

## A Multivariate Approach to Modeling Shapes of Individual Lactation Curves in Cattle

N. P. P. Macciotta,<sup>1</sup> D. Vicario,<sup>2</sup> Corrado Di Mauro,<sup>1</sup>  
and A. Cappio-Borlino<sup>1</sup>

<sup>1</sup>Dipartimento di Scienze Zootecniche,  
Università di Sassari, Via De Nicola 9, 07100 Sassari, Italy

<sup>2</sup>Italian Association of Simmental Breeders,  
Via Nievo 19, 33100 Udine, Italy

### ABSTRACT

Milk test-day records of 5728 lactations of Italian Simmental cows were analyzed with multivariate factor analysis in order to extract 2 common factors, whose scores were used as quantitative measures of 2 main features of lactation curve shape—i.e., the increasing rate of yield in the first part of lactation and the rate of decline of milk yield after the lactation peak. The 2 indices, objectively derived from the correlation matrix of original test-day records, showed a high discriminant power in separating lactation curves with different shapes. The weak correlation between the 2 factors (0.11), together with the high correlation of factors and the total 305-d yield (about 0.70), suggests that an increase in lactation yield could be achieved by acting only on one of the 2 factors related to lactation-curve shape, with the other kept constant at a medium or low value. The suitability of the 2 factors as descriptors of lactation patterns has been confirmed by the relationships found between factor scores and the main environmental effects known to affect the shape of the lactation curve, such as parity and season of calving.

**(Key words:** lactation curve, lactation peak, test day, multivariate factor analysis)

**Abbreviation key:** TD = test day.

### INTRODUCTION

The environmental and genetic control of main aspects characterizing lactation-curve shape represents an opportunity and a challenge for the dairy industry because of the great interest in traits that allow increased economic efficiency of the cow through reduced costs (Groen et al., 1997; Bichard, 2002). At present, attention is particularly focused on the cow's ability to

maintain production after peak (lactation persistency), the economic importance of which lies in its relationships to health status, reproduction efficiency, and feeding (Gengler, 1996; Pryce et al., 1997; Swalve, 2000; Bichard, 2002; Jakobsen et al., 2002). Actually, cows with flat curves are less susceptible to metabolic disorders, health and fertility problems, and have more consistent energy requirements throughout lactation, allowing for the use of cheaper feeds (Sölkner and Fuchs, 1987; Dekkers et al., 1998). Moreover, some studies reported a genetic correlation between lactation persistency and disease resistance (Jensen, 2001).

In spite of the great technical importance, there is no consensus on the most suitable measure of persistency and, in general, of main aspects of lactation-curve shape (Gengler, 1996; Jamrozik et al., 1998; Grossman et al., 1999; Rekaya et al., 2001). The modeling of lactation curves by analytical functions of time allows for estimating the increasing rate of production in the first part of lactation, the time at which lactation peak is attained, and the rate of decline of milk yield after the peak, as combinations of function parameters (France and Thornley, 1984). Such an approach is very efficient for averaging lactation curves of homogeneous groups of animals, but its usefulness becomes questionable when individual lactation patterns are fitted (Olori et al., 1999). The alternative methods based on combinations of test-day (TD) records taken at different stages of lactation, such as ratios between cumulated yields or measures of TD variation (Sölkner and Fuchs, 1987; Swalve, 1995), have been essentially aimed at expressing the shape of the lactation pattern by a single measure. However, they failed in characterizing persistency in a unique manner because they are not invariant with respect to the time period chosen (Rekaya et al., 2001). As a consequence, for example, genetic correlations between milk yield and persistency depend largely on what measure is used (Sölkner and Fuchs, 1987; Haile-Mariam et al., 2003). The major drawback of these empirical methods is the arbitrary assignment of relative weights to TD of different stages of lactation. This prob-

Received: June 9, 2003.

Accepted: October 6, 2003.

Corresponding author: N. P. P. Macciotta; e-mail: macciott@uniss.it.

lem remains when EBV estimated by random regression TD models are combined in a genetic index of lactation persistency (Schaeffer and Dekkers, 1994). In fact, when relative weights of TD are estimated from the genetic (co)variance matrix of random regression coefficients (Togashi and Lin, 2003), a priori genetic changes for different lactation stages have to be assumed.

