group compared to the standard lactation group. Decreased yields could be caused by fewer cows being in the peak lactation phase, during which milk production is markedly high. Conversely, Lehmann et al (2016) found that milk production per DIM was comparable between cows of standard length and lactation up to 17-19 months. Other studies have found a positive relationship between MY and lactation length, however these studies also implemented management strategies known to promote high persistency, such as increased milking frequency, nutritional support and bST administration (Osterman and Bertilsson, 2003; van Amburgh et al, 1997; Kolver et al, 2007; Mellado et al, 2016). Despite the positive relationship found among these studies, the intensity of implementing these management strategies to support extended lactations could make a producer hesitant of applying them on

commercial farms. Overall, because of the difference in methods and production systems among the studies, it is hard to draw a clear conclusion about the effect of extended lactations on farm MY. Despite the variation in results, one conclusion is apparent: when entire herds are managed for extended lactations, there is a significant portion of cattle who will not be able to sustain milk production beyond 305 DIM (Auldist et al, 2007; Bertilsson et al, 1997; van Amburgh et al, 1997).

However, when individual cows are selectively identified to be managed for longer lactations, as opposed to entire herds or groups, it results in increased production and improved economic returns (Arbel et al, 2001; Knight, 1997; Mellado et al, 2016, Kolver et al, 2007; Butler et al, 2010; Inchaisri et al, 2011). Studies were uniform when concluding which cows are suitable for extended lactations: high producing cattle that exhibit good persistency with PY that occur later in lactation (Nemeckova et al, 2015; Arbel et al 2001; Kolver et al, 2007).

Additionally, one study found that cattle with North American genotypes were more suitable to extended lactations, with 48% of these cattle milking until a final dry off date of 600 DIM as compared to 14% of cows with New Zealand genotypes (Kolver et al, 2007). Another study using Monte Carlo simulations identified that primiparous cattle have increased profitability when the VWP was increased by 28 days, with the opposite result being found for multiparous cows (Stangaferro et al, 2018).

Overall this indicates that there is a wide variation among cows regarding their ability to maintain successful production levels throughout an extended lactation (Bertilsson et al, 1997; van Amburgh et al, 1997) and their suitability is determined by many factors such as the shape of their curves, their genotype, parity and persistency. Even in countries with high average milk production levels, this large variation in suitability of cows remains (Haile-Mariam and Goddard,

2008). This implies that to maximize herd production, optimal breeding times need to be determined for individual cows.

Breeding decisions need to be made within early lactation to optimize lactation length on an individual cow basis. This highlights the importance of forecasting MY, as it provides insight into late lactation MY to allow a producer to determine when it is the most beneficial time to breed individual cows. Models can be used to identify high-producing cows and/or cows with high persistency to be selectively managed for extended calving intervals. This requires an accurate and flexible model that can predict extended lactations. Haile-Mariam and Goddard (2008) suggested that information from early lactation can help identify which cows will continue to produce at high levels if managed correctly. This suggest that one can use early lactation information as an input for models to help predict late lactation yields.

### **Modelling the Lactation Curve**

#### **Importance of Lactation Models**

The use of mathematical models to describe the lactation curve in dairy cows has been a successful application of modeling within the dairy industry, providing a useful tool for nutritionists, physiologist, biologists, geneticists and producers alike. The general objective of lactation models is to predict daily MY throughout a lactation while minimizing errors of prediction, which helps to reveal the basic pattern of milk production from gestation to dry-off (Olori et al, 1999).

For producers, models can be used as a tool to support management decisions such as feeding and breeding strategies for the herd. Models can describe seasonal patterns and help visualize yearly MY trends, supporting decisions such as breeding for genetic improvement, profit evaluation and yearly feed planning (Collard et al, 2000; Tekerli et al, 2000; Jensen, 2001). It is not only a tool for total herd management but also is a tool for managing individual cows.

Models can help identify outliers, either for selective culling on poor production or breeding high-production cows for genetic improvement. Estimates yielded by these models can also help assist producers with breeding decision in early lactation (Jensen, 2001). All these possible utilizations of models highlight the importance of a model to be not only accurate with its predictions, but also flexible to account for variation in individual cows.

### **Classification of Lactation Models**

Since the beginning of their development in 1920, many different types of lactation models have been created. These models can be categorized into two general groups; empirical and mechanistic models. Empirical models are created using mathematical data from lactations, resulting in parameters that simply describe the shape of lactation curves, but not its underlying biological mechanisms. Conversely, mechanistic models are characterized by their more complex theoretical assumptions that describe the causative biology behind the curves. The parameters of mechanistic models provide a physiological explanation of the ongoing cellular process of the mammary gland during the different phases of the lactation curve.

Both empirical and mechanistic models have their own distinct advantages and disadvantages. Due to the lack of fundamental physiological theories in empirical models, they tend to be simpler and more practical, allowing for rapid convergence and good fitness for a diverse range of lactations (France and Thornley, 1984). The main criticism of empirical models is their restriction of having a pre-determined curve for lactation, as well as the inability to explain parameters in relation to mammary gland physiology (Macciotta et al, 2005). Also, since these models are based on mathematical optimization and not theoretical assumption, the results of the models are limited in how they can be extrapolated and applied (France and Thornley 1984; Dijkstra and France 1995).

Conversely, some of the more complex mechanistic models are difficult to accurately fit to lactation curve, as well as tending to be over-parameterized (Lopez et al, 2015; Pollott, 2000). On average, mechanistic models tend to have a poorer fit than their empirical counterparts (Pollott, 2000). However, since their parameters are of biological significance, inferences can be made based on how various factors affect the model parameters. Choosing between models will depend on weighing on the model's ease of application and accuracy of fits against the need for biological interpretation

### **The Development of Lactation Models**

The initial lactation models were all empirical models. Preliminary accounts of applying math to describe lactation curves occurred in 1887, when Sturtevant published a study relating MY to time post calving. The first mathematical explanation of MY throughout a lactation was created when Brody et al (1923) proposed a model consisting of an exponential function to describe the declining phase of a lactation:

$$M_t = ae^{-ct}$$

where:

- $M_t$  = daily MY at time  $t$ ;
- $a$  = parameter representing estimated initial MY at day 1 of lactation;
- $c$  = declining slope parameter.

The main limitation of this model was that it failed to account for the initial ascending portion of the lactation curve; rather it simplified the sigmoidal shape of lactation into a monotonically decreasing curve. Therefore, this model was limited to explaining only the declining portion of lactation instead accurately depicting a lactation as a whole. The next lactation model that was proposed attempted to improve on this flaw by including a parameter relating the initial ascending phase of lactation (Brody et al, 1924). Their improved model

consisted of the difference of two exponential functions; one of the ascending and one of the descending phase:

$$M_t = ae^{-bt} - ae^{-ct}$$

This time including a new variable,  $b$ , to represent the ascending slope. Although the inclusions of the initial ascending phase allowed the model to be fitted to a whole lactation, the exponential functions were unrealistic since it worked under a false assumption that initial MY is equal to zero. This led to the model to underestimate mid-lactation yields and overestimate late lactation yields (Cobby and Le Du, 1978).

The Brody model was followed by a parabolic exponential model proposed by Sikka (1950):

$$M_t = ae^{(bt-ct^2)}$$

This model hoped to improve on the Brody model by limiting the ascending phase only to the initial phase of the lactation, like that in real lactation. However, this model yielded a bell-shape curve, creating symmetry around PYs. This restrictive shape led to poor fits, particularly for multiparous cows. Since primiparous cows have a peak occurring much later in lactation with both a moderate ascending and descending slope, this model still first primiparous cows reasonably well. Building on this model, Nelder (1966) created an inverse exponential polynomial model:

$$M_t = \frac{t}{(a + bt + ct^2)}$$

Not only did this model have an improved fit compared to the Sikka model, but it also allowed for calculations of characteristics influencing the shape of the curve, namely PY and PT. The ability to derive calculations from the model parameters to describe lactation characteristics

represented a useful tool for the study of lactation curves and was duplicated in later models. However, even with the improved fit, as compared to the Sikka model, it was still found to under predict MY around PT (Olori et al, 1999). The first lactation model to combat this problem and represent lactation curves with high accuracy was the Wood (WD) model (1967):

$$M_t = at^b e^{-ct}$$

Where:

- M<sub>t</sub> is the yield in week *t*;
- a* = scaling factor associated with the average yield;
- b* = inclining slope parameter up to PY;
- c* = declining slope parameter after PY.

This non-linear model, often referred to as the incomplete gamma function, modeled the standard lactation curve shape as a product of constant, a power function and an exponential decay function. This model had a high goodness of fit on a wide range of lactations and is relatively easy to compute. Each of its parameters allow for a simple description of their impact on the shape of the curve; *a* is related to the production level of individual animals; *b* quantifies the initial ascending phase and *c* the rate of decline after PY. It was also the first model that could fit a wider range of lactation curve shapes. As with the previous Nelder (1966) model, equations using the model parameters were derived to calculate additional information on the lactation curve, such as calculation of PY, PT, persistency and a measure of total MY at any given time.

Although this model was a distinct improvement from previous models, it still had it flaws. Using monthly test milk records instead of weekly resulting in poor fits (Kellogg et al, 1977). It also tended to overestimate MY in early and late lactations, while underestimating MY during peak production (Cobby and Le Du, 1978; Grossman and Koops, 1988). Since the basic

structure involves multiplication of the parameters, this leads to high correlation between parameters which decreases the model's flexibility (Macciotta et al, 2005). Although it was one of the first models to be commonly used, it still is a popular model due to its flexibility and accuracy (Scott et al, 1996).

After the 1980s, modelling in dairy cows shifted in focus from modelling curves in groups of cows to individual cows. Modelling individual lactations require more flexible models to deal with the larger amount of variation. To deal with these new model requirements, Wilmink (WL) (1987) proposed a combined exponential and linear model:

$$M_t = a + be^{kt} + ct$$

Where

- a = scaling factor relating to initial MY;
- b = rate of variation in initial ascending part of the curve;
- c = the slope of the straight line (declining phase;)
- k = fixed value.

The parameters of the WL model are an additive function, as opposed to the multiplicative function of Wood, making the model less rigid in comparison. Indeed, this was the original model used for describing lactation curves in the official genetic evaluation program for Canadian Dairy (Schaeffer et al, 2000). Similar to the WD model, the parameters can be used to easily explain the lactation curve shape by calculating values to PT and PY.

Ali and Schaeffer (1987) similarly created a new model in response to the need for a flexible model for individual curves, and proposed a new a polynomial regression model:

$$M_t = a_0 + a_1X + a_2x^2 + a_3 \log\left(\frac{1}{x}\right) + a_4 (\log(1/x))^2$$

Although this model had the flexibility to fit a wide variation of lactation curves, its parameters are not easily interpreted. These two models were the transitioning models between the limited

early models and more complex recent models. Since then, models continued to be developed which utilize more mathematical functions, such as splines and Legendre polynomials. Some of these models include the Ali-Schaeffer curve (Ali and Schaeffer, 1987), the multiphasic model (Grossman and Koops, 1988), the Guo-Swalve curve (Guo and Swalve, 1995), orthogonal polynomials (Olori et al., 1999), and natural cubic splines (White et al., 1999). Each of these models had their own advantages and disadvantages. However, the majority of these models being proposed were still empirical models of varying complexity. While all these empirical models were being developed, concurrent research was being done to elucidate the mammary gland cellular changes that occur throughout a lactation. This research would become the basis for the development of mechanistic lactation models.

The first mechanistic model was created by Neal and Thornley (1983). The model was based on a compartmentalized representation of the mammary gland, with the different compartments being made up of undifferentiated cells, differentiated cells arising from cell division and a compartment consisting of alveoli responsible for the milk production. As the rates of cell proliferation and cell apoptosis changed as lactation progressed, the number of cells in each compartment would also change – subsequently influencing the model's predicted milk yields. Although it was an accurate depiction of the mammary system, the model itself was complex and therefore limited for practical applications.

The first accurate mechanistic model that could be easily computed was the Dijkstra (DJ) et al model (1997). The model is based on the theory that cell division in the mammary gland experiences exponential trend among species, starting during pregnancy and extending into the first few weeks of lactation. Overall, the model is a set of differential equations using four parameters to represent cell proliferation and cell death during lactation:

$$M_t = M_0 \exp\{\mu_t[1 - \exp(-k^2t)]/k^2 - \lambda\}$$

Where

$M_0$  = theoretical initial milk production (kg day<sup>-1</sup>);  
 $\mu T$  = specific rate of mammary cell proliferation;  
 $k$  = decay parameter; and  
 $\lambda$  = the specific rate of mammary cell death.

This non-linear parametric function describes the lactation shape in relation to changes in mammary secretory cell numbers as determined by the opposing rates of cell death and proliferation throughout the lactation cycle. This last model has been shown to have high goodness of fit, as well as the ability to model lactations of various lengths (VanRaden et al, 2006). The largest advantage to this model is its ability to explain various characteristics of lactations curves relative to changes in mammary cell number through its mechanistic parameters.

