Use of Multivariate Analysis to Extract Latent Variables Related to Level of Production and Lactation Persistency in Dairy Cattle

N. P. P. Macciotta,^{*1} D. Vicario,[†] and A. Cappio-Borlino^{*}

^{*}Dipartimento di Scienze Zootecniche, Università di Sassari, Via De Nicola 9, 07100 Sassari, Italy

[†]Italian Association of Simmental Breeders, Via Nieve 19, 33100 Udine, Italy

ABSTRACT

Multivariate factor analysis and principal component analysis were used to decompose the correlation matrix of test-day milk yields of 48,374 lactations of 21,721 Italian Simmental cows. Two common latent factors related to level of production in early lactation and lactation persistency, and 2 principal components associated with the whole lactation yield and persistency were obtained. Factor and principal component scores were treated as new quantitative phenotypes related to prominent features of lactation curve shape. Genetic parameters were estimated by univariate and bivariate animal models. Estimates of heritability were moderately low for both latent factors (0.13 for persistency and yield early in lactation). Heritabilities of the principal component related to total lactation yield and 305-d yield were similar (0.19 and 0.20, respectively). Finally, heritability was quite low for the principal component related to lactation persistency (0.07). Repeatabilities between lactations were about 0.27 for both latent factors, around 0.4 for the first principal component and 305-d yield, and 0.11 for the second principal component. Moderate genetic correlation among common factors (0.26) and their high genetic correlation with total lactation yield (>0.60) suggest that selection can be used to change the shape of lactation curve as well as improve yield. Scores of the second principal component can be used to genetically improve persistency while maintaining constant total lactation yield.

Key words: lactation curve, lactation persistency, factor analysis, principal component analysis

INTRODUCTION

Genetic change of the shape of the lactation curve is of great interest for the dairy cattle industry for its technical and economic implications. The rate of ascent of milk yield to the lactation peak and the slope of the

curve in the second part of lactation have been widely investigated in dairy cattle. Persistency of lactation, that is, the ability of a cow to maintain a constant yield during lactation (Gengler, 1996), has an economic value of about 3.4% of that for the total lactation yield (Dekkers et al., 1998). Cows with daily yield distributed more uniformly over the lactation are subject to fewer metabolic disorders and health and fertility problems, and have more consistent energy requirements, allowing for the use of cheaper feeds (Jakobsen et al., 2002). Hence, the inclusion of lactation persistency in the breeding goal of selection programs for dairy breeds is desirable.

Different approaches for quantifying lactation persistency have been proposed (Gengler, 1996; Grossman et al., 1999; Togashi and Lin, 2004). Methods based on combinations of test-day (TD) records taken at different stages of lactation, such as ratios between cumulative yields or measures of TD variation (Sölkner and Fuchs, 1987; Swalve, 1995), or on the number of days at which a constant yield is maintained (Grossman et al., 1999) are simple and direct; however, they do not characterize persistency in a unique manner because they are not invariant with respect to the time period chosen (Rekaya et al., 2001). On the other hand, parametric measures based on coefficients of lactation curve functions (e.g., Shanks et al., 1981; van der Linde et al., 2000; Kamidi, 2005) are affected by the great variability among cows in the shape of the lactation curve (Olori et al., 1999a; Macciotta et al., 2005). Finally, measures of persistency based on random regression TD models not only require large computational effort, but also an a priori definition of which stages of lactation are the most important to define persistency (Schaeffer and Dekkers, 1994; Jamrozik et al., 1998; Togashi and Lin, 2003).

Different persistency measurements result in substantial variability of genetic parameter estimates for this trait. Heritability ranges from very low (0.01) to moderate values (0.30) (Jakobsen et al., 2003; Kistemaker, 2003; Muir et al., 2004). Genetic correlation among persistency and total lactation yield varies in magnitude and sign depending on the measure of per-

Received October 3, 2005.

Accepted February 13, 2006.

