J. Dairy Sci. 107:6888–6901
https://doi.org/10.3168/jds.2023-24224

Comparison of 3 mathematical models to estimate lactation performance in dairy cows

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

[1]Department of Animal Medicine, Production and Health (MAPS), University of Padova, 35020 Legnaro (PD), Italy
[2]Division of Animal and Human Health Engineering, Department of Biosystems, KU Leuven, Campus Geel, 2440 Geel, Belgium
[3]BioVism, Department of Data Analysis and Mathematical Modelling, Ghent University, 9000 Ghent, Belgium
[4]Animal Breeding and Genomics, Wageningen University and Research, 6700 AH Wageningen, the Netherlands
[5]Lactanet, Sainte-Anne-de-Bellevue, QC, H9X 3R4 Canada

ABSTRACT

Milk yield dynamics and production performance reflect how dairy cows cope with their environment. To optimize farm management, time series of individual cow milk yield have been studied in the context of precision livestock farming, and many mathematical models have been proposed to translate raw data into useful information for the stakeholders of the dairy chain. To gain better insights on the topic, this study aimed at comparing 3 recent methods that allow one to estimate individual cow potential lactation performance, using daily data recorded by the automatic milking systems of 14 dairy farms (7 Holstein, 7 Italian Simmental) from Belgium, the Netherlands, and Italy. An iterative Wood model (IW), a perturbed lactation model (PLM), and a quantile regression (QR) were compared in terms of estimated total unperturbed (i.e., expected) milk production and estimated total milk loss (relative to unperturbed yield). The IW and PLM can also be used to identify perturbations of the lactation curve and were thus compared in this regard. The outcome of this study may help a given end-user in choosing the most appropriate method according to their specific requirements. If there is a specific interest in the post–peak lactation phase, IW can be the best option. If one wants to accurately describe the perturbations of the lactation curve, PLM can be the most suitable method. If there is need for a fast and easy approach on a very large dataset, QR can be the choice. Finally, as an example of application, PLM was used to analyze the effect of cow parity, calving season, and breed on their estimated lactation performance.

Key words: lactation curve, milk loss, perturbation, precision livestock farming

Received September 22, 2023.
Accepted March 19, 2024.
*Corresponding author: giovanna.ranzato@phd.unipd.it

INTRODUCTION

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

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

Perturbations cause drops in MY that pull the fitted lactation curve downward. Many phenotyping tools have been proposed to estimate dairy cow expected milk production in the absence of perturbations. Thanks to high-frequency milk meter data and advanced computation, the most recent tools enable the study of lactation dynamics. Adriaens et al. (2021) proposed to iteratively

The list of standard abbreviations for JDS is available at adsa.org/jds-abbreviations-24. Nonstandard abbreviations are available in the Notes.

use the model by Wood (1967) to determine the expected production and characterize perturbations of the lactation curve, Ben Abdelkrim et al. (2021) implemented a lactation model with explicit representation of perturbations, and Poppe et al. (2020) suggested to use a fourth-order quantile regression (**QR**) model to make the resulting lactation curve fit less sensitive to drops in MY.

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

MATERIALS AND METHODS

Data Collection and Preprocessing

Back-up files of the management software of 14 automatic milking systems (**AMS**) were collected at 10 farms with Lely AMS (Lely Industries N.V., Maassluis, the Netherlands) and 4 farms with DeLaval AMS (DeLaval, Tumba, Sweden). Half of the farms housed Italian Simmental dairy cows in the north of Italy, whereas the remaining were Holstein dairy farms from Belgium and the south of the Netherlands. The data tables containing daily historical MY data, cow and lactation identifiers, and calving dates were extracted from the AMS software back-up files of each farm using an automated pipeline in Python (Gote et al., 2022). The MY data in DeLaval software back-ups differed from those of Lely: DeLaval sums the milk production from the milkings each 24-h period as the daily value, whereas Lely takes a 3-d moving average of the MY to adjust for variations in the number of milkings per day. Daily MY data from the DeLaval farms were therefore adjusted to match the Lely reporting style by replacing each daily MY value with the average of the current and previous days (Adriaens et al., 2021).

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

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

Iterative Wood Model

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

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

$$\mathrm{MY} = a \cdot \mathrm{DIM}^{b} \cdot \mathrm{e}^{-c \cdot \mathrm{DIM}} + \varepsilon, \quad [1]$$

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

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

(2) remove MY data below $\mathrm{ULC}_1 - 1.6 \times$ standard deviation (**SD$_1$**) of the residuals of the model estimated at $i = 1$;

(3) iteration $i > 1$: fit Wood's model (ULC_1) on the filtered MY data resulting from the previous iteration $i - 1$;

(4) remove MY data below $\mathrm{ULC}_i - 1.6 \times \mathrm{SD}_i$ of the residuals of the model estimated at i;

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

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

Starting from the ULC, it is possible to identify the perturbations in the actual milk production. After computing the residuals by subtracting ULC from the original MY data, a perturbation is defined as a period of at least 5 successive days of negative residuals with at least 1 d in which MY is below 80% of ULC. The start and end DIM of each perturbation correspond to the first and last residual below zero, respectively. The days before the largest negative residual are identified as the development phase of the perturbation, and the days afterward represent the recovery phase.

Perturbed Lactation Model

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

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

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

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

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

where DIM_p is the time of start of the pth perturbation. The total number of fixed parameters of the model is then equal to $4 + 4 \times P$ (i.e., a, b, c, P, $[k_{0p}, k_{1p}, k_{2p}, DIMp] \times P$). The fitting strategy ('nlsLM' function of the 'minpack.lm' package, Elzhov et al., 2023; 'nls_multstart' function of the 'nls.multsart' package, Padfield and Matheson, 2023) consists of 2 steps: (1) perform repeated fittings to estimate the most frequent number of detected perturbations, and (2) fix the number of perturbations to the value determined in the first step and estimate the remaining parameters of the model. Arbitrary values can be set in step (1) for the number of fittings to be performed and the maximum number of perturbations to be detected, as well as for the maximum number of iterations to be used in step (2) to reach the convergence of the estimation procedure. We performed 50 fittings in step (1) and set the maximum number of perturbations and the maximum number of iterations at 10 and 1,000, respectively. The PLM provides an explicit representation of the ULC and the perturbed lactation curve where perturbations can occur one inside another.

Quantile Regression

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

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

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

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

Comparison of Models

All the presented models depend on meta-parameters that can significantly change the performance based on their values: IW relies on the threshold for removing low MY values at each iteration of the algorithm, PLM depends on the initial number of fittings, the maximum number of perturbations to be detected, and the maximum number of iterations of the estimation procedure, and QR relies on the reference quantile of MY data points. Beyond the aforementioned values, on which all the results of this article are based, IW was also run using a threshold value of 2, PLM using 20 repeated fittings at the first step of the algorithm, 8 maximum perturbations to be detected, and 500 iterations to reach the convergence, while QR using a 0.6 reference quantile. As also analyzed by Adriaens et al. (2021), Ben Abdelkrim et al. (2021), and Poppe et al. (2020), a sensitivity assessment serves to show how the performance of the models change when the effect assigned to low MY data points is reduced.

