Modeling Extended Lactations of Dairy Cows

B. Vargas,*,† W. J. Koops,‡ M. Herrero,§,|| and J.A.M. Van Arendonk†

*Escuela de Medicina Veterinaria, Universidad Nacional de Costa Rica, PO Box 304-3000, Heredia, Costa Rica

†Animal Breeding and Genetics Group, Wageningen Institute of Animal Sciences, Wageningen Agricultural University, PO Box 338, 6700 AH Wageningen, The Netherlands

‡Animal Production Systems Group, Wageningen Institute of Animal Sciences, Wageningen Agricultural University, PO Box 338, 6700 AH Wageningen, The Netherlands

§Institute of Ecology and Resource Management, The University of Edinburgh, West Mains Road, Edinburgh EH9 3JH, United Kingdom

||International Livestock Research Institute, PO Box 30709, Nairobi, Kenya

ABSTRACT

Nine mathematical models were compared for their ability to predict daily milk yields (n = 294,986) in standard 305-d and extended lactations of dairy cows of Costa Rica. Lactations were classified by parity (first and later), lactation length (9 to 10, 11 to 12, 13 to 14, 15 to 16, and 16 to 17 mo), and calving to conception interval (1 to 2, 3 to 4, 5 to 6, 7 to 8, and 9 to 10 mo). Of the nine models, the diphasic model and lactation persistency model resulted in the best goodness of fit as measured by adjusted coefficient of determination, residual standard deviation, and Durbin-Watson coefficient. All other models showed less accuracy and positively correlated residuals. In extended lactations, models were also fitted using only test-day records before 305 d, which resulted in a different ranking. The diphasic model showed the best prediction of milk yield in standard and extended lactations. We concluded that the diphasic model provided accurate estimates of milk yield for standard and extended lactations. Interpretation of parameters deserves further attention because of the large variation observed. As expected, the calving to conception interval was found to have a negative effect on milk yield for cows with a standard lactation length. In extended lactations, these negative effects of pregnancy on milk yield were not observed.

(**Key words:** extended lactations, lactation curves, dairy cattle, milk yield)

Abbreviation key: CC = calving to conception, **DW** = Durbin-Watson, **LPM** = lactation persistency model, **MAD** = mean absolute deviation, **RLPM** = reduced lactation persistency model, **RSD** = residual standard deviation, **RSS** = residual sum of squares, **UNA** = Universidad Nacional de Costa Rica.

INTRODUCTION

Modeling of lactation curves has been a subject of extensive study during the past decade (2, 9, 10, 12, 16, 19, 26, 27). Different models have been evaluated for their ability to describe the pattern of milk yield as well as the ability to predict 305-d cumulative milk yield from partial records of lactation. Attention has focused on the 305-d lactation period, which implies that information collected after 305 d is usually ignored and that no attention is paid to milk yield in the period after 305-d in extended lactations. In most countries, many cows have lactations that are longer than 305 d. For example, in dairy herds of Costa Rica, more than 25% of cows are dried-off after 330 d of lactation, the average lactation length being about 328 d. Longer lactations partly result from failures to conceive at an early stage of lactation. Costs of prolonged calving interval have been demonstrated to greatly depend on milk yield in the latter part of lactation (6). This production depends on increase in lactation length and shape of lactation curve. Knowledge of lactation curves over the entire trajectory is a key element in determining optimum strategies for insemination and replacement of dairy cows (4, 5, 8, 21). An earlier study (9) has shown that lactation length has a significant effect on estimates of initial yield, peak yield, 305-d yield, time of peak and persistency. However, this analysis considered only lactation lengths less than 360 d.

Models to describe lactation have been classified into two main groups, i.e., as linear and nonlinear models (11). In linear models, parameters are linear functions of days in lactation or transformation of and can be

Received October 25, 1999.
Accepted December 10, 1999.
Corresponding author: Johan van Arendonk, Animal Breeding and Genetics Group, PO Box 338, 6700 AH Wageningen, The Netherlands; e-mail: johan.vanarendonk@alg.vf.wau.nl.

Table 1. General description of the dataset.

Herds	129
Cows	7608
Lactations	13,752
First lactations	3573
Test-day records	294,986
Records/lactation	21.5 ± 12.9
Daily milk yield (kg)	19.6 ± 8.6
Lactation length (d)	328 ± 61.4

easily computed by simple linear regression techniques. Nonlinear models cannot be expressed as linear functions of parameters and, therefore, need iterative techniques to be solved (11). These models have become more popular during recent years (2, 12, 16, 19, 26, 27), especially because they are able to describe a relatively wide range of shapes in lactation curves. Iterative procedures for fitting nonlinear regression implemented in statistical software have overcome the problem of model fitting. Many existing models show systematic deviations from actual milk yield, especially at the beginning and end of lactation (9). Multiphasic models have been suggested as an option to overcome these problems (9). These models were previously implemented with success to describe growth curves in mice and chickens and, more recently, to describe standard lactations in dairy cows (3, 20). Multiphasic model considers daily milk yield as the result of an accumulation from more than one phase of lactation, intrinsically reducing correlation between subsequent residuals.

The objective of this study was to compare existing models for their ability to provide consistent predictors of partial and total milk yield in normal and extended lactations and to subsequently analyze effect of lactation length and calving to conception (**CC**) interval on the lactation curve.

MATERIALS AND METHODS

Data Source

The analysis was performed on data provided by Universidad Nacional de Costa Rica (**UNA**) collected from 1987 to 1994 on dairy farms in Costa Rica. Farms participated in a project that focused on collection and analysis of data on health, milk yield, and reproduction performance to provide management support to farmers and to identify adequate management practices (7). Reproductive events, daily milk yield, and herd characteristics were entered into an improved version of VAMPP software package (14) by staff of UNA or directly by farmers.