In this work, individual lactation curve shapes of Italian Simmental cows are analyzed by means of factor analysis of TD data. This multivariate methodology is characterized by both the power of synthesis of mathematical modeling and direct reference to a raw data structure. These features allow factor analysis to express the 2 main components of lactation-curve shape—i.e., the rate of increase of yield toward the lactation peak, and the persistency of lactation by 2 linear combinations of TD whose relative weights are objectively derived from the correlation structure of original data without previous assumptions.

## MATERIALS AND METHODS

### Defining New Indices of Lactation Curve Shape

In the multivariate context, TD records taken at different intervals from parturitions (for example, each month) are considered different correlated traits. Previous studies showed that the use of multivariate factor analysis to model these traits results in the extraction of 2 latent common factors able to explain a relevant portion of the (co)variance of original data (Wilmink, 1987; Macciotta et al., 2002). These two new variables ( $\mathbf{X}_l$  and  $\mathbf{X}_p$ ) can be related to the 2 main aspects of the lactation curve shape—i.e., the rate of increase to the lactation peak and the rate of decline after the peak, respectively.

According to the factor analysis model, each of the  $m$  TD records can be represented as a linear combination of 2 common factors:

$$\begin{aligned} y_i &= b_{11}X_l + b_{12}X_p + e_i \\ &\dots\dots\dots \\ y_m &= b_{m1}X_l + b_{m2}X_p + e_m \end{aligned} \quad [1]$$

where

$y_i$  = TD records taken at different distance from calving,

$b_{ij}$  = factor coefficients (or loadings), i.e., correlations between the  $j$ th common factors and the  $i$ th TD record,

$X_l$  and  $X_p$  = latent common factors, and

$e_i$  = random residual.

The rationale of factor analysis is the modeling of the correlation matrix of TD records within lactation (S)

$$\mathbf{S} = \mathbf{B}\mathbf{B}' + \Psi \quad [2]$$

where

$\mathbf{B}$  = the matrix of the factor coefficients of model [1], and

$\Psi$  = residual (co)variance matrix (McDonald, 1985; Morrison, 1976).

The maximum-likelihood method estimates  $\mathbf{B}\mathbf{B}'$  by minimizing the residual matrix  $\Psi$ . Moreover, the factor rotation allows for the estimation of different  $\mathbf{B}$  matrices that, with the constraint of keeping  $\mathbf{B}\mathbf{B}'$  invariant, simplify the factor structure and facilitate their interpretation in terms of relationships with the 2 main aspects of lactation curve shape. The scores of 2 common factors for each lactation can be then calculated as follows:

$$\mathbf{x}' = \mathbf{y}'(\mathbf{B}\mathbf{B}' + \Psi)^{-1}\mathbf{B} \quad [3]$$

where

$$\mathbf{x}' = [\chi_l \quad \chi_p],$$

$\mathbf{y}'$  = row vector of standardized TD records, and

$(\mathbf{B}\mathbf{B}' + \Psi)^{-1}\mathbf{B}$  = scoring coefficients.

According to model [3], all TD records are considered in the calculation of the scores of both factors. Moreover, relative TD weights are not assigned a priori, but they are derived from the correlation structure of original data ( $\mathbf{B}\mathbf{B}' + \Psi$ ) and from the correlations between factors and the original variables ( $\mathbf{B}$ ).

### Data

Original data were milk TD yields recorded according to the A4 scheme (about 4 wk between 2 consecutive tests, respectively) in the period 1989 to 2002, supplied by the Italian Association of Simmental Cow Breeders. From these data, an archive of 5728 lactations of 4932 cows was extracted. The number of TD records per lactation was fixed at 7. Lactations with less than 7 TD records were discarded, whereas extra TD records for lactation with more than 7 records were deleted. Edits were also on DIM at which the first TD was recorded (<20), parity (1 to 6), lactation length (200 to 280 d), calving interval (300 to 600 d), and herd size (>15 cows).

The 7 TD records for each cow were regarded as different traits (MILK1, MILK2, . . . MILK7). Means and

**Table 1.** Means and standard deviations of DIM and milk yield for the test-day variables considered in the multivariate analysis.

