



Fitting mathematical functions to extended lactation curves and forecasting late-lactation milk yields of dairy cows

David J. Innes,^{1*} Linaya J. Pot,^{1*} Dave J. Seymour,^{1,2} James France,¹ Jan Dijkstra,³ John Doelman,² and John P. Cant^{1†}

¹Centre for Nutrition Modelling, Department of Animal Biosciences, University of Guelph, Ontario, N1G 2W1 Canada

²Trouw Nutrition R&D, 3800 AG Amersfoort, the Netherlands

³Animal Nutrition Group, Wageningen University and Research, 6700 AH Wageningen, the Netherlands.

ABSTRACT

A 305-d lactation followed by a 60-d dry period has traditionally been considered economically optimal, yet dairy cows in modern intensive dairy systems are frequently dried off while still producing significant quantities of milk. Managing cows for an extended lactation has reported production, welfare, and economic benefits, but not all cows are suitable for an extended lactation. Implementation of an extended lactation strategy on-farm could benefit from use of a decision support system, based on a mathematical lactation model, that can identify suitable cows during early lactation that have a high likelihood of producing above a target milk yield (MY) at 305 d in milk (DIM). Therefore, our objectives were (1) to compare the suitability of 3 commonly used lactation models for modeling extended lactations (Dijkstra, Wood, and Wilmink) in primiparous and multiparous cows under a variety of lactation lengths, and (2) to determine the amount of early-lactation daily MY data needed to accurately forecast MY at d 305 by using the most suitable model and determine whether this is sufficient for identifying cows suitable for an extended lactation before the end of a typical voluntary waiting period (50–90 d). Daily MY data from 467 individual Holstein-Friesian lactations (DIM >305 d; 379 ± 65-d lactation length [mean ± SD]) were fitted by the 3 lactation models using a nonlinear regression procedure. The parameter estimates of these models, lactation characteristics (peak yield, time to peak yield, and persistency), and goodness-of-fit were compared between parity and different lactation lengths. The models had similar performance, and differences between parity groups were consistent with previous literature. Then, data from only the first i DIM for each individual lactation, where

i was incremented by 30 d from 30 to 150 DIM and by 50 d from 150 to 300 DIM, were fitted by each model to forecast MY at d 305. The Dijkstra model was selected for further analysis, as it had superior goodness-of-fit statistics for $i = 30$ and 60. The data set was fit twice by the Dijkstra model, with parameter bounds either unconstrained or constrained. The quality of predictions of MY at d 305 improved with increasing data availability for both models and assisting the model fitting procedure with more biologically relevant constraints on parameters improved the predictions, but neither was reliable enough for practical use on-farm due to the high uncertainty of forecasted predictions. Using 90 d of data, the constrained model correctly classified 66% of lactations as being above or below a target MY at d 305 of 25 kg/d, with a probability threshold of 0.95. The proportion of correct classifications became smaller at lower targets of MY at d 305 and became greater when using more lactation days. Overall, further work is required to develop a model that can forecast late-lactation MY with sufficient accuracy for practical use. We envisage that a hybridized machine learning and mechanistic model that incorporates additional historical and genetic information with early-lactation MY could produce meaningful lactation curve forecasts.

Key words: mathematical model, lactation persistency, decision support

INTRODUCTION

Dairy cows in modern intensive, confined dairy systems are frequently dried off while still producing significant quantities of milk. Similarly, modern grazing dairy systems use seasonal calving or batch calving to manage fluctuations in environmental conditions or feed availability, as in Ireland, New Zealand, and Australia. Regardless of the system, a 305-d lactation followed by a 60-d dry period has traditionally been considered economically optimal (Strandberg and Oltenacu, 1989; Inchaisri et al., 2011). To maintain a 305-d lactation,

Received March 10, 2023.

Accepted July 31, 2023.

*These authors contributed equally to this work.

†Corresponding author: jcant@uoguelph.ca

dairy producers are tasked with maintaining a suitable voluntary waiting period (**VWP**), typically 50 to 90 d, before inseminating cows postpartum. This coincides with a period of high milk yield (**MY**) and usually negative energy balance of the cow, associated with high reproductive failure (Santos et al., 2009). In reality, over 50% of Holsteins in North America have lactations beyond 305 d (Tsuruta et al., 2005; VanRaden et al., 2006), with the mean lactation lengths in the US dairy cattle population estimated as 366 and 357 d for 2007 and 2017, respectively (Capper and Cady, 2020). This can be attributed to genetic and management factors and physiological challenges influencing reproduction, as reviewed in Walsh et al. (2011).

The mating period can be temporally separated from the period of high metabolic demand by intentionally delaying breeding beyond the typical VWP and subsequently extending the lactation. Such an extended lactation strategy may reduce the number of calves, decrease breeding costs, increase cow longevity within a herd, and improve animal welfare on-farm (Knight, 2001; Niozas et al., 2019b). However, the inherent variability between individual cows in their ability to maintain an extended lactation is well documented (Kolver et al., 2007; Butler et al., 2010; Niozas et al., 2019a; van Knegsel et al., 2022), and thus not all cows are suitable for an extended lactation. Some of the most significant types of data used by producers to evaluate herds or individual cows for extended lactations in the Netherlands and Denmark include BCS, MY, peak yield (**PY**), DIM, milking frequency data (in automatic milking systems), and persistency (Lehmann et al., 2017; Burgers et al., 2021b). Therefore, implementation of an extended lactation strategy on-farm could benefit from a decision support system that can identify suitable cows during early lactation that have a high likelihood of producing above a target MY at 305 DIM. Previously, MY in early lactation has been positively correlated with the MY in extended lactations (Lehmann et al., 2017). Machine learning models have been used to predict herd-level MY over different time horizons (Murphy et al., 2014) and to interpolate and predict missing MY data in individual lactations (Liseune et al., 2020) or predict the entire MY curve of individual cows based on previous lactation information (Liseune et al., 2021). Furthermore, discriminant analysis of lactation curve model parameters has been used to classify cows as having either a low (<20 kg/d) or high (>35 kg/d) MY at d 305 using daily MY data from the first 90, 120, or 150 DIM (Manca et al., 2020). However, these methods require large training data sets and generate predictions stochastically.

In contrast, deterministic models of lactation data are commonly used to predict future MY, persistency,

and PY of a lactation. Use of such models is common in various farm management software and genetic evaluation tools, even when only monthly test-day data are available. However, many commonly used lactation models have been created and optimized for traditional lactation lengths, and their ability to fit individual extended lactations needs to be evaluated. Previously, Grossman and Koops (2003) developed a specific multiphasic approach to modeling extended lactations, linking a first ascending phase of MY with a series of 3 descending phases of yield. Dematawewa et al. (2007) used test-day data from a large database to evaluate various types of extended lactation models fitted to herd data, whereas modeling individual extended lactations has previously required the combination of the Wilmink (1987) lactation model with a linear or squared function (Otwinowska-Mindur et al., 2021). Therefore, the first objective of this study was to compare the suitability of 3 commonly used lactation models for modeling extended lactations in primiparous and multiparous cows under a variety of lactation lengths. The Dijkstra et al. (1997) mechanistic model and the Wood (1967) and Wilmink (1987) empirical models—herein referred to as Dijkstra, Wood, and Wilmink, respectively—were selected for analysis due to their small numbers of parameters, frequency of use within literature, and flexibility and high accuracy to fit a wide range of lactation shapes. The second objective of our study was to determine the amount of early-lactation daily MY data needed to accurately forecast MY at d 305 by using the most suitable model from the first objective and determine whether this is sufficient for identifying cows suitable for an extended lactation before the end of a typical VWP (50 to 90 d).

MATERIALS AND METHODS

Because no human or animal subjects were used, this analysis did not require approval by an Institutional Animal Care and Use Committee or Institutional Review Board.

Data Source

Data were collected from lactating Holstein-Friesian cows housed in a freestall barn at the Trouw Nutrition Dairy Research Facility, located in Boxmeer, the Netherlands. Cattle were milked twice daily and fed a forage-based partial mixed ration *ad libitum* with additional access to concentrate feeders within the barn. The daily amount of concentrate allocated to each cow increased with higher milk production, and the allocation for primiparous cows began at a lower MY than multiparous cows. Cows were dried off 6 wk before calv-

Table 1. General description of the data sets used for full model fits and incremental model fits with forecasting

Variable	Fitting	
	Full	Incremental
Number		
Cows	262	229
Lactations	467	365
Primiparous	146	121
Multiparous	321	244
305 to 399 d in length	367	286
>400 d in length	100	79
Records per lactation		
Mean	310	310
SD	50	49
Lactation length (d)		
Mean	379	379
SD	65	65
Minimum	306	306
Maximum	670	669
Milk yield at d 305 ¹ (kg/d)		
Mean	21.9	22.0
SD	6.4	6.1
Minimum	2.0	5.6
Maximum	40.9	40.9

¹Calculated as the mean of daily milk yield from 303 to 307 DIM.

ing by removing concentrate allocations by 2 wk before dry off while maintaining twice-daily milking until dry off. The data set consisted of cow identification number, parity, DIM, and daily MY collected between June 2011 and October 2015. The data set was filtered to retain single lactations with a maximum DIM >305, MY records for >150 d of the lactation, and with at least 1 record before 10 DIM. Observations of MY with an absolute residual ≥ 3 standard deviations (**SD**) from the mean residual following smoothing (loess function from “utils” package in base R: span = 0.25 and surface = “direct”) were removed as extreme points. Overall, 467 lactation curves were retained with a maximum DIM >305 d (Table 1). The R scripts used for all analyses are deposited at <https://doi.org/10.5683/SP3/4QV100> (Innes et al., 2023) and maintained at <https://github.com/CNM-University-of-Guelph/forecasting-extended-lactation>, including a simulated lactation data set used as a reproducible example.