¹Corresponding author: macciott@uniss.it

sistency used (Swalve, 1995; Haile Mariam et al., 2003; Muir et al., 2004).

The main task in measuring persistency is to express the shape of the lactation curve by a single term (Sölkner and Fuchs, 1987). Methods based on the combination of phenotypic TD records can address this issue, are reasonably simple to implement, and consider the whole lactation. However, the problem is to define relative weights to assign to each TD record in the calculation of persistency. Data-reduction methods, such as multivariate factor analysis (MFA) and principal component analysis (PCA), are able to synthesize multivariate complex phenotypes in linear combinations of original data whose weights are objectively derived from the correlation (covariance) matrix of original variables (Kirkpatrick and Meyer, 2004). The MFA has been used to decompose the phenotypic correlation matrix of TD records deriving 2 latent variables that account for a large part of the variation of original variables (Wilmink, 1987) and whose scores are able to rank curves with different level of production in early lactation (X_1) and persistency (X_p ; Macciotta et al., 2004). In addition, PCA has been used by several authors to decompose the genetic (co)variance matrix of TD records estimated by random regression models: 2 eigenvectors able to explain most of genetic variation and that are interpreted as indicators of average lactation potential and lactation persistency have been highlighted (van der Werf et al., 1998; Olori et al., 1999b; Druet et al., 2005).

In this work, the ability of MFA and PCA to extract new variables related to the shape of the lactation curve from the phenotypic correlation matrix of TD records is compared. Moreover, genetic variability and relationship with total lactation yield of new variables are estimated.

MATERIALS AND METHODS

Data

Original data were milk test-day yields recorded according to the A4 or A6 scheme (approximately 4 or 6 wk between 2 consecutive tests, respectively) by the Italian Association of Simmental Cow Breeders. From these data, an archive of 48,374 lactations of 21,721 cows was extracted. The number of TD records per lactation was fixed at 7. The 7 TD records for each cow were regarded as different traits (MILK1, MILK2,..., MILK7). Lactations with fewer than 7 TD records were discarded, whereas extra TD records for lactations with more than 7 records were deleted. Edits were also performed on DIM at which the first TD was recorded (<30), parity (1 to 6), lactation length (200 to 280 d), calving interval (300 to 600 d), and herd size (>15 cows). Only cows that had the first lactation recorded were

considered. The pedigree file contained 101,752 individuals.

Latent Variables Extraction and Genetic Parameter Estimation

The PCA and MFA are multivariate dimension reduction techniques mainly aimed at synthesizing information contained in a set of n observed variables ($y_1, \dots, y_n$) by seeking a new set of p ($p < n$) variables ($X_1, \dots, X_p$), principal components and common latent factors, respectively.

The j th principal component is represented by the linear combination of the observed variables:

$$X_j = \alpha_{1j}y_1 + \dots + \alpha_{nj}y_n,$$

where coefficients α_{ij} are the elements of the eigenvector of the (co)variance (or correlation) matrix of observed variables corresponding to the j th eigen value. Principal component extraction is carried out in such a way as to maximize successively the amount of original variance explained (Morrison, 1976; Krzanowsky, 2003).

On the contrary, MFA starts from a linear modeling of observed variables in terms of a limited set of latent variables (or common factors; Morrison, 1976):

$$y_1 = b_{11}X_1 + \dots + b_{1p}X_p + e_1, \text{ and}$$

$$y_n = b_{n1}X_1 + \dots + b_{np}X_p + e_n,$$

where X_j is the j th common factor, b_{ij} are factor coefficients (or loadings); i.e., correlations between the j th common, e_i is the i th residual specific variable.

The rationale of factor analysis is the modeling of the correlation matrix of observed variables:

$$S = BB' + \Psi,$$

where B is the matrix of the factor coefficients and Ψ is a residual (co)variance matrix (Morrison, 1976; McDonald, 1985).

