Use of Lactation Curves for Analysis of Milk Production Data

T. A. SCOTT,[1] B. YANDELL,[2] L. ZEPEDA,[3]
R. D. SHAVER,[4] and T. R. SMITH[4]
University of Wisconsin, Madison 53706

ABSTRACT

Nonlinear equations were compared with categorical analysis to account for DIM effects on milk production. Five different models for lactation curves were evaluated. Derived from a multiphasic lactation curve, the selected lactation curve appeared to result in random residuals and performed more consistently than the multiphasic curve. Residuals from the fitting of lactation curves were then used for split-plot analysis (continuous model) to estimate treatment effects. Statistical performance of this model was compared with split-plot analysis based on a discrete model with regularly spaced intervals to account for DIM effects (discrete model).

The fitting of lactation curves accounted for herd, lactation number, and interaction effects of herd and lactation number and accounted for 34.1 and 44.3% of variance among cows for primiparous and multiparous cows, respectively. The continuous model detected interactions of genetic and management factors with treatment of multiparous cows that were not detected by the discrete model.

No statistically significant differences were detected between the two modeling approaches. The continuous model appeared to violate fewer assumptions regarding data distribution than did the discrete model, which reduced the risk of introducing bias during the estimation of treatment effects. The continuous model seemed to be more sensitive to subtle interactions of treatment and other factors.
(**Key words**: lactation curves, data analysis, milk production)

Abbreviation key: **BCSC** = body condition score at calving, **IG** = incomplete gamma (Wood's) lactation curve model, **IP** = inverse polynomial lactation curve model, **LCSP** = lactation curve fitted prior to split-plot model, **LHG** = lactation within a herd group, **M1** = monophasic lactation curve model, **M2** = diphasic lactation curve model, **MC** = monophasic lactation curve model with additive constant, **SCSP** = split-plot model using DIM subclasses.

INTRODUCTION

Production of milk and milk components varies with stage of lactation or DIM (8, 14, 21). Improper models for the effects of DIM may fail to remove autocorrelation from the residual error. When production among a group of cows is not measured using the same DIM, the use of discrete intervals to account for DIM effects may introduce biases in the estimation of treatment responses (13). Additionally, when cows are not simultaneously placed on trial, models for time-related effects can be complicated by the presence of multiple time series. In particular, time effects can confound the relationship among predictors and milk production (13).

The DIM account for a substantial amount of the variation in the production of milk and milk components within the lactation of a single cow. This time series may differ substantially among lactations and herds; however, DIM effects usually follow a pattern that is similar for all cows in the same lactation within a herd group (**LHG**) (10, 22, 23). Primiparous cows are more persistent, and lactation curves are flatter than those for multiparous cows (22). Lactation curves of multiparous cows are similar, except that the estimate of daily milk production is multiplied by a slightly higher factor for cows in third and greater lactations (10).

The incomplete gamma function (**IG**) is the predominant lactation curve used to model milk production of dairy cows; however, the inverse polynomial (**IP**) has outperformed IG in some situations (1). Both IG and IP have similar limitations: milk production at 0 DIM is forced to 0, milk production during early lactation is overpredicted, peak milk production is underpredicted (1, 8, 15), residuals are autocorrelated (7, 8), and model parameters are highly correlated, suggesting an overparameterized model (3).

Received September 21, 1995.
Accepted June 5, 1996.
[1]Current address: 398 Pleasant Hill Drive, Winona, MN 55987.
[2]Department of Statistics.
[3]Department of Agricultural Economics.
[4]Department of Dairy Science.

The multiphasic lactation curve (8) appears to overcome some of the problems inherent to both IG and IP. Initial milk production is not forced to 0, and, if the diphasic or triphasic variant of the model is used, residuals are not autocorrelated with one another. Both the diphasic and triphasic variants of the model have been shown to be the optimal version of the multiphasic model (2, 6, 8). The primary limitation of the multiphasic lactation curve is failure of the model to satisfy the convergence criteria when unadjusted data are analyzed.

Two additional time series, season of calving and season of the year in which each observation occurred, also contribute to the variance observed for daily milk production. Effects of season of calving might uniformly increase milk production throughout an entire lactation (9, 23) or also might influence the shape of the lactation curve (11). Seasonal effects, such as first-crop haylage and heat stress also might influence the shape of the lactation curve (23). Wood (23), but not Keown et al. (11), accounted for the variation in daily milk production that was associated with season of observation. Because season of calving and season of observation may be interdependent, unbalanced designs may not properly estimate the effects of season of calving and season of observation.