Model Fitting

The 3 lactation models described in Table 2 were fitted to individual lactation curves using a nonlinear regression procedure with the Levenberg-Marquardt algorithm, which was implemented using the *nlsmultstart* function from the “nls.multstart” package in R (Padfield and Matheson, 2020). Initial search grids for the parameters were based on ranges of the parameters found in the literature and are included in the supple-

Table 2. Equations for estimated daily milk yield (kg/d) and time to peak yield (d) for the 3 lactation curve models used¹

Model reference	Estimated daily milk yield	Time to peak yield
Wood, 1967	$at^b e^{-ct}$	b/c
Wilmink, 1987	$a - be^{-b_0 t} - ct$	$\ln(b \cdot b_0 / c) / b_0$
Dijkstra et al., 1997	$ae^{b(1-e^{-b_0 t})/b_0 - ct}$	$\ln(b/c) / b_0$

¹ t is DIM, and a , b , b_0 , and c are model parameters.

mentary code, with bounds put in place for the Dijkstra model to ensure the parameters remained biologically plausible: $a \geq 0$, $0 \leq b \leq 1$, $0 \leq b_0 \leq 1$, and $0 \leq c \leq 1$. 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 1.5×10^{-6} .

Daily MY and time to peak yield (**PT**) were calculated using the equations in Table 2, and PY was calculated by solving daily MY equations for $t = \text{PT}$. 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_{end}), divided by the time interval:

$$\text{Persistency (kg/d per d)} = \frac{(\text{MY}_{t_{\text{end}}} - \text{PY})}{(t_{\text{end}} - \text{PT})}$$

Missing daily MY values (15.7%) were linearly interpolated using the *na.approx* function from the “zoo” package in R (Zeileis and Grothendieck, 2005) for the purpose of estimating observed total 305-d MY and MY at d 305 independently from any model fitting. Total 305-d MY for each lactation was calculated as the sum of observed MY between 10 and 305 DIM, as there was very low frequency of observed MY data below 10 DIM. The MY at d 305 of lactation was estimated as the mean daily observed MY from 303 to 307 DIM.

Goodness-of-fit for the 3 models was assessed for each lactation curve with the square root of the mean square prediction error (**RMSPE**):

$$\text{RMSPE} = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}}$$

where n = number of observations within the lactation, O_i are observed MY values, and P_i are predictions. The RMSPE is indicative of the SD of the unexplained vari-

ance, with a smaller RMSPE indicating a better model fit. Additionally, the mean absolute error (MAE) was calculated as follows:

$$\text{MAE} = \frac{\sum_{i=1}^n |O_i - P_i|}{n}.$$

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. The RMSPE and MAE values were expressed as a proportion of mean observed MY values.

To assess both the accuracy and precision of model fits, Lin's (1989) concordance correlation coefficient (CCC) was calculated as $r \times C_b$ where r is the Pearson correlation coefficient and C_b is the bias correction factor. To test for the presence of first-order autocorrelation among residuals, the Durbin-Watson (DW) coefficient (Durbin and Watson, 1950) was calculated using the *durbinWatsonTest* function from the "car" package in R (Fox and Weisberg, 2019). This 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. After models were fitted 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 within a model were identified with Student's *t*-test.

Forecasting Milk Yield at 305 Days in Milk

To determine the most suitable model for forecasting, data from only the first i DIM for each individual lactation, where i was incremented by 30 d from 30 to 150 DIM and by 50 d from 150 to 300 DIM, were fitted by each model to forecast MY at d 305. Fitting attempts that did not converge on parameter estimates for any of the models using the full data set resulted in removal of the entire lactation from the data set of observations so that the same number of fits was compared across all i values and models. This data set was fitted by each model using parameter bounds matching those used above (unconstrained). It was determined that, although the Dijkstra and Wood model had similar 305-d MAE (%) and CCC values, the Dijkstra model performed better with the 30- and 60-DIM thresholds (Supplemental Figure S1, <https://doi.org/10.5683/SP3/4QVI00>;

doi.org/10.5683/SP3/4QVI00; Innes et al., 2023). Therefore, the Dijkstra model was used for subsequent analyses.

The full data set was re-filtered to keep only lactations where parameter estimates converged for the Dijkstra model, which resulted in a revised data set of 365 lactations (Table 1). This data set was fitted twice by the Dijkstra model for each incremented DIM threshold, first with parameter bounds matching those used above (unconstrained) and second with the b and b_0 parameter constrained to the mean \pm 1.0 SD for all lactations in the revised data set ($n = 365$) when all DIM were fitted, and the c parameter constrained to the mean \pm 1.5 SD after removal of outliers ($c > 0.006$): $1 \times 10^{-8} \leq b \leq 0.10924$; $0.01725 \leq b_0 \leq 0.08723$; and $0.00087 \leq c \leq 0.00449$. These thresholds were determined from visual inspection of histograms to represent the range of most meaningful values (Supplemental Figure S2, <https://doi.org/10.5683/SP3/4QVI00>; Innes et al., 2023).

To account for uncertainty in prediction, the probability (Prob) that MY at d 305 would be above a certain target X_{305} (kg/d) for each incrementally fitted lactation curve was estimated from a cumulative one-tailed t-distribution (TDIST) with $n - 4$ degrees of freedom to account for the 4 parameters of the Dijkstra model:

$$\text{Prob}(P_{305} \geq X_{305}) = 1 - \text{TDIST}\left(\frac{X_{305} - P_{305}}{s_{\text{pred}}}, n - 4\right),$$

where P_{305} is the predicted MY at d 305. The standard deviation of the prediction, s_{pred} , was estimated as RMSPE (kg/d) adjusted by $n - 4$ instead of n .

RESULTS

Extended Lactations Grouped by Parity

The Wood, Wilmink, and Dijkstra models were fitted to each individual lactation and then summarized by parity (primiparous or multiparous). The primiparous lactations had observed mean daily MY of 27.3 kg/d and total 305-d MY of 8,477 kg compared with 31.3 kg/d and 10,060 kg, respectively, for multiparous lactations, but both groups had similar mean lactation lengths of 367 and 371 d, respectively, and MY at d 305 of 22.3 and 21.7 kg/d, respectively.

The rise to PY is represented by parameters a , b , and b_0 in the fitted models, whereas the c parameter represents the declining phase of the lactation. The Wood, Wilmink, and Dijkstra models had mean MY scaling parameters, a , that were 23, 41, and 31%

Table 3. Mean parameters and statistics of fits of Wood, Wilmink, and Dijkstra models to individual lactation curves >305 d in length of primiparous (PP; n = 146) and multiparous (MP; n = 321) cows¹

Item	Wood		Wilmink		Dijkstra	
	PP	MP	PP	MP	PP	MP
<i>a</i>	15.8 ^a	20.4 ^b	37.4 ^a	63.0 ^b	19.3 ^a	28.1 ^b
<i>b</i>	0.23 ^a	0.25 ^b	18.7 ^a	38.7 ^b	0.11 ^a	0.15 ^b
<i>b</i> ₀			0.040 ^a	0.049 ^b	0.047 ^a	0.039 ^b
<i>c</i>	0.0030 ^a	0.0044 ^b	0.032 ^a	0.087 ^b	0.071 ^a	0.120 ^b
Peak yield (kg/d)	32.6 ^a	42.4 ^b	32.7 ^a	42.1 ^b	33.0 ^a	42.4 ^b
Time to peak (d)	76.0 ^a	57.1 ^b	75.9 ^a	57.6 ^b	74.6 ^a	61.3 ^b
Persistence (kg/d per day)	-0.044 ^a	-0.080 ^b	-0.051 ^a	-0.088 ^b	-0.047 ^a	-0.083 ^b
MAE (%)						
Mean	6.39 ^a	7.16 ^b	5.71 ^a	6.22 ^b	6.00 ^a	6.77 ^b
SD	1.58	1.75	1.18	1.31	1.34	1.56
RMSPE (%)						
Mean	8.11 ^a	9.10 ^b	7.35 ^a	7.99 ^b	7.68 ^a	8.63 ^b
SD	2.02	2.18	1.58	1.70	1.75	1.93
CCC	0.85 ^a	0.93 ^b	0.88 ^a	0.95 ^b	0.86 ^a	0.94 ^b
<i>r</i>	0.86 ^a	0.93 ^b	0.88 ^a	0.95 ^b	0.88 ^a	0.94 ^b
<i>C</i> _b	0.98 ^a	1.00 ^b	0.99 ^a	1.00 ^b	0.99 ^a	1.00 ^b
Durbin-Watson	0.68	0.72	0.80 ^a	0.90 ^b	0.75	0.79

^{a,b}Different superscripts within a model and row indicate significantly different means ($P \leq 0.05$).

¹*a*, *b*, *b*₀, and *c* = model parameters; MAE = mean absolute error (% of mean observed milk yield); RMSPE = square root of mean square prediction error (% of mean observed milk yield); CCC = concordance correlation coefficient; *r* = Pearson correlation coefficient (precision); *C*_b = bias correction factor (accuracy).

lower, respectively, in the primiparous compared with the multiparous lactations (Table 3). Similarly, initial rates of mammary cell proliferation, represented by parameter *b*, were 8, 52, and 27% lower, respectively. The mean decay parameter, *b*₀, was higher in the primiparous lactations compared with multiparous lactations in the Dijkstra model but was lower in the primiparous lactations in the Wilmink model. The Wood, Wilmink, and Dijkstra models had mean declining phase parameters, *c*, that were 32, 63, and 41% lower, respectively, in the primiparous compared with the multiparous lactations. The lower mean *a*, *b*, and *c* parameters in the primiparous lactations are consistent with their lower PY, longer PT, and higher persistency compared with multiparous lactations, from fits to all 3 models (Table 3). The mean estimated PY was consistently between 22 and 23% lower in primiparous lactations across all models.

The CCC was lower from primiparous lactations compared with multiparous lactations ($P \leq 0.05$), largely due to poorer correlation (*r*) between predicted and observed values, which is a measure of model precision. Despite this, the mean overall residual error, estimated by RMSPE and MAE and expressed as a percentage of observed MY, were lower for primiparous lactations ($P \leq 0.05$). Plots of the individual fits with the lowest (Figure 1a) and highest (Figure 1b) mean MAE show how CCC could be lower in lactations with perturbations (e.g., 752_4 in Figure 1b) but also higher when MAE was also high (e.g., 884_2 in Figure 1b).