The initial dataset consisted of 57,359 lactations of 26,072 cows. A subset of lactations was selected (Table 1), which included only Holstein cows with dates registered for conception and drying-off. Furthermore, each cow was required to have at least one test-day record in each of the following four periods: 1 to 60, 61 to 150, 151 to 240, and after 240 d in lactation. All test-day records between d 305 and the actual end of lactation were included in the analysis.

Our main interest was to find a model that provided a good description of the lactation curve for groups of cows with a range of lactation lengths and CC intervals. The results will be used in a bio-economic model to determine optimum insemination strategies. Given this objective, models were fitted to group mean yields rather than to individual lactations. Lactations were classified in two groups according to parity (first and later), five groups according to lactation length (9 to 10, 11 to 12, 13 to 14, 15 to 16, and 17 to 18 mo), and five groups according to CC interval during current lactation (1 to 2, 3 to 4, 5 to 6, 7 to 8, and 9 to 10 mo). Out of 50 possible groups ($2 \times 5 \times 5$), only those with more than 1000 test-day records were chosen for further analysis, which resulted in a total of 26 groups (Table 2).

The number of test-day records per group was highly variable because some combinations of lactation length and CC interval were less likely to occur. Test-day records within groups were classified in 2-wk periods, and for each 2-wk period, average DIM and milk yields were obtained and used in model fitting.

Model Fitting

Nine different models from the literature were analyzed (Table 3). The Wood model (27) is a gamma function, in which a approximates the initial milk yield after calving, b is the inclining slope parameter up to peak yield, and c is the declining slope parameter. The Cobby model (2) has the particularity that milk yield after peak is modeled as a linear decline function. The Rook model (16) describes lactation as a combination of a monotonically increasing growth function, in this case the Mistcherlich function, and a monotonically decreasing death function, which in this case is exponential. The Morant model (12) assumes that the change in milk yield after peak is not constant as implied in the exponential decline function (e^{-kt}). Wilmink model (26) is a modification of Cobby, and –0.05 is related to the moment of peak, which is about 50 d. Models based on the logistic function, such as the monophasic (9) and diphasic (9), were introduced to overcome problem of autocorrelation detected in models based on the gamma function. These models provided smaller and more random residuals (9). The lactation persistency model (**LPM**) (10) is also based on a logistic function. The LPM was developed to provide additional parameters to measure persistency, which is defined as number of

Table 2. Means, SD, and extreme values (Min. and Max.) for milk yield by parity, lactation length, and calving to conception interval (CC).

		Milk yield (kg/d)				
Lactation length (mo)[1]	CC (mo)[2]	Mean	n	SD	Min.	Max.
First parity						
10	2	17.5	9706	3.9	8.0	21.8
10	4	18.2	17,502	3.4	9.8	22.1
10	6	13.6	1035	2.8	6.4	17.4
12	4	17.4	16,297	4.1	7.6	22.2
12	6	17.2	9129	3.5	8.7	21.7
12	8	13.7	1444	3.9	4.6	18.6
14	6	17.1	6973	3.7	8.9	21.9
14	8	14.6	4492	3.3	7.4	19.2
14	10	12.7	1834	3.4	6.4	17.6
16	8	16.6	2803	4.2	7.0	22.3
16	10	16.2	4136	3.8	7.7	21.8
18	10	15.6	5013	3.2	9.5	20.4
Later parities						
10	2	19.6	21,983	6.1	7.5	27.1
10	4	20.5	53,723	5.8	8.9	27.6
10	6	16.8	5363	5.7	5.8	24.2
10	8	14.9	1401	5.1	5.6	21.9
12	4	20.2	39,565	6.7	8.1	29.0
12	6	19.7	29,624	6.2	8.1	28.1
12	8	17.1	4972	6.0	5.9	26.0
12	10	15.4	1929	5.7	6.5	24.0
14	6	19.3	13,912	6.4	7.3	28.3
14	8	19.0	15,348	6.4	8.3	28.8
14	10	15.7	5143	5.6	6.4	24.5
16	8	19.1	5123	6.0	9.0	28.3
16	10	17.6	8057	6.0	7.9	27.1
18	10	17.7	8479	6.2	8.8	28.0

[1]Lactation length classes: 10 = 9 and 10 mo, 12 = 11 and 12 mo, 14 = 13 and 14 mo, 16 = 15 and 16 mo, and 18 = 17 and 18 mo.

[2]Calving to conception interval classes: 2 = 1 and 2 mo, 4 = 3 and 4 mo, 6 = 5 and 6 mo, 8 = 7 and 8 mo, and 10 = 9 and 10 mo.

days during which peak production is maintained. The reduced LPM (**RLPM**) (10) is based on LPM, but the number of parameters is reduced from six to four.

Models were fitted to group mean yields by using a Gauss-Newton iterative method from the SAS software nonlinear regression NLIN procedure (18). Convergence was determined based on change (c) in residual sums of squares (**RSS**) between iteration i and iteration (i – 1) according to:

$$(RSS_{i-1} - RSS_i)/(RSS_i + 10^{-6}) = c \qquad [1]$$

if $c < 10^{-8}$, converging criterion is met, and iteration process stops.

Table 3. Description of models under analysis.

