GENETICS AND BREEDING

Analysis of Test Day Yield Data of Costa Rican Dairy Cattle

B. VARGAS,*,† E. PEREZ,* and J.A.M. VAN ARENDONK†
*Escuela de Medicina Veterinaria, Universidad Nacional de Costa Rica, PO Box 304-3000, Heredia, Costa Rica
†Animal Breeding and Genetics Group, Wageningen Institute of Animal Sciences, Wageningen Agricultural University, PO Box 338, 6700 AH Wageningen, The Netherlands

ABSTRACT

Estimates of variance components for test day records in an animal model that considered multiple traits over multiple lactations were calculated using REML methodology. Test day records were classified into 11 periods within first and later lactations. Missing ancestors in the relationship matrix were classified in genetic groups. Data were collected from Costa Rican dairy farms. Estimates of components for total and additive genetic variance were clearly heterogeneous during the lactation. Heritabilities for traits in later parities were slightly higher than those for traits in first parity. Heritabilities were highest for records of midlactation. Phenotypic and genetic correlations for adjacent test days were close to 1. Phenotypic correlations were lower than genetic correlations. Heterogeneity of variances during the lactation suggests the adequacy of the multiple-trait test day model to describe milk yield during the lactation. When missing ancestors were allocated to a single base population instead of genetic groups, the estimates of residual variance were lower, and the estimates of genetic variance and genetic correlations were higher. When standardized records were used instead of actual test day records, the estimates of residual and total variance were lower, and the estimates of genetic variance were higher. Consequently, estimates of heritability and genetic correlations were also higher. Use of standardized data obtained by interpolation procedures is not advised for estimation of genetic variance components in a test day model. (**Key words**: test day yields, genetic parameters, genetic groups)

Abbreviation key: **TDM** = test day model.

INTRODUCTION

The use of test day models (**TDM**) for the analysis of traits related to milk yield has received considerable attention during recent years (8, 10, 12, 17, 20, 21, 26). Test day models have been defined as a statistical procedure that considers all genetic and environmental effects directly on a test day basis (12). This methodology has several advantages over the traditional 305-d model. The TDM maximizes the amount of information to be gathered for each animal and avoids the use of factors to extend partial lactation records (26). Another important advantage is that TDM account for factors that are specific to each test day, such as management groups within a herd on a test day (5, 15). In addition, the problem of differences in the amount of information contributing to the 305-d prediction is overcome (15).

Various statistical models have been proposed for the analysis of test day records, and the most widely used model has been the repeatability model (12). Under this model, consecutive test day samples from the same lactation are considered as repeated observations on the same trait, and a permanent environmental effect accounts for environmental similarities between different test days within the same lactation. A random regression test day model has also been proposed (17). In that model, two sets of regressions of milk production on DIM are performed, one fixed for cows in the same subclass and one random, accounting for deviations of the cows with respect to the fixed parameters in the group. Multiple-lactation versions of this model have also been applied for the analysis of somatic cell score in Germany (14, 15). Recently, a model for multiple traits and multiple lactations has been suggested that integrates linear functions of test day observations into a canonical index (8, 26), and an alternative model used covariance functions (7).

Some problems of the TDM have not yet been elucidated clearly. A major disadvantage of the repeatability model is the heterogeneity of the residual variance during the lactation (12). The multiple-trait model has been proposed as a solution to this problem. However, the increase in the amount of information, which can be nearly 10 times higher

Received March 5, 1997.
Accepted August 14, 1997.

than with the traditional schemes, represents a large computational burden (8, 12, 26).

Estimates of genetic parameters for test day records have been presented previously (6, 8, 10, 11, 13, 26). Estimates of heritability were sometimes similar to those reported for 305-d milk yield, especially for midlactation records, and estimates were lower for records at the beginning and at the end of the lactation (6, 8, 26). In general, high phenotypic and genetic correlations have been reported and were close to 1 for consecutive records (6, 8, 11, 26). However, analysis frequently has been restricted to complete lactations with a minimum number of test days regularly distributed during the lactation (8, 10, 13). In some cases, standardized data have been calculated by the use of a test interval method, which allows the estimation of test yields at fixed times within the lactation (18). The effect of the use of selected and standardized data on the estimates of variance is not very clear. No heritability estimates have been presented using a multiple-trait approach for actual instead of standardized data without previous selection on a minimum number of records per lactation.