Split-plot models permit the partitioning of variance found in experimental observations for within-cow variance (DIM, season of observation) and among-cow variance [treatment, predicted producing ability, season of calving, herd, lactation number, management and nutrition effect, and body condition score at calving (**BCSC**)]. Proper allocation of experimental variance for the portions within and among cows improves the accuracy of estimation of whole-plot effects. Improper assignment of these variance components may result in overestimation or underestimation of whole-plot error and, thus, inaccurate hypothesis testing. (Whole-plot error is used for both F and t testing.) Additionally, improper partitioning of these variance components may result in nonrandomly distributed error terms, invalidating all hypothesis testing.

The objective of this research was to evaluate the use of nonlinear equations to account for DIM effects prior to analysis using general linear models. Improved estimation of treatment responses may result from decreases in standard errors of treatment means, the elimination of biases associated with considering continuous nonlinear effects as discrete intervals, or both. Nonlinear models for DIM effects may also help fulfill most of the statistical assumptions about error distribution that are inherent to hypothesis testing of parameter estimates.

MATERIALS AND METHODS

Data

Five commercial Holstein herds were selected based on rolling herd averages for milk production >9000 kg/yr per cow. All herds were milked three times daily. Each herd was scheduled to be sampled every 14 d by Wisconsin DHIA. Actual test intervals varied between 12 and 31 d; the mean interval was 15.6 d. Cows calving between July 1, 1989 and March 15, 1990 were on trial for the first 180 to 200 DIM (n = 443). Only 383 (86.5%) cows had ≥7 observations for milk production and composition. Four percent FCM was utilized as the production variable of choice because 4% FCM more closely reflects energy needs for milk production than does actual milk (19).

These data were obtained from a field trial designed to evaluate the effects of dietary calcium salts of long-chain palm oil fatty acids on production, health, and reproductive parameters (16). Cows were assigned randomly to control and treatment populations based on date of calving. Calcium salts of long-chain palm oil fatty acids were added to the control ration at the rate of 0.45 kg/d per cow for the first 180 to 200 DIM and were fed to 220 of the 443 cows on trial. Control rations contained 3.7 to 4.8% supplemental fat.

Lactation Curve Models

Models were fit to both individuals and each LHG. There were 15 LHG: five herds and three lactation numbers (first, second, and third or greater). The following three lactation curve models were compared.

IG (21):

$$y_t = At^b e^{-ct},$$

IP (14):

$$y_t = t(\beta_0 + \beta_1 t + \beta_2 t^2)^{-1},$$

and the multiphasic lactation curve model (8):

$$y_t = \sum_{i=1}^{n} a_i b_i [1 - \tanh^2 (b_i (t - c_i))]$$

where y_t = expected milk production on DIM t; A, b, c, β_0, β_i, β_2, a_i, b_i, and c_i are curve parameters; and n = number of phases in the multiphasic lactation curve.

Three variants of the multiphasic curve were compared. Model variants were diphasic (**M2**; n = 2), monophasic (**M1**; n = 1), and monophasic with an additive constant (**MC**). An additive constant was included in MC to allow increased model flexibility in selecting the appropriate segment of the hyperbolic tangent curve. The triphasic form of the multiphasic model was not evaluated because the 10 observations of 4% FCM production that were needed to fit the model were greater than the number available in this data file.

Statistical Analysis

The analysis was completed in three distinct phases. Lactation curve models were fitted by the Marquardt method using PROC NLIN of SAS (18). Starting grids were specified such that all solutions fell within the outer limits of the search grid. Models were not transformed to their linear approximations prior to iteration. The original models equally weighted all observations; however, the logarithmic transformation used to determine the linear approximation reduced the weight placed on high milk production (15). Individual lactation curves were estimated only for individuals with ≥7 DHIA sample days (383 of 443 cows in the data). All five models for lactation curves were also fit for pooled data for milk production within each of the 15 LHG.

Selection of the optimal model for lactation curves was based on error sums of squares, error range, variance of the errors, and the mean absolute value of errors. The Durbin-Watson test was also estimated; however, the usefulness of this statistic is severely limited by its lack of sensitivity when testing for autocorrelation for models based on <15 observations (4). Model coefficient of multiple determination is not a good criterion for evaluating goodness of fit for nonlinear models because residuals may be autocorrelated (8).

Using the residuals from the fitting of the optimal lactation curve to each LHG, lactation curves for split-plot models (**LCSP**) were developed. Results from the LCSP model were then compared with conventional split-plot analysis of the unadjusted data using DIM subclasses within the split plot to account for DIM effects (**SCSP**; subclass split plot).

Under SCSP, intervals of 5, 10, 15, and 20 d were analyzed to determine the effect of length of DIM interval on performance. Unless otherwise noted, the SCSP was based on 10-d intervals for DIM. The 10-d interval was selected as a compromise to prevent multiple observations of the same cow from falling within the same subclass and yet to maintain the highest possible number of observations within the subclass. The whole-plot error term was cow as a random effect.