Variable	DIM		Milk yield (kg)	
	Mean	SD	Mean	SD
MILK1	12	4	22.1	6.0
MILK2	46	9	22.3	6.5
MILK3	79	11	20.1	6.3
MILK4	112	14	19.2	6.1
MILK5	146	16	17.7	5.8
MILK6	180	17	16.1	5.6
MILK7	214	18	14.1	5.3

standard deviations for the 7 TD milk yields are reported in Table 1.

### Statistical Analysis

The 2 common factors related to the main aspects of lactation curve shape were extracted from the data by using a maximum likelihood procedure and a VARIMAX rotation technique (SAS, 1996). Factor scores were then calculated for each lactation according to equation [3].

To evaluate relationships between the 2 common factors  $X_1$  and  $X_p$  and the environmental effects known to influence the shape of the lactation curve, factor scores were analyzed with the following mixed linear model:

$$Y_{ijklmn} = \text{PAR}_i + \text{SEA}_j + \text{YEAR}_k + \text{PROD}_l + H_m + e_{ijklmn} \quad [4]$$

where

$Y = X_1$  or  $X_p$  scores,

$\text{PAR}_i$  = fixed effect of the parity class (1, 2, . . . , 6),

$\text{SEA}_j$  = fixed effect of calving season (1 = Jan–Feb, . . . , 6 = Nov–Dec),

$\text{YEAR}_k$  = fixed effect of year of calving (1 = 1989, . . . , 11 = 1999),

$\text{PROD}_l$  = covariable represented by the 305-d milk,

$H_m$  = random effect of the herd, and

$e_{ijklmn}$  = random residual.

### RESULTS AND DISCUSSION

In Table 2, Pearson and partial correlations are reported, the latter measuring relationships among each pair of variables controlling possible effects of other variables. A marked reduction of partial correlations in comparison with Pearson correlation can be noticed. An objective measure of such a reduction is the Kaiser

measure of sampling adequacy that was about 0.91 in this study, higher than the 0.80 usually assumed as the minimum threshold for the suitability of the data set for the factor analysis (Cerny and Kaiser, 1977).

In fact, the common factors  $X_1$  and  $X_p$  were able to explain about 86% of the original (co)variance of TD records, which is a relevant quota for a method that is very easy to apply. Models such as random regression or covariance functions are able to account for a larger amount of (co)variance but together with more theoretical and computational difficulties.

According to equation [2], the correlation matrix of original variables can be decomposed as follows (symm = symmetric)

$$\begin{bmatrix} 1 & 0.83 & 0.78 & 0.73 & 0.68 & 0.61 & 0.50 \\ & 1 & 0.89 & 0.85 & 0.79 & 0.72 & 0.62 \\ & & 1 & 0.90 & 0.86 & 0.79 & 0.68 \\ & & & 1 & 0.90 & 0.84 & 0.74 \\ & & & & 1 & 0.89 & 0.80 \\ & & & & & 1 & 0.86 \\ \text{symm} & & & & & & 1 \end{bmatrix} =$$

$$\begin{bmatrix} 0.73 & 0.81 & 0.79 & 0.75 & 0.68 & 0.60 & 0.49 \\ & 0.90 & 0.89 & 0.86 & 0.80 & 0.72 & 0.61 \\ & & 0.91 & 0.89 & 0.86 & 0.79 & 0.68 \\ & & & 0.89 & 0.88 & 0.84 & 0.75 \\ & & & & 0.90 & 0.89 & 0.81 \\ & & & & & 0.91 & 0.85 \\ \text{symm} & & & & & & 0.80 \end{bmatrix} +$$

$$\begin{bmatrix} 0.27 & -0.03 & -0.01 & -0.02 & 0.00 & 0.01 & 0.01 \\ & 0.10 & 0.00 & -0.01 & -0.0 & 0.00 & 0.01 \\ & & 0.09 & 0.01 & 0.00 & 0.00 & 0.00 \\ & & & 0.11 & 0.02 & 0.00 & -0.01 \\ & & & & 0.10 & 0.00 & -0.01 \\ & & & & & 0.09 & 0.01 \\ \text{symm} & & & & & & 0.20 \end{bmatrix}$$

The goodness of fit of the factor model is confirmed by the values of  $\Psi$  (second matrix of the right term of the equation). Off diagonal elements are the residual correlations, i.e., the part of the original correlations that have not been reconstructed by the factor model, whereas diagonal elements are the uniqueness of variables, i.e., the specific variance of each original variable plus the random error (Enevoldsen et al., 1996). Values

**Table 2.** Pearson (above the diagonal) and partial (under the diagonal) correlations among TDI yields recorded at different stages of lactation for the TD7 data set.