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

$$ML = 100 \cdot \left(1 - \frac{\sum_{DIM=min}^{max} MY_{DIM}}{\sum_{DIM=min}^{max} ULC_{DIM}} \right),$$

where $\sum_{DIM=min}^{max} ULC_{DIM}$ is the total estimated unperturbed production for that lactation over DIM, and $\sum_{DIM=min}^{max} MY_{DIM}$

is the respective observed total MY (where minimum [min] was below 5 DIM and maximum [max] between 200 and 305 DIM). Furthermore, every lactation was divided into early (DIM ≤ 60), mid (60 < DIM ≤ 150), and late (150 < DIM ≤ 305), and the sum of the estimated unperturbed daily production and percentage of ML were computed at each lactation stage (**ULC**$_{\mathbf{early}}$ and **ML**$_{\mathbf{early}}$, **ULC**$_{\mathbf{mid}}$ and **ML**$_{\mathbf{mid}}$, and **ULC**$_{\mathbf{late}}$ and **ML**$_{\mathbf{late}}$). Finally, IW and PLM were compared in terms of number of detected perturbations per lactation (**N**$_{\mathbf{P}}$), number of detected perturbations per lactation stage (**N**$_{\mathbf{P,early}}$, **N**$_{\mathbf{P,mid}}$, and **N**$_{\mathbf{P,late}}$), and average length of perturbations (**L**$_{\mathbf{P}}$, in days). The PLM estimates an ULC as well as a perturbed lactation curve with deviations from the ULC that can overlap in time; the IW, in contrast, comes with a criterion to identify individual perturbations. When the relative ML of an overlapping PLM perturbation was lower than 5%, that perturbation was considered as the same perturbation as the one it overlaps with and was not taken into account to determine the final number of perturbations.

To assess the overall model accuracy, the root mean squared error (**RMSE**) of the total unperturbed milk production was computed over all the lactations (n) of each farm

$$\left[\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{g=1}^{n}\left(\sum_{\mathrm{DIM=min}}^{\mathrm{max}}\mathrm{MY}_{\mathrm{DIM},g} - \sum_{\mathrm{DIM=min}}^{\mathrm{max}}\mathrm{ULC}_{\mathrm{DIM},g}\right)^2}\right],$$

representing the deviation of the observed total MY from the expected production, where g = 1, ..., n lactations. Then, the outcomes of interest from each lactation of each farm (i.e., ULC_{305}, ML, $\mathrm{N_P}$, and $\mathrm{L_P}$) were merged into one global table, keeping the information on cow breed, parity, and calving dates. To obtain a one-to-one comparison of models, Pearson correlation coefficients, together with the coefficient of determination (R^2) of the respective estimated linear regressions, were computed for ULC_{305} and ML. Pearson correlation coefficients revealed the degree of linear relationship between pairs of models, while the R^2 of the estimated regression lines shed light on systematic differences between them. Finally, a Kruskal-Wallis statistical test (Dodge, 2008) was applied to compare the outcomes estimated with the 3 models, and a post hoc Wilcoxon test with Holm correction (Holm, 1979) was performed if the median results differed among methods.

Application Example

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

$$y_{ghjlmq} = \mu + parity_h + calving\ \ season_j + breed_l + parity_h \cdot calving\ \ season_j + parity_h \cdot breed_l + calving\ \ season_j \cdot breed_l + cow_m + farm_q + \varepsilon_{ghjlmq}, \quad [2]$$

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

RESULTS

The final data consisted of 2,250 cows (65% Holstein; 35% Italian Simmental) and 4,441 individual lactations with parity ranging from 1 to 12 (33% parity 1; 26% parity 2; 41% parity ≥3). Descriptive statistics over farms are given in Table 1. Table 2 reports the average results of the outcomes (i.e., ULC_{305}, ML, $\mathrm{N_P}$, and $\mathrm{L_P}$) estimated using 2 different sets of meta-parameters for each of the 3 models; all the following results are based on set 1.

Comparison of Models

Unperturbed Lactation Curve. First, the 3 mathematical models were used to estimate the unperturbed MY for each lactation of the dataset; Table 3 reports the average RMSE of the total unperturbed yield over farms. An example of ULC obtained with each model for one randomly selected lactation is shown in Figure 1 (although this

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

Item	Mean ± SD over farms	Range over farms (minimum; maximum)
Time period covered (yr)	6.2 ± 1.0	4.3; 7.2
Lactations (no.)	317.2 ± 177.4	61; 674
Parity 1	104.9 ± 68.4	0; 254
Parity 2	82.4 ± 47.6	25; 193
Parity ≥3	129.9 ± 66.1	31; 247
Average daily MY[1] in first 305 d (kg)	30.6 ± 4.6	24.1; 38.4
Total sum of MY in first 305 d (kg)	9,119.3 ± 2,253.9	1,447.1; 16,704.8

[1]MY = milk yield.

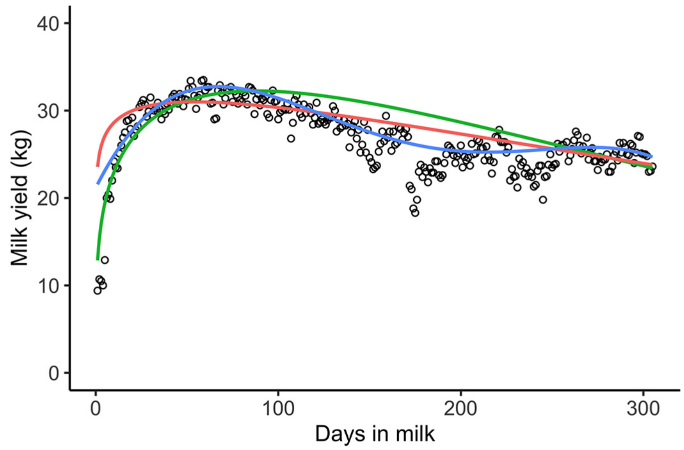

Figure 1. Unperturbed lactation curves estimated using iterative Wood model (pink), perturbed lactation model (green), and quantile regression (blue) for one lactation randomly selected (data points are daily observations from a single cow on a single farm).

specific example does not reflect all the results presented hereafter). Second, starting from the ULC, ULC_{305}, and ML were computed for every lactation. Figures 2 and 3 show, respectively, the scatterplots of ULC_{305} and ML estimated with different couples of models, together with linear regression statistics and Pearson correlation coefficients. Figure 4 instead compares the distributions of ULC_{305} and ML, respectively, across models. The median ULC_{305} were 9,632 kg, 9,693 kg, and 9,814 kg when estimated with IW, PLM, and QR, respectively. The median ML was 3.5% using IW, 4.3% using PLM, and 4.8% using QR. Based on Kruskal-Wallis statistical tests, the medians ULC_{305} and ML differed across methods ($P < 0.001$, respectively). Afterward, by performing post hoc Wilcoxon tests with Holm correction between pairs of methods, only the ULC_{305} median estimated with IW and with PLM did not result in a statistical difference ($P = 0.108$), whereas all the pairs of ML comparisons were statistically different ($P < 0.001$ in all scenarios).