Type III sums of squares (18) were used to determine significance. Probability values <10% were considered to be significant for whole-plot effects, and the significance level for split-plot effects was $P < 0.05$. General linear models were developed separately for 4% FCM production, milk production, and fat percentage. Identical models were derived from both forward selection and backward elimination stepwise procedures. All main effects and two-way interactions were considered when the final general linear models were developed for LCSP and SCSP. If a two-way interaction was significant for the final model, all three-way interactions involving the two-way interaction were also evaluated. No three-way interactions were significant.

RESULTS AND DISCUSSION

The discussion is divided into four sections. In the first section, the selection of the optimal model for the lactation curve is described. The impact of using lactation curves (LCSP), compared with using discrete intervals for DIM effects (SCSP), on the selection of significant main effects and interactions is considered in the second (whole-plot model) and third (split-plot model) sections. Finally, selection of the optimal model is discussed.

Lactation Curve Model

When fit to individuals, comparisons among the five models evaluated were similar, regardless of the error measurement examined. Only the mean absolute value of the errors and the variance of the error term are reported in Table 1. Ranking of the models, from best to worst, was M2, MC, IG, IP, and M1. The diphasic lactation curve always outperformed IG, IP, and M1 and outperformed MC in 94 to 97% of the lactation curves calculated, depending upon the error measurement used for comparison. The monophasic model with an additive constant always outperformed M1 and improved fit for 88.5% of individual lactation curves compared with either IG or IP; the exception was that MC always resulted in a better model than IP when the error sum of squares was used to determine fit.

The MC reduced the magnitude of the mean absolute value of errors by 7.3 to 10.8% over that of IG, IP, and M1. The diphasic curve resulted in an additional decrease of 8.2% in the mean absolute value of errors over MC. Curve fit, measured either by error sum of squares or variance of the error, improved an average

TABLE 1. Goodness of fit measurements for all lactation curve models fit to 4% FCM for each cow with at least 7 sample days.

	Lactation curve model				
	IG[1]	IP	M1	MC	M2
Mean absolute error, kg					
$\bar{X}$[2]	2.0[a]	2.04	2.10	1.87	1.73
SD[3]	0.41	0.44	0.39	0.35	0.40
Error variance, kg^2					
$\bar{X}$	8.6	8.8	9.2	7.4	6.7
SD	3.2	3.4	3.2	2.7	2.9
Direct comparison of individual models[4]					
M2 >[5]	100.0	100.0	100.0	94.0	
MC >	88.5	88.5	100.0		
M1 >	30.6	32.6			
IP >	52.2				

[1]IG = Incomplete gamma model, IP = inverse polynomial model, M1 = monophasic model, MC = M1 plus constant, and M2 = diphasic model.

[2]Mean of error measurement for each model across all individuals.

[3]Variance of error measurement for each model across all individuals.

[4]Based on mean absolute error.

[5]Percentage of 383 lactation curves in which the model in the left column reduced mean absolute error over the model in the column containing the percentage.

of 13.3 to 19.4% for MC over IG, IP, and M1; M2 reduced variance of the error by an additional 10.9% over MC.

The M2 was not always superior to MC. Regardless of the error measurement examined, MC performed more consistently than M2. The variance and maximum of each error measurement was smaller for MC than for M2. Thus, although MC resulted in a higher mean for each error measurement, this mean was more uniform across all cows compared. Conversely, M2 fit a number of cows very well without uniform improvement in fit across the entire population. At least a portion of these differences resulted from the ability of M2 to fit lactation curves with two peaks. Because treatment was initiated by calving, only a single peak was expected for the lactation curve, and curves with two peaks were considered to be aberrant. Curves with multiple peaks most often occurred from cows that had one or two occurrences of depressed FCM production, relative to preceding and subsequent FCM production, after peak production. Often those cows were clinically ill during the affected time period. Those cows that appeared to be healthy might have been afflicted with a subclinical health problem or might have had aberrant test day production. Therefore, we concluded that a single peaked curve was the desired model for lactation curves. Additionally, M2 failed to meet convergence criteria for 54% of the curves fitted to individual cows. Curve smoothing techniques might increase the probability of model convergence, but might also introduce bias when populations contain only a few observations (5).