	MILK1	MILK2	MILK3	MILK4	MILK5	MILK6	MILK7
MILK1	*	0.83	0.78	0.73	0.68	0.61	0.50
MILK2	0.47	*	0.89	0.85	0.79	0.72	0.62
MILK3	0.13	0.42	*	0.90	0.86	0.79	0.68
MILK4	0.02	0.16	0.35	*	0.90	0.84	0.74
MILK5	-0.02	0.01	0.17	0.36	*	0.89	0.80
MILK6	-0.01	-0.00	0.03	0.15	0.35	*	0.86
MILK7	-0.06	-0.01	-0.02	-0.00	0.18	0.54	*

of all elements, especially of those off diagonal, are very close to zero, thus evidencing the suitability of the factor model.

Factors  $X_1$  and  $X_p$  are mainly correlated with the first 3 TD and the last 3 TD of the lactation, respectively (Table 3). Because the maximum yield occurs within the first 3 TD,  $X_1$  scores can be considered as an index of the increasing rate of milk yield in the first phase of lactation. On the other hand,  $X_p$  scores can be proposed as a measure of the inverse of the rate of decrease of milk yield in the second part of lactation (persistency).

Factor scores can be calculated on the basis of equation [3] using scoring coefficients, i.e., the relative weights to be assigned to each standardized TD record, reported in Table 4. All TD records are used in the calculation of the 2 factors, even if with different relative weights and signs. As an example, the calculation of  $X_1$  and  $X_p$  scores for a cow is reported. Original values (16.6, 19, 16.8, 14, 16.2, 15, and 12.4 kg for MILK1, MILK2, MILK3, MILK4, MILK5, MILK6, and MILK7, respectively) are standardized using the overall means and standard deviations (Table 1).

$$\begin{aligned}
 & [\chi_l \quad \chi_p] = \\
 & \left[ \begin{array}{cccc} \frac{16.6 - 22.1}{6} & \frac{19 - 22.3}{6.5} & \frac{16.8 - 20.7}{6.3} & \frac{14 - 19.2}{6.1} \\ \frac{16.2 - 17.7}{5.8} & \frac{15 - 16.1}{5.6} & \frac{12.4 - 14.1}{5.3} & \end{array} \right] \\
 & \left[ \begin{array}{cc} 0.234 & -0.161 \\ 0.570 & -0.330 \\ 0.447 & -0.159 \\ 0.166 & 0.083 \\ -0.067 & 0.334 \\ -0.366 & 0.655 \\ -0.225 & 0.341 \end{array} \right] = [-0.761 \quad 0.018]
 \end{aligned}$$

Correlations among  $X_1$  and  $X_p$  scores and the total 305-d milk yield were 0.65 and 0.73, respectively, indicating that an increase of either the rate of increase to lactation peak or the lactation persistency, or both,

results in a higher 305-d milk. The low correlation between  $X_p$  and  $X_1$  scores (0.11) can be understood by examining the crossed frequency distribution of factor scores (Table 5). Actually, because factor scores are normal standardized variables, most of cows have values of both  $X_1$  and  $X_p$  scores around the mean, but the distribution of the factor scores within each of the levels of the other shows values in almost all the cells. It can be observed that, in the classes of average values of  $X_1$  ("−1 0" or "0 1"), there are 578 cows out of 4162 that have values of  $X_p$  higher than one. Thus about 14% of cows with a medium peak level show high persistency of lactation.

Relationships among  $X_1$  and  $X_p$  scores and the shape of the lactation curve can be inferred from Figures 1, 2, and 3, where average lactation patterns of groups of animals classified in different positions of Table 5 are shown. Lactation patterns, reported in Figure 1, show that an increase of both the 2 factors results in a parallel shift of the whole curve to higher values, i.e., in an increase of the total lactation yield, whereas increasing values of one factor for a constant value of the other result in patterns characterized by a progressively higher rate of increase to lactation peak or persistency (Figures 2 and 3, respectively). On the basis of these results,  $X_1$  and  $X_p$  scores can be proposed as criteria for detecting cows with the more suitable lactation-curve shape at a given 305-d yield level.