Name	Source	Equation
Wood	Wood (27)	$a \times t^{b} \times e^{-c \times t}$
Cobby	Cobby & Le Du (2)	$a - t \times b - a \times e^{(-c \times t)}$
Wilmink	Wilmink (26)	$a + t \times b + c \times e^{t \times -0.05}$
Morant	Morant and Gnanasakthy (12)	$a \times e^{(b1 \times \grave{t}2/2 + b2/t - c \times (1 + \grave{t}/2) \times \grave{t})}$, with $\grave{t} = (t - 21.4)/100$
Rook	Rook et al. (16)	$a \times (1 - b_1 \times e^{-b2 \times t}) \times e^{-c \times t}$
Monophasic	Grossman and Koops (9)	$a \times b\,(1 - \tanh^2(b \times (t - c))$
Diphasic	Grossman and Koops (9)	$a_1 \times b\,(1 - \tanh^2(b_1 \times (t - c_1)) + a_2 \times b_2 \times (1 - \tan^2(b_2 \times (t - c_2))$
LPM[1]	Grossman et al. (10)	$yp + b_1 \times (t - t_1) + r_1 \times (b_2 - b_1) \times \ln((e^{t/r1} + e^{t1/r1})/(1 + e^{t1/r1}))$ $+ r_2 \times (b_3 - b_2) \times \ln((e^{t/r2} + e^{(t1 + P)/r2})/(1 + e^{(t1 + P)/r2})) + r_3 \times (b_4 - b_3)$ $\times \ln((e^{(t/r3)} + e^{(t3/r3)})/(1 + e^{(t3/r3)}))$
RLPM[2]	Grossman et al. (10)	$yp/t_1 \times t - yp/t_1 \times \ln((e^{t} + e^{t1})/(1 + e^{t1} + b_3 \times \ln((e^{t} + e^{t1 + P})/(1 + e^{t1 + P}))$

[1]Lactation persistency model—extended.

[2]Reduced lactation persistency model.

Goodness of fit of models was evaluated according to following criteria:

1. Adjusted multiple coefficient of determination [R^2_{adj}, (13)]:

$$R^2_{adj} = 1 - (n - 1)/(n - p) \times (1 - R^2) \qquad [2]$$

where, R^2 = multiple coefficient of determination (=1 − (RSS/TSS)), RSS = residual sum of squares, TSS = total sum of squares, n = number of observations, and p = number of parameters in the model.

Note that R^2 is adjusted for the number of parameters in the model (p) to make a fair comparison of models. For simplicity, R^2_{adj} will be regarded only as R^2.

2. First-order positive autocorrelation among residuals was assessed by Durbin-Watson coefficient [**DW**; (13)]:

$$DW = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} \qquad [3]$$

where e_t = residual at time t, and e_{t-1} = residual at time t–1. The observed value of DW was evaluated against the tabulated critical value to test for positive autocorrelation. Negative autocorrelation was not tested because a negative autocorrelation coefficient implies that residuals fluctuate in a strict “up and down” way around the actual curve, which in the particular case of lactation curves was not a problem.

3. Residual standard deviation [**RSD**; (13)] was obtained by

$$RSD = \sqrt{RSS/(n - p)} \qquad [4]$$

For RSS, n and p, see Equation [2].

Models were categorized based on estimates of three criteria: RSD, adjusted R^2, and DW. Four categories were formed, two for models deviating less than one (‘+’ and ‘–’) SD from the mean of one criterion and two for models deviating more than one SD (‘++’ and ‘--’) where SD represents standard deviation across groups for each of three criteria within model.

Mean absolute deviation (**MAD**) across groups for partial and total milk yield was compared among models. The absolute difference between actual and predicted milk yield during the specified periods was calculated, and averaging these values over all groups resulted in MAD. Partial yields were calculated for periods 1 to 100 d, 101 to 200 d, 201 to 305 d, and 306 d to end of lactation. Lactation length was set to 305 d (10 mo), 365 d (12 mo), 425 d (14 mo), 486 d (16 mo), and 547 d (18 mo).

Actual daily milk yields within groups were calculated by smoothing actual milk yields with cubic splines (25) to interpolate actual records in different intervals within lactation. A maximum of 10 splines-knots, according to lactation length, were set at d 50, 100, 150, 200, 250, 305, 365, 425, 486, and 547. This procedure implied that a different cubic spline was fitted for every interval defined by the knots. Splines were required to have continuous first and second derivatives and discontinuous third derivatives.

Partial and total actual milk yields within lactation were further estimated by

$$MY_{i-n} = \sum_{t=1}^{n} y(t) \qquad [5]$$

where MY_{i-n} = milk yield, i = initial day within time period (1 or 101 or 201 or 306), n = final day within time period (100, 200, or 305) or end of lactation (365, 425, 486, or 547), and y(t) = yield at day t estimated by a spline function (piecewise cubic polynomial with 6 to 10 knots).

Predicted milk yields within group were also obtained for every model using Equation [5] with substitution of y(t) by the corresponding model equation (Table 3).

The model with best overall performance according to previous criteria was selected, and residuals were plotted for all groups. Additional measures of functions of parameters were obtained to evaluate effect of lactation length, CC interval, and parity on parameters of the model and estimates of milk yield.

RESULTS AND DISCUSSION

Comparison of Models

All models, except that of Rook et al. (16), achieved convergence for every lactation group. The Rook model failed to achieve convergence in 3 of the 26 groups. Problems with convergence for this model have also been mentioned previously (15). Values of adjusted R^2, RSD, and DW coefficient for each model were averaged over the 26 groups (Table 4). Goodness of fit was high in general, R^2 ranged between 0.957 and 0.987, and RSD ranged from 0.42 to 0.87 kg/d. This high level of accuracy has also been reported in previous studies fitting models on mean yields (1, 17). For all but two models, values for DW were less than 1, which indicated positive autocorrelation of residuals for the majority of models.

The greatest R^2 values were found for the diphasic model and LPM, whereas the Wood model ranked lowest (Table 4). Similar ranking was found for RSD. In this case, LPM had, on average, a lower value of RSD

Table 4. Comparison of models according to adjusted multiple correlation coefficient (R^2), residual standard deviation (RSD), and Durbin-Watson coefficient (DW) (mean ± SD between groups).