Since the Dijkstra model, more mechanistic models have been introduced. The Pollott (2000) model in its basic form is a 6-parameter model which also uses changes in mammary cell turnover to explain the lactation curve. Grossman and Koops (2003) also introduced a multiphasic model with 13 parameters. However, the addition of more parameters adds computational difficulties when the models are fit to a wide range of lactation curves. Additionally, the accuracies of these models are comparable to that of the Dijkstra model (Val-Arreola et al., 2004; Dematawewa et al, 2007). Therefore, among the mechanistic models, the Dijkstra model is recommended for practical use

### **Research Rationale and Study Objectives**

The purpose of this thesis is to take existing lactation models and determine the amount of early lactation data needed to accurately predict late lactation MY. This information will be

used to help optimize lactation lengths on an individual cow basis. To ensure that models can accurately predict extended lactations, three of the most commonly used lactation models will be compared to ensure their ability to accurately fit extended lactations

The explicit objectives set out in this thesis are:

1. To compare the ability of the most commonly used, practical and accurate models from current literature to model extended lactations from both primiparous and multiparous cows under a variety of lactation lengths, and
2. To determine the amount of early lactation data needed to accurately forecast late lactation MY to determine if a producer can identify animals suited to longer lactations within their VWP

## **CHAPTER 2: TEXT AND INVESTIGATION**

### **Introduction**

Although lactations of commercial dairy cattle have traditionally lasted 305 days, there has been a shift in the past decade to extend lactation lengths well beyond this. The main incentive to terminate one lactation and start another is the return to the high levels of daily milk production that occur during early lactation. However, modern cows are frequently being dried off while still producing significant quantities of milk, anywhere from 25 to 35 kg/d (Dingwell et al, 2001). An extended lactation allows cows to continue milking until production levels have declined to a degree that it is no longer a lost economic opportunity for the producer or uncomfortable for the cow to be dried off (Rajala-Schultz et al., 2018; Kolver et al, 2007). Benefits associated with extended lactations include a reduction in the amount of time spent in the high-risk phases around calving and PY (Lehmann et al, 2017), decreased numbers of calves, decreased insemination costs, greater cow longevity within a herd and improved animal welfare (Knight, 1997).

Research has shown that not all cows are suitable for extended lactation. If an entire herd is managed for extended lactations, there may be a decrease in overall herd production (Auld et al, 2007). Conversely, research has found that if only suitable cows within the herd are selected to be managed for longer lactations, there is increased total milk production and improved economic returns (Knight, 2001; Mellado et al, 2016, Kolver et al, 2007; Butler et al, 2010; Inchaisri et al, 2011). Cows best suited for extended lactations are those that exhibit good persistency and with high PY that occur later in lactation (Nemeckova et al, 2015; Arbel et al 2001; Kolver et al, 2007).

Implementation of an extended lactation strategy on farm involves delaying breeding for selected cows beyond the typical 60-90 day VWP. Thus, cows that are expected to be producing milk above a certain target value at 305 DIM must be identified during early lactation. Future MY can be forecasted with mathematical models. However, many commonly used lactation models were created and optimized for traditional lactation lengths and their ability to fit extended lactations needs to be evaluated. Therefore, the first purpose of this study was to compare three commonly used empirical and mechanistic models to find one that accurately fitted daily MY across a range of lactation lengths. The WD, WL and DJ models were selected for analysis due to their small number of parameters, frequency of use within literature, and flexibility and high accuracy to fit a wide range of lactation shapes. The second goal of our study was to assess the ability of the DJ model to forecast MY at 305 DIM from early lactation MY data, to determine the amount of data needed to identify animals suitable for longer lactations within the VWP.

## Materials and Methods

### Data Source

The data used in this analysis came from lactating Holstein-Friesian cattle housed in a free-stall barn at the Trouw Nutrition Ruminant Research Centre, located in Boxmeer, the Netherlands. Cattle were milked twice daily and fed ad libitum. The data set consisted of cow ID, parity, days in milk (DIM), morning and evening milk weights, daily concentrate and forage dry matter intakes (DMI), and body weights collected between June 2011 and October 2015. Lactation length varied from 105 to 780 d.

From the original dataset, a subset of data was selected based on the following criteria: the maximum day in milk must exceed 305 days, the amount of daily records per lactation must exceed 150 and daily milk records must begin at a maximum of 10 days post-partum.

Table 2.1 General description of the dataset

Number of cows	286
Number of complete lactations	651
Number of 1 <sup>st</sup> Parity Animals	203
Number of Multiparous Animals	448
Total Test-day Records	289386
Records per lactation $\pm$ SD	360.34 $\pm$ 91.34
Average $\pm$ SD daily milk yield (kg)	29.66 $\pm$ 9.052
Average $\pm$ SD days in milk (d)	382 $\pm$ 64.8
Mode of maximum days in milk (d)	350
Maximum lactation length (d)	719
Average $\pm$ SD 305-d milk yield (kg)	21.55 $\pm$ 6.524

## Statistical Analysis

The three lactation models described in Table 2.2 were fitted to individual lactation curves using the non-linear regression procedure with Marquardt method from SAS 9.4 (PROC NLIN; SAS Institute, 2008). Initial search grids for the parameters were based on ranges of the parameters found in literature, with bounds put in place for the DJ model to ensure the parameters remained biologically plausible;  $a$  greater to or equal to zero, and  $b$  and  $b_1$  not equal to zero. Convergence was determined based on changes in the residual sum of squares (RSS) between successive iterations, with the intent to minimize total RSS. Convergence was declared when changes per iteration yielded differences in the relative offset less than  $10^{-6}$ .

Table 2.2. Lactation curve models fitted to extended lactations

Model reference	MY <sup>1</sup> [1]	PT <sup>2</sup> [2]
Wood, 1967	$at^b e^{-ct}$	$b/c$
Wilmink, 1987	$a + be^{-0.05t} + ct$	$\ln[c/(0.05b)]/-0.05$
Dijkstra et al., 1997	$ae^{b(1-e^{-b_0t})/b_0-ct}$	$\ln(b/c)/b_0$

<sup>1</sup> MY is milk yield (kg/d),  $t$  is DIM, and  $a$ ,  $b$ ,  $b_0$  and  $c$  are parameters.

<sup>2</sup> PT is time to peak yield (d)

MY and PT were calculated using Equations 1 and 2 in Table 2.2, and PY was calculated by using values from PT entered into Equation 1. Because the post-peak decline in MY does not follow the same pattern in all 3 models, an overall persistency was calculated from curve fits as the difference in MY between PT and the end of lactation ( $t_{max}$ ), divided by the time interval:

$$persistency = \frac{(MY_{t_{max}} - PY)}{(t_{max} - PT)} \quad [3]$$

Total 305-d production was calculated as the sum of MY between 0 and 305 DIM, with missing daily MY values linearly interpolated for each lactation. Daily milk production at 305 DIM was estimated as the mean of daily MY from 303 to 307 DIM.

Goodness of fit for the three models was assessed for each lactation curve with the square root of the mean square prediction error (RMSPE):

$$RMSPE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \quad [4]$$

where  $n$  = number of observations within the lactation,  $O_i$  are observed MY values and  $P_i$  are predictions. RMSPE is indicative of the standard deviation of the unexplained variance, with a smaller RMSPE indicating a better model fit.

The adjusted multiple coefficient of determination ( $R^2_{adj}$ ) was also calculated to analyze goodness of fit between models:

$$R^2_{adj} = 1 - \frac{(n-1)}{(n-x)} \times (1 - R^2) \quad [5]$$

where  $R^2$  is the multiple coefficient of determination, calculated by:

$$R^2 = 1 - \left(\frac{RSS}{TSS}\right) \quad [6]$$

where RSS = residual sum of squares, TSS = total sum of squares,  $n$  = number of observations and  $x$  = number of parameters in the model.  $R^2$  is the amount of variance of daily MY explained by the lactation model, while the adjusted  $R^2$  allows for comparison between models with different numbers of parameters. Values for  $R^2_{adj}$  are between 0 and 1, with the fit of the model being satisfactory if  $R^2$  is close to 1.

To test for the presence of first-order autocorrelation among residuals, the Durbin-Watson coefficient (DW; Durbin and Watson, 1950) was calculated from the daily residuals from each lactation curve fit as

$$DW = \frac{\sum_i^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad [7]$$

where  $e_i = O_i - P_i$ . The autoregressive procedure in SAS (PROC AUTOREG with  $dw = 1$  and  $dwprob$  options) provided the DW statistic, autocorrelation coefficient, and the associated probabilities. The test statistic can range from 0 to 4, with a value of 2 indicating no autocorrelation and values toward 0 indicating positive correlation among consecutive residuals. For this study, a DW statistic of less than 1.5 was considered a positive correlation.

After models were fit to individual lactation curves, parameters and statistics were averaged and compared between primiparous and multiparous groups, and between lactations grouped by maximum DIM. Differences between group means were identified with Student's t-test.

## Forecasting MY at 305 DIM

The DJ model was used to predict MY at 305 DIM for each individual lactation after fits to data from only the first  $i$  DIM, where  $i$  was incremented from 30 to 300 DIM. Only fits that converged were used, creating a new dataset of 548 lactations (Table 2.3).

Table 2.3 General description of the of the data subset used for incremental fits

Number of cows	275
Number of complete lactations	548
Number of primiparous cows	167
Number of multiparous cows	371
Test-day Records	201021
Records per lactation $\pm$ SD	383 $\pm$ 65.7
Average $\pm$ SD daily milk yield (kg)	29.48 $\pm$ 9.042
Average $\pm$ SD maximum days in milk (d)	385 $\pm$ 65.6
Mode of maximum days in milk (d)	350
Maximum lactation length (d)	719
Average $\pm$ SD milk yield at 305 DIM (kg)	21.55 $\pm$ 6.524

Errors in prediction of MY at 305 DIM across all 548 lactations were assessed with RMSPE. The RMSPE was further decomposed into error due to a mean bias of prediction, error due to deviation of the regression slope from unity (slope bias) and error due to random variation (random error), as per Bibby and Toutenburg (1997). The mean bias, slope bias and random error were calculated as:

$$\text{Mean Bias} = (\bar{P} - \bar{O})^2 \quad [8]$$

$$\text{Slope Bias} = ((SD_P - (R)(SD_O))^2 \quad [9]$$

$$\text{Random Error} = (1 - R^2)(SD_O^2) \quad [10]$$

where  $\bar{P}$  and  $\bar{O}$  are the predicted and observed means, respectively,  $SD_P$  is the standard deviation of the predictions,  $SD_O$  is the standard deviation of observations and  $R$  is the Pearson correlation coefficient.

To assess both the accuracy and precision of the models, Lin's (1989) concordance correlation coefficient (CCC) was calculated as

$$CCC = R \times C_b \quad [aa]$$

with  $R$  being the Pearson correlation coefficient and  $C_b$  being the bias correction factor. The  $C_b$  variable was calculated as

$$C_b = \frac{2}{(v + (\frac{1}{v}) + \mu^2)} \quad [12]$$

where

$$v = SD_o / SD_p \quad [13]$$

$$\mu = \bar{P} - \bar{O} / (SD_o)(SD_p)^{\frac{1}{2}} \quad [14]$$

The CCC value helps quantify a measure of reliability based on covariation and correspondence from the model fit. For the CCC, the ideal value of 1 indicates perfect model accuracy and precision.

Due to the presence of large outliers, the mean absolute error (MAE) was calculated as

$$MAE = \frac{\sum_{j=1}^{538} |O_j - \hat{P}_j|}{n} \quad [15]$$

where  $O_j$  is observed MY at 305 DIM for curve  $j$  (estimated as the average MY from 303 to 307 DIM) and  $\hat{P}_j$  is predicted MY. Although it is similar to RMSPE in that it reports average prediction error on the same scale as the dependent variable and that a smaller value is indicative of an improved fit, the MAE does not place heavier weight on larger residuals through squaring.

Therefore MAE is a closer estimator of the expected prediction error on average that is not as heavily influenced by larger outliers.

When analyzing individual predictions, 95% prediction intervals (PI) were calculated for the forecasted MY at 305 DIM using

$$\hat{P}_j \pm (t_{crit})(SE_P) \quad [16]$$

Where the standard error of the prediction ( $SE_P$ ) at 305 d was calculated with

$$SE_P = s \sqrt{1 + \left(\frac{1}{n}\right) + \left(\frac{(O_j - \bar{O})^2}{s}\right)} \quad [17]$$

Both the PI and the  $SE_P$  were calculated twice; first using individual animal standard deviation ( $s$ ), as well as total herd  $s$ . Total herd  $s$  is a value that is more likely available to the average producer, who would not be likely to be able to account for deviation within individual animals.

To account for uncertainty in prediction, the probability that MY at 305 DIM would be above a certain target  $X$  (kg/d) for each lactation curve was estimated from the cumulative  $t$  distribution with  $n - 4$  degrees of freedom to account for the 4 parameters of the DJ model:

$$\Pr(y \geq X) = 1 - TDIST\left(\frac{X - P}{s_{pred}}, n - 4\right) \quad [16]$$

where  $P$  is the predicted MY at 305 DIM. The standard deviation of the prediction,  $s_{pred}$ , was calculated as  $s\sqrt{1 + h}$  where  $h$  is the tangential leverage outputted by PROC NLIN of SAS (SAS Institute Inc.) and  $s$  is the standard deviation estimated as RMSPE or MAE adjusted by  $n - 4$  instead of  $n$ .

## Results

### Effect of Parity on Lactation Curves and Model Fits

For this portion of the analysis, lactations were grouped into categories based on parity, with the first group consisting of primiparous lactations and the second consisting of second and greater lactations. Table 2.3 shows that mean daily MY and total 305-d MY were significantly lower in the primiparous group than multiparous group, but MY at 305 DIM was significantly higher in the primiparous group by almost 2 kg/d.