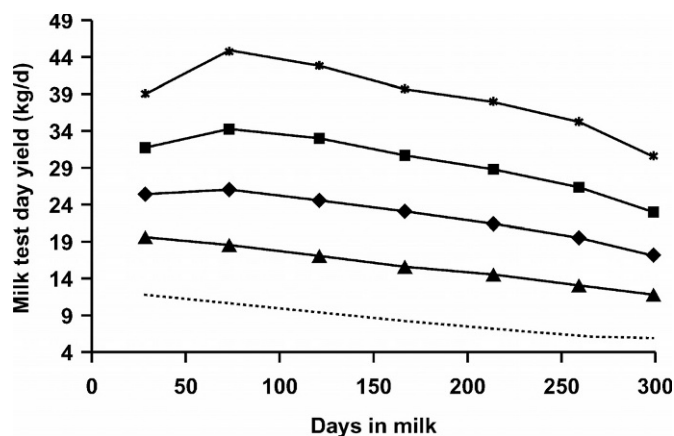
Results of mixed-model analysis carried out on  $X_1$  and  $X_p$  scores are reported in Tables 6 and 7. The 2 variables

**Table 3.** Common factor loadings and communalities for data set TD7.<sup>1</sup>

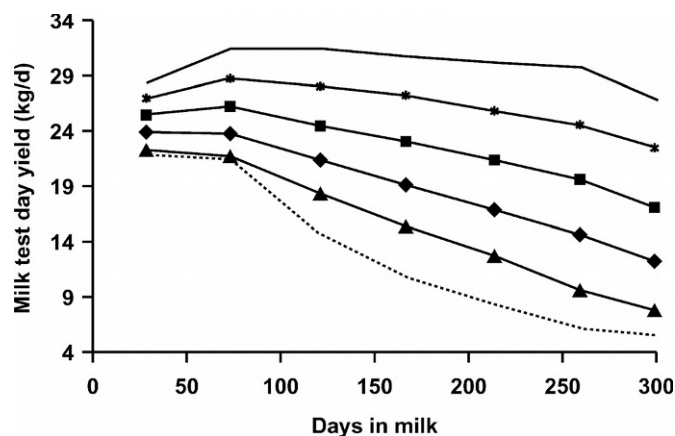
Variable	$X_1$	$X_p$
MILK1	0.81	0.29
MILK2	0.85	0.41
MILK3	0.79	0.53
MILK4	0.69	0.64
MILK5	0.58	0.75
MILK6	0.43	0.85
MILK7	0.31	0.84
Variance explained (%)	0.44	0.42

<sup>1</sup> $X_1$  = Latent common factor related to the lactation peak;  $X_p$  = latent common factor related to the lactation persistency.





**Figure 1.** Average lactation curves of groups of cows having values of both  $X_1$  and  $X_p$  scores of the same magnitude (---- =  $-2 \div -1$ ;  $\triangle$  =  $-1 \div 0$ ;  $\diamond$  =  $0 \div 1$ ;  $\blacksquare$  =  $1 \div 2$ ;  $\ast$  =  $>2$ ). ( $X_1$  = latent common factor related to the lactation peak;  $X_p$  = latent common factor related to the lactation persistency). Points are plotted for the average day in milk on each test day.



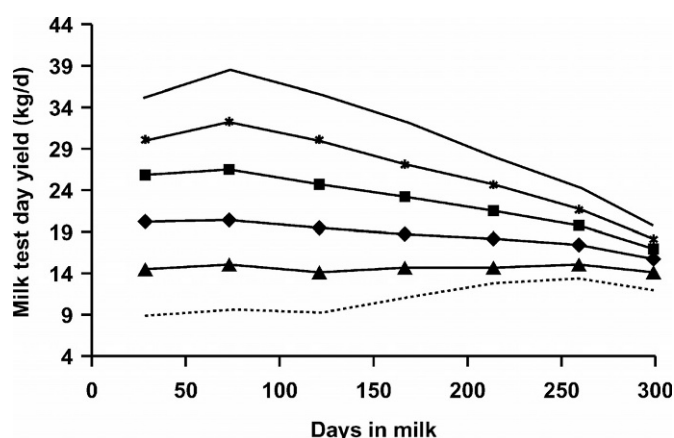
**Figure 3.** Average lactation curves of groups of cows with a constant value of  $X_1(0 \div 1)$  and increasing values of  $X_p$  scores (---- =  $< -2$ ;  $\triangle$  =  $-2 \div -1$ ;  $\diamond$  =  $-1 \div 0$ ;  $\blacksquare$  =  $0 \div 1$ ;  $\ast$  =  $1 \div 2$ ; --- =  $> 2$ ). ( $X_1$  = latent common factor related to the lactation peak;  $X_p$  = latent common factor related to lactation persistency). Points are plotted for the average DIM on each test day.