Model	R^2	Rank	RSD	Rank	DW	Rank	Cases DW > 0[1]
Wood	0.957 ± 0.03	--	0.87 ± 0.22	--	0.56 ± 0.27	--	23
Cobby	0.961 ± 0.04	-	0.78 ± 0.28	-	0.90 ± 0.48	-	15
Wilmink	0.968 ± 0.03	-	0.70 ± 0.22	-	0.81 ± 0.46	-	17
Morant	0.973 ± 0.02	+	0.66 ± 0.17	+	0.86 ± 0.42	-	16
Rook	0.961 ± 0.03	-	0.83 ± 0.19	-	0.74 ± 0.27	-	17
Monophasic	0.965 ± 0.02	-	0.80 ± 0.16	-	0.91 ± 0.29	-	16
Diphasic	0.987 ± 0.01	++	0.48 ± 0.13	++	1.74 ± 0.44	++	1
LPM[2]	0.985 ± 0.03	++	0.42 ± 0.26	++	1.79 ± 0.60	++	2
RLPM[3]	0.969 ± 0.03	-	0.69 ± 0.22	+	0.85 ± 0.48	-	15
Average	0.969 ± 0.01		0.69 ± 0.15		1.02 ± 0.44		

[1]Number of runs (of 26) with significant positive autocorrelation.

[2]Lactation persistency model—extended.

[3]Reduced lactation persistency model.

than diphasic; however, RSD values obtained for the latter showed a lower standard deviation (0.13 vs. 0.26), which reflects a better performance across groups.

Positive autocorrelation between residuals was detected in all models except diphasic and LPM. Other models presented problems of positive autocorrelation among residuals for more than half of the groups (Table 4). Problems with positive autocorrelation have already been reported for the Wood and monophasic models (9). Absence of autocorrelation for the diphasic model is in agreement with Grossman and Koops (9) for standard 305-d lactations.

Additional analysis was performed to check parameter estimates for stability across groups. Results indicated that LPM, even though with a high general goodness of fit, often resulted in atypical parameters, e.g., negative values for parameter P (persistency). All other models seemed to provide more reasonable estimates of parameters, even though they also presented a wide variation.

Table 5 shows MAD for different periods and models. Models with low MAD were ranked at the top (++). Predictive performance for most models was highly variable for different periods within lactation. Models with a consistently good performance over all periods were diphasic and LPM, which is in agreement with earlier results on standard 305-d lactations (20). The LPM was more accurate than diphasic for 1 to 100 d and 201 to 305 d, whereas diphasic was more accurate for 101 to 200 d and 306 d to the end of lactation. All other models performed irregularly, ranking poorly for one or more stages within lactation, which reinforces the fact that exponential models usually fail to model peak of lactation (9, 20). It is important to notice that differences in accuracy of prediction between LPM and diphasic compared with other models is especially large for last period of lactation (305 d to end). For this period, only diphasic, LPM, and Morant models rank positively. This finding reflects that multiphasic models are intrinsically more suitable to describe extended lactations.

Further analysis showed that standard deviations of MAD across groups were, in general, higher for LPM compared with the diphasic model (Table 5). This result might also be related to the lack of stability observed for estimates of parameters using LPM model and might indicate a serious drawback of LPM compared with the diphasic model.

Analysis of predicted milk yield using diphasic curves showed some systematic deviations with respect to actual milk yield. For 1 to 100 d all models, except diphasic and that of Wilmink, underestimated milk yield; for 101 to 200 d all models, except the Cobby and diphasic, overestimated milk yield. For 201 to 305 d all models except diphasic and LPM underestimated milk yield. For the final period (306 d to end) all models, except diphasic and LPM overestimated milk yield.

In extended lactations, an additional comparison of models was also performed, eliminating all test-day records after 305 d. Models were fitted again to group mean yields. As expected, R^2 were higher and RSD lower, in general, because of reduction in length of lactations. One important result was that the frequency of cases in which a positive autocorrelation was detected was reduced considerably (29 vs. 122). By considering only 305-d lactations the ranking of models based on MAD changed (results not shown). For 101 to 200 d and 201 to 305 d, diphasic and LPM were still better than the others, but relative differences were reduced substantially. For 1 to 100 d, the top 5 models were Rook (31.1), LPM (34.5), Wilmink (34.6), diphasic (36.3), and Wood (41.6), all of which scored +. The results of our model comparison, based on standard 305-d lactations, are very similar to earlier findings (20). From our analysis, it is clear that models that are suitable to describe

Table 5. Comparison of models according to mean absolute deviation (MAD)[1] in different parts of the lactation.[2]

	1 to 100 d		101 to 200 d		201 to 305 d		306 d to end of lactation	
Model	MAD (kg)	Rank	MAD (kg)	Rank	MAD (kg)	Rank	MAD (kg)	Rank
Wood	45.8 ± 14.9	+	42.7 ± 15.0	-	80.7 ± 41.0	--	115.9 ± 56.8	--
Cobby	100.0 ± 18.9	--	33.6 ± 18.7	+	66.8 ± 43.2	-	93.3 ± 59.0	-
Wilmink	38.2 ± 21.1	+	33.9 ± 12.8	+	63.4 ± 39.9	-	88.0 ± 57.9	-
Morant	58.2 ± 12.3	+	38.1 ± 16.1	-	53.3 ± 31.8	+	84.1 ± 47.2	-
Rook	61.0 ± 19.2	+	42.8 ± 15.9	-	68.1 ± 36.5	-	101.0 ± 54.7	-
Monophasic	85.5 ± 26.8	-	43.9 ± 14.9	--	54.3 ± 31.8	+	82.1 ± 47.2	+
Diphasic	46.1 ± 27.5	+	17.7 ± 8.2	++	26.8 ± 17.0	++	43.2 ± 28.2	++
LPM[3]	36.1 ± 30.5	++	18.3 ± 21.8	++	22.6 ± 16.5	++	50.4 ± 59.4	++
RLPM[4]	118.5 ± 21.6	--	36.7 ± 19.0	-	60.1 ± 36.7	-	86.0 ± 56.2	-
Average MAD	65.5 ± 29.2		34.2 ± 9.9		55.1 ± 19.1		82.7 ± 22.9	

[1]MAD is calculated as the sum of daily absolute deviations (predicted – actual milk yield) for the period specified.