Both IG and IP overpredicted milk production prior to peak (70 DIM) and after 160 DIM and underpredicted midlactation milk production (Figure 1). Although only the first 200 DIM were analyzed, these results were parallel to those reported in studies (1, 8, 15) evaluating residual patterns of these two models over an entire 305-d lactation. The residual pattern for M1 was opposite that of IG and IP, resulting in underprediction of milk production during early and late lactation and overprediction of midlactation milk production (Figure 2). This observation was in direct contradiction to the determination of Grossman and Koops (8) that M1 overpredicted early lactation, underpredicted midlactation, and accurately predicted late lactation. These differences may have resulted from differences in either the length or the persistency of the lactation curve. Because milk production >200 DIM declines linearly, the model would have been shifted further to the right on the hyperbolic tangent curve, resulting in flatter peaks (underprediction) and higher initial production (overproduction). Neither MC nor M2 displayed any apparent pattern for residual measurements over time. First-order autocorrelation of residuals within a single lactation was not detected for any model by the Durbin-Watson test; however, this statistic had a large inconclusive region when <15 observations were available (4). Individual cows had a mean of only 10

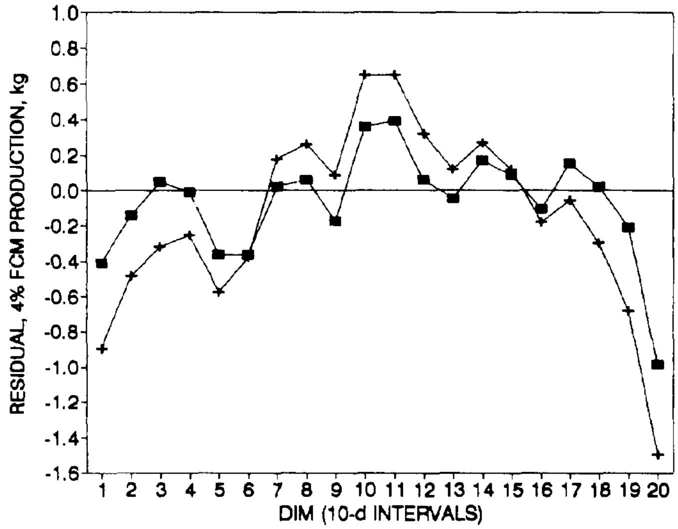

Figure 1. Mean residual plotted by DIM for the incomplete gamma (■) and inverse polynomial (+) models for lactation curves.

to 11 observations (range, 7 to 14) for 4% FCM production.

Residual plots for all five models for lactation curves suggested that a higher order autocorrelation process might have been present. Unfortunately, the asymptotic tests used to evaluate higher order autocorrelation processes required ≥30 observations before meaningful results could be calculated.

When data were grouped by LHG prior to fitting each model, differences in performance of IG, IP, M1, MC, and M2 were minimal. Each model accounted for 0 to 22% (mean of 11%) of the variability in the unadjusted data within each LHG. Lactation curves were expected to account for only a small portion of the variability within each LHG because the data were not adjusted for other variables prior to grouping. Based on the mean absolute value of errors, MC outperformed both IG and IP for 12 of the 15 LHG and always outperformed M1. The diphasic curve performed inconsistently when lactation curves were analyzed for each LHG. For three of the LHG, M2 led to no improvement over linear regression. The diphasic curve did not meet convergence criteria for any LHG, and MC failed to converge for 4 of the 15 LHG.

The three variants of the multiphasic model were also evaluated by F test to determine the significance of the added parameters within each LHG (17). The additive constant in MC contributed significant information ($P < 0.001$) for 14 of the 15 LHG. The additive constant was significant ($P < 0.05$) for the remaining LHG. Compared with M1, the second phase of M2 also contributed information ($P < 0.001$) for 13 LHG. Significance levels were $P < 0.05$ and $P < 0.10$ for the remaining two groups. The second phase of M2, compared with the additive constant of MC, was significant ($P < 0.05$) for 12 of the 15 LHG.

Based on these results, MC was regarded as the best model. Because the original design of this trial was to detect responses in milk production to supplementation of 0.45 kg of calcium salts of long-chain palm oil fatty acids/d per cow, the possible introduction of bias between control and treatment populations, secondary to fitting the data to MC, becomes important. There was no suggestion that MC resulted in any bias between treatment and control populations (Table 2). The model may have fit the treatment population slightly better, but the magnitude of this difference was small.

Combining Lactation Curves with General Linear Models

Whole plot. Whole-plot independent effects accounted for the variance among cows in the production of milk and milk components. The significance for each independent effect and interaction for 4% FCM models (LCSP and SCSP) is in Table 3. Continuous fixed effects included predicted producing ability (PPA) for 4% FCM (0.4 PPA for milk production + 15 PPA for fat production), body weight at calving, and BCSC. Effects of BCSC and body weight at calving on production data appeared to be linear