Extended Lactations Grouped by Lactation Length

The extended lactation fits by the Wood, Wilmink, and Dijkstra models were also summarized by lactation length: 305 to 399 DIM (shorter group, **S**; n = 367), and ≥ 400 DIM (longer group, **L**; n = 100). The group S lactations had an observed mean lactation length of 346 d, MY at d 305 of 20.5 kg/d, and total 305-d MY of 9,401 kg compared with 457 d, 26.8 kg/d, and 10,166 kg, respectively, for group L lactations. The 2 groups had similar mean daily MY of 30.1 and 30.0 kg/d, respectively. The *a*, *b*, and *c* parameters of group S in the Wood model were distinct from group L, whereas none of the parameters were different between group S and group L for the Wilmink and Dijkstra models (Table 4). Mean PY was not different between groups, but persistency was higher in group L for all models. The CCC indicated a poorer fit by the Wood model in group S, but there were no differences in MAE, RMSPE, and DW between groups for all models (Table 4).

Analysis of Residuals

The DW statistics (Tables 3 and 4) indicate that all fits contained strong positive autocorrelation among 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, indicating stages of

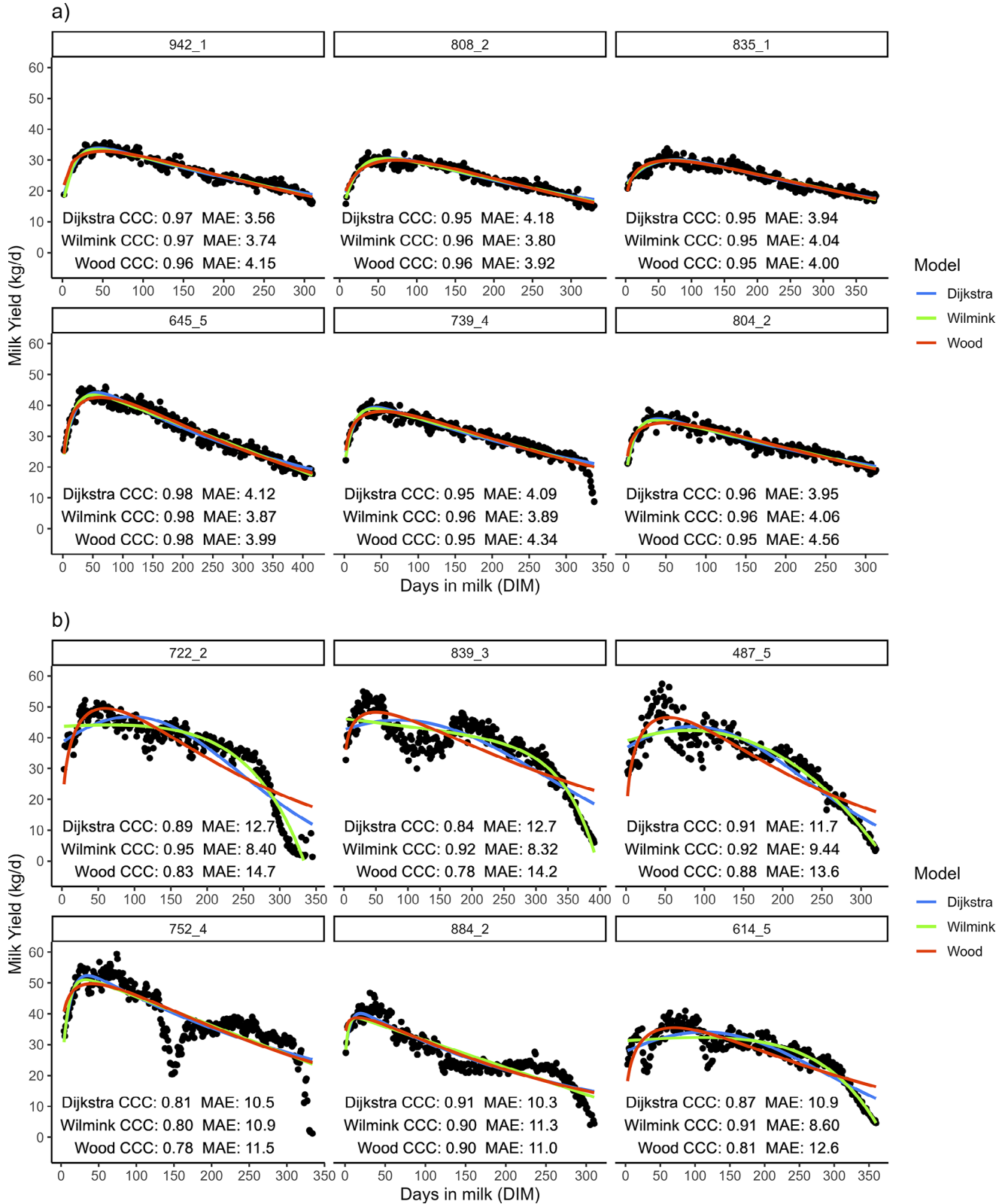


Figure 1. The (a) 6 best and (b) 6 worst individual lactations based on the mean of mean absolute error (MAE) values for the fitted Dijkstra, Wilmlink, and Wood models. Subplot titles are in the format “ID_parity” (e.g., 839_3 is cow 839 in her third lactation). The MAE (% of observed mean milk yield) and concordance correlation coefficient (CCC) for each model is displayed for each lactation. Lower values of MAE and higher values of CCC indicate a better goodness-of-fit.

Table 4. Mean parameters and statistics of fits of Wood, Wilmink, and Dijkstra models to individual lactation curves with a length of 305 to 399 d (group S; n = 367) and ≥ 400 d (group L; n = 100)¹

Item	Wood		Wilmink		Dijkstra	
	<399 d	≥ 400 d	<399 d	≥ 400 d	<399 d	≥ 400 d
<i>a</i>	18.5 ^a	20.8 ^b	57.7	44.9	25.2	25.7
<i>b</i>	0.25 ^a	0.22 ^b	35.2	22.4	0.15	0.12
<i>b</i> ₀			0.046	0.048	0.041	0.044
<i>c</i>	0.0042 ^a	0.0032 ^b	0.073	0.056	0.113	0.076
Peak yield (kg/d)	39.0	40.3	38.9	39.8	39.2	40.6
Time to peak (d)	61.9 ^a	67.2 ^b	59.6	81.2	65.2	66.6
Persistence (kg/d per day)	-0.071 ^a	-0.059 ^b	-0.080 ^a	-0.062 ^b	-0.075 ^a	-0.060 ^b
MAE (%)						
Mean	6.90	7.00	6.01	6.25	6.49	6.68
SD	1.79	1.52	1.30	1.24	1.57	1.41
RMSPE (%)						
Mean	8.77	8.88	7.73	8.02	8.29	8.51
SD	2.26	1.84	1.72	1.54	1.98	1.70
CCC	0.90 ^a	0.92 ^b	0.92	0.93	0.91	0.92
<i>r</i>	0.91 ^a	0.92 ^b	0.93	0.94	0.92	0.93
<i>C</i> _b	0.99 ^a	0.99 ^b	0.99	1.00	0.99	0.99
Durbin-Watson	0.72	0.68	0.88	0.82	0.79	0.73

^{a,b}Different superscripts within a model and row indicate significantly different means ($P \leq 0.05$).

¹*a*, *b*, *b*₀, and *c* = model parameters; MAE = mean absolute error (% of mean observed milk yield); RMSPE = square root of mean square prediction error (% of mean observed milk yield); CCC = concordance correlation coefficient; *r* = Pearson correlation coefficient (precision); *C*_b = bias correction factor (accuracy).

lactation when milk production was over- and under-predicted (Figure 2). Notably, the residuals in late lactation ($> \sim 200$ DIM) from the Wilmink model fits were closer to 0 than those from the other 2 models. All 3 models produced similar patterns that were repeated across parity and lactation-length groups except for at the very start of lactation, where the Wood residuals were positive (but declining), indicating underestimation of MY, and Wilmink and Dijkstra were negative, indicating overestimation of MY. After this, the residuals rapidly increased and became positive, indicating that MY was underestimated as cows approached PY. After PY was reached, the residuals declined and increased again, tending to be negative between 80 and 200 DIM and positive between 200 and 300 DIM. This repeated oscillation between positive and negative errors throughout the groups indicates that these models cannot capture all the local features of a lactation curve. The increased variability between days toward the end of the longer lactations can be attributed to the fewer number of lactations represented in each average point (Supplemental Figure S3, <https://doi.org/10.5683/SP3/4QV100>; Innes et al., 2023). However, it appears that the majority of residuals at the immediate ends of the lactation tend to be negative, indicating an overprediction of milk production from the models at this time. This is consistent with some of the lactations in Figure 1b that had a very fast decline in MY at the end.

Relationship Between Early-Lactation Information and Predictions of MY at Day 305

Overall, the goodness-of-fit statistics between the models were numerically and practically similar in Table 3 and 4. However, the Dijkstra model was selected for the forecasting analysis because it allowed for a mechanistic interpretation of its parameters and an initial comparison of forecasting 305-d MY with each model found that the 305-d MAE (%) and 305-d CCC were superior for the Dijkstra model when using 30- and 60-DIM thresholds (Supplemental Figure S2). When the parameters of the Dijkstra model were unconstrained, the predicted MY at d 305 was biologically realistic when the DIM threshold was ≥ 60 DIM (Table 5). However, the large values for mean *b* and *c* parameters (Supplemental Table S1, <https://doi.org/10.5683/SP3/4QV100>; Innes et al., 2023), as well as time to PY and forecasted persistency, indicated considerable error in the predictions, whereas these values were much closer to biologically realistic values for all DIM thresholds when the parameters of the model were constrained (Table 5). The predicted MY at d 305 was compared with observed using RMSPE, MAE, and CCC, which were all better for the constrained model when DIM threshold was ≤ 150 .