To analyze more in detail IW, PLM, and QR performance, the total unperturbed MY and ML were estimated for each stage of the lactation, obtaining ULC_{early} and ML_{early}, ULC_{mid} and ML_{mid}, and ULC_{late} and ML_{late}. Figure 5 shows the boxplots by stage of lactation for the total unperturbed MY and ML, respectively. The median ULC_{early} was 2,132 kg when estimated with IW, 2,092 kg when estimated with PLM, and 1,931 kg when estimated with QR; the median ULC_{mid} were 3,239 kg, 3,305 kg, and 2,896 kg, respectively; the median ULC_{late} were 4,229 kg, 4,294 kg, and 4,987 kg, respectively. The median ML_{early} was 3.1% using IW, 0.9% using PLM, and 0.0% using QR; the median ML_{mid} were 3.5%, 5.4%, and 0.0%, respectively; the median ML_{late} were 2.8%, 4.1%, and 19.4%, respectively. The medians of the total estimated unperturbed MY and ML at each lactation stage were significantly different across models (Kruskal-Wallis tests; $P < 0.001$ in all scenarios), and also between all pairs of models (post hoc Wilcoxon tests; $P < 0.05$ in all scenarios).

Perturbations of the Lactation Curve. The IW and PLM were used to detect the perturbations of every lactation of the dataset. Using the same MY data plotted in Figure 1, Figure 6A shows the ULC and the perturbations estimated with IW, whereas Figure 6B shows the ULC and the perturbed lactation curve estimated with PLM. The distribution of N_P across the 2 methods is represented in Figure 7. The mode of N_P was 3 when applying IW and 4 when applying PLM. To better compare IW and PLM performance, Figure 8 shows the distribution of the number of perturbations identified by the 2 methods at each lactation stage. The modes of $N_{P,early}$, $N_{P,mid}$, and

Table 2. Average estimated values of the total unperturbed milk production (ULC_{305}), the total milk loss (ML), the number of detected perturbations (N_P), and the length of perturbations (L_P) obtained using 2 different sets of meta-parameters for each different model[1]

Model	Meta-parameter	Value	ULC_{305} (kg)	ML (%)	N_P (no.)	L_P (d)
IW	Threshold	Set 1: 1.6	9,770	4.1	3.8	18.7
		Set 2: 2	9,579	2.2	3.4	16.8
PLM	Fittings, iterations, maximum perturbations	Set 1: 50, 1,000, 10	9,809	4.9	4.2	18.6
		Set 2: 20, 500, 8	9,724	4.6	3.5	21.0
QR	Quantile	Set 1: 0.7	9,942	5.1		
		Set 2: 0.6	9,781	3.0		

[1]IW = iterative Wood model; PLM = perturbed lactation model; QR = quantile regression.

Table 3. Average root mean squared error of the total unperturbed milk production estimated with iterative Wood model (IW), perturbed lactation model (PLM), and quantile regression (QR)

Model	All (mean ± SD over farms)	Parity 1 (mean ± SD over farms)	Parity 2 (mean ± SD over farms)	Parity ≥3 (mean ± SD over farms)
IW	850.2 ± 151.7	732.3 ± 316.5	790.0 ± 210.8	959.1 ± 152.3
PLM	865.5 ± 97.0	704.4 ± 114.9	791.3 ± 88.6	1,004.3 ± 146.6
QR	1,118.6 ± 185.0	875.0 ± 298.6	1,017.6 ± 239.4	1,322.3 ± 238.4

$N_{P,late}$ were 1 at each lactation stage using both IW and PLM. Finally, Figure 9 compares the distribution of L_P (in days) detected by the 2 models. On average, the perturbations identified by IW lasted 18.7 (±17.8, SD) days, while the ones identified by PLM lasted 18.6 (±23.7, SD) days.

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

The nonsignificant terms that resulted from the ANOVA tables were removed from the models on ULC_{305}, ML, and N_P, and the final estimated fixed effects are reported in Table 4. Multiparous cows of the same breed that calved in the same season potentially produce, on average, 1,500 kg (combined effect of main effect and interaction terms) of milk more during a whole lactation with respect to primiparous ones. High-parity cows

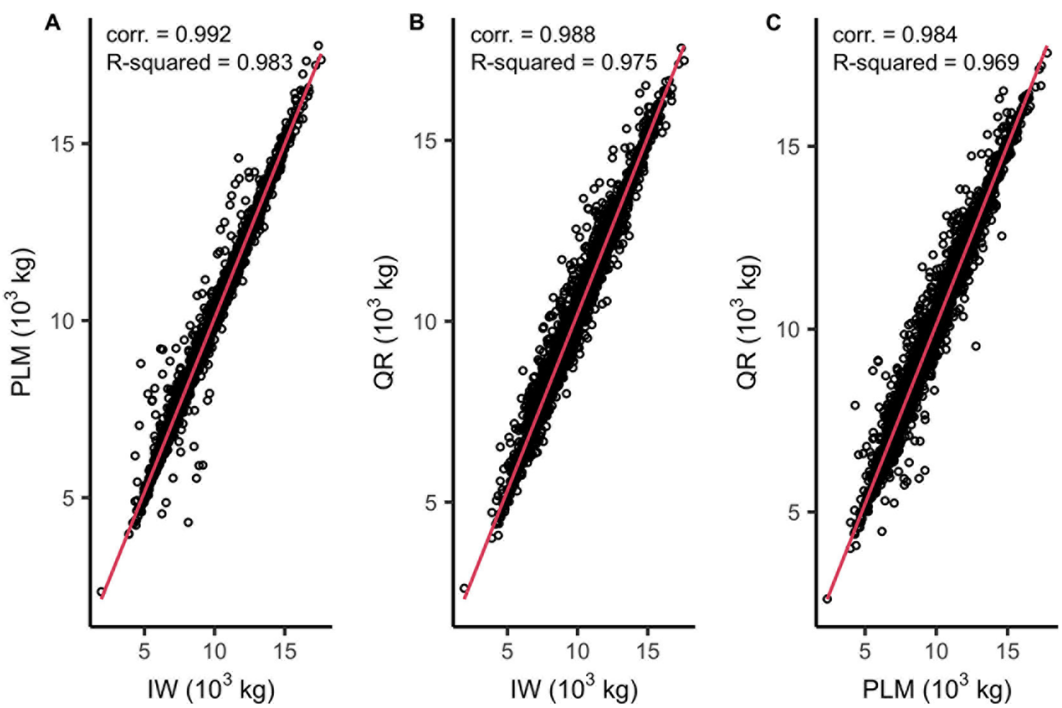

Figure 2. (A) Scatterplot of the total unperturbed milk production (in 10^3 kg) estimated with iterative Wood model (IW) versus perturbed lactation model (PLM); (B) scatterplot of the total unperturbed milk production estimated with IW versus quantile regression (QR); (C) scatterplot of the total unperturbed milk production estimated with PLM versus QR, together with the estimated regression line, related adjusted R^2, and Pearson correlation coefficient (corr.).