[2]Ranks are given according to relative performance by standardizing the criteria according to average MAD across groups.

[3]Lactation persistency model—extended.

[4]Reduced lactation persistency model.

standard 305-d lactations are not necessarily adequate to describe extended lactations.

In summary, the diphasic model was found to best describe normal and extended lactations, showing a high R^2, low RSD, uncorrelated errors according to DW test, more regular estimates of parameters for almost every group, and a similar performance along the whole lactation period. More detailed information on this model is given in next section.

Final Model

Estimates of residuals using the diphasic model were plotted for all groups of lactation length and CC interval within first parity (Figure 1). In general, residuals were randomly distributed. Residuals ranged between –1.5 and 1.5 kg/d. Significant positive autocorrelation was detected only for higher parity cows with a lactation length of 18 mo (Table 6). For this group, the model had problems in fitting the two phases, which resulted in inconsistent parameter estimates. A similar problem was found among first lactation cows with long lactations.

Estimated parameters for the diphasic model are given in Table 6. Values still showed a wide range of variation, which suggests that the shape of the curve greatly depends on lactation length, parity, and CC interval. Grossman and Koops (9) introduced the parameter duration of each phase, defined as days required to attain about 75% of asymptotic total yield during that phase and computed as $2b_i^{-1}$. For 305-d lactations of Dutch Black and White cows, they found duration of 198 and 415 d for the first and second phases, respectively. This finding agrees closely with our findings for cows with a lactation length of 12 mo. Duration of the second phase increased with lactation length and, to a much lesser extent, with CC interval. Duration of the first phase showed a different pattern in first and second lactation cows.

Based on criteria to measure goodness of fit, it can be concluded that the applied multiphasic concept for describing lactation also works well with extended lactations. However, parameters of the model still fluctuate too much with lactation length and, to a lesser extent, with CC interval. This fluctuation seriously limits application of the model and might be partly due to the choice of the logistic function, which is symmetric. Problems observed with very long lactations (Table 6) are likely caused by the symmetric nature of the applied function. Use of a nonsymmetric function, such as the Weibull model, could solve this problem. A second alternative would be to extend the number of phases of the model, which would be difficult to justify from a biological point of view. Finally, one could try to restrict some of the parameters to achieve more stable values for some of the parameters. For example, one might want to restrict duration of first peak. Results of this study might serve as a starting point for improving interpretation of parameters.

For cows with a 10-mo lactation, lactation curves with a CC interval of 2 and 4 mo showed a similar pattern (Figure 2a). Nevertheless, it could be observed that cows with 4-mo CC interval have higher production during the last part of lactation (after 200 d). The cumulative difference in estimates of 305-d milk yield is about 4% (Table 6). In previous research (1), pregnancy had an effect on parameters of the curve related to last part of the standard 305-d lactation, and a lower milk yield was found when CC interval was lower. This result is in line with results found in the present study for 2- and 4-mo CC. Lactation curves for cows with 6-mo CC are considerably lower and flatter than the others (Figure 2a). This result might be due to the way in which cows are grouped, i.e. we were looking at cows with a high CC but a relatively short lactation. In addition, differences in management strategies between herds might have influenced our results. For example, cows with high CC and relatively short lactation might be a reflection of a poorly managed herd.

For cows with a lactation of 12 mo, the increase in CC interval from 4 to 6 mo did not cause a major effect on milk yield (Table 6). On the contrary, when CC increased to 8 mo the curve was significantly lower (Figure 2b). For cows with longer lactations (14 mo and higher), the general trend was that milk yield (100 and 305 d) decreased as CC increased. It is certainly difficult to find a biological explanation for this reduction. The antagonistic relationship between milk yield and reproduction would lead to an increase rather than a decrease in milk yield. However, a negative effect of pregnancy on milk yield is only expected during the last part of the gestation period and, consequently, would only affect 305-d milk yield for cows that get pregnant during the first 3 mo of lactation (22, 23, 24), as observed in our study. Within lactation length, we did not find a negative effect of pregnancy, but we did find it across lactation lengths. The way data is presented certainly has an effect on the reduction observed for cows with large CC intervals. The number of test-day records for extended lactations or larger CC intervals was much lower (Table 2), as they were more unlikely to happen. Cows with specific health conditions or special treatments could be included in those groups and could certainly have an effect on the results. Also, differences in management strategies between farms might influence our results. Identification of such cases was not possible with this data set.

As expected, for cows with a given CC interval milk yield increased as lactation length increased (Table 6, Figure 3). Low producing cows with less persistent lactations are likely to be dried off earlier than high producing cows with persistent lactations. Consequently, lactation curves for cows with shorter lactations within a given CC interval tend to be lower (Figure 3).

Milk yield beyond 305 d for cows with different lactation lengths and CC intervals was very similar in first and later parity cows (Table 6). This finding reflects the effect of a flatter and more persistent lactation curve during first lactation, which has also been mentioned in earlier studies (19). Cows with a lactation length of 16 mo produced, on average, 2200 kg of milk after 305 d, which corresponds to as much as 26% of total milk yield. This finding raises the question of what the effects are of increased lactation length on lactation revenues. Economic consequences of a prolonged calving interval greatly depend on persistency of production and increase in lactation length with an increase in CC interval (6). These factors can be manipulated to some extent by breeding or feeding strategies. The effect on milk yield during next lactation must be also taken into account. Based on data from individual cows we observed an increase in lactation of 0.62 and 0.56 d for each