The distribution of the observed MY at d 305 of the extended lactation data used for the incremental fit analysis is represented by the blue violin plots in Fig-

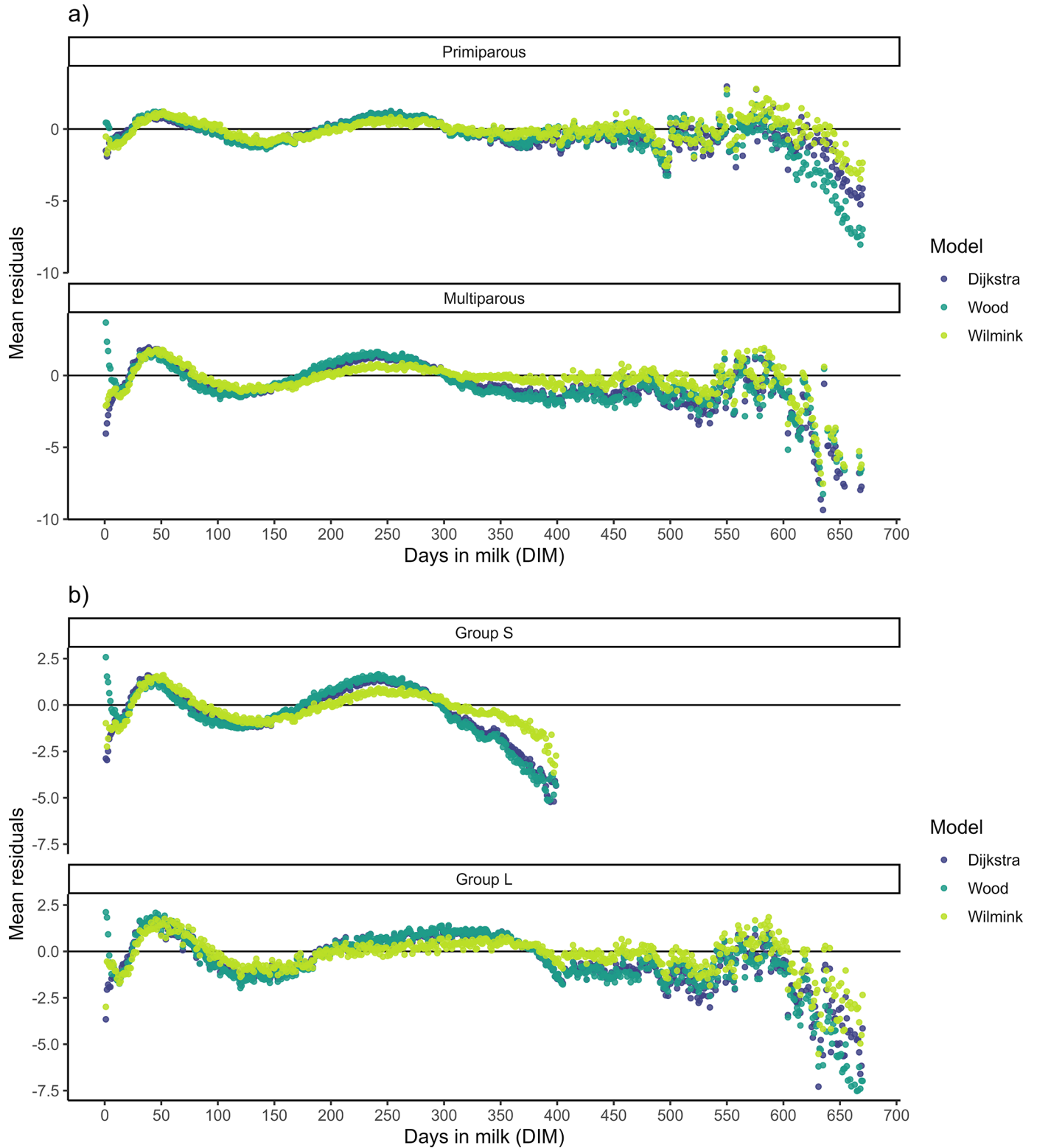


Figure 2. Mean daily residuals (observed milk yield [MY] minus predicted MY; kg/d) from individual lactation curves fitted by the Dijkstra, Wood, and Wilmink models and grouped by (a) parity: primiparous ($n = 146$) and multiparous lactations ($n = 321$); or (b) lactation length: group S (305 to 399 DIM; $n = 367$) and group L (≥ 400 DIM; $n = 100$).

Table 5. Mean statistics of fits of Dijkstra model to individual lactation curves ($n = 365$) with data from 1 to i DIM, where i is the DIM threshold with either unconstrained or constrained parameters¹

Item ²	DIM threshold							
	30	60	90	120	150	200	250	300
Unconstrained parameter								
Peak yield (kg/d)	$8.3 \times 10^{+51}$	41.2	40.5	40.2	40.0	39.8	39.7	39.6
Time to peak (d)	$3.02 \times 10^{+8}$	$4.58 \times 10^{+7}$	$4.43 \times 10^{+6}$	62.8	59.0	54.3	53.5	55.1
Persistency (kg/d per d)	$2.70 \times 10^{+47}$	-0.107	-0.094	-0.083	-0.076	-0.067	-0.063	-0.066
305-d mean predicted MY (kg/d)	444	20.6	17.2	19.4	21.0	23.0	23.7	23.2
305-d RMSPE (%)	7,350	258	65.7	42.9	34.3	24.9	19.6	13.9
305-d MAE (%)	1,990	103	52.0	33.2	25.8	19.1	14.9	10.2
305-d CCC	0.0003	0.002	0.12	0.25	0.34	0.56	0.71	0.86
Constrained parameter								
Peak yield (kg/d)	47.5	41.0	40.2	40.1	40.0	39.8	39.6	39.6
Time to peak (d)	89.2	61.1	55.4	54.8	54.5	54.2	54.3	55.4
Persistency (kg/d per day)	-0.059	-0.080	-0.080	-0.077	-0.073	-0.066	-0.064	-0.065
305-d mean predicted MY (kg/d)	33.7	20.9	20.1	20.7	21.5	23.1	23.6	23.3
305-d RMSPE (%)	114	51.1	38.3	32.3	28.7	23.6	19.6	14.4
305-d MAE (%)	84.8	40.8	31.2	25.5	22.3	18.3	14.8	10.6
305-d CCC	0.03	0.04	0.22	0.33	0.42	0.58	0.71	0.84

¹The b , b_0 , and c parameters were either unconstrained or constrained as follows: $1 \times 10^{-8} \leq b \leq 0.10924$; $0.01725 \leq b_0 \leq 0.08723$; and $0.00087 \leq c \leq 0.00449$.

²The 305-d concordance correlation coefficient (CCC), mean absolute error (MAE), and square root of mean square prediction error (RMSPE) compare the observed milk yield (MY) at d 305 and predicted MY at d 305 (both as the mean daily MY from 303 to 307 DIM), with the MAE and RMSPE expressed as a percentage of the mean observed MY at d 305 (22.0 kg/d).

ures 3a and 3b. When compared with the incremental fits using only i number of days of data (orange violin plots), a clear, progressive improvement was revealed in the predicted values toward a tighter dispersion as i increased for both the unconstrained (Figure 3a) and the constrained fits (Figure 3b). The predicted MY at 200-, 250-, and 300-DIM thresholds also had a tighter dispersion than the observed MY, which reflects the inability of the models to account for all of the variation in the observed MY even when sufficient data are available. When the parameters were not constrained, the 30-, 60-, and 90-DIM thresholds exhibited a high density of predicted MY at d 305 at or around 0 kg/d, as well as some very high predicted MY (Figure 3a). These outliers arose because, when the Dijkstra model was unable to fit the data well, it could decrease to 0 kg/d very quickly or increase rapidly without achieving a meaningful PY, as observed in the example lactations in Figure 4a. In total, 117 (32%), 138 (38%), 46 (13%), 8 (2%), and 2 (1%) predicted MY at d 305 were less than 1 kg/d from 30, 60, 90, 120, and 150 d of data, respectively, whereas the fits from 200, 250, and 300 d of data had none. In contrast, no predicted MY at d 305 were less than 1 kg/d when the parameters were constrained, and distributions of predictions were similar to those observed when using 90 d of data or more (Figure 3b).

To further assess the relationship between early-lactation information and MY at d 305, the constrained Dijkstra model was re-fitted using data incremented by 5 d from 30 to 305 DIM. The 305-d MAE (%) had

a steep improvement between 30 and 65 d, followed by a more gradual, approximately linear improvement until 305 d (Figure 5a), whereas 305-d CCC had a clear improvement between 60 and 65 d followed by an approximately linear improvement until 305 d (Figure 5b). Although MY at d 305 was the observation of interest used to calculate goodness-of-fit in the incremental analysis (Table 5), the mean fitted MY curves up to 305 d appeared to improve with increasing data availability for the unconstrained (Figure 6a) and constrained models (Figure 6b). For the unconstrained model, the predicted curve increased off the chart when only up to 30 d were used (Figure 6a), which was also noted in the individual lactation examples (Figure 4a), whereas the constrained model predictions were more realistic (Figure 4b and 5b).

Probability of Day-305 Predictions Exceeding Targets

In addition to the 305-d goodness-of-fit statistics (Table 5) and distributions (Figure 3), the probability that MY at d 305 is greater than or equal to a target MY is a useful way to consider the risk associated with selecting cows in early lactation for an extended lactation strategy. Target MY of 15, 20, and 25 kg/d were selected to represent a useful range of MY at d 305, and different levels of risk were represented by a probability of 0.50, 0.75, or 0.95 (Table 6). The proportions of lactations with observed MY at d 305 greater than or equal to 15, 20, or 25 kg/d were 88, 62,