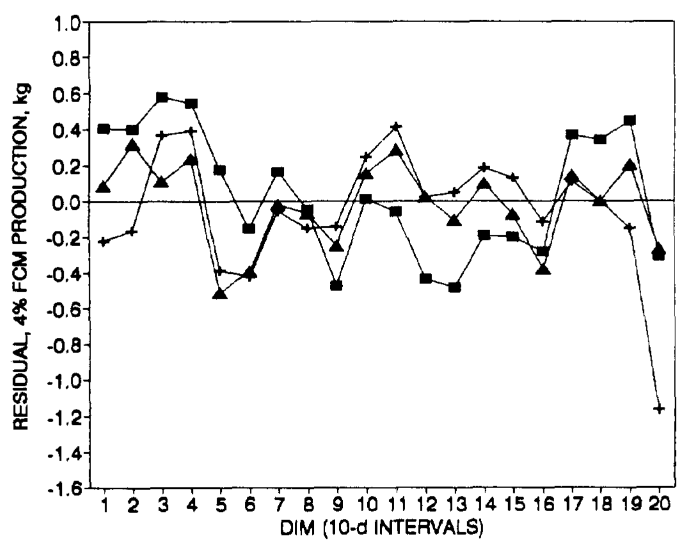

Figure 2. Mean residual plotted by DIM for the three variants of the multiphasic lactation curve: monophasic (■), monophasic with an additive constant (+), and diphasic (▲).

TABLE 3. Comparison of models for 4% FCM whole plots resulting from the lactation curve method of analysis (LCSP) versus conventional analysis (SCSP).

Effect	Lactation 1		Lactation ≥2	
	LCSP	SCSP	LCSP	SCSP
	P[1]			
Predicted producing ability (PPA)	<0.01	<0.01	<0.001	<0.0001
Season of calving (SC)	0.07	0.08	<0.0001	<0.0001
Herd	. . .[2]	<0.0001	. . .	<0.0001
Management and nutrition (MNE)	0.09	. . .	0.63[3]	. . .
Herd within MNE	<0.001	. . .	. . .	. . .
Fresh BW	<0.0001	<0.0001	. . .	. . .
Body condition score at calving (BCSC)	. . .	. . .	<0.0001	<0.0001
Treatment (T)	0.08	0.07	<0.01	0.02
PPA × T	. . .	. . .	0.10	. . .
PPA × BCSC	. . .	. . .	0.09	. . .
MNE × T	. . .	. . .	<0.01	. . .
MNE × SC	. . .	. . .	0.03	. . .
MNE × BCSC	0.09	. . .	. . .	. . .
Herd × SC	. . .	. . .	. . .	0.08
Lactation number	. . .	. . .	. . .	0.37[3]
Lactation number × BCSC	. . .	. . .	. . .	0.06

[1]For Type III sums of squares.

[2]Not significant ($P < 0.10$) for fully reduced model.

[3]Main effect included in model because interactions including main effect were significant ($P < 0.10$).

(not shown). Discrete fixed effects included season of calving and season of observation (2-mo intervals), herd, lactation number (first, second, and third or greater), treatment, management and nutrition effect (blocking variable to separate herds based upon magnitude of treatment response), and DIM (10-d intervals). Predicted producing ability, season of calving, BCSC, and body weight at calving had similar coefficients and significance, regardless of analysis technique.

For multiparous cows, the process of fitting lactation curves to each LHG accounted for herd, lactation number, and all interactions of these fixed effects with other whole-plot variables. Therefore, herd, lactation number, and all interactions with herd and lactation number were eliminated from the whole-plot model. Because the generation of each lactation curve removed one degree of freedom from the whole-plot error term, the degrees of freedom associated with herd, lactation number, and interactions of herd with lactation number were not available to be added to error degrees of freedom. Herd, lactation number, and their interactions with other variables, particularly treatment, were significant in the general linear model with SCSP.

For multiparous cows, the use of lactation curves prior to split-plot analysis separated the five herds into two populations, based on the observed response to calcium salts of long-chain palm oil fatty acids. Four of the five herds did not respond to calcium salts of long-chain palm oil fatty acids (low response), but, in one herd, 4% FCM production increased by 2.88 kg/d per cow when calcium salts of long-chain palm oil fatty acids were fed (high response). This factor was accounted for by the inclusion of a blocking factor, effect of management and nutrition, in the whole-plot model for multiparous cows. Because only one herd (experimental unit) was in the high response group, factors contributing to differences in treatment response among herds could not be determined. Possible causes of the discrepancy in estimates of treatment response among herds are discussed by Scott et al. (16). The single herd in the high response block for effect of management and nutrition could not be differentiated from the other four herds when SCSP was utilized.

TABLE 2. Comparison of error measurements, peak 4% FCM production, and days to peak production for the monophasic model with an additive constant for control cows versus cows fed calcium salts of long-chain fatty acids (Ca-LCFA).

	Control	Ca-LCFA
Mean absolute error, kg of 4% FCM	1.88	1.86
Error variance, $(kg)^2$ of 4% FCM	7.43	7.31
Peak of 4% FCM production		
kg	41.9	42.5
d	70.8	75.4

The advantage of including blocking for effect of management and nutrition with LCSP must be qualified. Because only one herd was included in the high response block for effect of management and nutrition, the ability of LCSP to detect the effect of management practices may simply have resulted from one herd being markedly different from the remaining four herds. Because only one experimental unit was included in one level of the predictor variable for effect of management and nutrition, the variable might not represent a true cause and effect relationship.