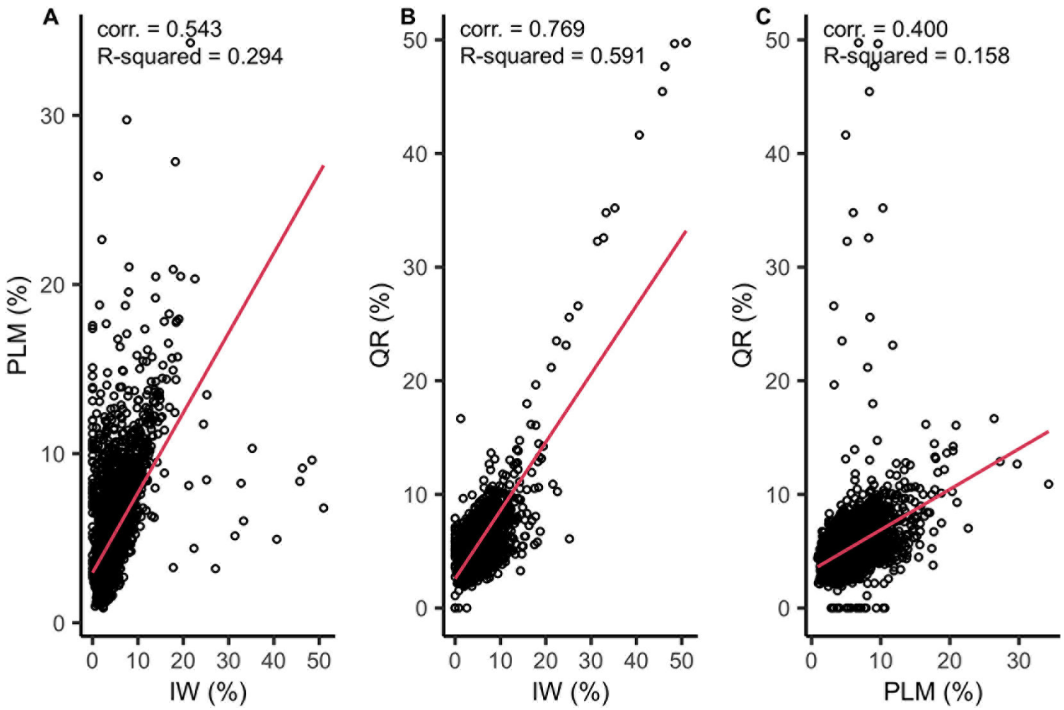

Figure 3. (A) Scatterplot of the total milk loss (in percentage) estimated with iterative Wood model (IW) versus perturbed lactation model (PLM); (B) scatterplot of the total milk loss estimated with IW versus quantile regression (QR); (C) scatterplot of the total milk loss estimated with PLM versus QR, together with the estimated regression line, related adjusted R^2, and Pearson correlation coefficient (corr.).

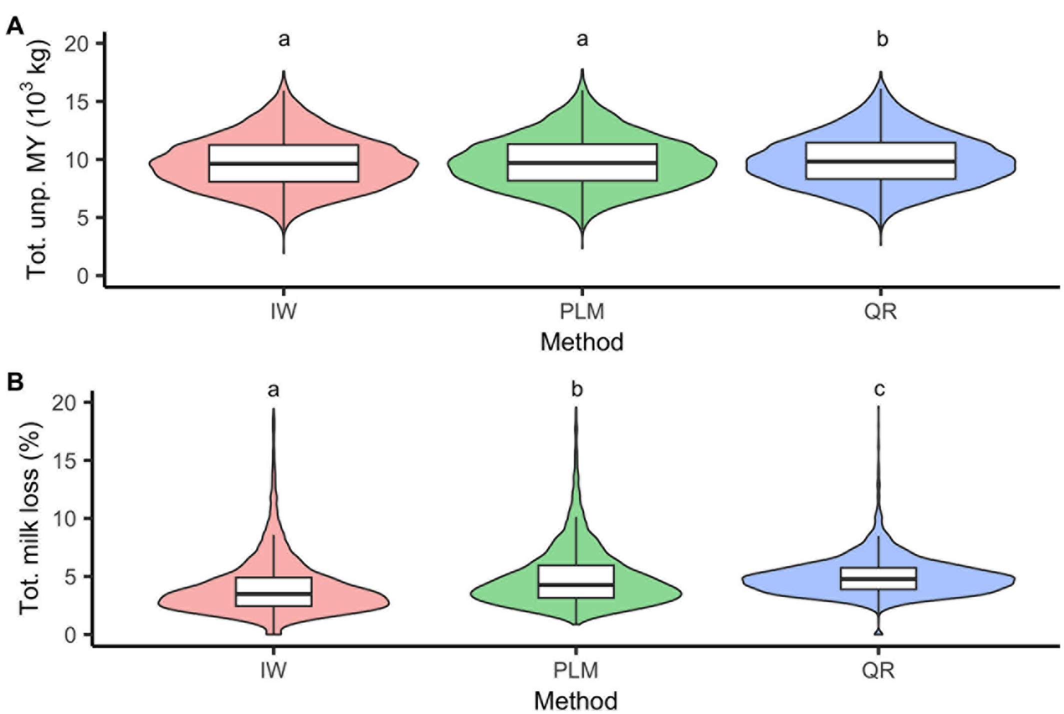

Figure 4. (A) Distribution (boxplot and violin plot) of the estimated total unperturbed milk yield (tot. unp. MY) across different estimation methods (iterative Wood model [IW], perturbed lactation model [PLM], quantile regression [QR]); (B) distribution (boxplot and violin plot) of the estimated total milk loss (tot. milk loss) across different estimation methods (top limit of the y-axis is reduced). Distributions with different letters differ ($P < 0.05$). Upper edge: third quartile. Lower edge: first quartile. Midline: median. Upper whisker: maximum. Lower whisker: minimum. Shaded area: probability density of the data.