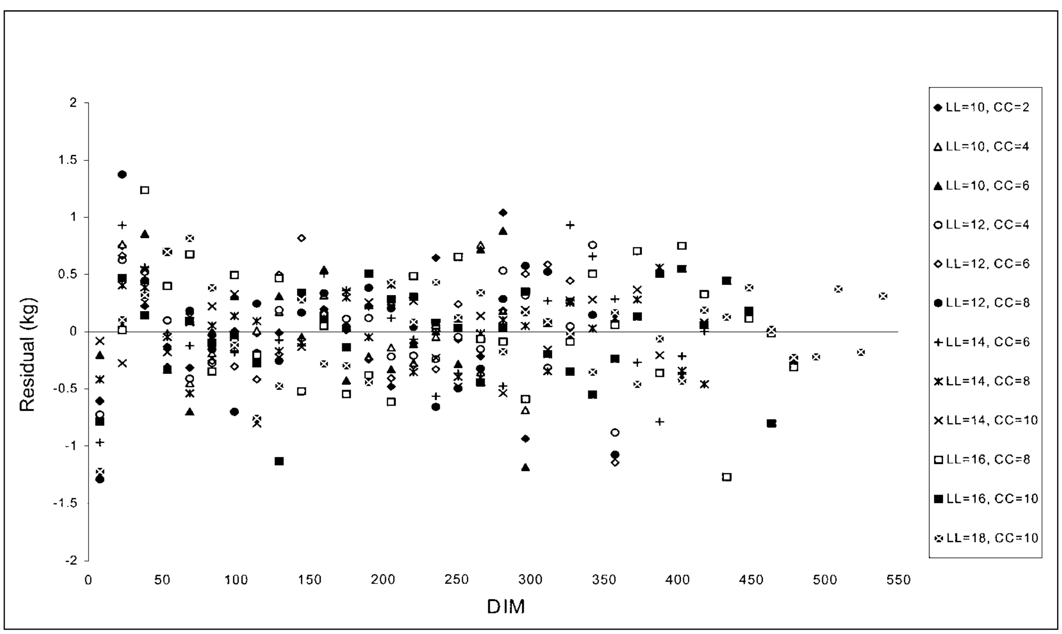

Figure 1. Residuals obtained from fitting diphasic model to mean group yields of first parity lactations grouped by length (LL; mo) and calving to conception interval (CC; mo).

Table 6. Goodness of fit, estimated parameters, and cumulative milk yield predicted with diphasic model for cows with different lactation length (LL), calving to conception interval (CC), and parity.[1]

Group		Goodness of fit			Estimated parameters[2]								Predicted milk yield (kg)[3]		
					First phase				Second phase						
LL (mo)	CC (mo)	R^2	RSD	DW	a_1	b_1	c_1	DUR_1	a_2	b_2	c_2	DUR_2	1 to 100 (%)	1 to 305 (%)	306 to end (kg)
First parity															
10	2	0.981	0.53	2.36	2091.5	0.0093	46.0	216	1453.8	0.0092	210.3	217	2085	5310	. . .
10	4	0.981	0.47	1.54	2398.5	0.0083	45.2	241	1492.8	0.0088	225.0	226	2	4	. . .
10	6	0.953	0.60	2.20	3011.8	0.0054	30.6	372	564.0	0.0107	230.6	187	–22	–22	. . .
12	4	0.988	0.45	1.89	2420.3	0.0079	43.2	254	1916.3	0.0072	242.7	278	2	8	592
12	6	0.977	0.53	1.36	1086.1	0.0114	33.9	175	3333.5	0.0050	203.7	399	–1	6	670
12	8	0.972	0.65	1.77	2301.9	0.0075	50.6	268	1122.7	0.0083	256.0	241	–15	–14	431
14	6	0.980	0.53	1.17	1799.7	0.0087	36.4	229	3035.7	0.0053	249.1	380	1	9	1369
14	8	0.990	0.33	1.94	1479.0	0.0086	25.9	233	2970.1	0.0045	244.1	443	–12	–7	1222
14	10	0.989	0.35	2.08	2777.0	0.0056	29.7	359	1468.1	0.0055	298.9	361	–19	–17	1014
16	8	0.972	0.71	1.75	3790.6	0.0051	39.3	394	2183.7	0.0055	325.7	363	1	11	2188
16	10	0.985	0.47	1.59	3628.2	0.0053	37.9	376	2203.0	0.0055	340.5	364	–1	6	2219
18	10	0.980	0.44	1.36	–4435.6	0.0045	194.4	441	12,117.8	0.0031	167.0	643	–7	3	3096
Later parities															
10	2	0.996	0.37	2.14	2130.8	0.0104	27.9	193	1931.2	0.0082	180.1	243	2575	5940	. . .
10	4	0.996	0.39	2.15	2086.9	0.0105	29.8	191	2180.5	0.0076	184.8	263	2	4	. . .
10	6	0.991	0.52	1.96	2077.0	0.0096	25.6	209	1579.9	0.0077	182.8	260	–11	–14	. . .
10	8	0.994	0.41	2.22	3709.1	0.0056	1.9	359	654.6	0.0077	206.5	261	–21	–24	. . .
12	4	0.997	0.35	1.48	1296.5	0.0124	26.3	161	3694.1	0.0056	163.6	360	7	14	559
12	6	0.996	0.37	1.32	1466.0	0.0112	25.8	179	3621.1	0.0051	173.1	391	4	10	602
12	8	0.993	0.50	1.27	888.6	0.0136	29.5	148	3726.5	0.0046	140.0	437	–6	–4	504
12	10	0.992	0.51	1.90	64.0	0.0501	38.5	40	5287.8	0.0040	48.3	504	–15	–13	426
14	6	0.996	0.43	1.53	2281.4	0.0089	28.2	224	3220.8	0.0052	219.7	382	5	14	1223
14	8	0.995	0.43	1.17	876.3	0.0132	34.4	152	5152.0	0.0039	152.6	510	5	14	1304
14	10	0.995	0.38	2.74	1183.5	0.0100	26.1	200	3981.1	0.0038	163.6	532	–11	–6	1072
16	8	0.993	0.52	1.48	617.2	0.0128	31.6	156	7147.6	0.0030	127.2	665	5	17	2168
16	10	0.995	0.41	2.08	691.7	0.0123	38.2	163	6495.7	0.0030	133.9	666	0	10	2023
18	10	0.979	0.89	0.87	–∞	8698.4	–∞	0	16,259.7	0.0018	–135.8	1092	3	16	2928

[1]RSD = residual standard deviation; DW = Durbin-Watson.