Most of the research on TDM has been carried out in countries with well-established breeding programs, official milk recording schemes, and accurate pedigree information. In Costa Rica, the total number of dairy cattle is relatively small. A local breeding program has not yet been established successfully, and genetic improvement has relied mainly on the importation of semen. Use of local unproven bulls is still frequent. Official milk recording schemes have only been implemented in a small proportion of the population, pedigree information is not always available, and breed variation is high at the farm level. However, a considerable increase in milk yield of 111.4 kg/yr per lactation has been reported for Holstein cattle from 1979 to 1992 (24). This increase has been achieved mainly by the improvement of nutritional and management conditions. The genetic component seems to have played a less important role (19, 24). The TDM have been suggested as the method of choice for the analysis of milk yield traits in order to maximize the use of all available information (12). This result becomes even more important in countries with few cattle and without well-established milk recording schemes. In addition, the use of a genetic grouping strategy provides a powerful tool to deal with cases for which accurate pedigree information is not available (25).

The objective of this analysis is to determine genetic and environmental factors affecting daily milk yield of Costa Rican dairy cattle using an animal model for multiple traits and multiple lactations. Given the characteristics of the data file, the effects are also investigated using a genetic grouping strategy and actual instead of standardized data on the final estimates of variance components and genetic parameters.

MATERIALS AND METHODS

Data Source

Analyses were performed on data provided by Universidad Nacional de Costa Rica. Those data were collected from 1980 to 1996 on dairy farms from five regions in Costa Rica. Those farms participated in a project that focuses on the collection and analysis of information about health, production, and reproduction (2). Test day milk yields were entered using the software package VAMPP (9) by the university staff or directly by farmers. Cows included in the analysis were of the following breed types: Holstein, Jersey, Guernsey, Brown Swiss, and combinations of *Bos taurus* × *Bos indicus* and *Bos taurus* × *Bos taurus* breeds.

Statistical Model

The following animal model for multiple traits and multiple lactations was used to analyze first parity test day records.

$$[y_1 - y_{11}]_{ijkl} = \mu + HYS_i + b_1(age_j) + b_2(age_j)^2 + b_3(day_k) + A_l + E_{ijkl} \quad [1]$$

where

$[y_1 - y_{11}]_{ijkl}$ = test day records for 11 milk production traits analyzed by defining 11 periods within the lactation such that trait y_1 represented samples obtained between d 4 and 16 in the lactation, trait y_2 represented samples obtained between d 15 and 31 in the lactation, and traits y_3 to y_{11} represented samples obtained in subsequent periods of 30.4 d;
μ = population mean;
HYS_i = fixed effect of herd-year-season i in which the sample was taken;
b_1, b_2 = linear and quadratic effects of age at calving j (age_j) on test day yield;
b_3 = linear effect of day of sampling k (day_k);
A_l = random animal effect for which relationship matrix was used; and
E_{ijkl} = random residual.

Two seasons were defined according to the ecological region where the farms were located (4). The length of the seasons ranged from 4 to 8 mo accounting for climatic characteristics of the region.

For later parities, records in different lactations for the same period were treated as repeated records for the same trait (i.e., y_{12} to y_{22}). To account for repeated samples on the same animal, a random permanent environmental effect was added to Model [1] for test day records in later parities.

Data Editing

Given the aforementioned definition of traits and effects, the following editing procedures were undertaken. All data from a cow were removed when the breed type could not be classified in one of the predefined groups. Data from one lactation was removed when the previous gestation length was >295 d or <240 d and when age at calving was <18 mo or >42 mo for the first parity group. Lactation data were also removed when age at calving in later parities was <28 or >150 mo and when the previous dry period was <15 d. A maximum of 10 lactations per cow was considered.