Table 2.4 shows the mean parameter estimates calculated from model fits to individual lactation curves within each parity group. The rise to PY is represented by parameters  $a$ ,  $b$  and  $b_0$ . All three models produced an MY scaling ‘ $a$ ’ parameter that was 22 to 29% lower, on average, in the primiparous group than multiparous, caused by lower MY values. Similarly, initial rate of mammary cell proliferation represented by parameter  $b$  in the DJ model was 18% lower in primiparous versus multiparous cows ( $P = 0.048$ ). Mean incline parameter  $b$  from all three models differed less than 12% between parity groups ( $P = 0.046$ ). The declining phase parameter  $c$  was 25 to 42% lower ( $P < 0.05$ ) in primiparous than multiparous curves.

The lactation curves generated from these parameter estimates are plotted in Figure 2.1. It is clearly visible that the primiparous group has an average lactation curve that is smoother and flatter than that of the multiparous group. The difference in shapes can be attributed to differences in their lactation curve characteristics, which are also summarized in Table 2.5. Not only are PY from the primiparous group significantly lower (32 kg vs 40 to 42 kg), PT occurs an estimated 13 to 18 d later. The higher persistency exhibited in primiparous cows supports higher daily MY at 305 DIM despite over 10 kg/d less MY at peak lactation. The predictions from all three models yielded significant differences in PY between groups. All the predictions from all three models correctly identified the trend of primiparous cattle having a later PT and an

improved persistency, with the WD and WL models also producing a significant difference in PT between parity groups.

Table 2.3. Characteristics of lactations grouped into primiparous or multiparous groups<sup>1</sup>

Group	Number of Lactations	Mean Daily MY (kg)	SD	Mean Lactation Length (d)	SD	Total 305D MY (kg) <sup>2</sup>	SD	MY at 305-d (kg/d) <sup>3</sup>	SD
Primiparous	203	26.94 <sup>a</sup>	6.077	386	70.9	8461 <sup>a</sup>	1120.1	19.48 <sup>a</sup>	4.404
Multiparous	445	30.75 <sup>b</sup>	9.746	395	73.2	10018 <sup>b</sup>	1480.4	17.60 <sup>b</sup>	5.258

<sup>1</sup> Means within columns that have no superscript in common are significantly different from each other ( $P < 0.05$ )

<sup>2</sup> The sum of total daily milk yields from 0 to 305 DIM

<sup>3</sup> Represented as the average of the daily milk yields from 303 to 307 DIM

Table 2.4. Comparison of mean model parameter estimates yielded from three lactation models fitted to individual lactations grouped by parity<sup>1</sup>

Model	Parameters							
	<i>a</i>	SD	<i>b</i>	SD	<i>b</i> <sub>0</sub>	SD	<i>c</i>	SD
Primiparous								
Dijkstra	18.65 <sup>a</sup>	5.778	0.137	0.2216	0.049	0.0375	0.094	0.233
Wood	15.56 <sup>a</sup>	4.798	0.23 <sup>a</sup>	0.090			0.003 <sup>a</sup>	0.0011
Wilmink	36.29 <sup>a</sup>	4.672	-20.38	6.932			-0.046 <sup>a</sup>	0.0163
Multiparous								
Dijkstra	26.28 <sup>b</sup>	8.155	0.169	0.2728	0.045	0.0432	0.127	0.2823
Wood	19.89 <sup>b</sup>	6.941	0.26 <sup>b</sup>	0.118			0.004 <sup>b</sup>	0.0017
Wilmink	46.39 <sup>b</sup>	4.609	-21.42	9.812			-0.080 <sup>b</sup>	0.0213

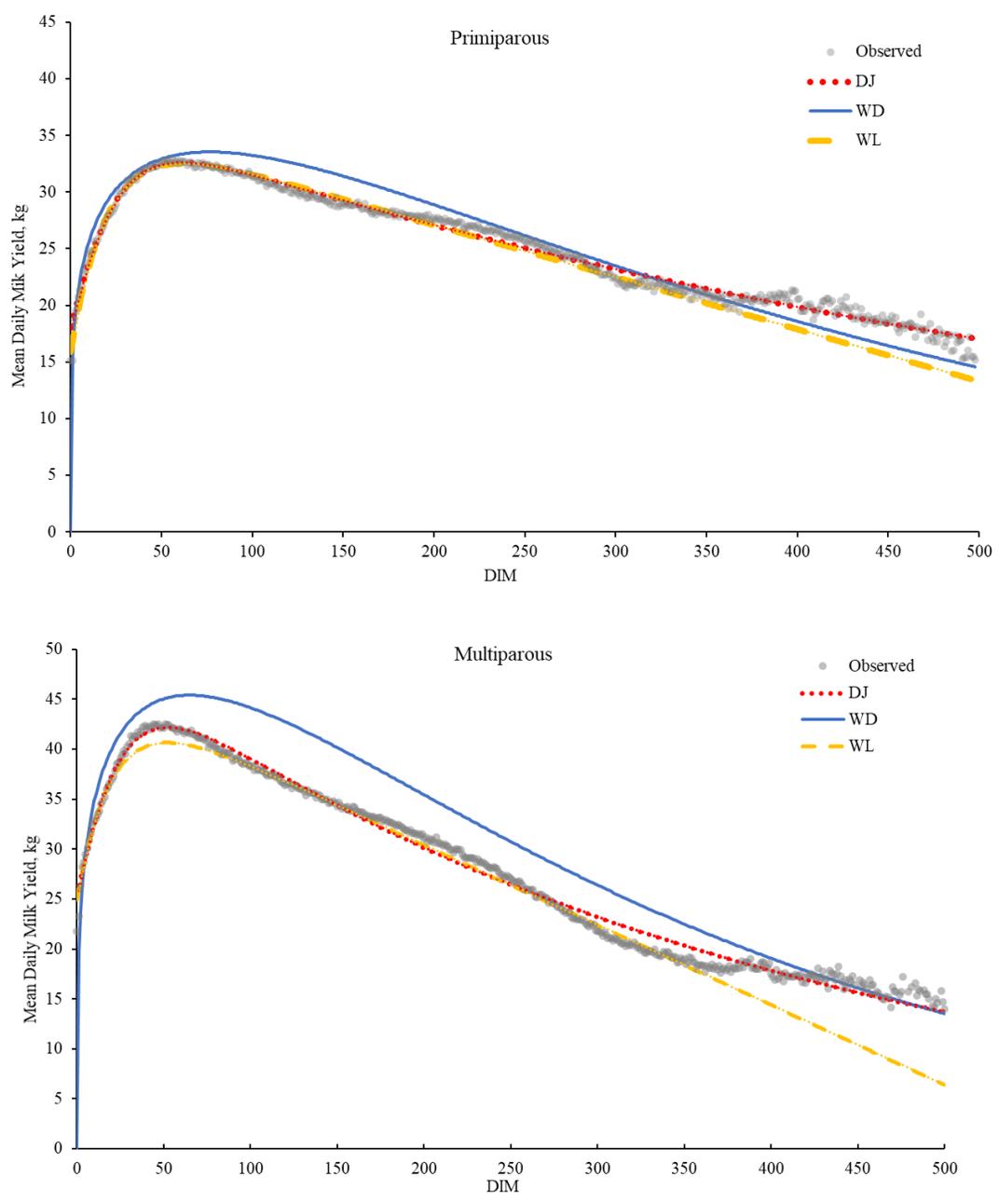
<sup>1</sup> Means within columns and models that have no superscript in common are significantly different from each other ( $P < 0.05$ )

Table 2.5. Comparison of mean predicted lactation curve characteristics for primiparous and multiparous cattle <sup>1,2</sup>

Model	Peak Yield		Time to Peak		Persistency	
	(kg)	SD	(d)	SD	(kg/d)	SD
Primiparous						
Dijkstra	32.63 <sup>a</sup>	4.213	73	30.3	-0.046 <sup>a</sup>	0.0179
Wood	32.16 <sup>a</sup>	3.890	76 <sup>a</sup>	18.1	-0.044 <sup>a</sup>	0.0148
Wilmink	32.52 <sup>a</sup>	3.983	62 <sup>a</sup>	8.7	-0.042 <sup>a</sup>	0.0151
Multiparous						
Dijkstra	42.07 <sup>b</sup>	5.730	60	25.9	-0.085 <sup>b</sup>	0.0245
Wood	41.73 <sup>b</sup>	5.324	58 <sup>b</sup>	15.6	-0.081 <sup>b</sup>	0.0224
Wilmink	40.85 <sup>b</sup>	40.90	50 <sup>b</sup>	11.2	-0.074 <sup>b</sup>	0.0197

<sup>1</sup> Means within columns and models that have no superscript in common are significantly different from each other ( $P < 0.05$ )

<sup>2</sup> Calculated using equations 1,2 and 3



**Figure 2.1.** Mean milk yields (observed) over days in milk plotted with lactation curves yielded from the Dijkstra (DJ), Wood (WD) and Wilmink (WL) models using the mean parameter estimates generated by model fits to individual to test day yields grouped by parity

### **Comparison of Fit Statistics**

There were no significant differences in mean goodness of fit measurements between models in either parity group ( $P = 0.05$ ). All models showed strong positive autocorrelation among residuals, which aligns with results found in other studies (Vargas et al, 2000; Dematawewa et al, 2007; Grossman and Koops, 2003). Although there was no significant difference between the means of the three models, numerically the DJ model produced marginally better fit statistics in both parity groups, with slightly lower values for RMSPE and slightly higher  $R^2_{adj}$ , followed by the WL model and then the WD model.

Table 2.6. Mean goodness of fit criteria calculated from individual lactation curves grouped by lactation lengths fitted to Dijkstra, Wood and WL models

	Mean Fit Statistic Criteria <sup>1</sup>						Proportion DW <1.5 <sup>2</sup>
	R <sup>2</sup> <sub>adj</sub>	SD	RMSPE	SD	DW	SD	
Primiparous							
Dijkstra	0.740	±0.1453	2.21	±0.704	0.87	±0.378	0.96
Wood	0.717	±0.1497	2.30	±0.695	0.83	±0.394	0.95
Wilmink	0.725	±0.1504	2.20	±0.748	0.84	±0.414	0.94
Multiparous							
Dijkstra	0.860	±0.1026	2.82	±0.810	0.91	±0.430	0.94
Wood	0.852	±0.1094	2.93	±0.851	0.85	±0.433	0.95
Wilmink	0.854	±0.0996	2.84	±0.867	0.90	±0.462	0.95

<sup>1</sup>R<sup>2</sup><sub>adj</sub> = the coefficient of determination, RMSPE = root mean square prediction error; DW = Durbin Watson test statistic

<sup>2</sup>The proportion of lactations that exhibit significant positive autocorrelation with a p value (P=0.01)

### **Effect of Lactation Length on Lactation Curves and Model Fits**

For the next comparison, lactations were put into three groups based on their length: Group 1 consisted of lactations finishing between 305 and 399 d, Group 2 consisted of lactations lasting 400 to 499 d and Group 3 with lactations 500 d and greater. There was no significant difference in mean daily MY between groups, but total 305-d MY and MY at 305 DIM were significantly higher in Group 3.

2.7. Lactation characteristics of lactations grouped together by lactation length<sup>1</sup>

Group	Number of Lactations	% of 1 <sup>st</sup> Parity Animals	Mean Daily MY (kg)	±SD	Mean Lactation Length (d)	±SD	Total 305 MY (kg) <sup>2</sup>	±SD	MY at 305-d (kg/d) <sup>3</sup>	±SD
Group 1	448	33.1	29.72	8.656	350	26.8	9312 <sup>a</sup>	1524.5	17.48 <sup>a</sup>	4.999
Group 2	157	27.4	29.46	9.436	438	26.4	9967 <sup>a</sup>	1493.6	19.44 <sup>a</sup>	4.888
Group 3	37	22.2	28.63	9.480	572	62.5	10405 <sup>b</sup>	1417.5	20.84 <sup>b</sup>	4.958

<sup>1</sup> Means within columns that have no superscript in common are significantly different from each other (  $P < 0.05$  )

<sup>2</sup> The sum of total daily milk yields from 0 to 305 DIM

<sup>3</sup> Represented as the average of the daily milk yields from 303 to 307 DIM

## Lactation Curve and Parameters Estimates

Table 2.8 summarizes the mean parameter estimates obtained by model fits to three different lactation curve lengths. The  $a$  parameter for the Dijkstra and WL model was not different between groups, while in the WD model, this parameter increased with increasing lactation length and was significantly lower in Group 1. From the WD model, the  $b$  and  $c$  parameters were significantly higher in Group 1. There were no significant differences between the Dijkstra parameters, but numerically the  $b$  and  $b_0$  parameters decreased as lactation length increased. The WL model parameter  $c$  was significantly lower in group 1, and numerically decreased as lactation length increased.

Table 2.8. Comparison of mean lactation model parameters estimates yielded from three lactation models fitted to individual lactations grouped by lactation length<sup>1,2</sup>

Model	Parameters							
	$a$	SD	$b$	SD	$b_0$	SD	$c$	SD
Group 1								
Dijkstra	23.64	7.942	0.160	0.2649	0.121	0.2777	0.0450	0.03770
Wood	17.72 <sup>a</sup>	6.221	0.262 <sup>a</sup>	0.1128	-		0.0042 <sup>a</sup>	0.00168
Wilmink	42.92	6.716	-21.535	8.6231	-		-0.0720 <sup>a</sup>	0.02598
Group 2								
Dijkstra	24.39	9.355	0.171	0.2578	0.119	0.2603	0.0495	0.05065
Wood	20.12 <sup>b</sup>	7.145	0.231 <sup>b</sup>	0.0982	-		0.0037 <sup>b</sup>	0.00139
Wilmink	43.80	6.427	-19.765	10.0072	-		-0.0658 <sup>b</sup>	0.02339
Group 3								
Dijkstra	23.94	7.596	0.082	0.1039	0.037	0.1092	0.0516	0.04253
Wood	21.54 <sup>b</sup>	7.377	0.208 <sup>b</sup>	0.1033	-		0.0031 <sup>b</sup>	0.00143
Wilmink	44.42	5.244	-21.611	8.8925	-		-0.0570 <sup>b</sup>	0.02025

<sup>1</sup>Groups consist of lactations grouped together by maximum days in milk (DIM); Group 1 lactations being 305-400 DIM (n=448), Group 2 between 400-500 DIM (n=157) and Group 3 > 500 DIM (n=37)

<sup>2</sup>Means within columns and models that have no superscript in common are significantly different from each other (P<0.05)

PY was lowest in Group 1 and highest in Group 3, although there was only a significant difference between Groups 1 and 2. PT was comparable between groups, although Group 3 had the latest PT with all three models. There was a numeric trend for persistency to increase with increasing lactation length with all models. The predicted average daily decline from the DJ model was numerically lowest in Group 3 but similar between Groups 1 and 2.