considered were particularly affected by parity and calving season ( $P < 0.0001$ ), as evidenced by the change in signs of least squares means across levels of each factor. Calving year also had a significant effect, even if to a lesser extent ( $P < 0.01$ ), as evidenced by the absence of changes in sign of least squares means across years.

Peak factor scores were higher for cows of third and greater parities, whereas the lowest value was observed for first-parity cows (Table 6). Spring calvings showed the highest peak value due to the favorable environ-

mental conditions (pastures, climate). Finally, a slight, even if irregular, tendency of peak factor to increase across years can be observed. This result could be ascribed to the effect of the selection program carried out on this dual-purpose breed.

First-parity cows showed the highest values of persistency that tends to decrease in older cows (Table 7), in agreement with previous figures obtained for cattle (Shanks et al., 1981; Sölkner and Fuchs, 1987). The higher persistency of younger animals is usually explained with the maturation process, which is still in progress in young animals and that counteracts the normal decline in milk yield in the second part of lactation (Stanton et al., 1992). Cows calving in fall and winter were the most persistent, probably because they had the end of lactation in the spring, i.e., when environmental conditions are better. This result agrees with previous findings on persistency calculated with the mean standard deviation of TD yields along the 305-d lactation (Sölkner and Fuchs, 1987). The lowest persistency has been observed for cows calving in spring. No



**Figure 2.** Average lactation curves of groups of cows with a constant value of  $X_p(0 \div 1)$  and increasing values of  $X_1$  scores (---- =  $< -2$ ;  $\triangle$  =  $-2 \div -1$ ;  $\diamond$  =  $-1 \div 0$ ;  $\blacksquare$  =  $0 \div 1$ ;  $\ast$  =  $1 \div 2$ ; --- =  $> 2$ ). ( $X_1$  = latent common factor related to the lactation peak;  $X_p$  = latent common factor related to the lactation persistency). Points are plotted for the average DIM on each test day.

**Table 4.** Scoring coefficients used for the calculation of factor scores.<sup>1</sup>

Variable	$X_1$	$X_p$
MILK1	0.234	-0.161
MILK2	0.570	-0.330
MILK3	0.447	-0.159
MILK4	0.166	0.083
MILK5	-0.067	0.334
MILK6	-0.366	0.655
MILK7	-0.225	0.341

<sup>1</sup> $X_1$  = latent common factor related to the lactation peak;  $X_p$  = latent common factor related to the lactation persistency.

**Table 5.** Crossed distributions of lactations between  $X_l$  and  $X_p$  score classes.<sup>1</sup>

$X_l$ class <sup>2</sup>	$X_p$ class <sup>2</sup>						Total
	I	II	III	IV	V	VI	
I Absolute frequencies	0	10	18	13	1	2	44
Relative frequencies		0.17	0.31	0.23	0.02	0.03	0.77
II	0	136	361	190	37	6	730
		2.37	6.30	3.32	0.65	0.10	12.74
III	9	216	936	815	221	50	2247
	0.16	3.77	16.34	14.23	3.86	0.87	39.23
IV	23	220	717	648	240	67	1915
	0.4	3.84	12.52	11.31	4.19	1.17	33.43
V	17	93	194	193	101	34	632
	0.3	1.62	3.39	3.37	1.76	0.59	11.03
VI	9	20	51	38	29	13	160
	0.16	0.35	0.89	0.66	0.51	0.23	2.79
Total	58	695	2277	1897	629	172	5728
	1.01	12.13	39.75	33.12	10.98	3.00	

<sup>1</sup> $X_l$  = latent common factor related to the lactation peak.  $X_p$  = latent common factor related to the lactation persistency.