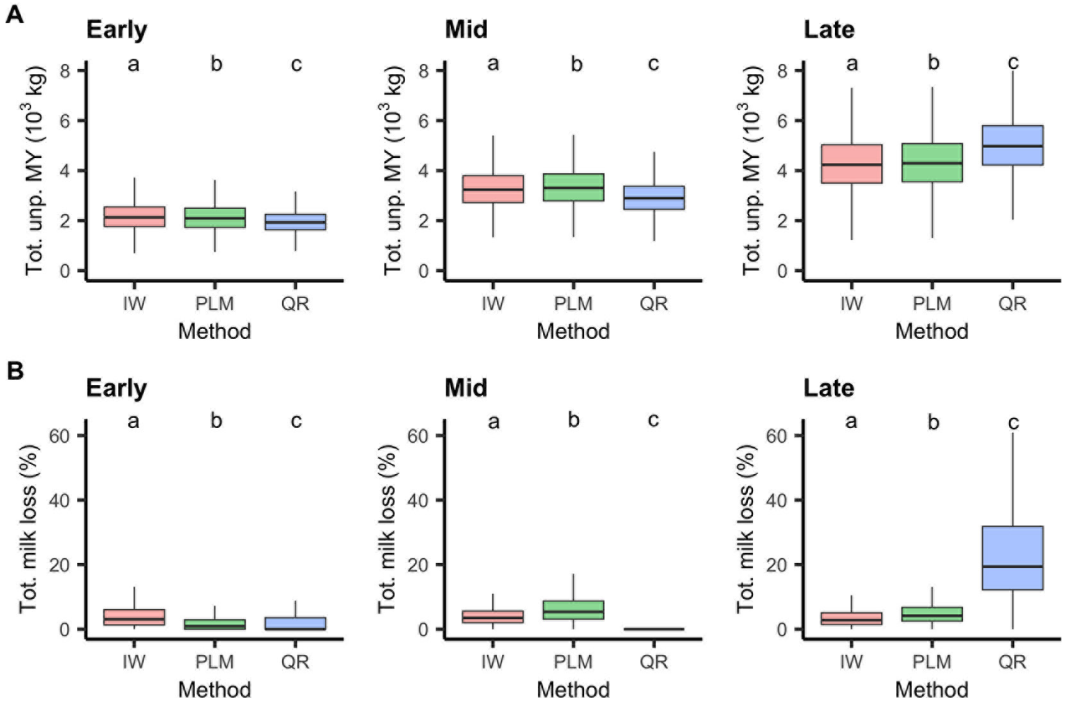

Figure 5. (A) Boxplot of the estimated total unperturbed milk yield (tot. unp. MY) per lactation stage (early: DIM ≤ 60, mid: 60 < DIM ≤ 150, late: 150 < DIM ≤ 305) using different estimation methods (iterative Wood model [IW], perturbed lactation model [PLM], quantile regression [QR]); (B) boxplot of the estimated total milk loss (tot. milk loss) per lactation stage using different estimation methods. Distributions with different letters within lactation stage differ ($P < 0.05$). Upper edge: third quartile. Lower edge: first quartile. Midline: median. Upper whisker: maximum. Lower whisker: minimum.

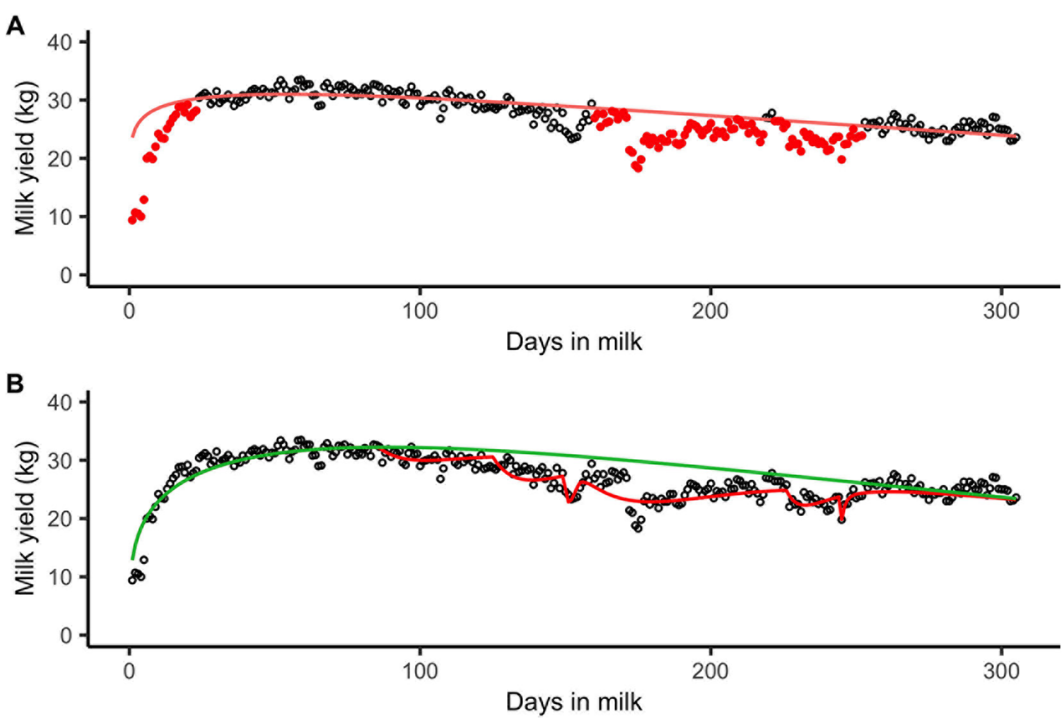

Figure 6. (A) Unperturbed lactation curve (pink line) estimated with iterative Wood model and perturbations identified (red points). (B) Unperturbed (green line) and perturbed (red line) lactation curves estimated with perturbed lactation model.

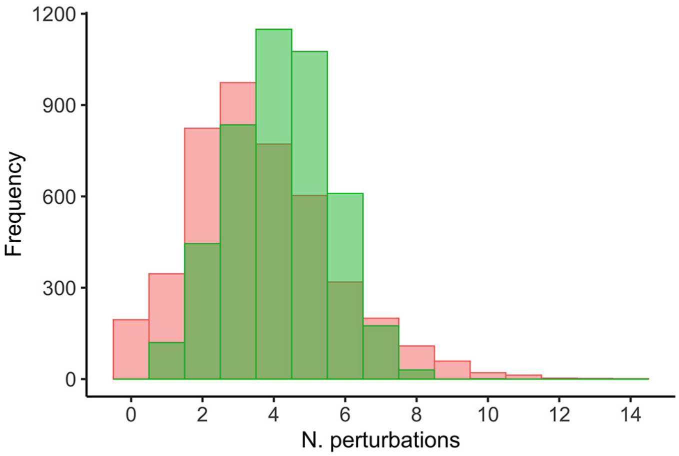

Figure 7. Histogram of the number of perturbations identified across lactations using iterative Wood model (pink) and perturbed lactation model (green).

(parity ≥3) calving in winter produce 700 kg (combined effect of main effect and interaction term) of milk more than the ones calving in summer in the absence of perturbations of the lactation curve; high-parity Holstein cows potentially produce around 2,000 kg (combined effect of main effect and interaction term) of milk more than high-parity Italian Simmental ones. Multiparous cows have more relative milk losses with respect to the unperturbed production than primiparous ones ($e^{0.13}$% = 1.1% × primiparous ML); the cows calving in winter have fewer relative milk losses than the cows calving in summer ($e^{-0.09}$% = 0.9% × summer ML). Last, summer-calving cows have fewer perturbations in the lactation curve than spring-calving cows (0.25 perturbations less), whereas the animals at parity 2 have fewer perturbations than the primiparous ones (0.25 perturbations less).

DISCUSSION

Comparison of Models

This work compared 3 existing mathematical models (i.e., IW, PLM, and QR) to estimate lactation performance in dairy cows. The IW, PLM, and QR were found to be valuable methods to obtain cow expected production and ML, each one with advantages and disadvantages depending on the specific application (see Table 5 for a summary).