Table 2.9. Comparison of mean predicted lactation curve characteristics from extended lactations grouped by length<sup>1,2,3</sup>

Model	Peak yield (kg)	SD	Time to Peak (d)	SD	Persistency (kg/d)	SD
Group 1						
Dijkstra	38.56 <sup>a</sup>	6.677	64	27.6	-0.075	0.0294
Wood	38.19 <sup>a</sup>	6.683	63	17.2	-0.071	0.0265
Wilmink	38.16	6.435	54	22.8	-0.066	0.0241
Group 2						
Dijkstra	40.28 <sup>b</sup>	6.564	65	28.5	-0.070	0.0287
Wood	39.96 <sup>b</sup>	6.569	65	21.2	-0.068	0.0274
Wilmink	39.92	5.784	59	21.1	-0.061	0.0217
Group 3						
Dijkstra	40.97 <sup>b</sup>	5.141	67	28.9	-0.061	0.0243
Wood	40.16 <sup>b</sup>	4.837	68	19.8	-0.056	0.0223
Wilmink	40.98	4.954	60	25.6	-0.054	0.0199

<sup>1</sup>Groups consist of lactations grouped together by maximum days in milk (DIM); Group 1 lactations being 305-400 DIM (n=448), Group 2 between 400-500 DIM (n=157) and Group 3 > 500 DIM (n=37)

<sup>2</sup>Means within columns and models that have no superscript in common are significantly different from each other (P<0.05)

<sup>3</sup>Calculated with equations 1,2 and 3

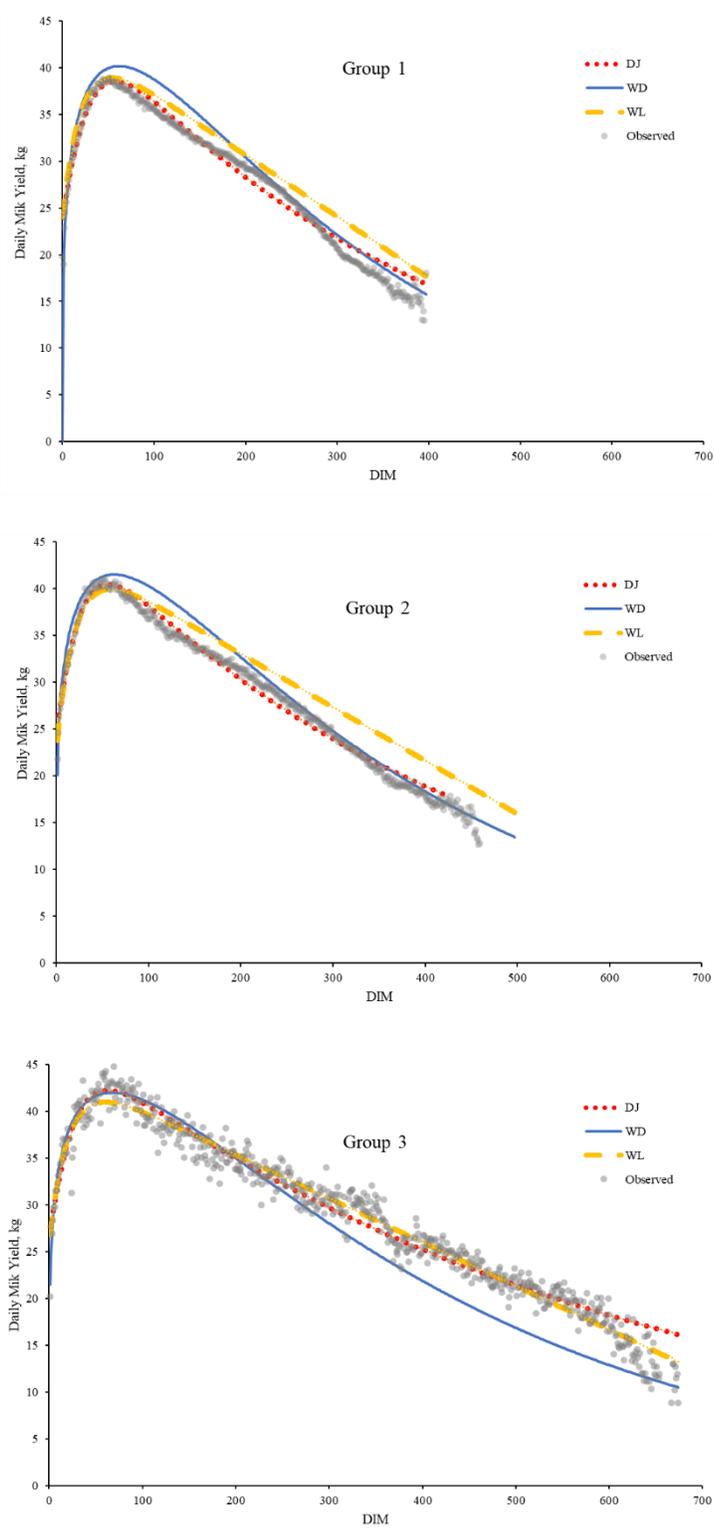


Figure 2.2 Mean milk yields (observed) over days in milk (DIM) plotted with three lactation curves yielded from the Dijkstra et al, Wood and Wilmink models fitted to test day yields from lactations grouped by length: Group 1 (305-400 DIM; n =448), Group 2 (400-500 DIM, n=157) and Group 3 (>500 DIM; n=37)

## Parameter Comparison

Similar to the parity group comparison, there were no significant differences between the mean fit statistics within lactation length groups. Again, there was strong positive autocorrelation among residuals, especially in Group 3. As lactation length increased, autocorrelation similarly increased. The DJ model produced slightly better fit statistics in each lactation length group.

Table 2.10. Mean goodness of fit criteria calculated from individual lactation curves grouped by lactation lengths fitted to Dijkstra, Wood and WL models<sup>1</sup>

Model	Fit Statistic Criteria <sup>2</sup>						
	R <sup>2</sup> <sub>adj</sub>	SD	RMSPE	SD	DW	SD	Proportion DW > 1.5 <sup>3</sup>
Group 1							
Dijkstra	0.813	0.1267	2.57	0.821	0.92	0.414	0.93
Wood	0.799	0.1317	2.67	0.868	0.86	0.425	0.94
Wilmink	0.802	0.1350	2.62	0.839	0.90	0.445	0.94
Group 2							
Dijkstra	0.822	0.1484	2.79	0.840	0.85	0.408	0.97
Wood	0.817	0.1275	2.86	0.830	0.80	0.417	0.96
Wilmink	0.815	0.1367	2.84	0.866	0.84	0.461	0.95
Group 3							
Dijkstra	0.868	0.0953	2.68	2.143	0.75	0.193	1.00
Wood	0.849	0.0941	2.77	2.276	0.71	0.205	1.00
Wilmink	0.858	0.0926	2.71	0.817	0.79	0.312	0.97

<sup>1</sup>Groups consist of lactations grouped together by maximum days in milk (DIM); Group 1 lactations being 305-400 DIM (n=448), Group 2 between 400-500 DIM (n=157) and Group 3 > 500 DIM (n=37)

<sup>2</sup> R<sup>2</sup><sub>adj</sub> = the coefficient of determination, RMSPE = root mean square prediction error; DW = Durbin Watson test statistic

<sup>3</sup>The proportion of lactations that exhibit significant positive autocorrelation with a p value of 0.01

## **Analysis of Residuals**

The mean daily residuals from all groups and all models exhibited a consistent cyclic pattern, oscillating between negative and positive values throughout the lactation cycle. Notably, the residuals in early-lactation from WL model fits were much closer to zero than those from the other two models.

The analysis of residuals in Figures 2.3 and 2.4 indicates the stages of lactation when milk production may be over- or under-predicted. All three models show similar patterns that are repeated across parity and lactation length groups. Residuals are negative at the very start of lactation, indicating over-estimation of MY at the time. After this, the residuals rapidly increase and become positive, indicating an underestimation in MY as the cows are experiencing peak milk production. This indicates a lack of flexibility of the models to model the early lactation MY, as some of the largest errors and extreme oscillations occur at this time. Perhaps the parameters that shape the inclining phase of the lactation curve cannot model how quickly MY rises in the beginning of the lactation. The overestimations in early lactation and the under estimations at PT have been previously noted in other studies, particularly with the WL and WD models (Cobby and Le Du, 1978; Grossman and Koops, 1988; Dematawewa et al, 2007; Dijkstra et al, 2010).

After PY is reached, the residuals drop and rise again, tending to be negative around 100-200 DIM and positive between 200 and 300 DIM. This repeatable oscillation between positive and negative errors throughout the groups indicates a lack of randomness in the errors. This is supported by a high number of lactations exhibiting positive autocorrelation with the DW test in Table 2.6 and 2.10. This cyclic pattern of residuals has been noted in other studies (Silvestre et al, 2006; Druet et al 2003). For most of the lactation, the majority of the residuals are within 2 kg

of observed MY, indicating an accurate model fit. However, towards the end of lactation, MY become more variable, exceeding 2 kg in error and reaching average errors up to 4 kg. This too has been seen in other studies and has been related to the distribution and decreased frequency of test day records towards the end of lactations (Silvestre et al, 2006; Torshizi et al, 2011), a problem that also exists within the data used in this analysis. The amount of variation in MY towards the end of lactation makes the residual patterns less consistent and harder to interpret. One trend is clear however – the majority of residuals at the immediate ends of the lactation tend to be overwhelming negative, indicating an over prediction of milk production from the models at this time. Over prediction by the WL and WD models at this stage in the lactation was also found by Torshizi et al. (2011). However, overprediction could perhaps be explained by the implementation of dry off procedures on farms at this time, inducing a rapid dry off in milk production due to less nutrient availability that the model cannot account for. The presence of negative residuals towards the end of lactations indicates that producers may have to be conservative when using milk forecasts generated by these models.

The residuals from the DJ and WD model follow very similar patterns. The WL residuals deviate from the other two, particularly in the primiparous group. For the WL model, the early lactation MY seems to be under-predicted instead of over-predicted, and during PT, the residuals are much closer to 0. This pattern is also seen in Group 1 of the lactation length groups, although it seems less pronounced. It is unclear whether the WL model is better at fitting the MY in early lactation; or if residuals are more varied between positive and negatives at this time, resulting in a mean residual closer to zero.

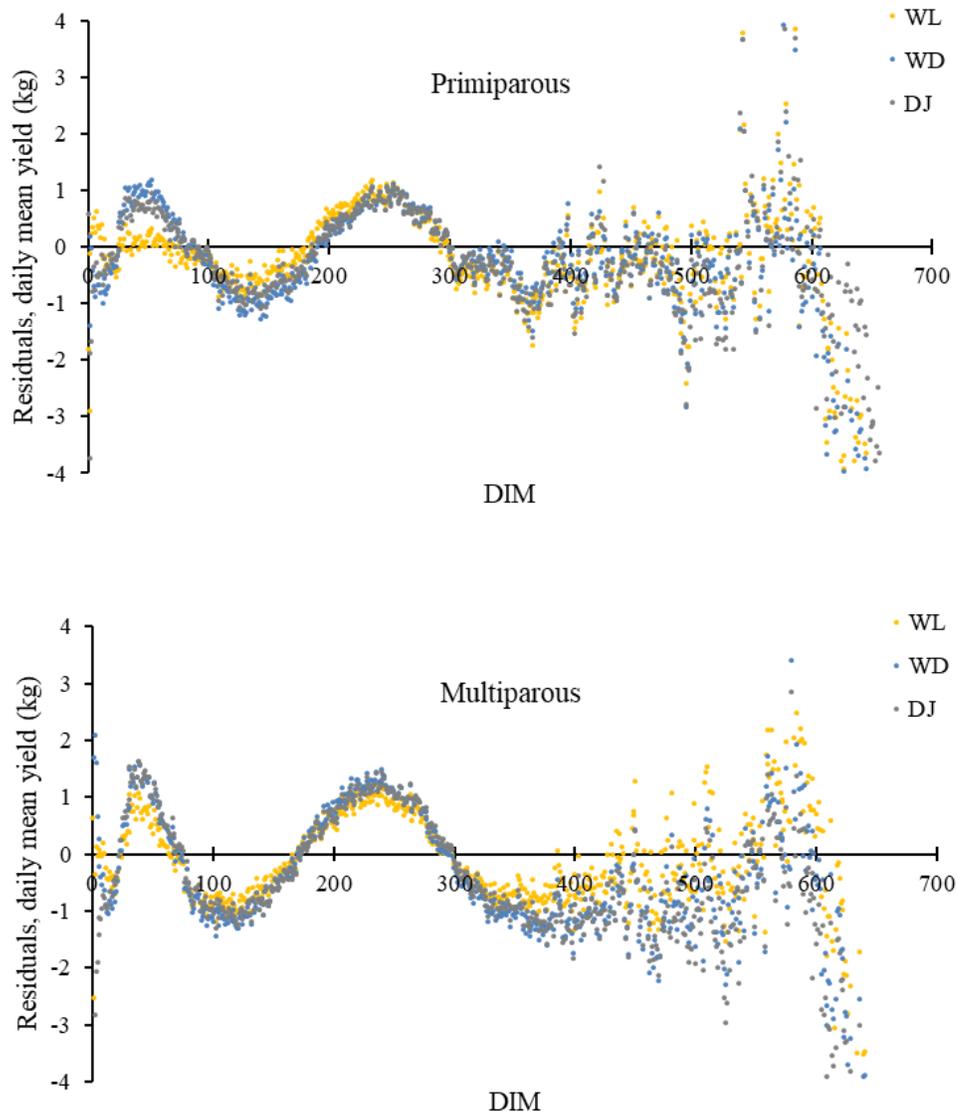


Figure 2.3 Mean daily residuals from individual lactation curves in the primiparous (n=203) or multiparous group (n=445) fitted to the Wood, Wilmink and Dijkstra et al. models