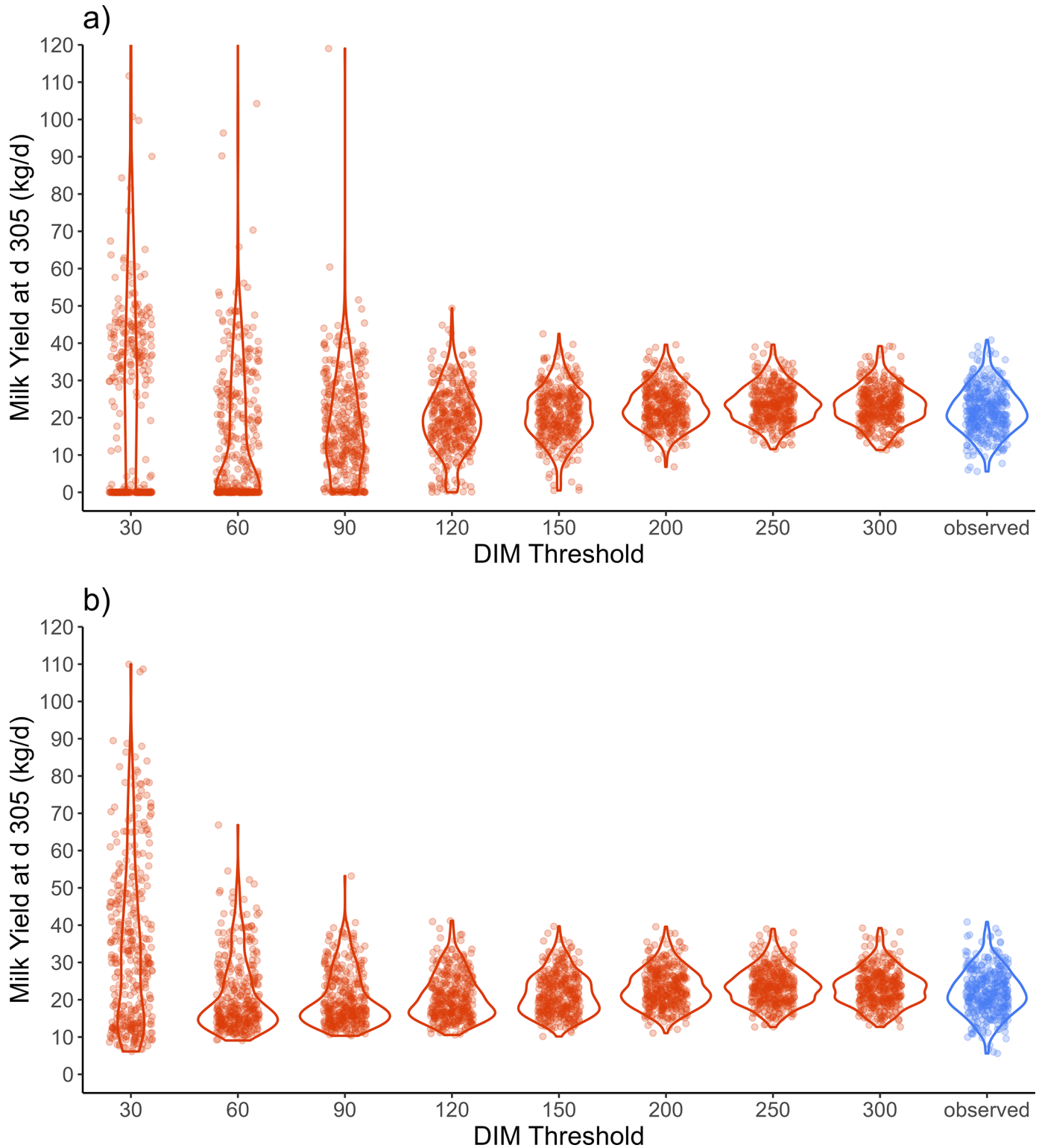


Figure 3. Distribution of predictions of milk yield at d 305 compared with the observed milk yield at d 305 (blue) after fitting the Dijkstra model to lactations with up to i DIM (where i is the DIM threshold; orange) and with either (a) unconstrained or (b) constrained parameters. Frequency of points at a given milk yield at d 305 (kg/d) is represented by the width of the violin plot line. The individual data points are also plotted but are partially transparent and randomly spread to a fixed width to help interpret density. Scale of y-axis limited to 120 kg/d, excluding 82 points in panel a, ranging from 121 to 15,751 kg/d, due to poor model fit. The b , b_0 , and c parameters of the Dijkstra model were either unconstrained or constrained as follows: $1 \times 10^{-8} \leq b \leq 0.10924$; $0.01725 \leq b_0 \leq 0.08723$; and $0.00087 \leq c \leq 0.00449$.

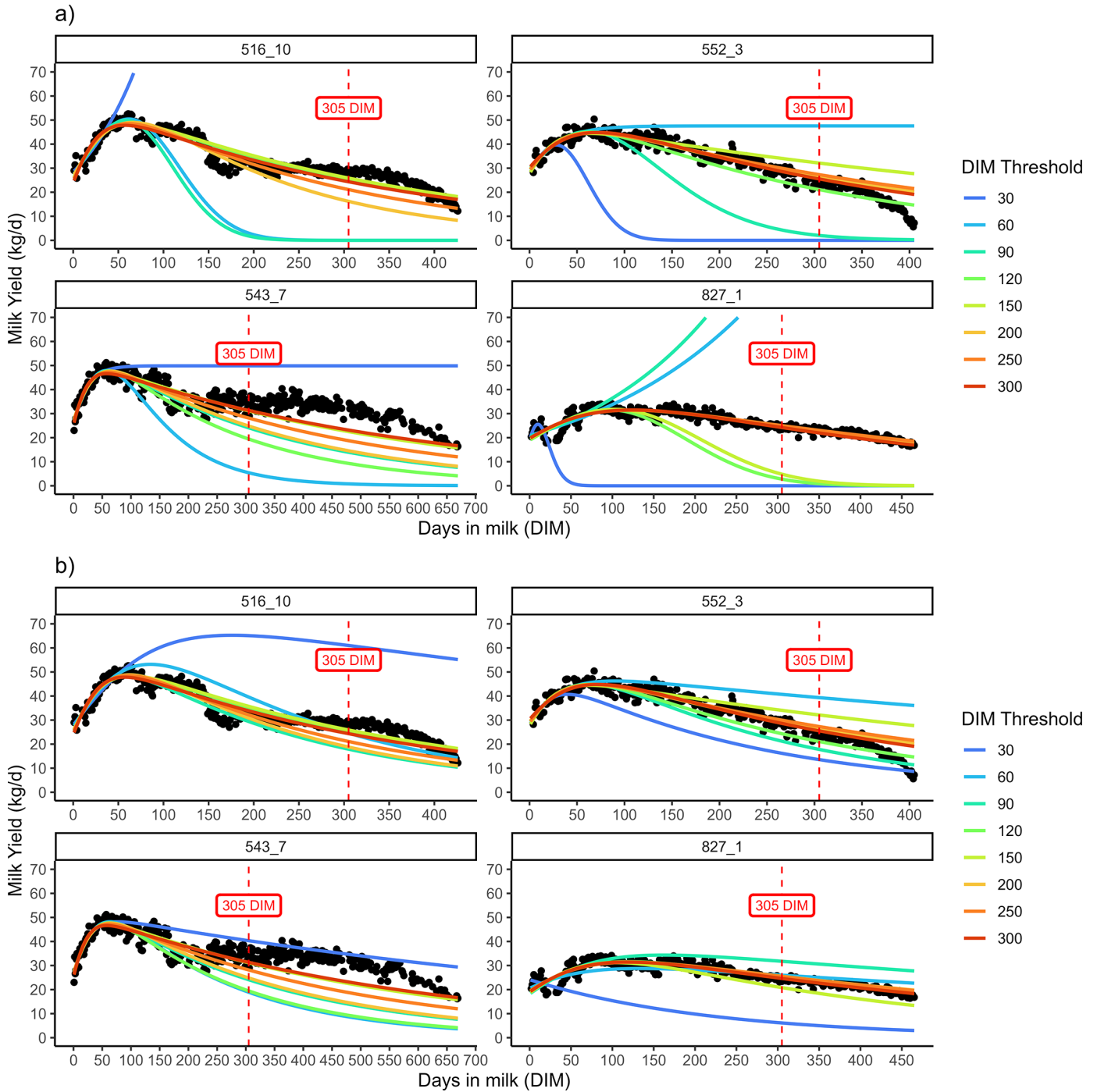


Figure 4. Four individual lactations with predicted milk yield lines from fitting the Dijkstra model with up to i DIM (where i is the DIM threshold) and with (a) unconstrained or (b) constrained parameters. Lactations were selected to represent different ways the Dijkstra model can respond when only 30 d of data are fitted. Subplot titles are in the format “ID_parity”; e.g., 516_10 is cow 516 in her 10th lactation. The b , b_0 , and c parameters of the Dijkstra model were either unconstrained or constrained as follows: $1 \times 10^{-8} \leq b \leq 0.10924$; $0.01725 \leq b_0 \leq 0.08723$; and $0.00087 \leq c \leq 0.00449$.

and 30%, respectively (Table 6). Within each target level, the proportion of lactations with a probability of predicted MY at d 305 greater than or equal to the target generally decreased as the probability threshold

increased from 0.50 to 0.95 but increased with data availability from 90 to 300 DIM (Table 6). However, these predicted proportions are not directly comparable to the observed proportions, as the individual lacta-

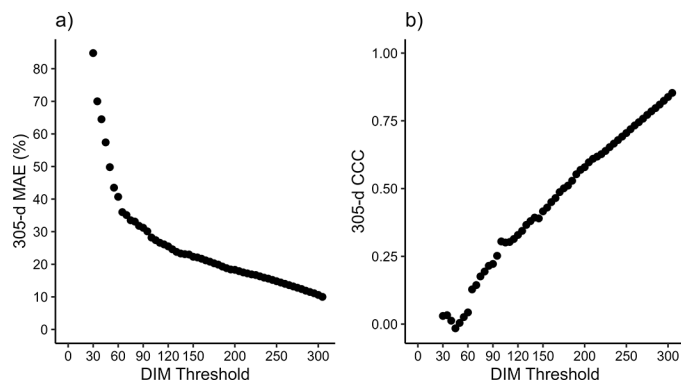


Figure 5. The (a) 305-d mean absolute error (MAE; %) and (b) 305-d concordance correlation coefficient (CCC) after fitting the Dijkstra model to individual lactations ($n = 365$) with up to i DIM (where i is the DIM threshold incremented by 5 d from 30 to 305 DIM). The MAE and CCC compare the observed milk yield (MY) at d 305 and predicted MY at d 305 (both as the mean daily MY from 303 to 307 DIM), with the MAE expressed as a percentage of the mean observed MY at d 305 (22.0 kg/d).

tions with a predicted MY at d 305 greater than or equal to target may have had an observed MY at d 305 less than target. Therefore, an accuracy was calculated as the proportion of all lactations that were correctly predicted; that is, when predicted and observed were both greater than or equal to target or both less than target. Overall, accuracy improved with increasing data availability from 90 to 300 DIM, and these were highest for the 15 kg/d target (Table 6). However, at a probability threshold of 0.95 we found higher accuracy in the 25 kg/d target compared with the 15 kg/d target at 90 and 120 DIM data availability. At a risk level of 0.95 with 90 d of data and a 15 kg/d target, we detected a 50.1% accuracy, indicating only half of the lactations were correctly characterized as above or below the 15 kg/d target, whereas this accuracy was 65.8% for the 25 kg/d target.