Two significant interactions, predicted producing ability with treatment and BCSC with treatment, were detected by LCSP. Milk production and fat percentage did not demonstrate a significant interaction of predicted producing ability with treatment ($P <$ 0.25 and $P <$ 0.77, respectively), but 4% FCM ($P <$ 0.10) and fat production did ($P <$ 0.10). Milk production and fat percentage exhibited significant interactions of BCSC with treatment ($P <$ 0.07 and $P <$ 0.02, respectively). Because decreased BCSC was associated with increased milk production and decreased fat percentage when cows were fed calcium salts of long-chain palm oil fatty acids, this interaction was not significant when 4% FCM was analyzed. These interactions were not significant under SCSP. The difference between LCSP and SCSP might have resulted from either confounding or colinearity of variables under SCSP.

Few differences were detected between LCSP and SCSP when data from primiparous cows were analyzed. Because an interaction of management and nutrition with herd was still present under LCSP, the inclusion of effects of management and nutrition and of herd within management and nutrition in the LCSP model was nearly the same as the use of herd as a blocking effect with SCSP. The only difference in models for 4% FCM that were generated by the two systems was the ability of LCSP to detect an interaction of management and nutrition effect with BCSC that was not detected with SCSP ($P <$ 0.09). Because calcium salts of long-chain palm oil fatty acids were not fed to these cows prior to calving, this interaction reflects a difference in BCSC between the high response herd for management and nutrition effect and the remaining four herds rather than a treatment effect.

Combining Lactation Curves with General Linear Models

Split plot. With general linear models for analysis of production data, the split-plot portion of the model accounts for variance within cows, such as DIM and season of observation effects. When 4% FCM was analyzed for multiparous cows, LCSP successfully accounted for all DIM effects and interactions prior to fitting the general linear model. Season of observation, and its interactions with other independent variables, was the only parameter to retain significance under LCSP. Interactions of season of observation with herd and lactation number were not eliminated by LCSP. These interactions were not expected to be accounted for by lactation curves because the models that were used accounted for effects of DIM and not season of observation. These effects were assumed to be randomly distributed across treatments within each LHG and to have little impact on estimation of treatment response.

When milk production or fat percentage was analyzed with LCSP, LCSP did not completely account for all DIM effects within the split-plot model. Two factors might have influenced the differences in LCSP performance between 4% FCM and milk production and fat percentage. First, the lactation curve model used was developed specifically for 4% FCM production and might not have been the optimal model for lactation curves for either milk production or fat percentage. Second, lactation curves for 4% FCM production were more consistent among cows in each LHG than were lactation curves for either milk or fat percentage. Production of 4% FCM converged because of the negative correlation between milk production and fat percentage (12). Production of 4% FCM was more similar for two individual cows than either milk production or fat percentage because cows with high milk production tended to have lower fat percentage than cows with lower production. Lactation curves for 4% FCM were typically flatter than lactation curves for milk because the lowest fat percentage frequently occurred near peak milk production.

Selection of the Optimum Model

Differences in the ability of LCSP to fit data for primiparous and multiparous cows appeared to be related to differences in the lactation curves of these two groups. Although a lactation curve fit to each LHG accounted for one-third of the split-plot variance for multiparous cows, the lactation curve only slightly outperformed a straight line for primiparous cows (Table 4), either because of the lack of well-defined peaks (high persistency) or because of the high degree of variability in time to peak production for primiparous cows. Multiparous cows peaked from 40 to 100 DIM, but several primiparous cows had not yet peaked by the end of the study (200 DIM). Primiparous cows in one herd peaked at 145 DIM, on average. Patterns were similar for milk production: 27% (primiparous cows) and 41% (multiparous cows) of split-plot variance was explained by lactation curves.

For 4% FCM and milk, lactation curves accounted for 34% (primiparous cows) and 40 to 45% (multiparous cows) of whole-plot variance (Table 4). Whole-plot variance was accounted for by the elimination of differences in mean production for each LHG and by the removal of DIM effects unique to each LHG. Because a unique curve was fit to each LHG, most of the interactions of herd and lactation number with other variables were accounted for also.