From a mathematical and computational point of view, PLM is quite complex. Depending on the number of iterations imposed and the computational power of the machine, the estimation procedure of 4 + 4 × P parameters can take several hours to fit the individual curves of an entire set of lactation data. For instance, setting 1,000 iterations and using a simple machine with a 1.8 GHz Intel Core i5 dual-core processor and a RAM of 8 GB

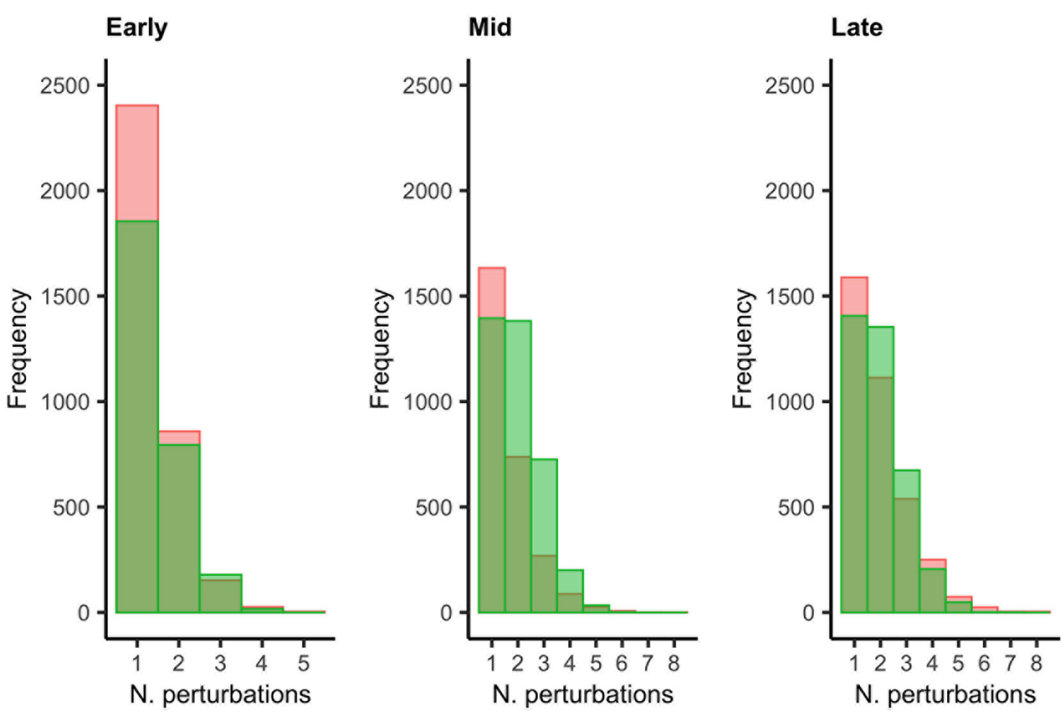

Figure 8. Histogram of the number of perturbations identified across lactation stages (early: DIM ≤ 60, mid: 60 < DIM ≤ 150, late: 150 < DIM ≤ 305) with iterative Wood model (pink) and perturbed lactation model (green).

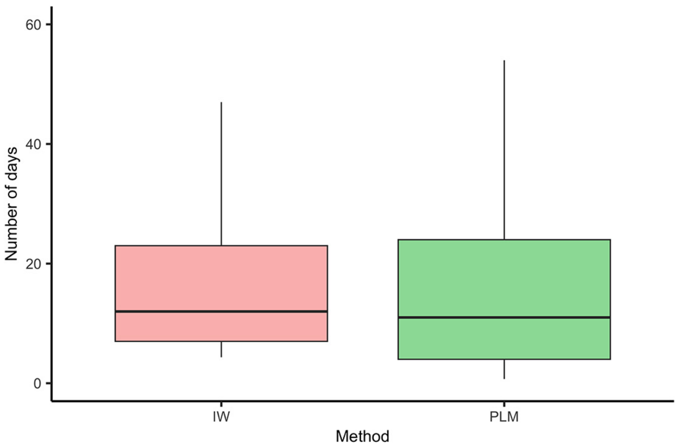

Figure 9. Boxplot of the length of the perturbations (in days) identified across lactations using iterative Wood model (IW) and perturbed lactation model (PLM). Upper edge: third quartile. Lower edge: first quartile. Midline: median. Upper whisker: maximum. Lower whisker: minimum.

takes around 5 h for a set of 300 lactations, compared to about 20 seconds for IW and 10 seconds for QR. In addition, PLM requires a maximum number of perturbations to be set, with the computational burden of the algorithm raising as this number increases. We set a maximum of 10 perturbations to be identified per lactation, knowing that higher numbers may occur, especially in case of severe metabolic disorders or chronic mastitis (Hostens et al., 2012). Nonetheless, the resulting unperturbed and perturbed lactation curves obtained with PLM fit the data well, as shown in the example of Figure 4B. The IW and QR are mathematically easier, yet they also depend on meta-parameters; both Adriaens et al. (2021) and Poppe et al. (2020) selected the meta-parameters through visual observation of numerous fitted lactation curves. Through a sensitivity assessment (Table 2), when the effect assigned to low MY data points was reduced (set 2), all the models produced lower average values of ULC_{305} and ML, and IW and PLM detected fewer perturbations, as expected. The average length of the perturbations instead increased with PLM, contrary to IW. This could be explained by a lower number of overlapping perturbations identified by PLM, increasing the global length of individual perturbations.

The average RMSE of the total unperturbed MY, a measure of deviation of the total observed milk production from the total expected milk production, was higher for QR, meaning that QR tends to produce higher ML compared to the other 2 methods. In addition, as demonstrated in Adriaens et al. (2021), the RMSE increased with parity, suggesting a higher difference between actual and expected milk production for higher parities. The correlation coefficients between the different models, as well as the adjusted R^2 of the estimated linear regressions, revealed that the methods were strongly linearly related in the case of ULC_{305} estimation, whereas they were only

Table 4. Estimates of the fixed effects of the mixed models on the total unperturbed milk yield (ULC_{305}), the logarithm of the total milk loss (ML), and the number of perturbations (N_P) estimated with the perturbed lactation model (PLM), with relative SE and observed levels of significance[1]

	ULC_{305} (kg)			log(ML) (%)			N_P		
Fixed effect[2]	Estimate	SE	*P*-value	Estimate	SE	*P*-value	Estimate	SE	*P*-value
Intercept	9,088	412	***	1.4	0.07	***	4.2	0.09	***
Parity 2	1,329	108	***	0.06	0.02	**	−0.25	0.08	**
Parity ≥3	1,777	104	***	0.13	0.02	***	−0.09	0.07	NS
Calving season autumn	227	103	*	0.002	0.02	NS	−0.20	0.08	*
Calving season winter	176	105	NS	−0.09	0.03	***	0.06	0.08	NS
Calving season spring	127	105	NS	0.002	0.03	NS	0.25	0.09	**
Breed IS	−1,516	578	*	0.15	0.10	NS	0.21	0.09	*
Parity 2 × calving autumn	53	140	NS	—[3]	—	—	—	—	—
Parity ≥3 × calving autumn	431	134	**	—	—	—	—	—	—
Parity 2 × calving winter	375	142	**	—	—	—	—	—	—
Parity ≥3 × calving winter	580	133	***	—	—	—	—	—	—
Parity 2 × calving spring	318	147	*	—	—	—	—	—	—
Parity ≥3 × calving spring	377	141	**	—	—	—	—	—	—
Parity 2 × breed IS	−137	97.7	NS	−0.03	0.04	NS	—	—	—
Parity ≥3 × breed IS	−543	98.8	***	−0.10	0.03	**	—	—	—
Calving autumn × breed IS	—	—	—	−0.08	0.04	*	—	—	—
Calving winter × breed IS	—	—	—	0.02	0.04	NS	—	—	—
Calving spring × breed IS	—	—	—	0.02	0.04	NS	—	—	—

[1]Referent categories of the variables (parity 1; calving season summer; breed Holstein) are not reported; the intercept represents a parity 1 Holstein cow that calved in summer.