DISCUSSION

Comparison of Models

Three traditional lactation models (Wood, Wilmink, and Dijkstra) were fitted to extended lactation data from high-producing Holstein-Friesian cows in the Netherlands. On average, each model showed a lower mean residual error in the primiparous lactations than in the multiparous lactations, but a poorer quality of fit (CCC) in the primiparous lactations. Overall, the mean MAE ranged from 5.7 to 6.4% (1.56–1.75 kg/d) of the mean daily MY for primiparous lactations, and 6.2 to 7.2% (1.95–2.24 kg/d) for multiparous lactations. The mean residuals showed a cyclic pattern, indicating the models overestimated MY between 80 and 200 DIM

and underestimated MY between 200 and 300 DIM. A similar pattern was also observed by Dematawewa et al. (2007) when fitting various models to an extended lactation data set, including the Wood, Wilmink, and Dijkstra models. Some of the Wilmink and Dijkstra fits did not converge, but their data are still included in Tables 3 and 4 to avoid bias in evaluating model performance. For example, the mean c parameter of the Dijkstra model appears higher than previously reported (Dijkstra et al., 2010), which can be attributed to large outliers from fits that did not converge (Supplemental Figure S2).

Parity

Lactation yields in primiparous cows are normally lower than in multiparous cows. Lactation models pertaining to primiparous cows also exhibit a flatter, more persistent curve with lower PY across various breeds (Wood, 1969; Cole et al., 2009; Dijkstra et al., 2010). Likewise, primiparous cows had a total 305-d MY that was 15.7% lower than for multiparous cows (8,477 vs. 10,060 kg) in the current analysis. Primiparous cows also showed lower PY, longer PT, and higher persistency, which is consistent with the flatter shape of the mean daily observed MY curves (Supplemental Figure S3). In addition, the a parameter of primiparous lactations was significantly lower than for the multiparous lactations in all 3 models. Although the Dijkstra model is the only mechanistic model of the 3, the a parameter from each model is still relatable to the scale of initial MY, indicating that primiparous cows had a lower initial MY at the beginning of their lactations.

The Wilmink model had numerically lower RMSPE and MAE and higher CCC compared with the Dijkstra and Wood models in both the primiparous and multiparous groups (Table 3); however, these mean values represent a broad range of lactation curves. The lactations with the lowest mean MAE (Figure 1a) showed model predictions that were almost identical, but lactations with the highest mean MAE had greater day-to-day variability and differences between the predictions from the 3 models (Figure 1b). Therefore, this poses a question when highly variable lactation data are used to fit a lactation model, as is common in on-farm data: is the most useful model the one that is forced to look like a typical lactation curve or one that best minimizes the RSS? To use lactations 722_2 and 839_3 in Figure 1b as examples, the Wood model, with only 3 parameters, remains in a similar shape regardless of the variability in the data, whereas the Wilmink model, with 4 parameters, has the flexibility to better minimize RSS in later lactation at the expense of not showing a typical PY. The Dijkstra model may be a useful intermediate,

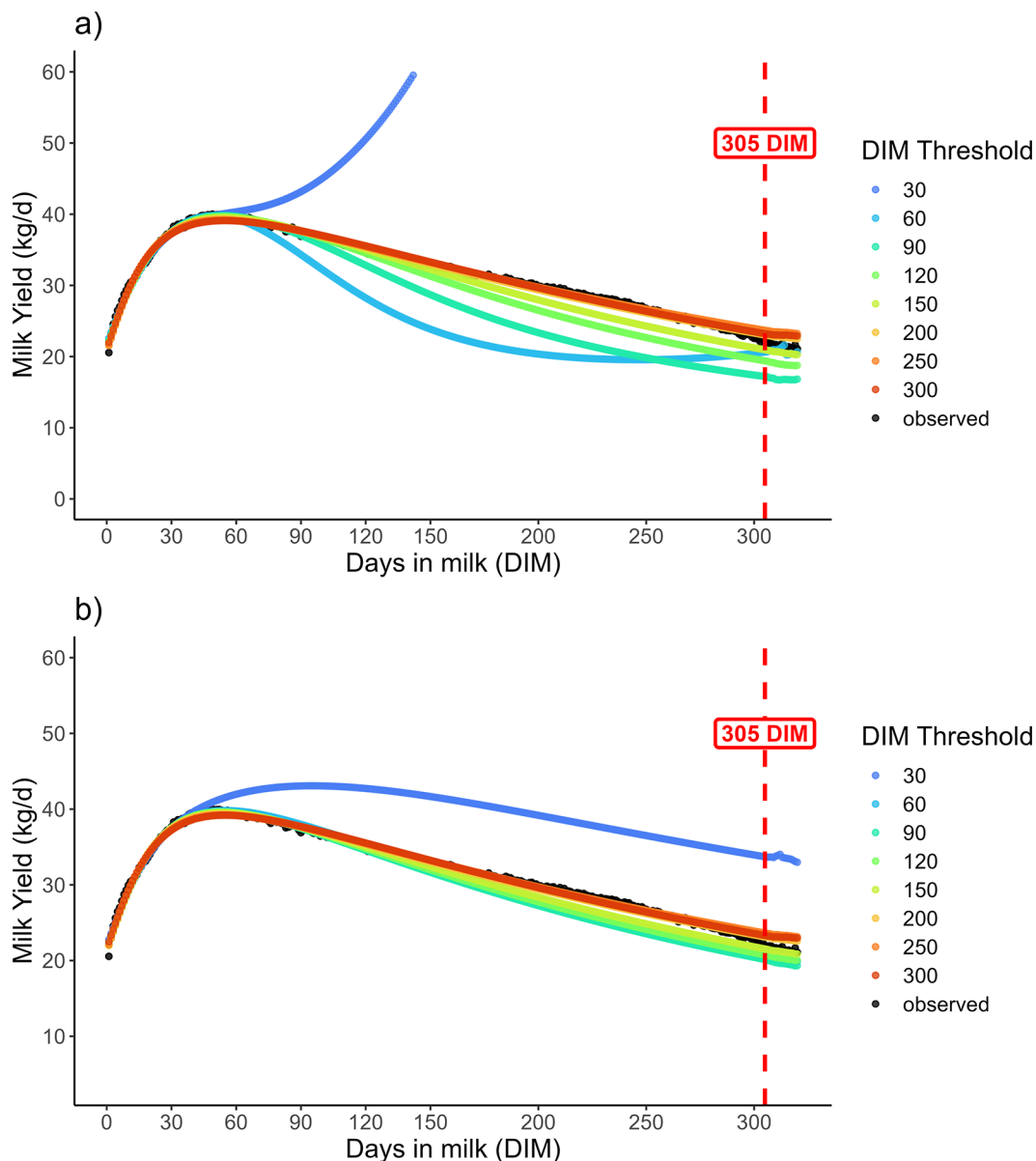


Figure 6. Mean predicted milk yield after fitting the Dijkstra model to lactations with up to i DIM (where i is the DIM threshold) and with either (a) unconstrained or (b) constrained parameters. The y-axis is constrained to show 0 to 60 kg/d, focusing on the lactation curve but hiding the very high predicted values for the 30-DIM predictions. The b , b_0 , and c parameters of the Dijkstra model were constrained as follows: $1 \times 10^{-8} \leq b \leq 0.10924$; $0.01725 \leq b_0 \leq 0.08723$; and $0.00087 \leq c \leq 0.00449$.

showing aspects of both models with the added benefit that it is a mechanistic model based on secretory cell dynamics.

An interesting problem also exists for lactation curves with multiple peaks, such as in lactation 839_3 in Figure 1b. The Dijkstra model appears to fit a PY between the 2 peaks, whereas the Wood model fits the PY on the first peak and the Wilmink model has the highest MY at the first day of lactation. In this case the Wilmink model may show a better prediction at

d 305, but it is unlikely that this would be accurately forecasted when using only early-lactation data. Multiple peaks in lactation curves are particularly relevant when modeling extended lactations in grazing systems where seasonal changes in diet quality can cause a second peak, as in New Zealand (Kolver et al., 2007) and Ireland (Butler et al., 2010). Lactation curves are also influenced by clinical and subclinical health events. It was recently found that only 1% of more than 50,000 lactations had 0 perturbations, defined as 5 d of nega-

Table 6. Proportion (%) of individual lactation curves ($n = 365$) with an observed or predicted milk yield at d 305 greater than or equal to a target of 15, 20, or 25 kg/d at different probabilities (prob) and with accuracy¹ of predictions

Item	DIM threshold ²			
	90	120	200	300
Target = 15 kg/d				
Observed	88.2	88.2	88.2	88.2
Predicted with prob >0.50	73.4	84.1	97.0	97.3
Accuracy	69.3	79.5	89.6	90.4
Predicted with prob >0.75	61.6	72.3	91.2	93.2
Accuracy	61.4	72.1	87.7	92.3
Predicted with prob >0.95	46.6	56.2	80.8	84.9
Accuracy	50.1	59.7	80.5	92.3
Target = 20 kg/d				
Observed	61.6	61.6	61.6	61.6
Predicted with prob >0.50	38.9	47.1	70.4	74.2
Accuracy	54.8	59.2	71.0	83.0
Predicted with prob >0.75	32.9	40.3	59.5	61.6
Accuracy	52.6	56.2	69.9	83.6
Predicted with prob >0.95	31.0	31.0	45.5	45.5
Accuracy	51.2	52.9	67.4	78.9
Target = 25 kg/d				
Observed	30.1	30.1	30.1	30.1
Predicted with prob >0.50	23.6	24.1	35.3	34.5
Accuracy	66.0	69.9	74.0	85.8
Predicted with prob >0.75	20.3	18.9	24.7	23.3
Accuracy	65.5	70.1	76.4	87.7
Predicted with prob >0.95	16.7	14.0	14.5	12.3
Accuracy	65.8	70.7	76.2	81.6

¹Accuracy is percentage of predictions that were correct.

²Data from 1 to i DIM, where i is the DIM threshold, fitted by the Dijkstra model with the b , b_0 , and c parameters constrained as follows: $1 \times 10^{-8} \leq b \leq 0.10924$; $0.01725 \leq b_0 \leq 0.08723$; and $0.00087 \leq c \leq 0.00449$.

tive residuals with at least 1 d where the daily MY is below 80% of the unperturbed fitted lactation model following Wood's model (Adriaens et al., 2021). The authors chose these thresholds to identify health events such as mastitis during a lactation, but not day-to-day errors due to equipment or estrus. In the context of forecasting extended lactations, these perturbations will almost always occur, and it is important to quantify them, but forecasting them would rely on stochastic rather than deterministic models. Therefore, this risk factor must be incorporated into the decision process when attempting to select cows best suited for extended lactation. Moreover, in addition to forecasted milk production, other variables related to physiological stage, including BCS and nonesterified fatty acid and insulin-like growth factor 1 blood levels, may help to select the most suitable cows (Sehested et al., 2019).