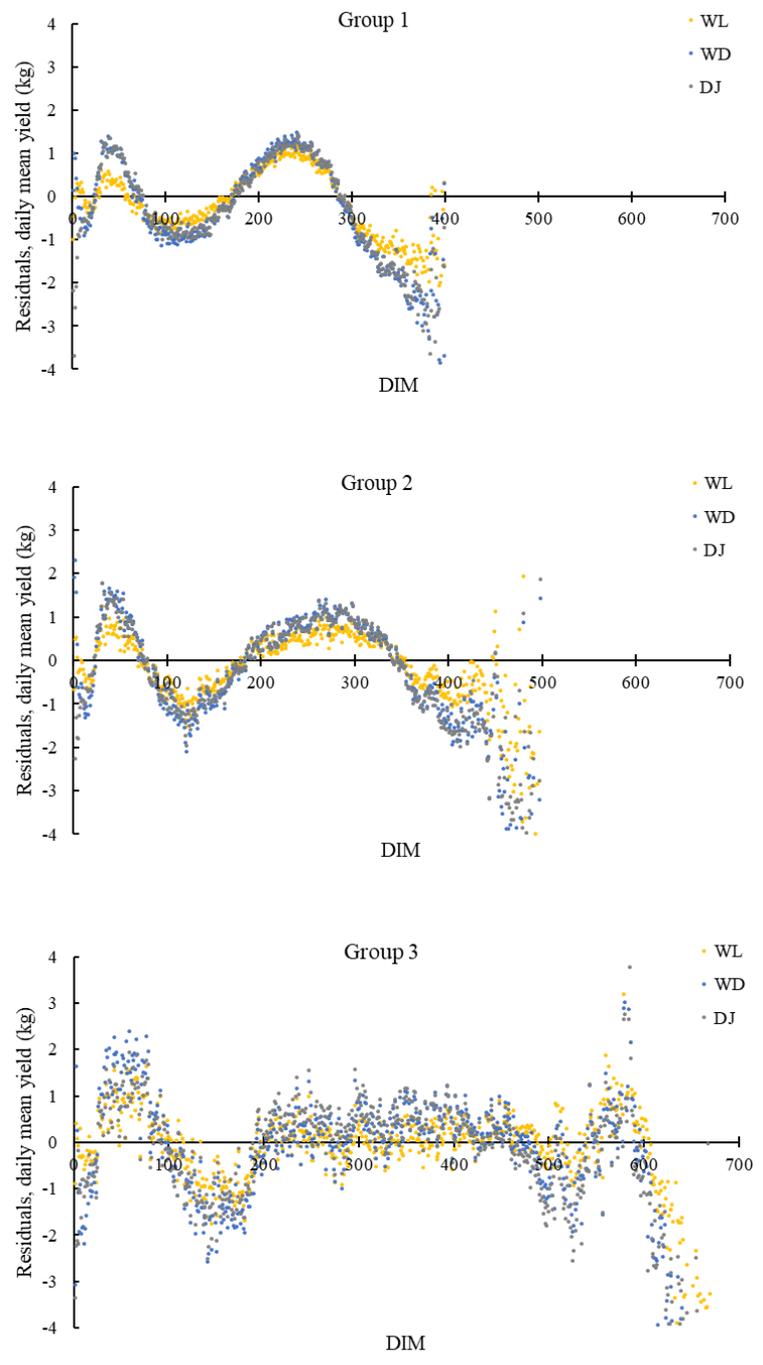


Figure 2.4 Mean daily residuals from lactation length Group 1 (305-400 DIM; n=448), Group 2 (400-500 DIM; n=157) and Group 3 (> 500 DIM; n=37) fitted to the Wood, Wilmink and Dijkstra et al. models

### **Relationship between early lactation information and 305-d MY Predictions**

Effects on predicted MY at 305 DIM from curve fits within the first 30 to 300 DIM were evaluated. As seen in Figure 2.5 and Table 2.11, there is a positive relationship between the amount of data available to the model and the goodness of fit; as the RMSPE and MAE improve with increasing data availability.

The mean RMSPE and MAE for predicted MY at 305 DIM based on fits from the first 30 and 60 DIM are very high. The decomposition of RSMPE from these time points indicates that almost all of the error arises from slope bias, with no error being contributed from the random error term. Additionally, the CCC for both the 30 DIM and the 60 DIM fits were remarkably low; again, indicating poor model precision and accuracy. Only when the model used 90 DIM of data did the fits yield acceptable values for the CCC, along with mean RMSPE and MAE values below the 305-d MY average. A visual interpretation is depicted in Figure 2.5, which plots the log of the RMSPE against the amount of data available to the model. This figure shows the large reduction in the fit statistic between the 30- to 90-DIM window, after which the addition of more data had a lesser impact on the declining RMPSE values, as apparent from the slope starting to approach zero as more data were used.

The CCC also improved with more data; however, the highest CCC values were found using data up to 150 DIM. When the data available increased beyond 150 DIM, the CCC decreased in value until 240 DIM, after which the values again increased. The percentage of error contributed by mean bias and random error tended to increase as the amount of data available increased, while the proportion of slope bias decreased.

Figure 2.6 shows the average lactation curve generated by the DJ model after fits to subsections of early lactation data. When only 30 DIM of data were used, the typical skewed-bell

shaped lactation curve was not correctly modelled. Instead, the resulting curve increased to a maximum early in the lactation and maintained this PY indefinitely. This inability was likely due to the absence of a natural peak or descending milk yield within the 30-d timeframe of the observed data, therefore the model was not able to detect or simulate the descending phase of a lactation curve. The poor fits resulted in 305-d predictions being grossly overestimated, causing the very high RSMPE and MAE, the low CCC, and the 97% slope error seen in Table 2.11. Therefore, fits generated from this time point were neither accurate nor precise and should not be used for predictive forecasting.

When 60 DIM of data were available to the model, the typical shape of the predicted lactation curve became recognizable, as seen by the grey line in Figure 2.6. The modeled curve did correctly exhibit a descending phase of the lactation cycle; however, the slope of the descending portion of the curve was very steep compared to the mean observed curve. This resulted in predictions being underestimated. Despite the drastic reduction in the RSMPE and MAE values from the 30 DIM fitting, they were still too high to indicate an accurate fit. Additionally, the slope bias still constituted the vast majority of the error, and the CCC was still remarkably low.

As more data were included, notably data that represented the declining phase after PY, the model was able to more accurately fit the descending slope. As seen in Figure 2.6, for every 30 day of additional data used, the steepness of slopes post-PY decreased and began to more closely match the observed MY patterns. Table 2.11 shows that the addition of another 30 DIM, to a total of 90 DIM, again drastically reduced the RSMPE and MAE by almost 97%. It is with the use of 90 DIM of data that the RSMPE and MAE were below the mean average 305-d MY of 18.12 kg. Due to the presence of large individual residuals, the RSMPE remained very high and

close to this mean average 305-d MY. The MAE, which is less influenced by large outliers, was a more manageable 10.52 kg. There was also a large increase in the CCC values and a reduction in the slope bias, indicating an improvement in the accuracy of predictions. Importantly, random error due to deviation also started to contribute to the error term at this time.

The increase in the CCC value to 0.625 and the decrease of the RMSPE and MAE to below the MY average indicate that the DJ model needs, at a very minimum, 90 DIM of early lactation data to begin to forecast MY at 305-D. However, the descending portion of the curve as seen in Figure 2.6 is still quite steep and the mean RSMPE yielded from these fits remains very large, with the majority of the error term still consisting of slope bias. This suggests that the forecasting ability is still limited at best.

Since 90 DIM was determined to be the minimum amount of data needed for the DJ model to forecast MY at d-305, PIs and the  $SE_p$  were calculated from the fits of the 90 DIM and onwards. Table 2.12 shows that the highest  $SE_p$  and largest PI interval size were found using 90 DIM (Table 2.12), and the lowest at 300 DIM of data. The PI and the  $SE_p$  calculated from the herd standard deviation followed the same patterns, although they are larger due to the increased variability within the total herd.

Using 120 DIM allows for another large reduction in the RSMPE and the MAE and a more accurate portrayal of the downward slope. For the 120 DIM fit, the RMSPE and MAE were 9.85 kg and 7.71 kg, respectively. This is a reduction of 56.7% and 64.3% compared to the fit statistics of the 90 DIM fits. The CCC value increased even higher from the 90 DIM fit to 0.915, close to the ideal value of 1. Additionally, when 120 DIM of data were used for the fits, the % of error attributed to random error finally took over as the majority error term.

The  $SE_P$  calculated with individual animal  $s$  from fits to 90 DIM of data was 6.9 kg (Table 2.12). When 120 DIM were used, the  $SE_P$  decreased to 3.6 kg, or 19.97% of the MY average. Both the PI and the  $SE_P$  were almost halved by increasing DIM of data from 90 to 120.

The log plot of RMSPE in Figure 2.5 shows a breakpoint between 90 and 120 DIM, beyond which RMSPE is less affected by inclusion of more data in the curve fits, as indicated by the evenness of the slope beyond 120 DIM of data. After the 120 DIM fits, a plateau in the reduction of the fit statistic is almost reached, when the log of the RMSPE plotted against the DIM levels off and becomes more gradual, except for a slight dip at 300 DIM of data. Indeed, an addition of another 180 days of data only reduces the fit statistics by approximately 38%.

As expected, the best fit statistics and the lowest PIs and  $SE_P$  were found with the 300-DIM fit. The RMSPE was reduced to 3.75 kg, with the PI and  $SE_P$  decreasing in size to 4.7 and 2.4 kg, respectively.

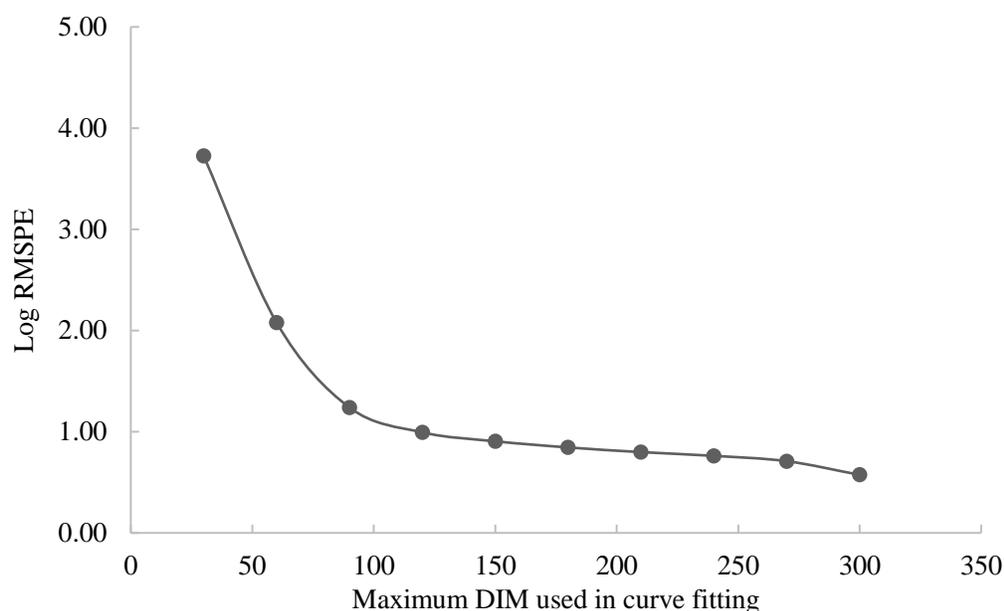


Figure 2.5 Relationship between the log residual mean square predicted error (RMSPE) of predicted MY (n=548) at 305 DIM and amount of days in milk (DIM) data available to the DJ model

Table 2.11. Root mean square prediction error (RMSPE), log RSMPE, the decomposition of RSMPE into the mean bias, slope bias and random error components, mean absolute error (MAE) and the concordance correlation coefficient (CCC) calculated from individual 305-DIM (n=548) early lactation MY records increasing by 30-day increments