[2]a_i, b_i, c_i = parameters for phase i (i = 1,2) of diphasic model $E(a_i b_i - \tanh^2(b_i(t - c_i)))$; DUR_i = duration (d) of phase i ($2b_i^{-1}$) (9).

[3]Change in cumulative milk yield of a group for first group (LL = 10 and CC = 2) is given as fraction (%) of milk yield in the first group within parity. Absolute milk yield (kg) is given for first group. Milk yield beyond 305 d is given in kilograms.

additional day open in first and later parity cows, respectively. Further research must evaluate the profitability of extended lactations by using the results produced in the present study. Our findings could also have implications for the application of test-day models for the genetic evaluation of milk production; nevertheless, that is out of the scope of the present study.

CONCLUSION

Our results show that the diphasic model adequately fits lactations with variable length and variable CC interval. Accurate estimates of milk yield at later stages within lactation can be obtained. These results will be used in a bio-economic model to determine optimum insemination strategies, taking into account variation in lactation length between cows. Modifications are needed to improve consistency of parameters over a range of lactation lengths. Ranking of models changed when standard 305-d records only were analyzed, which supports the fact that further research is needed on modeling of extended lactations. As expected, CC interval was found to have a negative effect on milk yield for cows with a standard (10 mo) lactation length. In extended lactations, these negative effects of pregnancy on milk yield are no longer observed.

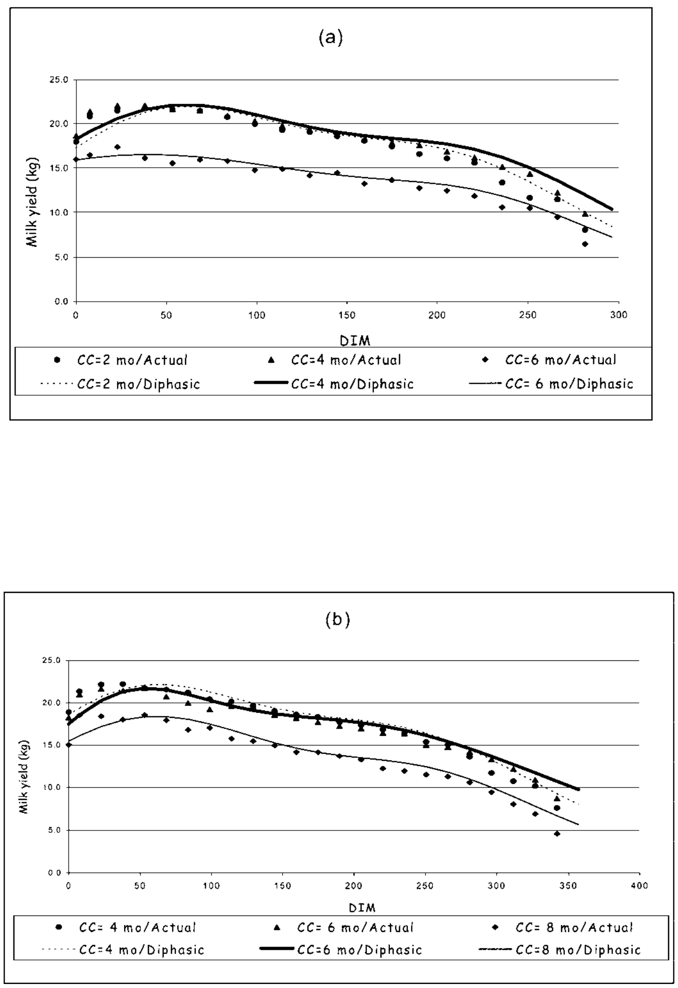

Figure 2. Actual yields vs. diphasic curves for first parity lactations grouped by calving to conception interval (CC) of cows with (a) 10-mo lactation of (b) 12-mo lactation.

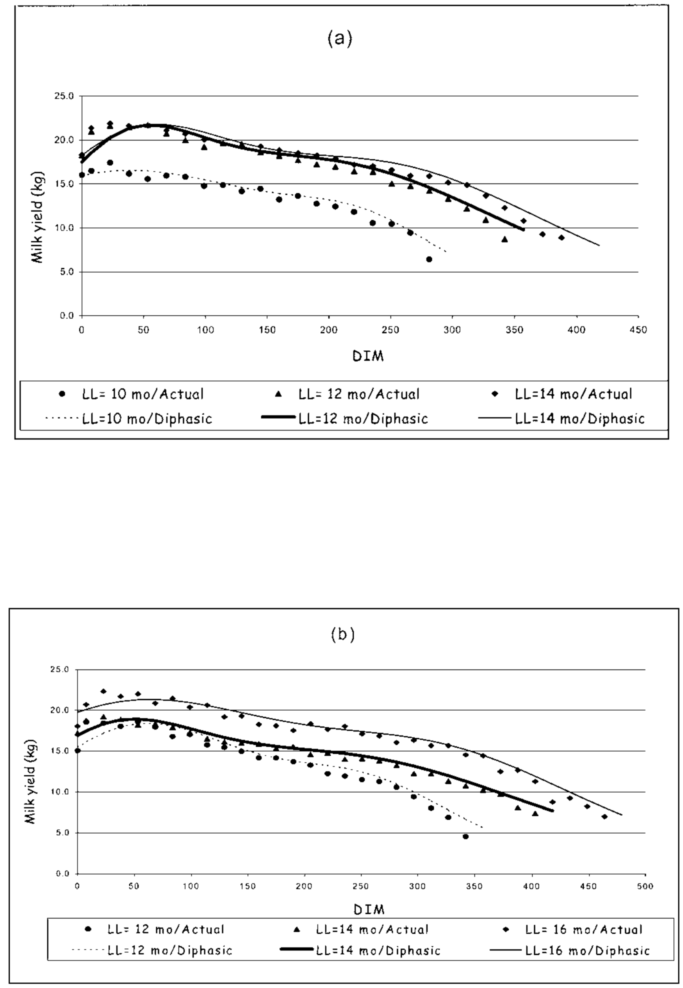

Figure 3. Actual yields vs. diphasic curves for first-parity lactations grouped by lactation length (LL) for cows with (a) 6-mo or (b) 8-mo calving to conception interval.