Moreover, PCA leads to a unique set of latent variables, whereas MFA is able to extract different sets of common latent factors via the rotation technique (McDonald, 1985). Such flexibility allows the simplification of factor structure and facilitates their interpretation in terms of relationships with original variables.

The PCA was developed to calculate linear combinations of the 7 traits (MILK1 to MILK7) by extracting eigen values and eigenvectors of the phenotypic correlation matrix of TD records. The MFA was performed by ML estimation and a VARIMAX orthogonal rotation technique (SAS Institute, 1996). Suitability of scores of

Table 1. Correlations between common latent factors and consecutive test-days (MILK1 to MILK7)

Variable ¹	Common latent factor ²	
	X _l	X _p
MILK1	0.83	0.17
MILK2	0.89	0.32
MILK3	0.82	0.45
MILK4	0.72	0.59
MILK5	0.59	0.73
MILK6	0.41	0.84
MILK7	0.13	0.69
Variance explained (%)	46	34

¹MILK1 to MILK 7 = milk tests recorded at different DIM (on average 17, 60, 103, 145, 189, 232, and 267, respectively).

²X_l = common latent factor related to level of production in early lactation; X_p = common latent factor related to the lactation persistency.

variables extracted both by MFA and PCA to represent main traits of lactation curve shape was tested by comparing average lactation patterns of cows grouped according to different score classes.

Scores of new variables and total lactation yield (LY) were then analyzed with the following univariate animal model to estimate variance components:

$$y_{ijklmno} = \mu + H_i + Y_j + S_k + \text{Par}_l + a_m + p_n + e_{ijklmno} \quad [1]$$

where y = scores of factors, or PC, or 305-d yield; μ = overall mean; H_i = fixed effect of herd (1,772 levels); Y_j = fixed effect of calving year (11 levels, 1989 to 1999); S_k = fixed effect of calving season (1 = Jan to Feb, ..., 6 = Nov to Dec); Par_l = fixed effect of parity (1 to 6); a_m = random additive effect of animal; p_n = random effect of permanent environment; and $e_{ijklmno}$ = random residual.

Model [1] was solved by using a REML procedure with an average information algorithm (Gilmour et al., 2000). Genetic correlations between LY, factor, and principal component scores were estimated with bivariate animal models with the same structure of model [1]; variance components estimates obtained with the univariate analysis were used as starting values.

RESULTS AND DISCUSSION

The MFA was able to extract from the phenotypic correlation matrix of TD records two latent factors that explained 80% of variance of original data. Factors were correlated with TD records in the first (X_l) and in second part (X_p) of lactation, respectively (Table 1). These results are in agreement with those of Wilmink (1987), who found 2 latent factors interpreted as indicators of

Table 2. Scoring coefficients used for the calculation of factor scores

Variable ²	Common latent factor ²	
	X _l	X _p
MILK1	0.242	-0.187
MILK2	0.525	-0.323
MILK3	0.366	-0.105
MILK4	0.141	0.132
MILK5	-0.066	0.407
MILK6	-0.292	0.651
MILK7	-0.099	0.168

¹MILK1 to MILK 7 = milk tests recorded at different DIM (on average 17, 60, 103, 145, 189, 232, and 267, respectively).

²X_l = common latent factor related to level of production in early lactation; X_p = common latent factor related to the lactation persistency.

peak yield and lactation persistency at the phenotypic level.

Factor scores were obtained as linear combinations of the standardized original variable (MILK_i) multiplied by scoring coefficients (Table 2). All TD records were used, with positive (negative) weights given to MILK1 to MILK4 and negative (positive) to MILK5 to MILK7 in the calculation of X_l (X_p) scores. The approach for calculation of X_p follows the suggestion of Cole and VanRaden (2005), according to which cows with high persistency tend to milk less than of expected at the beginning of lactation and more than expected at the end. Factor scores follow the standardized normal distribution, with the average lactation curve (represented by average values of MILK1 to MILK7) having value 0 for both X_l and X_p.