DIM Subset	RMSPE	Log (RSMPE)	MAE	Mean Bias (%)	Slope Bias (%)	Random Error (%)	CCC
30	5315.08	3.73	913.80	2.86	97.14	0.00	0.002
60	119.91	2.08	28.81	0.30	99.42	0.28	0.104
90	17.36	1.24	11.99	4.11	86.42	13.10	0.625
97	14.75	1.17	10.54	3.36	78.62	18.05	0.720
104	11.71	1.07	9.29	5.28	66.65	28.12	0.826
111	10.79	1.03	8.46	4.14	63.03	32.90	0.868
120	9.85	0.99	7.71	2.52	58.32	39.24	0.915
150	8.04	0.91	6.06	0.07	43.79	56.26	0.981
200	6.29	0.8	4.75	15.38	12.23	72.57	0.907
250	5.77	0.76	4.32	23.08	5.11	72.04	0.878
300	3.75	0.57	2.87	22.86	0.01	77.67	0.935

Table 2.12 The mean standard error of predictions ( $SE_p$ ) and mean prediction interval (P.I) size of 305-D milk yield predictions (n=548) resulting from fits  $i$  days in milk (DIM) data to the Dijkstra et al. model calculated using individual animal and total herd standard deviations

DIM used for model fits	Individual $s$				Total Herd $s$			
	Mean $SE_p$	$\pm SE$	Mean PI ( $\pm$ kg/d)	$\pm SE$	Mean $SE_p$	$\pm SE$	Mean PI ( $\pm$ kg/d)	$\pm SE$
90	6.50	18.605	12.92	36.986	14.26	23.684	28.35	47.083
97	5.52	12.480	10.97	24.783	12.39	16.209	24.61	32.189
104	4.28	4.309	8.50	8.549	10.22	6.893	20.29	13.766
111	3.72	3.501	7.37	6.940	8.94	4.056	17.72	8.040
120	3.62	2.364	7.18	4.685	8.82	3.852	17.48	7.502
200	2.40	0.782	4.73	1.543	5.43	0.044	10.75	0.871
300	2.37	0.685	4.68	1.348	5.17	0.016	10.24	0.031

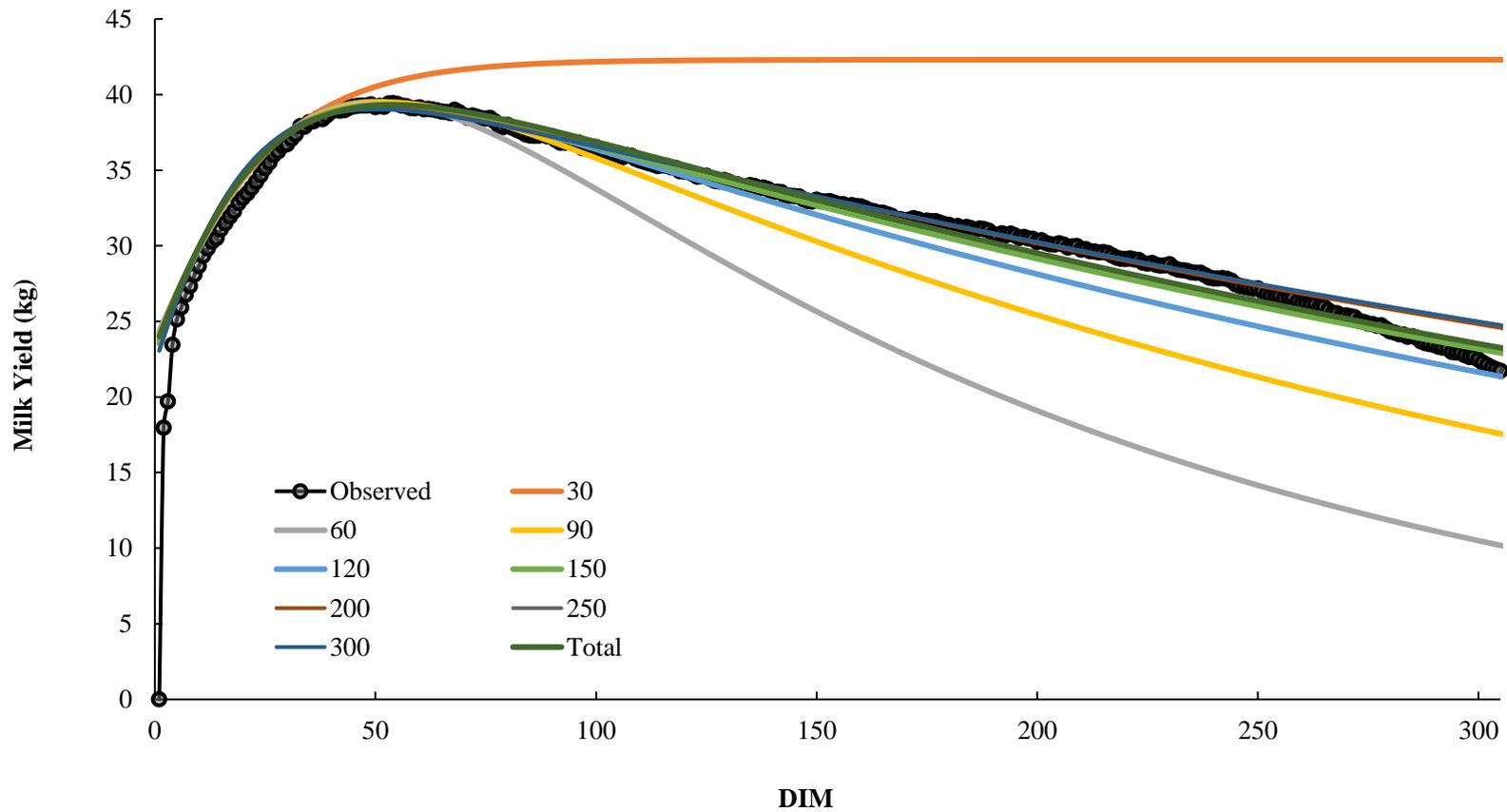


Figure 2.6 Lactation curves generated by parameters obtained from the Dijkstra et al model fit to increasing amounts of early lactation data compared to lactation curve from the average observed daily milk production

### **Probability of 305-D MY Predictions Exceeding Targets**

Probabilities of MY at 305 DIM actually being equal to or greater than a specified MY target were estimated from a one-tailed cumulative t-distribution as a way to consider the risk associated with selecting cows in early lactation for extended lactation. Dingwell et al (2001) determined an average MY at dry off from Ontario DHI records of 16.6 kg, with 20% of cattle exceeding 22 kg of milk at dry off. However, the National Mastitis Council (NMC) recommends drying off cows when production drops below 15 kg (NMC, 2006). Based on these values, target MY at 305 DIM were set to 10, 15, 20, 25 and 30 kg/d in order to capture the recommended dry-off yields, as well as the higher milk yields at day 305 that would be needed to sustain longer lactations. Table 2.13 reports means of the individual probabilities of a cow's estimated MY meeting or exceeding the aforementioned targets for each *i* DIM used for model fits.

Within each *i* DIM group, the mean probability of meeting the target decreased as the 305-D target MY increased. When using 90 DIM of data, the mean probability of meeting the 10 kg/d goal was 0.66. Since the MY average at 305 DIM is over 18 kg, this indicates that predicted 305 MY is being underestimated. Indeed, when 300 DIM of data were used and the predictions were more accurate, the mean probability of meeting this 10 kg target increased to 0.99.

Within the 10, 15 and 20 kg/d targets, the mean probability increased as the amount of data used increased, with the probabilities between 200 and 300 DIM being either equal or very similar. This pattern broke for the 25 and 30 kg/d targets- which was not unexpected due to the target milk yield being higher than the observed daily milk yield of 18 kg at day 305. For the 25 kg/d target, the mean probability was similar between 90 and 120 DIM fits (0.31 vs 0.30), higher for the 200 DIM fit (0.40) but reduced to a mean probability of 0.37 for the 300 DIM fit. For the

30 kg/d target, mean probability was highest in the 90 DIM fit group and lowest for the 300 DIM fit group.

Table 2.14 compares how estimated MY at 305 DIM compared to that observed within the herd. The percentage of cattle whose estimated MY had probabilities greater than 50, 70 or 95% of meeting 15, 20 or 25 kg/d targets at 305 DIM were calculated and compared to the percentage of the herd whose observed MY met or exceeded those same targets at 305 DIM. Over 73% of the herd was observed to have MY at 305 DIM  $\geq$  15 kg/d, which then dropped to 35.9% and 8.2%, when the MY targets were increased to 20 and 25 kg/d, respectively.

When using individual animal SD to calculate the probabilities, the percentage of fits with  $> 50\%$  probability of exceeding the 15 kg/d target at 305 DIM increased from 53.7% to 93.5% as the amount of DIM data increased (Table 2.14). Similar relationships between the percentage of the herd and the amount of data used were seen with  $> 75\%$  and  $> 95\%$  probabilities for the 15 kg/d target. However, as the probability increased and the t-threshold narrowed, the proportion of fits meeting the target decreased. Compared to the proportion of observed MY meeting the target, the predicted proportion was lower for all the probabilities using 90 and 120 DIM, and higher using 200 and 300 DIM of data for the model predictions.

When the target was increased to 20 kg/d, the same positive relationship between amount of data available and percentage of fits meeting the target was found. When using probabilities of 50% or 75%, the proportion of fits meeting the 20 kg/d target tended to be overestimated. Using 95% probability led to underestimation, unless 300 DIM of data were used. When the target was further increased to 25 kg/d, the percentage of fits meeting the target was much higher than observed for all probabilities.

When these calculations were replicated using observed herd SD at 305 DIM instead of individual SD of curve fits to estimate prediction errors, there was no effect on the proportion of fits meeting the targets when a probability of 50% was used, due to the symmetry of the t-distribution. However, when the probability was increased to 75% or 95%, the proportion of fits that met the target was heavily reduced, compared to the results using individual animal SD. The effect of DIM remained the same with herd SD, though, where there was a positive relationship between the amount of data and the proportion of fits meeting both the 15 and 20 kg/d targets. At 95% probability, there was an underestimation of the observed percentage of fits meeting any of the three targets, regardless of the amount of data available.

Table 2.13 The mean probability of predicted MY (n=548) at 305 DIM (n=548) meeting or exceeding various target levels according to fits of the Dijkstra et al. model to MY data up to *i* DIM

<i>i</i> DIM used for model fits	305-D MY Targets									
	10kg	±SE	15kg	±SE	20kg	±SE	25kg	±SE	30kg	±SE
90	0.66	0.397	0.54	0.420	0.42	0.416	0.31	0.382	0.20	0.317
120	0.85	0.290	0.71	0.370	0.50	0.410	0.30	0.376	0.15	0.283
200	0.99	0.081	0.92	0.193	0.72	0.357	0.40	0.401	0.16	0.286
300	0.99	0.047	0.92	0.199	0.70	0.356	0.37	0.382	0.12	0.285

Table 2.14 Percent of cows (n=548) with observed 305-D MY over a certain target value (kg) compared to the percent of individual predictions over the target value based on probabilities of either 50, 75 or 95 % from fits of *i* DIM data to the Dijkstra et al. model

	<i>i</i> DIM data used for model fits							
	Individual Animal s				Total Herd s			
	90	120	200	300	90	120	200	300
<b>≥15kg Target</b>								
% Observed	73.8	73.8	73.8	73.8	73.8	73.8	73.8	73.8
Pr>50	53.7	70.6	94.4	93.5	53.7	70.6	94.4	95.5
Pr>75	44.4	63.0	89.6	89.4	30.7	48.1	80.1	82.5
Pr>95	33.1	48.3	79.6	79.7	11.9	19.5	45.0	48.0
<b>≥20kg Target</b>								
% Observed	34.9	34.9	34.9	34.9	34.9	34.9	34.9	34.9
Pr>50	42.6	49.3	72.7	73.0	42.6	49.3	72.7	73.0
Pr>75	32.9	38.5	63.8	59.9	20.3	27.0	47.8	48.5
Pr>95	21.4	26.8	26.8	43.7	8.6	8.7	18.0	14.7
<b>≥25kg Target</b>								
% Observed	8.2	8.2	8.2	8.2	8.2	8.2	8.2	8.2
Pr>50	30.1	27.7	40.1	35.3	30.1	27.7	40.1	35.3
Pr>75	21.6	21.4	29.9	25.5	11.3	13.0	19.1	14.9
Pr>95	12.6	12.3	17.8	13.4	4.3	3.0	3.9	3.3

## CHAPTER 3: DISCUSSION AND CONCLUSION

### Discussion

The objective of this research was to determine how the number of consecutive days of MY data available to lactation models would impact their ability to predict MY at d 305. Forecasting MY could help producers identify which cattle would be suitable to be managed for extended lactations by identifying those with high estimated MY at d 305. To do this, the ability of lactation models to both fit extended lactations and predict late lactation MY needs to be assessed.

Three of the most commonly used lactation models were assessed for their ability to fit extended lactations. Convergence was obtained for all individual curves. When fitting curves, F-values provided from the Proc NLIN anova tables showed that all three models were able to provide highly significant fits ( $P < 0.01$ ) to individual curves with a mean root mean square prediction error (RMSPE) of 9.8% or less of the MY average. Parameter estimates were able to accurately reproduce the observed shapes of the lactation curves.

Parity is a known influencer of the shape of lactation curves, in which first parity cows exhibit flatter curves due to lower initial MY, resulting in a less defined peak in daily MY (Scott et al, 1996). The data generated in this study supports this; with each model exhibiting a higher 'a' parameter for the multiparous group. Although the DJ model is the only mechanistic model of the three, this parameter from each model is still relatable to the scale of initial MY at  $t = 0$ , indicating that multiparous cattle have higher initial MY at the beginning of their lactations. In addition to higher initial MY, peak milk production was higher in the multiparous group. Furthermore, the multiparous group from this analysis reached PY earlier than the primiparous group, similar to what was reported by Cole et al (2009). This short time to PY, combined with

higher initial and PY, contributes to multiparous cattle having a more defined peak in early lactation than the primiparous cattle, as exhibited by the representative lactation curves presented in Figure 2.1.