Lactation curves accounted for similar amounts of split-plot variance for both primiparous and multiparous cows (41 and 36%, respectively) when fit to fat percentage. However, only a small portion of whole-plot variance was explained by lactation curves (12 and 10%, respectively), suggesting that LCSP did little to increase the sensitivity of analysis to detect treatment differences in fat percentage. This result might have been related to smaller differences in fat percentage between LHG than 4% FCM and milk production. Whole-plot variance accounted for only 51% (primiparous cows) and 39% (multiparous cows) of the total variance in fat percentage, but whole-plot variance accounted for 60 to 67% of the total variance in 4% FCM and milk production. Because the model for the lactation curve was developed based upon 4% FCM, another possible explanation for these differences was that the curve utilized was not an adequate model for fat percentage.

Predicted treatment responses for both 4% FCM and milk production differed substantially for multiparous cows between LCSP and SCSP. Conventional split-plot analysis using DIM subclasses overestimated 4% FCM treatment response by 50% (1.00 and 0.66 kg/d per cow for SCSP and LCSP, respectively) and milk treatment response by 110% (0.86 and 0.41 kg/d per cow, respectively) compared with those for LCSP. At least a portion of the differences in the estimation of treatment response was a function of two distinct, nonrandom distributions found within the data.

Based on individual lactation curves, peak 4% FCM production increased 0.6 kg/d for the treatment population (Table 2). This value was similar to the treatment response of 0.66 kg of 4% FCM/d estimated by LCSP (Table 4). No differences in persistency were detected between the control and treatment populations. Differences in mean of 200-d cumulative amount of FCM production, estimated by fitting MC

TABLE 4. Comparison of variance explained by general linear models for 4% FCM resulting from lactation curve method of analysis (LCSP) versus conventional analysis (SCSP).

	Lactation 1		Lactation ≥2	
	LCSP	SCSP	LCSP	SCSP
Estimate of treatment difference, kg FCM/d[1]	1.01	0.96	0.66[2]	1.00
P[3]	0.06	0.06	0.001	0.02
Between cow variance (BCV) explained by				
	(%)			
Lactation curve	34.1	. . .[4]	44.3	. . .
General linear model	16.2	48.5	35.2	79.5
Error	49.7	51.5	20.5	20.5
BCV MSE[5]	106.1	100.1	85.7	83.5
BCV adjusted R^2, %	41.0	44.3	76.8	76.8
Total variance explained by BAV	67.0	67.0	60.5	60.5
Within-cow variance (WCV) explained by				
Lactation curve	6.2	. . .	33.0	. . .
General linear model	15.5	27.8	10.7	43.2
Error	78.3	72.2	56.3	56.8
WCV MSE	7.9	7.7	15.1	15.5
WCV adjusted R^2, %	18.7	20.7	42.1	40.6

[1]At population means for all other independent effects.

[2]Weighted equally for all five herds to facilitate comparison with SCSP.

[3]Treatment difference.

[4]Not applicable.

[5]Mean square error.

to each individual, between treatment and control populations were 290 and 109 kg for primiparous and multiparous cows, respectively. These results compared favorably with the estimated daily treatment response from LCSP multiplied by 200 d (202 and 132 kg of 4% FCM for primiparous and multiparous cows, respectively). Based on daily treatment responses as estimated by SCSP, differences in 200 d cumulative amount of 4% FCM production were 192 and 200 kg, respectively.

Observations from the treatment group were clustered near peak milk production (45 to 104 DIM), and control observations were clustered during early (<15 DIM) and late (>174 DIM) lactation. This unbalanced distribution caused SCSP to estimate the effect of DIM improperly, resulting in inflated estimates of treatment effects and biased estimates of production.

The distribution of observations within each DIM subclass further affected the estimation of treatment effect. Prior to peak production (75 to 84 DIM), cows in the treatment population were observed, on average, 0.1 d later in lactation than were control cows. Conversely, after 114 DIM, treated cows were observed on average 0.4 d earlier than were control cows. Thus, the treatment group had the benefit, on average, of an extra 0.1 d to increase milk production prior to peak production; also, these cows were observed 0.4 d earlier after peak production when daily production was declining. During midlactation, 4% FCM production declined 0.1 to 0.2 kg/d.

The use of discrete intervals to analyze nonlinear continuous effects might have also introduced other biases into the estimation of treatment response. Not all cows were observed within each DIM interval because of ongoing recruitment, illness, culling, uneven sampling intervals, and other factors beyond experimental control, which might have contributed to inaccurate estimation of treatment effects because different populations were used to calculate the effect of DIM within each interval.

Similarities in the magnitude of split-plot error sums of squares for both LCSP and SCSP were related to the length of the interval selected for each DIM subclass (10 d).

CONCLUSIONS

For fitting 4% FCM production over the first 200 DIM, the best model for lactation curves was MC. The MC consistently outperformed IG, IP, and M1. The M2 generally outperformed MC; however, MC performed more uniformly across the entire population. Because lactation curves are not symmetrical near peak production, replacement of DIM with the fourth root of DIM might be beneficial for derivatives of the model of Grossman and Koop (20).