Lactation Length

The Wood model had a lower mean a parameter and higher mean b and c parameters in group S (<399 d) compared with group L (≥ 400 d; Table 4). The a and b

parameters can be related to the initial rate of increase and decay, respectively, of the lactation curve, and the c parameter can be related to the rate of decline of the lactation curve. Therefore, these differences in parameters for the Wood model are consistent with the shorter PT (higher b parameter) and lower persistency (higher c parameter) in group S (Table 4). The additional parameters in the Wilmink and Dijkstra models are likely responsible for the greater variability between individual fits and, therefore, a lack of statistical difference between mean parameters, despite both showing greater persistency in group L (Table 4). Overall, the persistency values are consistent with data from the Netherlands (Burgers et al., 2021a), where cows that were assigned to a 200-d VWP had better persistency than cows assigned to a 50-d VWP (-0.05 vs. -0.07 kg/d per day). However, the extended VWP of the cows in the current analyses were not due to intentional management decisions.

Forecasting Milk Yield at Day 305

Forecasting daily MY at d 305 is an important first step in developing a decision support tool for dairy producers. To be most useful, this forecasting should use MY data available during the VWP. However, in our attempt using unconstrained parameters, meaningful predictions of MY at d 305 using the Dijkstra equation were not possible until at least 120 d of data were fitted. Large variation in the predictions occurred when <120 d of data were available to the model (Figure 3a), despite the realistic mean predicted MY values at d 305 for the 60- and 90-DIM thresholds (Table 5). However, these values were improved when the parameters of the Dijkstra model were constrained, and the similarities in values (Table 5) and distributions (Figure 3b) from 90 to 300 d of data availability suggests that using 90 d of data to forecast MY at d 305 has potential to be useful. A machine learning model that used the parameters from various lactation models fitted to the first 90, 120, or 150 d of daily MY data of cows was able to correctly classify lactations as having a MY at d 305 <20 kg/d or >32 kg/d; however, no indication of the variation in the lactation model parameters was reported, and the reported errors of the predictions did not represent all possible errors that could occur (Manca et al., 2020). We also observed that when the Dijkstra model was fitted every 5 d, the 305-d MAE and CCC improved approximately linearly with increasing data availability from 65 d of data onward (Figure 5). This suggests that the largest improvements in the model could be achieved with 65 d of data; however, further refinements to the model would be required to improve the highly variable prediction of MY at d 305 for this threshold.

Despite the overall improvement in mean parameters (Supplemental Table S1), fit statistics (Table 5), and fitted values (Figure 6b) using a constrained Dijkstra model, it is difficult to estimate the likelihood that a prediction for MY at d 305 for an individual lactation is correct when only early-lactation data are available. Here we calculated RMSPE and MAE for the MY at d 305, but these calculations rely on knowing the real observed values, which are unknown when forecasting. Instead, if we were to calculate RMSPE or MAE using only the data available—for example, 90 d of data—then the residual error embedded in the first 90 d is not a realistic estimate of the expected residual error at d 305. Likewise, the SD within the first i DIM is not a meaningful estimate of the expected SD at d 305. Therefore, calculating a prediction interval for the estimated MY at d 305 without knowing the true SD of the predictions may not be an appropriate representation of the likelihood that the forecasted values are correct. Instead, we have attempted to estimate the probability that predicted values were above a target value (kg/d) using a Student's t -distribution calculated with an adjusted RMSPE of the available data as the estimate of the SD (Table 6). Although this still has the same problems with the unknown forecasted SD, it allows an accuracy to be estimated for different levels of risk (probability threshold of 0.50, 0.75, or 0.95) at different target MY (15, 20, or 25 kg/d). For example, if a producer wanted to select cows that would have a high probability (>0.95) of producing over 25 kg of milk on d 305, the current model would correctly classify about 66% of all lactations if only 90 d of data were fitted, but this increases to 76% if 200 d of data were used. In other words, 30% of lactations had an observed MY at d 305 ≥ 25 kg, but 34% of lactations were also incorrectly classified as being under or over this threshold when predicted using 90 d of data, suggesting that level of accuracy in the current model is unlikely to be sufficient for on-farm management decisions.

Challenges with Forecasting Beyond Day 305

Overall, using early-lactation data to predict late-lactation MY is not yet reliable enough for practical use on-farm due to the high uncertainty of forecasted predictions. However, our successful attempt to assist the model fitting procedure with more biologically relevant constraints on parameters in the Dijkstra model (Table 5) suggests that statistical approaches could further refine parameter bounds for individual herds that could improve the reliability of the models. One of the key problems for the model when using only 30 or 60 DIM of data was that the PY and start of the descending phase of the lactation were not always included within

that period. The PY in the Dijkstra model is calculated from PT which is a function of the b , b_0 , and c parameters. Therefore, it might be possible to improve model predictions if reasonable estimates of these parameters could be derived from previous data on either the individual or the herd, although the repeatability of these parameters between consecutive lactations within a cow is not known. This approach is also challenging for primiparous cows with no prior lactation data, but a recent machine learning approach to estimating MY from the MY of dams and paternal siblings of primiparous cows in a herd appears promising (Zhang et al., 2022). Machine learning models have also been used to forecast MY at a herd level (Murphy et al., 2014) and individual cow level (Liseune et al., 2021), but these do not allow for biologically relevant interpretation of model parameters and have goodness-of-fit values similar to those of the current analysis. In addition, the parameters of the Wood model for conventional lactations have been associated with temporal, geographic, and management factors using a mixed effects model approach (Li et al., 2022). Those authors also described a difference in second-lactation cows compared with third-lactation cows, which means actual parity number may make an important contribution to parameter estimation. Assessing the contribution of these other factors to the variance of forecasted parameters should also be included in future developments. These other factors could also be incorporated into a nonlinear mixed effects model of the herd's data while also incorporating the individual cow variation. For example, including the random effect of cow improved fits by 3 (Wood, Milkbot, and diphasic) nonlinear lactation models (Piccardi et al., 2017). However, these statistical approaches of predicting future MY still require training with sufficient historical data. Overall, we envisage that the hybridization of machine learning and mechanistic models, where a machine learning prediction framework may be augmented by parameters generated by a mechanistic model, or where biological bounds may be placed on parameters within a mechanistic model prediction framework with the mechanistic model for individual cows being parameterized using data driven methodology (Ellis et al., 2020), has the most potential for forecasting late-lactation MY.

An additional concern is that some of the lactations in the current analysis had a very fast decline in MY toward dry off, most noticeably in the worst-fitting models (Figure 1b). This phenomenon was also observed in test-day records of extended lactations in Polish cows, where Otwinowska-Mindur et al. (2021) fitted a Wilmink function to data from 0 to 305 DIM and a linear or squared function to the data >305 DIM to account for the faster decline in MY. Although this technique

may improve overall goodness-of-fit of historical data, it serves little use in forecasting the likelihood of a faster decline in MY at the end of a lactation. Future work will need to incorporate a correction for cows that may dry off more abruptly than others. Ideally, a biological mechanism explaining this phenomenon should be incorporated into an updated mechanistic model unique to extended lactations. Finally, modeling milk composition during extended lactations should also be incorporated into a decision support system, as the changes in composition of milk fat, protein, and lactose may have economic implications in the context of the whole-herd production.

CONCLUSIONS

The Wood, Wilmlink, and Dijkstra models were fitted to individual extended lactations with similar mean goodness-of-fit statistics, but the Dijkstra model was preferred because it is mechanistic and flexible enough to fit various lactation curve shapes that are typical on-farm. The differences in lactation characteristics and model parameters between parity groups were consistent with previous literature. When forecasting using early-lactation data, the Dijkstra model with parameters constrained within biologically relevant bounds improved the goodness-of-fit and predictions of MY at d 305 compared with an unconstrained model. Using 90 d of data, the model correctly classified 66% of lactations as being above or below a target MY at d 305 of 25 kg, with a probability threshold of 0.95. The proportion of correct classifications became smaller at lower target MY at d-305 levels and became greater when using more lactation days. Overall, further work is required to develop a model that can forecast late-lactation MY with sufficient accuracy for practical use. We envisage that a hybridized machine learning and mechanistic model that incorporates additional historical and genetic information with early-lactation MY could produce meaningful lactation curve forecasts.

ACKNOWLEDGMENTS








The authors thank J. B. Daniel (Trouw Nutrition R&D, Amersfoort, the Netherlands) for his contributions to preparing the initial dataset. This work was supported by the Canada First Research Excellence Fund (CFREF-2015-00004; Ottawa, Canada) under grant number THE2-020 and by the Canada Research Chairs program, grant number 045867 (Natural Sciences and Engineering Research Council of Canada, Ottawa). The authors have not stated any conflicts of interest.