<sup>2</sup>I =  $X_l$  or  $X_p < -2$ ; II =  $-2 < X_l$  or  $X_p < -1$ ; III =  $-1 < X_l$  or  $X_p < 0$ ; IV =  $0 < X_l$  or  $X_p < 1$ ; V =  $1 < X_l$  or  $X_p < 2$ ; VI =  $X_l$ , or  $X_p > 2$ .

**Table 6.** Least squares means of  $X_l$  factor scores estimated with model [4].

Parity			Calving season			Calving year		
Mean	SE		Mean	SE		Mean	SE	
1	−0.48	0.02	1	0.04	0.02	1989	0.03	0.05
2	0.03	0.02	2	0.34	0.02	1990	−0.01	0.05
3	0.15	0.02	3	0.27	0.03	1991	0.03	0.04
4	0.17	0.03	4	−0.02	0.03	1992	0.08	0.04
5	0.18	0.03	5	−0.20	0.03	1993	0.07	0.04
6	0.20	0.04	6	−0.20	0.02	1994	0.05	0.03
						1995	0.07	0.03
						1996	0.00	0.03
						1997	0.04	0.03
						1998	0.04	0.03
						1999	0.11	0.02

**Table 7.** Least squares means of  $X_p$  factor scores estimated with model [4].

Parity			Calving season			Calving year		
Mean	SE		Mean	SE		Mean	SE	
1	0.21	0.02	1	0.02	0.02	1989	-0.09	0.05
2	-0.15	0.02	2	-0.40	0.02	1990	-0.10	0.04
3	-0.24	0.02	3	-0.50	0.02	1991	-0.16	0.04
4	-0.25	0.02	4	-0.30	0.03	1992	-0.22	0.04
5	-0.25	0.03	5	0.01	0.02	1993	-0.19	0.04
6	-0.31	0.04	6	0.11	0.02	1994	-0.16	0.03
						1995	-0.22	0.03
						1996	-0.13	0.03
						1997	-0.19	0.03
						1998	-0.22	0.02
						1999	-0.14	0.02

detectable trends of persistency were evidenced across years of calving.

## CONCLUSIONS

Problems in defining indices able to synthesize most economically and technically important aspects of evolution of milk yield over time are due to the poor fitting of analytical functions to individual lactation patterns and to the arbitrariness of the more empirical methods based on the direct combination of TD records. The multivariate factor analysis applied to TD records taken at different intervals from parturition is a simple and effective method for extracting the latent factors  $X_1$  and  $X_p$ , related respectively to the rate of increase of yield to the lactation peak and to the persistency of lactation. Scores of the 2 factors are calculated as linear combinations of all TD records considered, with objective relative weights derived from the correlation structure of original data.

The 2 latent factors allow for the separation of individual lactation curves characterized by different shapes. At the phenotypic level,  $X_1$  and  $X_p$  are lowly correlated, whereas each of them is highly related to the 305-d milk. If confirmed at a genetic level, such a result could be of great interest for the genetic improvement of lactation yield by selecting animals on the basis of persistency.

Results of the mixed linear model analysis highlight relationships among  $X_1$  and  $X_p$  and some environmental factors known to affect lactation-curve shape, such as calving season and parity.

Finally, a major limitation of the multivariate methodology, i.e., the requirement of a fixed number of observation for each subject, can be overtaken by combining factor analysis with one of the several extrapolation techniques proposed in the literature.

## ACKNOWLEDGMENTS

Research was funded by the Italian Ministry of University and Research (grant ex 40%).