ACKNOWLEDGMENTS

The authors thank the farmers and staff at Universidad Nacional de Costa Rica for collecting the data and the financial support from Stichting Samenwerkings Verband IO-instellingen & LUW (SAIL), Delft, The Netherlands.

REFERENCES

1 Batra, T. R. 1986. Comparison of two mathemathical models in fitting lactation curves for pureline and crossline dairy cows. Can. J. Anim. Sci. 66:405–414.

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

3 De Boer, J. A., J. I. Weller, T. A. Gipson, and M. Grossman. 1989. Multiphasic analysis of milk and fat yield curves of Israeli Holstein. J. Dairy Sci. 72:2143–2152.

4 Dekkers, J.C.M. 1991. Estimation of economic values for dairy cattle breeding goals: bias due to sub-optimal management policies. Livest. Prod. Sci. 29:131–149.

5 De Lorenzo, M. A., T. H. Spreen, G. R. Bryan, D. K. Beed, and J.P.T.M. van Arendonk. 1992. Optimizing model: insemination, replacement, seasonal production, and cash flow. J. Dairy Sci. 75:885–896.

6 Dijkhuizen, A. A., J. A. Renkema, and J. Stelwagen. 1985. Economic aspects of reproductive failure in dairy cattle. II. The decision to replace animals. Prev. Vet. Med. 3:265–276.

7 Dwinger, R. H., E. Cappella, E. Pérez, M. Baaijen, and E. Muller. 1994. Application of a computerized herd management and production control program in Costa Rica. Trop. Agric. (Trinidad) 71:74–76.

8 Freeze, B., and T. Richards. 1992. Lactation curve estimation for use in economic optimization models in the dairy industry. J. Dairy Sci. 75:2984–2989.

9 Grossman, M., and W. J. Koops. 1988. Multiphasic analysis of lactation curves in dairy cattle. J. Dairy Sci. 71:1598–1608.

10 Grossman, M., W. J. Koops, and S. M. Hartz. 1999. Persistency of lactation production: a novel approach. J. Dairy Sci. 82:2192–2197.

11 Masselin, S., P. Sauvant, D. Chapoutot, and D. Milan. 1987. Les modeles d'ajustment des courbes de lactation. Ann. Zootech. 36:171–206.

12 Morant, S. V., and A. Gnanasakthy. 1989. A new approach to the mathematical formulation of lactation curves. Anim. Prod. 49:151–162.

13 Neter, J., W. Wasserman, and M. H. Kutner. 1985. Applied Statistical Methods. 2nd ed. Richard D. Irwin, Inc., Homewood, IL.

14 Noordhuizen, J.P.T.M, and J. Buurman. 1984. Veterinary automated management and production control program for dairy farms (VAMPP). The application of MUMPS for data processing. Vet. Q. 6:62. Vet. Q. 6:62–77.

15 Pérochon, L., J. B. Coulon, and F. Lescouret. 1996. Modeling lactation curves of dairy cows with emphasis on individual variability. Anim. Sci. 63:189–200.

16 Rook, A. J., J. France, and M. S. Dhanoa. 1993. On the mathematical description of lactation curves. J. Agric. Sci. (Camb.) 121:97–102.

17 Schaeffer, L.R., C. E. Minder, I. McMillan, and E. B. Burnside. 1977. Nonlinear techniques for predicting 305-days lactation production of Holsteins and Jerseys. J. Dairy Sci. 60:1636–1644.

18 SAS® User's Guide: Statistics, Version 6.0 Edition 1990. SAS Inst., Inc., Cary, NC.

19 Shanks, R. D., P. J. Berger, A. E. Freeman, and F. N. Dickinson. 1981. Genetic aspects of lactation curves. J. Dairy Sci. 64:1852–1860.

20 Sherchand, L., R. W. McNew, D. W. Kellog, and Z. B. Johnson. 1995. Selection of a mathemathical model to generate lactation curves using daily milk yields of Holstein cows. J. Dairy Sci. 78:2507–2513.

21 Van Arendonk, J.A.M. 1988. Management guides for insemination and replacement decisions. J. Dairy Sci. 71:1050–1057.

22 Vargas, B., and C. Solano. 1995. Cálculo de factores de corrección para producción diaria de leche en ganado lechero de Costa Rica. Arch. Latinoam. Prod. Anim. 3:131–148.

23 Vargas, B., and C. Solano. 1995. Factores de proyección y de corrección para producción por lactancia en vacas lecheras de Costa Rica. Arch. Latinoam. Prod. Anim. 3:149–164.

24 Vargas, B., T. van der Lende, M. Baaijen, and J.A.M. van Arendonk. 1998. Event-time analysis of reproductive traits of dairy heifers. J. Dairy Sci. 81:2881–2889.

25 White, I.M.S., R. Thompson, and S. Brotherstone. 1999. Genetic and environmental smoothing of lactation curves with cubic splines. J. Dairy Sci. 82:632–638.

26 Wilmink, J.B.M. 1987. Comparison of different methods of predicting 305-day milk using means calculated from within herd lactation curves. Livest. Prod. Sci. 17:1–17.

27 Wood, P.D.P. 1967. Algebraic model of the lactation curve in cattle. Nature (Lond.) 216:164–165.