REFERENCES

- Adriaens, I., I. Van Den Brulle, L. D'Anvers, J. M. E. Statham, K. Geerinckx, S. De Vliegher, S. Piepers, and B. Aernouts. 2021. Milk losses and dynamics during perturbations in dairy cows differ with parity and lactation stage. *J. Dairy Sci.* 104:405–418. <https://doi.org/10.3168/jds.2020-19195>.
- Burgers, E. E. A., A. Kok, R. M. A. Goselink, H. Hogeveen, B. Kemp, and A. T. M. Van Knegsel. 2021a. Effects of extended voluntary waiting period from calving until first insemination on body condition, milk yield, and lactation persistency. *J. Dairy Sci.* 104:8009–8022. <https://doi.org/10.3168/jds.2020-19914>.
- Burgers, E. E. A., A. Kok, R. M. A. Goselink, H. Hogeveen, B. Kemp, and A. T. M. Van Knegsel. 2021b. Fertility and milk production on commercial dairy farms with customized lactation lengths. *J. Dairy Sci.* 104:443–458. <https://doi.org/10.3168/jds.2019-17947>.
- 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. *J. Dairy Sci.* 93:1283–1295. <https://doi.org/10.3168/jds.2009-2407>.
- Capper, J. L., and R. A. Cady. 2020. The effects of improved performance in the U.S. dairy cattle industry on environmental impacts between 2007 and 2017. *J. Anim. Sci.* 98:skz291. <https://doi.org/10.1093/jas/skz291>.
- Cole, J. B., D. J. Null, and P. M. Vanraden. 2009. Best prediction of yields for long lactations. *J. Dairy Sci.* 92:1796–1810. <https://doi.org/10.3168/jds.2007-0976>.
- Dematawewa, C. M. B., R. E. Pearson, and P. M. Vanraden. 2007. Modeling extended lactations of Holsteins. *J. Dairy Sci.* 90:3924–3936. <https://doi.org/10.3168/jds.2006-790>.
- Dijkstra, J., J. France, M. S. Dhanoa, J. A. Maas, M. D. Hanigan, A. J. Rook, and D. E. Beaver. 1997. A model to describe growth patterns of the mammary gland during pregnancy and lactation. *J. Dairy Sci.* 80:2340–2354. [https://doi.org/10.3168/jds.S0022-0302\(97\)76185-X](https://doi.org/10.3168/jds.S0022-0302(97)76185-X).
- Dijkstra, J., S. Lopez, A. Bannink, M. S. Dhanoa, E. Kebreab, N. E. Odongo, M. H. Fathi Nasri, U. K. Behera, D. Hernandez-Ferrer, and J. France. 2010. Evaluation of a mechanistic lactation model using cow, goat and sheep data. *J. Agric. Sci.* 148:249–262. <https://doi.org/10.1017/S0021859609990578>.
- Durbin, J., and G. S. Watson. 1950. Testing for serial correlation in least squares regression. *Biometrika* 37:409–428. <https://doi.org/10.2307/2332391>.
- Ellis, J. L., M. Jacobs, J. Dijkstra, H. Van Laar, J. P. Cant, D. Tulpán, and N. Ferguson. 2020. Review: Synergy between mechanistic modelling and data-driven models for modern animal production systems in the era of big data. *Animal* 14:s223–s237. <https://doi.org/10.1017/S1751731120000312>.
- Fox, J., and S. Weisberg. 2019. *An {R} Companion to Applied Regression*. 3rd ed. Sage.
- Grossman, M., and W. J. Koops. 2003. Modeling extended lactation curves of dairy cattle: A biological basis for the multiphasic approach. *J. Dairy Sci.* 86:988–998. [https://doi.org/10.3168/jds.S0022-0302\(03\)73682-0](https://doi.org/10.3168/jds.S0022-0302(03)73682-0).
- 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. *J. Dairy Sci.* 94:3811–3823. <https://doi.org/10.3168/jds.2010-3790>.
- Innes, D., L. Pot, D. Seymour, J. France, J. Dijkstra, J. Doelman, and J. Cant. 2023. Supplementary material for: Fitting mathematical functions to extended lactation curves and forecasting late-lactation milk yields of dairy cows. *Borealis*. <https://doi.org/10.5683/SP3/4QVIO0>.
- Knight, C. H. 2001. Lactation and gestation in dairy cows: Flexibility avoids nutritional extremes. *Proc. Nutr. Soc.* 60:527–537. <https://doi.org/10.1079/PNS2001115>.
- 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. *J. Dairy Sci.* 90:5518–5530. <https://doi.org/10.3168/jds.2007-0324>.

- Lehmann, J. O., L. Mogensen, and T. Kristensen. 2017. Early lactation production, health, and welfare characteristics of cows selected for extended lactation. *J. Dairy Sci.* 100:1487–1501. <https://doi.org/10.3168/jds.2016-11162>.
- Li, M., G. J. M. Rosa, K. F. Reed, and V. E. Cabrera. 2022. Investigating the effect of temporal, geographic, and management factors on US Holstein lactation curve parameters. *J. Dairy Sci.* 105:7525–7538. <https://doi.org/10.3168/jds.2022-21882>.
- Lin, L. I. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45:255–268. <https://doi.org/10.2307/2532051>.
- Liseune, A., M. Salamone, D. Van den Poel, B. Van Ranst, and M. Hostens. 2020. Leveraging latent representations for milk yield prediction and interpolation using deep learning. *Comput. Electron. Agric.* 175:105600. <https://doi.org/10.1016/j.compag.2020.105600>.
- Liseune, A., M. Salamone, D. Van den Poel, B. van Ranst, and M. Hostens. 2021. Predicting the milk yield curve of dairy cows in the subsequent lactation period using deep learning. *Comput. Electron. Agric.* 180:105904. <https://doi.org/10.1016/j.compag.2020.105904>.
- Manca, E., A. Cesarani, N. P. P. Macciotta, A. S. Atzori, G. Pulina, and C. Dimauro. 2020. Use of discriminant statistical procedures for an early detection of persistent lactations in dairy cows. *Comput. Electron. Agric.* 176:105657. <https://doi.org/10.1016/j.compag.2020.105657>.
- Murphy, M. D., M. J. O'Mahony, L. Shalloo, P. French, and J. Upton. 2014. Comparison of modelling techniques for milk-production forecasting. *J. Dairy Sci.* 97:3352–3363. <https://doi.org/10.3168/jds.2013-7451>.
- Niozas, G., G. Tsousis, C. Malesios, I. Steinhöfel, C. Boscós, H. Bollwein, and M. Kaske. 2019a. Extended lactation in high-yielding dairy cows. II. Effects on milk production, udder health, and body measurements. *J. Dairy Sci.* 102:811–823. <https://doi.org/10.3168/jds.2018-15117>.
- Niozas, G., G. Tsousis, I. Steinhöfel, C. Brozos, A. Römer, S. Wiedemann, H. Bollwein, and M. Kaske. 2019b. Extended lactation in high-yielding dairy cows. I. Effects on reproductive measurements. *J. Dairy Sci.* 102:799–810. <https://doi.org/10.3168/jds.2018-15115>.
- Otwinowska-Mindur, A., E. Ptak, J. Makulska, and O. Jarnecka. 2021. Modelling extended lactations in Polish Holstein–Friesian cows. *Animals (Basel)* 11:2176. <https://doi.org/10.3390/ani11082176>.
- Padfield, D., and G. Matheson. 2020. nls.mulstart: Robust non-linear regression using AIC scores. R Package version 1.2.0. <https://CRAN.R-project.org/package=nls.mulstart>.
- Piccardi, M., R. Macchiavelli, A. C. Funes, G. A. Bó, and M. Balzarini. 2017. Fitting milk production curves through nonlinear mixed models. *J. Dairy Res.* 84:146–153. <https://doi.org/10.1017/S0022029917000085>.
- Santos, J. E. P., H. M. Rutigliano, and M. F. S. Filho. 2009. Risk factors for resumption of postpartum estrous cycles and embryonic survival in lactating dairy cows. *Anim. Reprod. Sci.* 110:207–221. <https://doi.org/10.1016/j.anireprosci.2008.01.014>.
- Sehested, J., C. Gaillard, J. O. Lehmann, G. M. Maciel, M. Vestergaard, M. R. Weisbjerg, L. Mogensen, L. B. Larsen, N. A. Poulsen, and T. Kristensen. 2019. Review: Extended lactation in dairy cattle. *Animal* 13(Suppl. 1):s65–s74. <https://doi.org/10.1017/S1751731119000806>.
- Strandberg, E., and P. A. Oltenacu. 1989. Economic consequences of different calving intervals. *Acta Agriculturae Scandinavica* 39:407–420. <https://doi.org/10.1080/00015128909438534>.
- Tsuruta, S., I. Misztal, and T. J. Lawlor. 2005. Changing definition of productive life in US Holsteins: Effect on genetic correlations. *J. Dairy Sci.* 88:1156–1165. [https://doi.org/10.3168/jds.S0022-0302\(05\)72782-X](https://doi.org/10.3168/jds.S0022-0302(05)72782-X).
- van Kneegsel, A. T. M., E. E. A. Burgers, J. Ma, R. M. A. Goselink, and A. Kok. 2022. Extending lactation length: Consequences for cow, calf, and farmer. *J. Anim. Sci.* 100:skac220. <https://doi.org/10.1093/jas/skac220>.
- VanRadon, P. M., C. M. B. Dematawewa, R. E. Pearson, and M. E. Tooker. 2006. Productive life including all lactations and longer lactations with diminishing credits. *J. Dairy Sci.* 89:3213–3220. [https://doi.org/10.3168/jds.S0022-0302\(06\)72596-6](https://doi.org/10.3168/jds.S0022-0302(06)72596-6).
- Walsh, S. W., E. J. Williams, and A. C. O. Evans. 2011. A review of the causes of poor fertility in high milk producing dairy cows. *Anim. Reprod. Sci.* 123:127–138. <https://doi.org/10.1016/j.anireprosci.2010.12.001>.
- Wilmink, J. B. M. 1987. Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. *Livest. Prod. Sci.* 16:335–348. [https://doi.org/10.1016/0301-6226\(87\)90003-0](https://doi.org/10.1016/0301-6226(87)90003-0).
- Wood, P. D. P. 1967. Algebraic model of the lactation curve in cattle. *Nature* 216:164–165. <https://doi.org/10.1038/216164a0>.
- Wood, P. D. P. 1969. Factors affecting the shape of the lactation curve in cattle. *Anim. Sci.* 11:307–316. <https://doi.org/10.1017/S0003356100026945>.
- Zeileis, A., and G. Grothendieck. 2005. zoo: S3 infrastructure for regular and irregular time series. *J. Stat. Softw.* 14:1–27. <https://doi.org/10.18637/jss.v014.i06>.
- Zhang, F., K. A. Weigel, and V. E. Cabrera. 2022. Predicting daily milk yield for primiparous cows using data of within-herd relatives to capture genotype-by-environment interactions. *J. Dairy Sci.* 105:6739–6748. <https://doi.org/10.3168/jds.2021-21559>.

ORCID

- David J. Innes  <https://orcid.org/0000-0002-6246-0489>
 Linaya J. Pot  <https://orcid.org/0009-0008-6612-941X>
 Dave J. Seymour  <https://orcid.org/0000-0001-6208-8720>
 James France  <https://orcid.org/0000-0003-2923-7098>
 Jan Dijkstra  <https://orcid.org/0000-0003-3728-6885>
 John Doelman  <https://orcid.org/0000-0002-3328-8042>
 John P. Cant  <https://orcid.org/0000-0003-2021-1764>