[2]IS = Italian Simmental.

[3]The effect was not significant and was removed from the model.

**P* < 0.05; ***P* < 0.01; ****P* < 0.001; NS = nonsignificant (*P* ≥ 0.05).

Table 5. Comparison of iterative Wood model (IW), perturbed lactation model (PLM), and quantile regression (QR) in terms of advantages and disadvantages of each method

Model	Advantage	Disadvantage
IW	Relatively fast and easy to implement.	Possible overestimation of ML[1] in the early lactation stage.
	Based on Wood model (stability, good description of MY[2] curve in the first 305 DIM).	Separated characterization of perturbations required.
	Detection of perturbations.	
	It can operate with less frequent MY data.	
PLM	Based on Wood model (stability, good description of MY curve in the first 305 DIM).	Not fast and not easy to implement.
	Detection of perturbations.	High computational burden.
	Characterization of perturbations within characterization of ULC.[3]	Possible overestimation of number and amplitude of perturbations.
	Characterization of overlapping perturbations.	Maximum number of perturbations to be detected must be set.
	Parametrization of perturbations.	Not certain it can operate with less frequent MY data.
QR	Very fast and easy to implement.	Estimation of ULC can be too influenced by low MY values.
	Flexible polynomial model.	Possible overestimation of ML in the late lactation stage.
	It can operate with less frequent MY data.	No criteria to detect perturbations.

[1]ML = milk loss.

[2]MY = milk yield.

[3]ULC = unperturbed lactation curve.

moderately correlated in the case of ML estimation. In particular, IW estimated lower ML compared to PLM; the opposite happened for QR compared with the other 2 methods, especially in correspondence of higher values estimated with IW and of lower values estimated with PLM. The distribution of ULC_{305} was very similar among models, but the same did not apply for ML especially when using QR compared to the other 2 methods. The QR, indeed, produces less-variable ML estimates compared to IW and PLM because it tends to follow the observed MY data more closely. This is due both to the intrinsically higher sensitivity of QR to very large perturbations from the expected lactation curve, and to the fourth-order polynomial regression that can take more variable shapes (Lever et al., 2016). The 3 models performed differently according to the lactation stage. The IW estimated the ML to be higher during the early stage of the lactation with respect to the other 2 methods. As shown in Figure 1, IW performs less well in capturing the ascending phase of the ULC, although we set it to never exclude MY values below ULC – 1.6 × SD when DIM <30 in the estimation process. The PLM, conversely, produces higher ML estimates during the mid-stage of the lactation, whereas QR follows more closely the time trend of MY and tends to drag down the curve when there is a dip in mid lactation. Additionally, QR results in a higher ULC in late lactation compared to the other methods, likely also contributing to the ML. The IW and PLM were compared also in terms of estimated MY perturbations. Both the distributions of N_P estimated using the 2 methods were right-skewed, but PLM detected on average more perturbations than IW. The IW detected more perturbations during the early stage of the lactation compared to PLM, which is mostly due to the poor adaptation of ULC estimated with IW on the first days after calving. In contrast, PLM tended to find more perturbations during the mid and late stages of the MY curve, mainly because multiple perturbations can overlap, also affecting all the previous perturbations in the estimation process. The 2 methods detected the same average length, around 19 d, of a perturbation (also Adriaens et al., 2021, found a mean of 19.8-d length), but differed in the variability of L_P. Because PLM can identify overlapping perturbations, the global length of a perturbation reaches a higher number of days compared to IW. On the other side, PLM does not set a limit for the minimum length of a perturbation, contrary to IW (i.e., 5 d).

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

Possible Improvements

The QR could also be used for the detection of perturbations using a criterion such as the one of IW (i.e., at least 5 d of negative residuals with at least 1 d of MY below 80% of ULC), bearing in mind that QR does not produce ULC with robust shape, running the risk of not identifying relevant perturbations.

Specific adaptations of the 3 models could further improve their performance, in addition to further tuning their meta-parameters. Optimizing the weights of specific phases of the lactation during the model fitting could enhance the performance of both IW and QR. For example, the conditional quantile of QR could change based on DIM, to avoid possible overestimation or underestimation of ULC during specific stages. As an alternative to the fourth-order polynomial, a linearized Wood model could be used in combination with QR, to keep at the same time the ease of a regression and the stability of shape produced by the Wood model. Finally, PLM and the iterative approach of Adriaens et al. (2021) could be tested by substituting the Wood model with other typical lactation models, such as the one developed by Wilmink (1987). Although the Wood model is the most-used function for modeling lactation curves, it has, in fact, many limitations (Bouallegue and M'Hamdi, 2019). According to Macciotta et al. (2011), main limitations of the Wood model are an overestimation of daily MY in the first part of the curve and an underestimation around and after the peak of lactation. In addition, by construction of the model itself, the production on calving day is constrained to be zero, which does not necessarily reflect lactation biology. Finally, the model is characterized by a high correlation between its parameters which results in a higher sensitivity to data distribution (Silvestre et al., 2006).

Practical Implications

Dairy practitioners and researchers can choose the model that best fits their specific needs. Farmers, veterinarians, or technicians may need a tool for phenotyping purposes; they generally require a fast and rough approach on large datasets with minimal computational effort, putting QR forward as the most suitable method for the estimation of the expected total milk production. Farm technology suppliers may be interested in implementing one of these methods to forecast the expected MY in the upcoming days of lactation or for cow monitoring through real-time perturbation detection (e.g., after 4 mo of registered MY values from the beginning of the lactation). If dairy practitioners are interested in perturbation detection, IW is a fast and rather precise method for identifying individual perturbations of the herd lactation curves. Researchers that need a precise fitting of the expected production or want to characterize (i.e., parametrize) perturbations to link them with external factors or farm management practices might be more interested in PLM. When high-frequency milk meter data are not available (e.g., a farm is not equipped with an AMS), then PLM should be avoided, because it has not yet been tested in combination with less frequent than daily MY data.