The fitting of a unique lactation curve to each LHG to determine the effect of DIM prior to conventional split-plot analysis reduced biases introduced by nonrandom distribution of production data across DIM and other predictor variables. When compared with the classification of DIM effects into a number of discrete intervals, lactation curves fit production data in a manner more representative of actual patterns of lactation production. The elimination of bias should result in more accurate estimation of treatment effects.

For multiparous cows, the fitting of lactation curves prior to analysis increased the ability of general linear models to detect interactions of treatment response with independent variables by repartitioning the variance from herd and lactation number. By accounting for the effects of herd and lactation number and their interactions with other effects in the model, the use of a unique lactation curve for each LHG increased the ability of general linear models to detect important interactions.

Improvement of the general linear model for primiparous cows was minimal when lactation curves were fitted to analysis, which might have resulted from the more persistent lactation curves for production during first lactation.

The development of models capable of analyzing effects of both DIM and season of observation through the use of continuous nonlinear equations should help reduce biases that are secondary to chance imbalances in data distribution across predictor variables. When combined with lactation curves, such equations should more accurately predict treatment effects from production data from field trials. The use of such equations might also permit the detection of smaller treatment differences or the use of fewer observations for the detection of significant treatment effects.

REFERENCES

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

2 DeBoer, J. A., J. I. Weller, T. A. Gipson, and M. Grossman. 1989. Multiphasic analysis of milk and fat yield curves in Israeli Holsteins. J. Dairy Sci. 72:2143.

3 Dhanoa, M. S. 1981. A note on an alternative form of the lactation model of Wood. Anim. Prod. 32:349.

4 Durbin, J. 1970. An alternative to the bounds test for testing for serial correlation in least-squares regression. Econometrica 38:422.

5 Gipson, T. A., R. L. Fernando, and M. Grossman. 1990. Effects of smoothing data on estimation of parameters for multiphasic lactation curves of dairy goats. Livest. Prod. Sci. 24:205.
6 Gipson, T. A., and M. Grossman. 1989. Diphasic analysis of lactation curves in dairy goats. J. Dairy Sci. 72:1035.
7 Goodall, E. A., and D. Sprevak. 1984. A note on a stochastic model to describe the milk yield of a dairy cow. Anim. Prod. 38: 133.
8 Grossman, M., and W. J. Koops. 1988. Multiphasic analysis of lactation curves in dairy cattle. J. Dairy Sci. 71:1598.
9 Grossman, M., A. L. Kuck, and H. W. Norton. 1986. Lactation curves of purebred and crossbred dairy cattle. J. Dairy Sci. 69: 195.
10 Hyde, G. M. 1975. Milk yield as a factor in dairy herd management by exception. Ph.D. Diss., Univ. Illinois, Urbana-Champaign.
11 Keown, J. F., J. W. Everett, N. B. Empet, and L. H. Wadell. 1986. Lactation curves. J. Dairy Sci. 69:769.
12 Larson, B. L. 1985. Biosynthesis and cellular secretion of milk. Chapter 4 *in* Lactation, A Comprehensive Treatise. B. L. Larson, ed. Iowa State Univ., Ames.
13 Milliken, G. A., and D. E. Johnson. 1984. Analysis of messy data. Volume 1: Designed Experiments. Van Nostrand Reinhold, New York, NY.
14 Nelder, J. A. 1966. Inverse polynomials, a useful group of multifactor response functions. Biometrics 22:128.
15 Rowlands, G. J., S. Lucey, and A. M. Russell. 1982. A comparison of different models of the lactation curve in dairy cattle. Anim. Prod. 35:135.
16 Scott, T. A., R. D. Shaver, L. Zepeda, B. Yandell, and T. R. Smith. 1995. Effects of rumen-inert fat on lactation, reproduction, and health in high producing Holstein herds. J. Dairy Sci. 78:2435.
17 Snedecor, G. W., and W. G. Cochran. 1980. Statistical Methods. 7th ed. Iowa State Univ. Press, Ames.
18 SAS User's Guide: Statistics, Version 6.03 Edition. 1988. SAS Inst. Inc., Cary, NC.
19 Tyrrell, H. F., and J. T. Reid. 1965. Prediction of the energy value of cows milk. J. Dairy Sci. 48:1215.
20 Weigel, K. A., B. A. Craig, T. R. Bidwell, and D. M. Bates. 1992. Modification of diphasic lactation curve models under bovine somatotropin administration. J. Dairy Sci. 75:580.
21 Wood, P.D.P. 1967. Algebraic model of the lactation curve in cattle. Nature 216:164.
22 Wood, P.D.P. 1969. Factors affecting the shape of the lactation curve in cattle. Anim. Prod. 11:307.
23 Wood, P.D.P. 1972. A note on seasonal fluctuations in milk production. Anim. Prod. 15:89.