The other contributor to the different lactation curve shape seen in the primiparous group is the significantly higher persistency. The persistency values calculated from the parameters of each lactation model were significantly different between groups, with the primiparous cattle having significantly reduced rates of decline in daily MY. Although the cause of the higher persistency remains unclear, the impact of the higher persistency on 305-d MY is easily seen. This higher persistency allows the primiparous group to have significantly higher MY at day 305 than the multiparous group, despite lower initial and PY. This supports the commonly viewed idea that primiparous cattle may be suitable for a longer VWP, as they will have an improved ability to sustain higher MY into late lactation.

The parameter estimates and lactation curve characteristics of cows with different lactation lengths can also provide information on which cows are suitable to be managed for longer lactations. The Dijkstra and WL models both exhibited a trend of having a larger  $a$  parameter, related to initial MY, within the longer lactation length groups; with the WD model having a significantly higher  $a$  parameter in Group 2 and 3, as compared to Group 1. This, combined with the statistically higher PY in Group 2 and 3, as calculated from parameters of the Dijkstra and WD models, indicates that high initial MY and PY indicate a proclivity for an extended lactation. This agrees with results of Lehman et al (2017), who found that cows with the most successful extended lactations also had the highest PY. Since PY generally be identified within early lactation data, , this could be included in selection criteria when trying to identify cattle who could be managed for longer lactations.

Regardless of the difference in residuals between the WL model and the other models, there was no significant difference among the means of the individual fit statistics between the parity or the lactation length groups. Despite the lack of significant difference between the means of the fit statistics, there was a numerical trend for the DJ model to exhibit better fit statistics within all groups. This suggests that the DJ model is most suitable to model extended lactations of the three models. The high  $R^2_{adj}$  and relatively low RMSPE indicate the model can accurately fit extended lactations from various parities and lactation lengths. Most importantly, the parameters from the DJ model allow for biological interpretations. Due to this flexibility, the DJ model was used for the remainder of the study to determine how the amount of data available to the model affects its ability to predict late lactation MY.

Currently, a 60 day VWP is the industry standard to maintain a conventional lactation length. Ideally, within 60 DIM would be the timeframe for a producer to decide which cows to delay inseminating and which ones would be suitable for longer lactations. However, the predictions yielded from 60 DIM fits have a RSMSPE and MAE of 119 and 28 kg, respectively. These high error estimates combined with the low CCC (0.104) indicate that the fits from this time frame are not accurate- therefore the producer should not base breeding decisions solely from these fits. The lactation curve from the 60 DIM fits (Figure 2.6), while correctly capturing the general shape of a lactation curve, had a declining phase much steeper than what was observed on average.

More accurate fits were obtained within the 90- to 120-DIM window, specifically past 100 DIM, when the RMPSE and MAE were below 10 kg and the CCC rose above 0.80. However, if the producer delayed a breeding decision until there was over 100 DIM of early lactation data, this would create a lactation length of 380-400 DIM at a minimum. Since not all

cattle could successfully milk out for a lactation of this length, the delay in breeding could cause a loss of income for the producer if this VWP were applied to the entire herd.

One suggestion would be to delay rebreeding in select cows that a producer has identified as being potentially suitable for longer lactations. These cattle could be selected based on a combination of several factors; parity, level of PY, previous lactation yield and persistency, and estimated breeding values. There would be less risk in delaying a breeding decision within this subgroup of cattle, until there were enough data to start more accurately predicting late lactation MY. Within these low-risk cattle, a producer could comfortably delay rebreeding until 90 DIM and then use the model to make individual predictions to support an individualized VWP.

Past the 90 DIM window, the largest weekly proportional reduction of the SE and PI occurred between 97 and 104 DIM, accounting for a 22% decrease. The smallest difference in SE and PI occurred between 111 and 120 DIM (2.6% decrease). This indicates that if a producer was hesitant about delaying rebreeding past 90 DIM, one could wait an additional 2 weeks until 104 DIM to help improve the accuracy of predictions. After 104 DIM, the addition of MY data does not substantially improve the accuracy of prediction, so the decision to breed or delay breeding should be done within this window.

The PI of  $\pm 8.50$  kg for MY at d 305 is too large to be singularly used for breeding decisions. Therefore, the probability calculations in Table 2.13 and 2.14 should also be used to help make informed decisions. The data from Table 2.14 could be used to suggest a minimum probability threshold to help producers to select cattle. For example, to identify a similar number of animals compared to the number of observed animals of reaching the 20 kg target, a cut off probability value of 75% or higher could be used to accept or reject an animal. If the target is higher, or the producer wishes to be more conservative, the minimum probability threshold could

be increased to 95%. A recommended probability threshold would be dependent on the target value at day 305 and the level of accuracy the producer would be comfortable with.

Accuracy of the predictions will also depend on the data available to the producer; if individual cow data are available the predictions will be much more accurate. But if the herd average standard deviation is the only information available to the producer, predictions and PI will be much larger and make the milk forecasting more difficult. Using the herd deviation in Table 2.14 significantly narrowed the number of animals with a minimum of 75 or 95% probability for the target values of 20 and 25 kg targets, indicating that if more variability is introduced into the calculations by using total herd deviation, the minimum probability threshold would have to be decreased to relax the cut-off range to ensure the correct number of animals is selected.

However, production is not the only factor that should determine a cow's optimal VWP. Extending the VWP is shown to have positive impact on animal welfare and fertility factors (Borman et al, 2004; De Vries 2006). When comparing the total 305D MY and average daily milk production between the lactation length groups of the animals in this study, the production values are very closely matched between groups. This indicates that cows can be successfully managed for longer lactations without negatively impacting performance, while capturing the welfare, nutritional and health benefits associated with extended lactation (Lehmann et al, 2017).

### **Conclusions**

Compared to the Wood and WL models, the DJ model provided the best fit for extended lactations under a wide range of parity and lactation lengths, as indicated by numerically higher mean fit statistics of individual curves. Since the DJ model also has the advantage of being mechanistic in nature, it is recommended for modeling extended lactations.

The model can provide satisfactory fit statistics from being fit to data from 90 or more DIM. Recent literature suggests that the majority of high producing cattle are able to sustain adequate MY until 365 days. Therefore, if a producer is willing to extend lactation lengths in an attempt to capture the animal welfare benefits, longevity and/or breeding success associated with a longer lactation, producers can lengthen the VWP in select cows that would be likely to be successful in an extended lactation, ensuring there is enough data to predict late lactation MY. An addition of two weeks beyond 90 DIM generates more accurate fits, but beyond 104 DIM, more data have a proportionally decreasing effect on the accuracy of predictions. Using the 305 day estimates, prediction probabilities and the size of the PIs, producers can make educated breeding decisions for individual cows. High-producing cows that are suitable to be milked for much longer should be easy to identify. If data on individual cows are not accessible, predictions can also be made using the herd standard error although these predictions will be much less accurate.

## References

- Akers, R. M. 2006. Major advances associated with hormone and growth factor regulation of mammary growth and lactation in dairy cows. *Journal of Dairy Science* 89(4):1222-1234.
- Ali, T. E. and L. R. Schaeffer. 1987. Accounting for covariances among test day milk yields in dairy-cows. *Canadian Journal of Animal Science* 67(3):637-644.
- Arbel, R., Y. Bigun, E. Ezra, H. Sturman, and D. Hojman. 2001. The effect of extended calving intervals in high- yielding lactating cows on milk production and profitability. *Journal of Dairy Science* 84(3):600-608.
- Auldist, M. J., G. O'Brien, D. Cole, K. L. Macmillan, and C. Grainger. 2007. Effects of varying lactation length on milk production capacity of cows in pasture-based dairying systems. *Journal of Dairy Science* 90(7):3234-3241.
- Auran, T. 1974. Studies on monthly and cumulative monthly milk yield records. II. The effect of calving interval and stage in pregnancy. *Acta Agriculturae Scandinavica* 24(4):339-348.
- Bachman, K. C., M. J. Hayen, D. Morse, and C. J. Wilcox. 1988. Effect of pregnancy, milk-yield, and somatic-cell count on bovine milk fat hydrolysis. *Journal of Dairy Science* 71(4):925-931.
- Bates, A. J. and I. Dohoo. 2016. Risk factors for peri-parturient farmer diagnosed mastitis in New Zealand dairy herds: findings from a retrospective cohort study. *Preventive Veterinary Medicine* 127:70-76.
- Bertilsson, J., B. Berglund, G. Ratnayake, K. Svennersten-Sjaunja, and H. Wiktorsson. 1997. Optimizing lactation cycles for the high-yielding dairy cow. A European perspective. *Livestock Production Science* 50(1-2):5-13.
- Bibby, J., and Toutenburg, H. 1977. Prediction and improved estimation in linear models, New York: John Wiley
- Borman, J. M., K. L. Macmillan, and J. Fahey. 2004. The potential for extended lactations in Victorian dairying: a review. *Australian Journal of Experimental Agriculture* 44(6):507-519.
- Brody, S., A. C. Ragsdale, and C. W. Turner. 1923. The rate of decline of milk secretion with the advance of the period of lactation. *Journal of General Physiology* 5(4):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. *Journal of General Physiology* 6(5):541-545.

- Butler, S. T., L. Shalloo, and J. J. Murphy. 2010. Extended lactations in a seasonal-calving pastoral system of production to modulate the effects of reproductive failure. *Journal of Dairy Science* 93(3):1283-1295.
- Boutinaud, M., Guinard-Flament, J. and Jammes, H. 2004. The number and activity of mammary epithelial cells, determining factors for milk production. *Reproduction Nutrition Développement* 44(5):499-508.
- Capuco, A. V., D. L. Wood, R. Baldwin, K. McLeod, and M. J. Paape. 2001. Mammary cell number, proliferation, and apoptosis during a bovine lactation: relation to milk production and effect of bST. *Journal of Dairy Science* 84(10):2177-2187.
- Capuco, A. V., S. E. Ellis, S. A. Hale, E. Long, R. A. Erdman, X. Zhao, and M. J. Paape. 2003. Lactation persistency: insights from mammary cell proliferation studies. *Journal of Animal Science* 81(Suppl.3):18-31.
- Chapinal, N., G. Zobel, K. Painter, and K. E. Leslie. 2014. Changes in lying behavior after abrupt cessation of milking and regrouping at dry-off in freestall-housed cows: A case study. *Journal of Veterinary Behavior: Clinical Applications and Research* 9(6):364-369
- Cobby, J. M. and Y. L. P. Le Du. 1978. On fitting curves to lactation data. *Animal Production* 26(2):127-133.
- Cole, J. B. and P. M. VanRaden. 2006. Genetic evaluation and best prediction of lactation persistency. *Journal of Dairy Science* 89(7):2722-2728.
- Cole, J. B., D. J. Null, and P. M. VanRaden. 2009. Best prediction of yields for long lactations. *Journal of Dairy Science* 92(4):1796-1810.
- Collard, B. L., P. J. Boettcher, J. C. M. Dekkers, D. Petitclerc, and L. R. Schaeffer. 2000. Relationships between energy balance and health traits of dairy cattle in early lactation. *Journal of Dairy Science* 83(11):2683-2690.
- Dekkers, J. C. M., J. H. Ten Hag, and A. Weersink. 1998. Economic aspects of persistency of lactation in dairy cattle. *Livestock Production Science* 53(3):237-252.
- Dematawewa, C. M. B., R. E. Pearson, and P. M. VanRaden. 2007. Modeling extended lactations of Holsteins. *Journal of Dairy Science* 90(8):3924-3936.
- De Vries, A. 2006. Economic value of pregnancy in dairy cattle. *Journal of Dairy Science* 89(10):3876-3885.
- Dijkstra, J. and J. France. 1995. Modelling and methodology in animal science. In: Danfær, A. and P. Lescoat. (eds), *Proceedings of the Fourth International Workshop on Modelling Nutrient Utilisation in Farm Animals*. National Institute of Animal Science, Foulum, Denmark, pages 9–18.

- 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. *Journal of Dairy Science* 80(10):2340-2354.
- Dingwell, R. T., D. F. Kelton, K. E. Leslie, and V. L. Edge. 2001. Deciding to dry-off: does level of production matter? Proceedings of the National Mastitis Council Annual Meeting. Natl. Mastitis Council, Inc., Arlington, VA (2001). Pages 69–79.
- Dossing, F. 1994. Clinical mastitis in the dry period. *Dansk Veterinaertidskrift* 77(8):353-359.
- Druet, T., F. Jaffrezic, D. Boichard, and V. Ducrocq. 2003. Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. *Journal of Dairy Science* 86(7):2480-2490.
- Durbin, J. and G. S. Watson. 1950. Testing for serial correlation in least squares regression. I. *Biometrika* 37(3/4):409-428.
- Ferris, T. A., I. L. Mao, and C. R. Anderson. 1985. Selecting for lactation curve and milk-yield in dairy cattle. *Journal of Dairy Science* 68(6):1438-1448.
- France, J. and Thornley, J. H. M. 1984. *Mathematical models in agriculture. A quantitative approach to problems in agriculture and related sciences.* Butterworths, London, 1984.
- Gonzalez-Recio, O., M. A. Perez-Cabal, and R. Alenda. 2004. Economic value of female fertility and its relationship with profit in Spanish dairy cattle. *Journal of Dairy Science* 87(9):3053-3061.
- Grossman, M. and W. J. Koops. 2003. Modeling extended lactation curves of dairy cattle: A biological basis for the multiphasic approach. *Journal of Dairy Science* 86(3):988-998.
- Grossman, M. and W. J. Koops. 1988. Multiphasic analysis of lactation curves in dairy cattle. *Journal of Dairy Science* 71(6):1598-1608.
- Guo, Z. and H. H. Swalve. 1995. Modelling of the lactation curve as a sub-model in the evaluation of test day records. INTERBULL open meeting, 7-8 September 1995, Prague, Czech Republic
- Hadsell, D., J. George, and D. Torres. 2007. The declining phase of lactation: peripheral or central, programmed or pathological? *Journal of Mammary Gland Biology and Neoplasia* 12(1):59-70.
- Hagnestam-Nielsen, C., U. Emanuelson, B. Berglund, and E. Strandberg. 2009. Relationship between somatic cell count and milk yield in different stages of lactation. *Journal of Dairy Science* 92(7):3124-3133.