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

Application Example: Effect of Parity, Calving Season, and Breed on Lactation Performance

The results from the mixed-effects models on cow lactation performance estimated with PLM were consistent with the literature. The expected production was higher as parity progressed: heifers having not yet reached adult BW, nor having fully developed udders, produce less milk than multiparous cows (Wathes et al., 2007; Siewert et al., 2019). Nonsummer calving led to a higher MY potential compared with summer calving, especially for multiparous cows. It has been widely demonstrated that lactations that start during the hot season usually have lower-than-average production performance as an effect of thermal stress on the animals (Elahi Torshizi, 2016; Li et al., 2022). For the same reason, ML was higher for the cows calving in summer compared to the ones calving in winter. Moreover, ML was higher for multiparous cows than for primiparous ones. The same result was found by Carvalho et al. (2019) and Adriaens et al. (2021), and it could be explained by a higher incidence rate of diseases during higher parities (Lee and Kim, 2006). The number of perturbations was larger in first than second lactations, which could be explained by a higher susceptibility to stressors due to management changes (e.g., regrouping or milking) of first-parity cows (Proudfoot and Huzzey, 2022). The number of perturbations increased also for spring calvings compared to summer calvings and this could be linked with heat stress episodes the animals calving in spring have to face during the first critical months of the lactation (McNamara, 2002). Last, as expected (Knob et al., 2023), Holstein cows had higher production potential than Italian Simmental ones, especially at high parities (parity ≥3).

CONCLUSIONS

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

NOTES

The authors thank the Livestock Technology group of KU Leuven (Geel, Belgium) and Associazione Nazionale Allevatori Pezzata Rossa Italiana (ANAPRI; Udine, Italy) for providing AMS data. No human or animal subjects were used, so this analysis did not require approval by an Institutional Animal Care and Use Committee or Institutional Review Board. The authors have not stated any conflicts of interest.

Nonstandard abbreviations used: AMS = automatic milking systems; IW = iterative Wood; L_P = length of perturbations; ML = milk loss; MY = milk yield; N_P = number of detected perturbations; PLM = perturbed lactation model; QR = quantile regression; RMSE = root mean squared error; ULC = unperturbed lactation curve.

REFERENCES

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

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.

Bach, A., N. Valls, A. Solans, and T. Torrent. 2008. Associations between nondietary factors and dairy herd performance. J. Dairy Sci. 91:3259–3267. https://doi.org/10.3168/jds.2008-1030.

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

Bates, D., and M. Maechler. 2017. Linear mixed-effects models using "Eigen" and S4. Accessed Jul. 9, 2023. http://cran.nexr.com/web/packages/lme4/index.html.

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

Bouallegue, M., and N. M'Hamdi. 2019. Mathematical modeling of lactation curves: A review of parametric models. Pages 21–142 in Lactation in Farm Animals—Biology, Physiological Basis, Nutritional Requirements, and Modelization. N. M'Hamdi, ed. IntechOpen, London, UK.

Carvalho, M. R., F. Peñagaricano, J. E. P. Santos, T. J. DeVries, B. W. McBride, and E. S. Ribeiro. 2019. Long-term effects of postpartum clinical disease on milk production, reproduction, and culling of dairy cows. J. Dairy Sci. 102:11701–11717. https://doi.org/10.3168/jds.2019-17025.

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.

Dodge, Y. 2008. Kruskal-Wallis test. Pages 288–290 in The Concise Encyclopedia of Statistics. Springer, New York, NY.

Elahi Torshizi, M. 2016. Effects of season and age at first calving on genetic and phenotypic characteristics of lactation curve parameters in Holstein cows. J. Anim. Sci. Technol. 58:8. https://doi.org/10.1186/s40781-016-0089-1.

Elzhov, T. V., K. M. Mullen, A.-N. Spiess, and B. Bolker. 2023. Minpack. Lm: R Interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK, plus support for bounds. Accessed May 15, 2023. https://cran.r-project.org/web/packages/minpack.lm/minpack.lm.pdf.

Fitzmaurice, G. M., N. M. Laird, and J. H. Ware. 2004. Applied Longitudinal Analysis. John Wiley & Sons, Hoboken, NJ.

Gote, M. J., L. D'Anvers, D. Meuwissen, B. Aernouts, and I. Adriaens. 2022. An automated pipeline for real-time data integration, health monitoring, and validation for dairy cows. Page 549 in 73rd Annual Meeting of the European Federation of Animal Science, Porto, Portugal.

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

Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 6:65–70.

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

Knob, D. A., A. M. Scholz, L. Perazzoli, B. P. B. Mendes, R. Kappes, D. R. M. Alessio, Â. F. Rech, and A. Thaler Neto. 2023. Feed efficiency and physiological parameters of Holstein and crossbred Holstein ×

Simmental cows. Animals (Basel) 13:1668. https://doi.org/10.3390/ani13101668.

Koenker, R. 2005. Quantile Regression. Econometric Society Monographs. Cambridge University Press, New York, NY.

Koenker, R. 2023. Quantreg: Quantile regression. Accessed May 16, 2023. https://cran.r-project.org/web/packages/quantreg/index.html.

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

Lever, J., M. Krzywinski, and N. Altman. 2016. Model selection and overfitting. Accessed Aug. 31, 2023. https://www.nature.com/articles/nmeth.3968.

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.

Macciotta, N. P. P., C. Dimauro, S. P. G. Rassu, R. Steri, and G. Pulina. 2011. The mathematical description of lactation curves in dairy cattle. Ital. J. Anim. Sci. 10:e51. https://doi.org/10.4081/ijas.2011.e51.

McNamara, J. P. 2002. Body condition: Effects on health, milk production and reproduction. Pages 168–173 in Encyclopedia of Dairy Sciences. Academic Press, Cambridge, MA.

Padfield, D., and G. Matheson. 2023. Robust non-linear regression using AIC scores. Accessed Dec. 18, 2023. https://cran.r-project.org/web/packages/nls.multstart/nls.multstart.pdf.

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

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

Ranzato, G., I. Adriaens, I. Lora, B. Aernouts, J. Statham, D. Azzolina, D. Meuwissen, I. Prosepe, A. Zidi, and G. Cozzi. 2022. Joint models to predict dairy cow survival from sensor data recorded during the first lactation. Animals (Basel) 12:3494. https://doi.org/10.3390/ani12243494.

Ranzato, G., I. Lora, B. Aernouts, I. Adriaens, F. Gottardo, and G. Cozzi. 2023. Sensor-based behavioral patterns can identify heat-sensitive lactating dairy cows. Int. J. Biometeorol. 67:2047–2054. https://doi.org/10.1007/s00484-023-02561-w.

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

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

Silvey, S. D. 1975. The likelihood-ratio test and alternative 'large-sample' equivalents of it. Pages 95–110 in Statistical Inference. 1st ed. Chapman & Hall/CRC, Boca Raton, FL.

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. Domest. Anim. Endocrinol. 33:203–225. https://doi.org/10.1016/j.domaniend.2006.05.004.

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.

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

ORCIDS

G. Ranzato https://orcid.org/0000-0001-6534-8110
B. Aernouts https://orcid.org/0000-0001-6266-3019
I. Lora https://orcid.org/0000-0001-8725-8597
I. Adriaens https://orcid.org/0000-0001-9768-2308
A. Ben Abdelkrim https://orcid.org/0000-0002-5721-0944
M. J. Gote https://orcid.org/0000-0003-4153-1943
G. Cozzi https://orcid.org/0000-0003-0408-1082