A MULTIPLE TRAIT APPROACH TO MODELLING THE LACTATION CURVE

D. Trus and L.G. Buttazzoni
Associazione Italiana Allevatori
Via G. Tomassetti 9
00161 Roma, Italia

SUMMARY

Description of the lactation curve as a series of correlated traits expands the usefulness of lactation data for genetic evaluation and management purposes. A fixed effects model was applied and unobserved records predicted from estimated parameters and functions of the observed records. Performance of the method is shown for first lactation Italian Simmental cows.

INTRODUCTION

Projection of unobserved lactation records and estimation of 305-day milk production from test day yields are currently considered two separate tasks. However, in both cases it is desirable to make maximum use of the predictability of the lactation curve and to minimize the error of prediction from a sample of daily records. Many mathematical models have been proposed to model the lactation curve (Aleandri et al., 1988) and factors which influence the magnitude and shape of the lactation curve have been suggested by many authors (e.g., Danell, 1981; Kachman and Everett, 1989). A multiple trait approach is proposed which will minimize errors of estimation and combine into a single procedure the projection of unobserved records and estimation of 305-day milk traits.

MATERIAL AND METHODS

General Method

Collection schemes for milk recording are primarily based on equally spaced intervals of 14, 21, 30 or 42 days. The A4 (30 days) method, considered the standard method by the ICRPMA, and the A6 (42 days) method are currently used in Italy. If the lactation curve is divided into fixed days-in-milk (DIM) intervals corresponding to the sampling method these intervals can be treated as correlated traits.

Within a DIM interval the curve can be modelled based on field data because testing of individual cows is often well distributed within intervals. A fixed, multiple trait generalized least squares model is applied on an iterative basis beginning with a prior, positive definite R matrix. Solutions to the model are used to get $E(y)=X\beta^{o}$, and missing(m) residuals are predicted from observed(p) residuals and the estimated residual variance-covariance matrix.

Using notation of Henderson (1984), let $Var\begin{bmatrix} e_p \\ e_m \end{bmatrix} = \begin{bmatrix} R_{pp} & R_{pm} \\ R'_{pm} & R_{pp} \end{bmatrix}$

Then $\hat{e}_m = R'_{pm}R_{pp}^{-1}\hat{e}_p$ where, $\hat{e}_p=y-E(y)$ a vector of observed residuals, R_{pp} is a submatrix of R corresponding to intervals with observations, and R'_{pm} is a submatrix of R with rows corresponding to DIM intervals without observations and columns corresponding to intervals with observations. Therefore, $\hat{y}_i = E(y_i) + \hat{e}$ can be used to predict any i^{th} daily observation. A model must be chosen such that the desired function of the solutions is estimable.

The model

The model used in this analysis for each DIM interval was,

$$y_{ijk} = P_i + G_j + M_k + b_{ij1}DPREG + b_{ij2}DPREG^2 + b_{j3}DIM + b_{j4}DIM^2 + b_5AGE + b_6AGE^2 + b_7YR + e_{ijk}$$

where:

- y_{ijk} is the ijk^{th} daily milk observation for a breed-lactation;
- P_i is the fixed effect of the i^{th} pregnancy status with i=1 for not pregnant and i=2 for pregnant;
- G_j is the fixed effect of the j^{th} production group with j=1,2,3
- M_k is a fixed month of sample effect with k=1,...,12;
- b_{ij1}, b_{ij2} are coefficients for linear and quadratic regression of y_{ijk} on days pregnant (DPREG) within the ij^{th} pregnancy status x group subclass, only when i=2;
- b_{j3}, b_{j4} are coefficients for linear and quadratic regression of y_{ijk} on DIM within the j^{th} production group;
- b_5, b_6 are coefficients for linear and quadratic regression of y_{ijk} on age of cow at calving (AGE) expressed in months;
- b_7 is a linear coefficient for regression of y_{ijk} on year of birth of the cow (YR); and
- e_{ijk} is a random residual error with assumed homogeneous (co)variances within DIM intervals and heterogeneous (co)variances across.

In matrix notation the equations to be solved are,

$$X'R^{*-1}X\beta = X'R^{*-1}y$$

where R^* is R with rows and columns deleted corresponding to DIM intervals without an observation for a given cow. A new R matrix is calculated at each round of iteration from cows having observations in every interval.

The Data

Data for this analysis were first lactation milk yield records collected from 36 685 Simmental cows tested between 1979 and 1988. Edits were made for valid dates, reasonableness of the data, and amount of information available. For parameter estimation, only completed lactations were accepted and date of the next calving were required to be known to calculate days pregnant. Edit requirements were relaxed for the projections and estimation of 305-day yield. Days-in-milk intervals were set to 43 days for the A6 method for a total of 7 intervals over the 305 days. There were 13 483 cows with observations in every interval. Cows were allocated to production groups with divisions determined by the phenotypic mean for first sample yield adjusted to a 30 day basis (16.41 kg), plus or minus 0.75 of the standard deviation (3.525 kg).

RESULTS

The correlations among the residuals in each DIM interval are shown in table 1. The greatest contribution to the prediction of missing residuals was always made by residuals in adjacent intervals. Highest correlations between adjacent residuals occurred toward the centre of the curve, while residuals at either end appeared to be more independent, for this model.