Relationships among scores of X_l and X_p and the shape of the lactation curve can be inferred from Figures 1 and 2 where average lactation patterns of ani-

Figure 1. Average lactation curves of groups of cows of different X_l score classes (····· = -2 to -1; ▲ = -1 to 0; ◆ = 0 to 1; ■ = 1 to 2; ● = >2). X_l = latent common factor related to the level of production in early lactation. Points are plotted for the average day in milk on each test day. Vertical bars represent the standard errors of the mean.

Figure 2. Average lactation curves of groups of different X_p score classes ($\cdots$ = -2 to -1; $\blacktriangle$ = -1 to 0; $\blacklozenge$ = 0 to 1; $\blacksquare$ = 1 to 2; $\bullet$ = >2). X_p = latent common factor related to lactation persistency. Points are plotted for the average day in milk on each test day. Vertical bars represent the standard errors of the mean.

mals grouped according to different factor score classes are reported. An increase of scores of X_1 or X_p results in an increase of the average level of production in early lactation or persistency, respectively.

Phenotypes of X_1 and X_p were weakly correlated (0.11). Actually, this low correlation was a result of the rotation of factors that, from a geometrical standpoint, corresponds to a rotation of the axis of the factor space to obtain the pattern of loadings that is most easily interpreted (Krzanowsky, 2003). Consequently, the complete independence between factors is lost. However, if only small increases in correlations are produced, factors may be regarded as being independent and free of the inference problems caused by large correlations (Sieber et al., 1988). Both factors show a high correlation with LY (0.80 and 0.63 for X_1 and X_p , respectively).

Eigenvectors and eigen values of the 2 leading principal component (87% of original variation) are reported in Table 3. The first principal component (**PC1**) had a similar relationship with all MILK traits, whereas the second (**PC2**) was negatively associated with the first and positively with the second part of lactation. These results agree with previous research on the decomposition of genetic covariance matrix for TD milk (van der Werf et al., 1998; Druet et al., 2003; Druet et al., 2005). Average lactation patterns for groups of animal ranked according to PC1 and PC2 scores (Figures 3 and 4) confirm their meaning as phenotypic indicators of level of production for the whole lactation and lactation persistency, respectively.

Phenotypes of PC1 and PC2 were uncorrelated, as principal components are orthogonal by definition. The PC1 had a very high correlation with LY (0.95) whereas

Table 3. Leading eigenvectors and associated eigen values of the correlation matrix of milk tests estimated by principal component analysis

Variable ¹	Eigenvectors ²	
	PC1	PC2
MILK1	0.345	-0.461
MILK2	0.396	-0.308
MILK3	0.411	-0.181
MILK4	0.415	-0.046
MILK5	0.409	0.123
MILK6	0.383	0.360
MILK7	0.261	0.717
Eigen values (%)	73	14

¹MILK1 to MILK 7 = milk tests recorded at different DIM (on average 17, 60, 103, 145, 189, 232, 267, respectively).

²PC1 = principal component related to the average lactation yield; PC2 = principal component related to lactation persistency.

PC2 was essentially uncorrelated with total yield (-0.05). The negligible correlation between PC2 and LY fits the suggestion of some authors to use measures of persistency uncorrelated from total lactation yield (Gengler, 1996; Jamrozik et al., 1998).

All factors included in model [1] affected the analyzed variables significantly. In particular, variables related to lactation persistency (X_p and PC2) were greater for first-calving cows and cows calving in fall (results not reported for brevity). Estimates of genetic parameters for X_1 , X_p , PC1, PC2, and LY from the univariate animal model [1] are reported in Table 4. Estimated h^2 for X_1 scores was lower than values reported in previous studies for peak yield in other populations of cattle (Shanks et al., 1981; Rekaya et al., 2001). The moderate value of heritability of X_p is in agreement with results

Figure 3. Average lactation curves of groups of cows of different PC1 score classes ($\cdots$ = -2 to -1; $\blacktriangle$ = -1 to 0; $\blacklozenge$ = 0 to 1; $\blacksquare$ = 1 to 2; $\bullet$ = >2). PC1 = principal component related to level of production for the whole lactation. Points are plotted for the average day in milk on each test day. Vertical bars represent the standard errors of the mean.