- Haile-Mariam, M. and M. E. Goddard. 2008. Genetic and phenotypic parameters of lactations longer than 305 days (extended lactations). *Animal* 2(3):325-335.
- Inchaisri, C., R. Jorritsma, P. L. A. M. Vos, G. C. van der Weijden, and H. Hogeveen. 2011. Analysis of the economically optimal voluntary waiting period for first insemination. *Journal of Dairy Science* 94(8):3811-3823.
- Jensen, J. 2001. Genetic evaluation of dairy cattle using test-day models. *Journal of Dairy Science* 84(12):2803-2812.
- Kellogg, D. W., N. S. Urquhart, and A. J. Ortega. 1977. Estimating Holstein lactation curves with a gamma-curve. *Journal of Dairy Science* 60(8):1308-1315.
- Knight, C. H. 1997. Biological control of lactation length. *Livestock Production Science* 50(1):1-3.
- Knight, C. H. 2001. Lactation and gestation in dairy cows: flexibility avoids nutritional extremes. *Proceedings of the Nutrition Society* 60(4):527-537.
- Knight, C. H. and M. Peaker. 1984. Mammary development and regression during lactation in goats in relation to milk secretion. *Quarterly Journal of Experimental Physiology and Cognate Medical Sciences* 69(2):331-338.
- Knight, C. H. and C. J. Wilde. 1987. Mammary growth during lactation-implications for increasing milk yield. *Journal of Dairy Science* 70(9):1991-2000.
- Knight, C. H., Docherty, A. H. & Peaker, M. 1984. Milk yield in rats in relation to activity and size of the mammary secretory cell population. *Journal of Dairy Research* 51(1):29-35.
- Kocak, O. and B. Ekiz. 2008. Comparison of different lactation curve models in Holstein cows raised on a farm in the south-eastern Anatolia region. *Archiv Fur Tierzucht* 51(4):329-337.
- Kolver, E. S., J. R. Roche, C. R. Burke, J. K. Kay, and P. W. Aspin. 2007. Extending lactation in pasture-based dairy cows: I. Genotype and diet effect on milk and reproduction. *Journal of Dairy Science* 90(12):5518-5530.
- Larsson, B. and B. Berglund. 2000. Reproductive performance in cows with extended calving interval. *Reproduction in Domestic Animals* 35(6):277-280.
- Lee, J. K., P. M. VanRaden, H. D. Norman, G. R. Wiggans, and T. R. Meinert. 1997. Relationship of yield during early lactation and days open during current lactation with 305-day yield. *Journal of Dairy Science* 80(4):771-776.
- Lehmann, J. O., J. G. Fadel, L. Mogensen, T. Kristensen, C. Gaillard, and E. Kebreab. 2016. Effect of calving interval and parity on milk yield per feeding day in Danish commercial dairy herds. *Journal of Dairy Science* 99(1):621-633.

- Lehmann, J. O., L. Mogensen, and T. Kristensen. 2017. Early lactation production, health, and welfare characteristics of cows selected for extended lactation. *Journal of Dairy Science* 100(2):1487-1501.
- Lin, I. L. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45(1):255-268
- Lopez, S., J. France, N. E. Odongo, R. A. McBride, E. Kebreab, O. AlZahal, B. W. McBride, and J. Dijkstra. 2015. On the analysis of Canadian Holstein dairy cow lactation curves using standard growth functions. *Journal of Dairy Science* 98(4):2701-2712.
- Lucy, M.C. 2001. Reproductive loss in high-producing dairy cattle: Where will it end? *Journal of Dairy Science* 84(6):1277-1293.
- Macciotta, N. P. P., D. Vicario, and A. Cappio-Borlino. 2005. Detection of different shapes of lactation curve in dairy cattle by empirical mathematical models. *Journal of Dairy Science* 88(3):1178-1191.
- Marett, L. C., M. J. Auldist, C. Grainger, W. J. Wales, D. Blache, K. L. Macmillan, and B. J. Leury. 2011. Temporal changes in plasma concentrations of hormones and metabolites in pasture-fed dairy cows during extended lactation. *Journal of Dairy Science* 94(10):5017-5026.
- Mellado, M., J. M. Flores, A. de Santiago, F. G. Veliz, U. Macias-Cruz, L. Avendano-Reyes, and J. E. Garcia. 2016. Extended lactation in high-yielding Holstein cows: Characterization of MY and risk factors for lactations > 450 days. *Livestock Science* 189:50-55.
- Mellenberger, R. W., D. E. Bauman, and D. R. Nelson. 1973. Metabolic adaptations during lactogenesis. Fatty acid and lactose synthesis in cow mammary tissue. *Biochemical Journal* 136(3):741-748.
- Miller, N., L. Delbecchi, D. Petitclerc, G. F. Wagner, B. G. Talbot, and P. Lacasse. 2006. Effect of stage of lactation and parity on mammary gland cell renewal. *Journal of Dairy Science* 89(12):4669-4677.
- Muir, B. L., J. Fatehi, and L. R. Schaeffer. 2004. Genetic relationships between persistency and reproductive performance in first-lactation Canadian Holsteins. *Journal of Dairy Science* 87(9):3029-3037.
- Nemeckova, D., L. Stadnik, and J. Citek. 2015. Associations between milk production level, calving interval length, lactation curve parameters and economic results in Holstein cows. *Mljekarstvo*. 65(4):243-250.
- Niozas, G., G. Tsousis, I. Steinhofel, C. Brozos, A. Roemer, S. Wiedemann, H. Bollwein, and M. Kaske. 2019a. Extended lactation in high-yielding dairy cows. I. Effects on reproductive measurements. *Journal of Dairy Science* 102(1):799-810.

- Niozas, G., G. Tsousis, C. Malesios, I. Steinhofel, C. Boscas, H. Bollwein, and M. Kaske. 2019b. Extended lactation in high-yielding dairy cows. II. Effects on milk production, udder health, and body measurements. *Journal of Dairy Science* 102(1):811-823.
- NMC (National Mastitis Council). 2006. NMC Factsheet - Dry Cow Therapy. National Mastitis Council, New Prague, MN.
- Neal, H. and J. H. M. Thornley. 1983. The lactation curve in cattle – a mathematical model of the mammary gland. *Journal of Agricultural Science* 101(2):389-400.
- Nelder, J. A. 1966. Inverse polynomials, a useful group of multi-factor response functions. *Biometrics* 22(1):128-141.
- Olori, V. E., S. Brotherstone, W. G. Hill, and B. J. McGuirk. 1997. Effect of gestation stage on MY and composition in Holstein Friesian dairy cattle. *Livestock Production Science* 52(2):167-176.
- Olori, V. E., S. Brotherstone, W. G. Hill, and B. J. McGuirk. 1999. Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd. *Livestock Production Science* 58(1):55- 63.
- Osterman, S., and J. Bertilsson. 2003. Extended calving interval in combination with milking two or three times per day: effects on milk production and milk composition. *Livestock Production Science* 82(2-3):139-149.
- Pollott, G. E. 2000. A biological approach to lactation curve analysis for milk yield. *Journal of Dairy Science* 83(11):2448-2458.
- Rajala-Schultz, P. J., J. S. Hogan, and K. L. Smith. 2005. Short communication: association between milk yield at dry-off and probability of intramammary infections at calving. *J. Dairy Sci.* 88(2):577–579.
- Rajala-Schultz, P. J., P. N. Gott, K. L. Proudfoot, and M. Schuenemann. 2018. Effect of milk cessation method at dry-off on behavioral activity of dairy cows. *Journal of Dairy Science* 101(4):3261-3270.
- Roche, J. R. 2003. Effect of pregnancy on milk production and bodyweight from identical twin study. *Journal of Dairy Science* 86(3):777-783.
- Roemer, A. 2011. Investigations on longevity in German Holstein cows. *Zuchtungskunde* 83(1):8-20.
- Safayi, S., P. K. Theil, V. S. Elbrond, L. Hou, M. Engbaek, J. V. Norgaard, K. Sejrsen, and M. O. Nielsen. 2010. Mammary remodeling in primiparous and multiparous dairy goats during lactation. *Journal of Dairy Science* 93(4):1478-1490.

Scheu, A., A. Powell, R. Bollongino, J. D. Vigne, A. Tresset, C. Cakirlar, N. Benecke, and J. Burger. 2015. The genetic prehistory of domesticated cattle from their origin to the spread across Europe. *BMC Genetics* 16:11.

Schaeffer, L. R., J. Jamrozik, G. J. Kistemaker, and B. J. Van Doormaal. 2000. Experience with a test-day model. *Journal of Dairy Science* 83(5):1135-1144.

Schindler, H., S. Eger, M. Davidson, D. Ochowski, E. C. Schermerhorn, and R. H. Foote. 1991. Factors affecting response of groups of dairy-cows managed for different calving-conception intervals. *Theriogenology* 36(3):495-503.

Scott, T. A., B. Yandell, L. Zepeda, R. D. Shaver, and T. R. Smith. 1996. Use of lactation curves for analysis of milk production data. *Journal of Dairy Science* 79(10):1885-1894.

Sikka, L.C. 1950. A study of lactation as affected by heredity and environment. *Journal of Dairy Research* 17(3):231-252.

Silvestre, A. M., F. Petim-Batista, and J. Colaco. 2006. The accuracy of seven mathematical functions in modeling dairy cattle lactation curves based on test-day records from varying sample schemes. *Journal of Dairy Science* 89(5):1813-1821.

Stangaferro, M. L., R. Wijma, M. Masello, M. J. Thomas, and J. O. Giordano. 2018. Economic performance of lactating dairy cows submitted for first service timed artificial insemination after a voluntary waiting period of 60 or 88 days. *Journal of Dairy Science* 101(8):7500-7516.

Strandberg, E. and P. A. Oltenacu. 1989. Economic consequences of different calving intervals. *Acta Agriculturae Scandinavica* 39(4):407-420.

Tekerli, M., Z. Akinci, I. Dogan, and A. M. Akcan. 2000. Factors affecting the shape of lactation curve of Holstein cows from the Balikesir Province of Turkey. *Journal of Dairy Science* 83(6):1381-1386.

Torshizi, M. E., A. A. Aslamenejad, M. R. Nassiri, and H. Farhangfar. 2011. Comparison and evaluation of mathematical lactation curve functions of Iranian primiparous Holsteins. *South African Journal of Animal Science* 41(2):104-115.

Tsuruta, S., I. Misztal, and T. J. Lawlor. 2005. Changing definition of productive life in US Holsteins: effect on genetic correlations. *Journal of Dairy Science* 88(3):1156-1165.

Val-Arreola, D., E. Kebreab, J. Dijkstra, and J. France. 2004. Study of the lactation curve in dairy cattle on farms in Central Mexico. *Journal of Dairy Science* 87(11):3789-3799.

van Amburgh, M. E., D. M. Galton, D. E. Bauman, and R. W. Everett. 1997. Management and economics of extended calving intervals with use of bovine somatotropin. *Livestock Production Science* 50(1-2):15-28.

- VanRaden, P. M., C. M. B. Dematawewa, R. E. Pearson, and M. E. Tooker. 2006. Productive life including all lactations and longer lactations with diminishing credits. *Journal of Dairy Science* 89(8):3213-3220.
- Vargas, B., W. J. Koops, M. Herrero, and J. A. M. van Arendonk. 2000. Modeling extended lactations of dairy cows. *Journal of Dairy Science* 83(6):1371-1380.
- Vetharaniam, I., S. R. Davis, T. K. Soboleva, P. R. Shorten, and G. C. Wake. 2003. Modeling the interaction of milking frequency and nutrition on mammary gland growth and lactation. *Journal of Dairy Science* 86(6):1987-1996.
- Wathes, D. C., Z. Cheng, N. Bourne, V. J. Taylor, M. P. Coffey, and S. Brotherstone. 2007. Differences between primiparous and multiparous dairy cows in the inter-relationships between metabolic traits, milk yield and body condition score in the periparturient period. *Domestic Animal Endocrinology* 33(2):203-225.
- Weller, J.I. and Y. Folman. 1990. Effects of calf value and reproductive management on optimum days to 1st breeding. *Journal of Dairy Science* 73(5):1318-1326.
- White, I. M. S., R. Thompson, and S. Brotherstone. 1999. Genetic and environmental smoothing of lactation curves with cubic splines. *Journal of Dairy Science* 82(3):632-638.
- Wilde, C. J., A. J. Henderson, and C. H. Knight. 1986. Metabolic adaptations in goat mammary tissue during pregnancy and lactation. *Journal of Reproduction and Fertility* 76(1):289-298.
- Wilmink, J. B. M. 1987. Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. *Livestock Production Science* 16(4):335-348.
- Wood, P. D. P. 1967. Algebraic model of the lactation curve in cattle. *Nature* 216(5111):164-165.