## REFERENCES

- Bichard, M. 2002. Genetic improvement in dairy cattle—an outsider's perspective. *Livest. Prod. Sci.* 75:1–10.
- Cerny, B. A., and H. F. Kaiser. 1977. A study of a measure of sampling adequacy for factor analytic correlation matrices. *Multiv. Behav. Res.* 12: 43–47.
- Dekkers, J. C. M., J. H. Ten Hag, and A. Weersink. 1998. Economic aspects of persistency of lactation in dairy cattle. *Livest. Prod. Sci.* 53:237–252.
- Enevoldsen, C., J. Hindhede, and T. Kristiansen. 1996. Dairy herd management types assessed from indicators of health, reproduction, replacement and milk production. *J. Dairy Sci.* 79:1221–1236.
- France, J., and J. H. M. Thornley. 1984. *Mathematical Models in Agriculture*. Butterworths, London, UK.
- Gengler, N. 1996. Persistency of lactation yields: A review. *INTERBULL Bull.* 12:87–96.
- Groen, A. F., T. Steine, J. J. Colleau, J. Pedersen, J. Pribyl, and N. Reinsch. 1997. Economic values in dairy cattle breeding, with special reference to functional traits. Report of an EAAP-working group. *Livest. Prod. Sci.* 49:1–21.
- Grossman, M., S. M. Hartz, and W. J. Koops. 1999. Persistency of lactation yield: A novel approach. *J. Dairy Sci.* 82:2192–2197.
- Haile-Mariam, M., P. J. Bowman, and M. E. Goddard. 2003. Genetic and environmental relationship among calving interval, survival, persistency of milk yield and somatic cell count in dairy cattle. *Livest. Prod. Sci.* 80:189–200.
- Jamrozik, J., G. Jensen, L. R. Schaeffer, and Z. Liu. 1998. Analysis of persistency of lactation calculated from a random regression test day model. *INTERBULL Bull.* 16:64–68.
- Jakobsen, J., P. Madsen, J. Jensen, J. Peersen, L. G. Cristiansen, and D. A. Sorensen. 2002. Genetic parameters for milk production and persistency for Danish Holsteins estimated in random regression models using REML. *J. Dairy Sci.* 85:1607–1616.
- Jensen, J. 2001. Genetic evaluation of dairy cattle using test day models. *J. Dairy Sci.* 84:2803–2812.
- Macciotta, N. P. P., D. Vicario, G. Pulina, and A. Cappio-Borlino. 2002. An index of lactation persistency based on multivariate factor analysis. Pages 67–70 in *Proc. 7th World Congress of Genetics Applied to Livestock Production*, Montpellier, France.
- McDonald, R. P. 1985. *Factor analysis and related methods*. Lawrence Erlbaum Associates Publishers, Mahwah, NJ.
- Morrison, F. 1976. *Multivariate Statistical Methods*. McGraw-Hill, New York, NY.
- Olori, V. E., S. Brotherstone, W. G. Hill, and B. J. McGurk. 1999. Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd. *Livest. Prod. Sci.* 58:55–63.
- Pryce, J. E., R. F. Veerkamp, R. Thompson, W. G. Hill, and G. Simm. 1997. Genetic aspects of common health disorders and measures of fertility in Holstein Friesian dairy cattle. *Anim. Sci.* 65:353–360.
- Rekaya, R., K. A. Weigel, and D. Gianola. 2001. Hierarchical nonlinear model for the persistency of milk yield in the first three lactations of Holsteins. *Livest. Prod. Sci.* 68:181–187.
- SAS. 1996. *User's Guide: Statistics*. SAS Inst., Cary, NC.
- Schaeffer, L. R., and J. C. M. Dekkers. 1994. Random regressions in animal models for test day production in dairy cattle. *Proc. 5th World Congr. Genet. Appl. Livest. Prod.* 18:443–446.
- 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.
- Sölkner, J., and W. Fuchs. 1987. A comparison of different measures of persistency with special respect to variation of test day yields. *Livest. Prod. Sci.* 16:305–319.
- Stanton, T. L., L. R. Jones, R. W. Everett, and S. D. Kachman. 1992. Estimating milk, fat and protein lactation curves with a test day model. *J. Dairy Sci.* 75:1691–1700.
- Swalve, H. H. 1995. Genetic relationship between dairy lactation persistency and yield. *J. Anim. Breed. Genet.* 112:303–311.
- Swalve, H. H. 2000. Theoretical basis and computational methods for different test day genetic evaluation methods. *J. Dairy Sci.* 83:1115–1124.
- Togashi, K., and C. Y. Lin. 2003. Modifying the lactation curve to improve lactation milk and persistency. *J. Dairy Sci.* 86:1487–1493.
- Wilmink, J. B. M. 1987. Comparison of different methods of predicting 305-day milk yield using means calculated from within-herd lactation curves. *Livest. Prod. Sci.* 17:1–17.