Figure 4. Average lactation curves of groups of cows of different PC2 score classes (····· = -2 to -1 ; ▲ = -1 to 0 ; ◆ = 0 to 1 ; ■ = 1 to 2 ; ● = >2). PC2 = principal component related to lactation persistency. Points are plotted for the average day in milk on each test day. Vertical bars represent the standard errors of the mean.

reported by several authors for lactation persistency (Gengler, 1996; Jamrozik et al., 2001; Jakobsen et al., 2002).

Heritability values of PC1 and LY were similar (Table 4). This was an expected result, because PC1 was equally related to all milk tests. Heritability of LY was quite low in comparison with reference values reported for 305-d yield, but of the same order of recent estimates obtained by Gengler et al. (2001).

Estimated h^2 for PC2 (Table 4) was lower than results previously reported for measures of persistency corrected for total yield (Swalve, 1995). However, the estimate was of a similar magnitude to the second eigen value of the genetic correlation matrix of TD records associated to lactation persistency (Olori et al., 1999b; Druet et al., 2003).

Repeatability of X_p was equal to the value reported by Gengler (1996) for both real and apparent persistency

Table 5. Genetic correlations (SE in parentheses) estimated with the bivariate analysis¹

Trait 1	Trait 2	r_g
X_l	X_p	0.26 (0.067)
X_l	LY	0.67 (0.08)
X_p	LY	0.61 (0.09)
PC1	PC2	0.05 (0.001)
PC1	LY	0.78 (0.01)
PC2	LY	0.10 (0.008)
X_l	PC1	0.81 (0.01)
X_p	PC1	0.76 (0.01)
X_l	PC2	-0.39 (0.006)
X_p	PC2	0.62 (0.01)

¹ X_l = common latent factor related to level of production in early lactation; X_p = common latent factor related to the lactation persistency; PC1 = principal component related to the average lactation yield; PC2 = principal component related to lactation persistency; LY = 305-d milk yield; r_g = genetic correlation.

(Table 4). These estimates, together with the results of PC2, agree with the conclusion that persistency of milk yield should be considered as a different trait in different parities (Gengler, 1996; Jamrozik et al., 1998; van der Linde et al., 2000).

Genetic correlations among the different multivariate measures and LY, estimated by bivariate analysis are reported in Table 5. The weak linkage between latent factors is confirmed. Both X_l and X_p were highly correlated with total lactation yield. As expected, PC1 and PC2 were unrelated and had high and low correlations with LY, respectively.

Genetic relationships between different multivariate measurements were also of interest (Table 5). A high correlation was observed between variables related to the level of production; i.e., X_l and PC1. Finally, whereas X_p and PC1 were highly related, a negative correlation was observed between the latent factor associated with the level of production in the first part of lactation (X_p) and the principal component related to lactation persistency (PC2).

Table 4. Genetic parameters (SE in parentheses) estimated with the univariate analysis

Parameter ¹	Trait ^{2,3}				
	X_l	X_p	PC1	PC2	305-d yield
σ^2_A	0.07329 (0.007)	0.05387 (0.005)	0.4427 (0.038)	0.05264 (0.007)	168,689 (1438)
σ^2_p	0.08143 (0.007)	0.05116 (0.004)	0.4654 (0.030)	0.02954 (0.006)	214,446 (1155)
σ^2_e	0.39748 (0.003)	0.29527 (0.002)	1.4005 (0.012)	0.66050 (0.005)	448,220 (3821)
h^2	0.13 (0.0129)	0.13 (0.0130)	0.19 (0.016)	0.07 (0.009)	0.20 (0.016)
r_p	0.28 (0.006)	0.26 (0.0062)	0.39 (0.006)	0.11 (0.006)	0.43 (0.006)

¹Parameters: σ^2_A = additive genetic variance; σ^2_p = permanent environmental variance; σ^2_e = residual variance; r_p = repeatability.