Table 1. Residual variances (on the diagonal), covariances (above diagonal) and correlations (below diagonal) for 1st lactation Simmental milk yield among 7 equally spaced 43-day intervals.

DIM Interval	DIM Interval 1	2	3	4	5	6	7
1	3.079	2.517	2.186	1.952	1.782	1.643	1.460
2	0.557	6.619	5.006	4.545	4.177	3.863	3.389
3	0.475	0.742	6.875	5.339	4.934	4.585	4.073
4	0.421	0.669	0.771	6.968	5.588	5.130	4.524
5	0.379	0.606	0.702	0.790	7.187	5.849	5.116
6	0.340	0.546	0.636	0.706	0.793	7.572	6.001
7	0.287	0.454	0.535	0.591	0.658	0.752	8.419

For a cow with observations in each interval, all residuals are realized rather than predicted. Subsequent estimates of 305-day yield should thus be the most accurate indicators of true yield. Therefore, the performance of the prediction method was tested by comparing estimates from the same cows with different amounts of information missing. Results are presented in table 2.

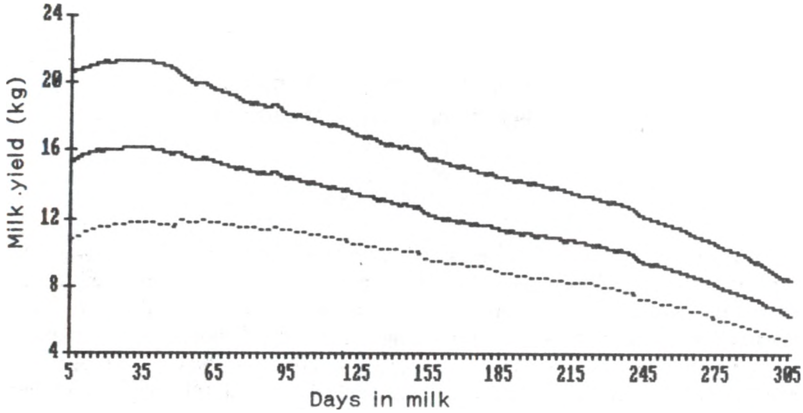
Table 2. Measures of performance of the prediction method comparing milk yield estimates derived with varying amounts of information. (N = 13 483)

Information No.	Code*	Deviation (kg) Mean	Std Dev	Correlations With Case No. 0	1	2	3	4	5	6	7
0	1111111	0.0	0.0	1.000							
1	1111110	-8.3	70.8	0.997	1.00						
2	1111100	-21.4	130.8	0.989	.99	1.00					
3	1111000	-31.3	193.7	0.977	.98	.99	1.00				
4	1110000	-41.7	266.7	0.955	.96	.97	.98	1.00			
5	1100000	-55.0	354.8	0.919	.92	.93	.94	.96	1.00		
6	1011111	0.4	65.7	0.997	.99	.99	.97	.95	.90	1.00	
7	1001100	-21.3	181.1	0.980	.98	.99	.97	.93	.90	.98	1.00

* Information code contains 1 if an observation occurred in an interval or a 0 otherwise. Deviations are with respect to case 0 (mean = 4070.6 kg).

All deviations about the "true" (case no. 0) 305-day yield were normally distributed. However, a minimal downwards bias is indicated by the mean deviation, causing underestimation of total yield for some cows.

Figure 1. Plot of E(y) for daily milk yield (kg) for a low, medium and high production cow born August 30,1982, with first freshening occurring Feb. 1,1985 and second conception at 90 days in milk.



DISCUSSION

A multiple trait approach to extending lactations treats observations within discrete intervals of the lactation curve as separate traits and data are analyzed with a fixed effects model. The estimated residual covariances contain covariances due to temporary environmental effects, some permanent environmental effects and most genetic covariances. Unobserved records at any point on the curve can be predicted from the estimated population parameters and functions of observed records. The sum of the realized and predicted daily yields gives the 305-day yield.

There are many implications for genetic evaluations based on this method of modelling the lactation curve. First, estimates of 305-day production traits should be more accurate, therefore, allowing for more accurate estimates of breeding value. Second, cumulative yield traits can be preadjusted for factors which influence the shape of the curve such as day of conception or known treatments with hormones such as Bovine Somatotropin (BST), which are difficult to account for in a genetic evaluation model. Third, the residuals themselves can be used to select on different parts of the lactation curve.

The model also expands possible uses of the data for management purposes. Projections will be more accurate and solutions can be entered into a spreadsheet to allow alternative case analysis. For example, figure 1 compares the lactation curve of each production group; or the effect of earlier conception on total yield can be studied. Economic projections are also simplified and solutions from the model can be used on farm.

REFERENCES

ALEANDRI, R., BACHELOR, D. and MOCQUOT, J. 1988. Rep. in Prog. of the ICRPMA Wrkg. Grp.: Computation methods of lactation curve. ICRPMA 26th Session.

DANELL, B. 1981. Thesis. Swedish Univ. of Agric. Sci. Uppsala. (Report 51).

HENDERSON, C.R. 1984. Applications of linear models in animal breeding. University of Guelph. Ontario. Canada.

KACHMAN, S.D. and EVERETT, R.W. 1989. J. Anim. Sci. 67(1):60. [Abstract 144].