²Traits: X_l = common latent factor related to level of production in early lactation; X_p = common latent factor related to the lactation persistency; PC1 = principal component related to the average lactation yield; PC2 = principal component related to lactation persistency.

³Variances of X_l , X_p , PC1, and PC2 are dimensionless; 305-d yield variance is expressed in kg^2 .

The results reported above confirm that the main traits of lactation curve shape can be altered by selection. Moreover, the different multivariate measurements offer several options of improvement over previously proposed approaches for selection to alter the shape of the lactation curve. The low genetic correlation between common latent factors suggests that the initial and final parts of lactation curve can be changed independently, at least to a certain extent, and in both cases an improvement of total yield can be obtained, due to the high correlations between total yield and both of these factors. The use of PC2 allows for the genetic modification of lactation persistency, although to a lesser extent than X_p , while keeping constant total lactation yield.

CONCLUSIONS

Multivariate techniques used in this study were able to derive 2 new variables related to prominent aspects of lactation curve shape from the phenotypic correlation matrix of TD milk yields. The amount of original variation explained by extracted variables was similar for MFA and PCA. However, differences were found in their distributions, with variance more equally partitioned in MFA than in PCA, and in the meaning of new variables. The MFA approach yielded 2 indexes of main aspects of lactation curve shape, 1) daily yield in early lactation, and 2) persistency. In PCA, the first variable expressed level of production for the whole lactation and accounted for most of the variance; the second variable was related to persistency, but accounted only for a small amount of variation.

Common latent factors had moderate heritabilities, thus confirming the possibility of being able to manipulate the shape of the lactation curve via genetic selection. Correlations between factors were small and all factors were highly correlated with total lactation yield. These features suggest that these variables can be used for developing different selection strategies. A common goal in many dairy-breeding systems is to increase total lactational yield without increasing occurrence of diseases or reproductive failure (Muir et al., 2004). This result can be achieved by selecting on the factor associated with lactation persistency and keeping constant (or selecting against) the factor related to level of production in early lactation. On the other hand, in pasture-based seasonal calving systems, where the key breeding objective is to achieve the highest pregnancy rate in the shortest period (Haile-Mariam et al., 2003; Horan et al., 2005), more emphasis could be put on selection against the X_1 factor.

The choice of the PC1 as breeding goal is expected to yield similar results as selection for total lactation yield,

because genetic correlation between the 2 traits is high. Persistency can be improved without modifying total lactation yield by selecting on PC2.

The structure of the multivariate indicators, i.e., relationships between TD yields and scores they are assigned for the calculation of latent variables, originates from the correlation matrix of observed TD records. Specific values estimated in this work are likely to reflect effects of selection and the farming system of the breed considered. Additional investigation is required to check robustness of the variables related to lactation curve shape in other breeds and in other farming systems.

ACKNOWLEDGMENTS

Research was funded by the Italian Ministry of University and Research (grant PRIN 2005).

REFERENCES

- Cole, J., and P. VanRaden. 2005. Genetic evaluation and best prediction of lactation persistency. *J. Dairy Sci.* 88(Suppl. 1):379. (Abstr.)
- 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.
- Druet, T., F. Jaffrezic, D. Boichard, and V. Ducrocq. 2003. Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. *J. Dairy Sci.* 86:2480–2490.
- Druet, T., F. Jaffrezic, and V. Ducrocq. 2005. Estimation of genetic parameters for test day records of dairy traits in the first three lactations. *Genet. Sel. Evol.* 37:257–271.
- Gengler, N. 1996. Persistency of lactation yields: A review. *Interbull Bull.* 12:87–96.
- Gengler, N., A. Tijani, G. R. Wiggans, and J. C. Philpot. 2001. Indirect estimation of (co)variance functions for test-day yields during first and second lactations in the United States. *J. Dairy Sci.* 84:542.
- Gilmour, A. R., B. R. Cullis, S. J. Welham, and R. Thompson. 2000. ASREML manual. New South Wales Dept. Agric., Orange, Australia.
- 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.
- Horan, B., P. Dillon, D. P. Berry, P. O'Connor, and M. Rath. 2005. The effect of strain of Holstein-Friesian, feeding system and parity on lactation curves characteristics of spring-calving dairy cows. *Livest. Prod. Sci.* 95:231–241.
- 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.
- Jakobsen, J. H., R. Rekaya, J. Jensen, D. A. Sorensen, P. Madsen, D. Gianola, L. G. Cristiansen, and J. Peersen. 2003. Bayesian estimates of covariance components between lactation curve parameters and disease liability in Danish Holstein cows. *J. Dairy Sci.* 86:3000–3007.
- Jamrozik, J., D. Gianola, and L. R. Schaeffer. 2001. Bayesian estimation of genetic parameters for test day records in dairy cattle using linear hierarchical models. *Livest. Prod. Sci.* 71:223–240.

- 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.
- Kamidi, R. E. 2005. A parametric measure of lactation persistency in dairy cattle. *Livest. Prod. Sci.* 96:141–148.
- Kirkpatrick, M., and K. Meyer. 2004. Direct estimation of genetic principal components. Simplified analysis of complex phenotypes. *Genetics* 168:2295–2306.
- Kistemaker, G. J. 2003. Comparison of persistency definitions in random regression test day models. *Interbull Bull.* 30:96–98.
- Krzanowsky, W. J. 2003. *Principles of multivariate analysis.* Oxford University Press Inc., New York, NY.
- Macciotta, N. P. P., D. Vicario, and A. Cappio-Borlino. 2005. Detection of different shapes of lactation curve in dairy cattle by empirical mathematical models. *J. Dairy Sci.* 88:1178–1191.
- Macciotta, N. P. P., D. Vicario, C. Dimauro, and A. Cappio-Borlino. 2004. A multivariate approach to modelling shapes of individual lactation curves in cattle. *J. Dairy Sci.* 87:1092–1098.
- 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.
- Muir, B. L., J. Fatehi, and R. L. Schaeffer. 2004. Genetic relationships between persistency and reproductive performance in first-lactation Canadian Holsteins. *J. Dairy Sci.* 87:3029–3037.
- Olori, V. E., S. Brotherstone, W. G. Hill, and B. J. McGuirk. 1999a. 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.
- Olori, V. E., W. G. Hill, B. J. McGuirk, and S. Brotherstone. 1999b. Estimating variance components for test day milk records by restricted maximum likelihood with a random regression animal model. *Livest. Prod. Sci.* 61:53–63.
- 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 Institute. 1996. *User's Guide: Statistics.* SAS Inst., Inc., 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.
- Sieber, M., A. E. Freeman, and P. N. Hinz. 1988. Comparison between factor analysis from a phenotypic and genetic correlation matrix using linear type traits of Holstein dairy cows. *J. Dairy Sci.* 71:477–484.
- 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.
- Swalve, H. H. 1995. Genetic relationship between dairy lactation persistency and yield. *J. Anim. Breed. Genet.* 112:303–311.
- Togashi, K., and C. Y. Lin. 2003. Modifying the lactation curve to improve lactation milk and persistency. *J. Dairy Sci.* 86:1487–1493.
- Togashi, K., and C. Y. Lin. 2004. Development of an optimal index to improve lactation yield and persistency with the least selection intensity. *J. Dairy Sci.* 86:3047–3052.
- van der Linde, R., A. Groen, and J. de Jong. 2000. Estimation of genetic parameters of milk production in dairy cattle. *Interbull Bull.* 25:113–116.
- van der Werf, J. H. J., M. E. Goddard, and K. Meyer. 1998. The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. *J. Dairy Sci.* 81:3300–3308.
- 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.