A minimum of 5 lactations within each herd-year-season of calving class was required. When this number was lower, an attempt was made to join adjacent seasons of consecutive years. A maximum of three seasons, when necessary, were joined. If the final number of lactations in the newly formed herd-year-season class was still <5, the respective lactations were removed.

The number of samples per lactation was not restricted. Test day records within a lactation were eliminated when the day of sampling was <6 or >304 and when milk yield was <4 or >70 kg. When additional samples were available within a period and within a parity, only the first sample was used in the analysis, and the others were removed.

Analytical Procedures

Because of incomplete data and pedigree records, a procedure for genetic grouping was used. Missing ancestors were classified in genetic groups according to selection path, breed type, and estimated birth year (25). Animals included in the relationship matrix were cows having own information (n = 28,417) and sires (n = 1161). Sires included 656 non-AI sires, 353 AI sires with ≥5 daughters in the data file, and 152 sires of sires. Not all sires of sires had an identified sire themselves. Four generations of AI sires were included in the relationship matrix, when available.

TABLE 1. Groups of test day records[1] of first and later parities in the analysis.

First parity		Later parities	
Group	Traits	Group	Traits
1	y_1 and y_7	7	y_{12} and y_{18}
2	y_2 and y_8	8	y_{13} and y_{19}
3	y_3 and y_9	9	y_{14} and y_{20}
4	y_4 and y_{10}	10	y_{15} and y_{21}
5	y_5 and y_{11}	11	y_{16} and y_{22}
6	y_6	12	y_{17}

[1]Traits y_1 and y_{12} are test day records between d 4 and 16 of first and later parities, respectively. Traits y_2 and y_{13} are test day records between d 15 and 31 of first and later parities, respectively. Traits y_3 to y_{11} and y_{14} to y_{22} are test day records between d 30 and 306 of first and later parities, divided into 30.4-d periods.

Missing sires or sires with <5 daughters in the data file were allocated to genetic groups. Dams of cows without own information on milk production and dams of sires, although identified in some cases, were also allocated to genetic groups. Following this strategy, a total of 208 genetic groups were formed.

The variance-covariance matrix for the 22 traits was calculated by REML using a superlinearly converging quasi-Newton algorithm with exact analytical derivatives as implemented in the REML-VCE software (3). Because of the high number of equations and the limited computing resources, the following steps were performed to estimate all the elements of the variance-covariance matrix.

Step 1. In order to get starting values for REML-VCE, phenotypic correlations were calculated using SAS least squares analysis (16).

Step 2. Subsequently, all traits were analyzed separately using univariate analysis and Model [1]. In this way, estimates for residual, genetic, and permanent environmental variance components were obtained for each trait separately.

Step 3. Six groups of traits within parity group were formed (Table 1); five of them had two traits, and the sixth group had one trait. Thirty different REML-VCE runs were performed following a strategy combining two different groups of traits within parity level. For every run, starting values for the estimates of residual and genetic variance components were specified based on estimates of variance calculated from the phenotypic correlations and the estimates of variance components obtained from the univariate analysis. Five heritability estimates were obtained for every trait using this strategy. Similarly, five estimates of genetic correlation were obtained for traits in the same group and one estimate for traits in different groups. Standard errors of the estimates

were obtained based on the approximated Hessian matrix produced by the quasi-Newton optimizer implemented in REML-VCE.

Step 4. All estimates of heritability and correlations were pooled. When more than one estimate of heritability or genetic correlation was available, as specified previously, the respective mean is reported in the final variance-covariance matrix.

Two additional analyses were carried out using a subset of data that included only midlactation first parity traits, y_4 to y_7. One analysis was performed in order to evaluate the effect of genetic groups on the estimates of variance components. Variance components for the traits y_4 to y_7 were calculated for both cases. For the first case, the genetic groups were coded following the strategy previously explained, and, for the second case, unidentified ancestors were coded as missing (i.e., genetic grouping was ignored), and consequently all missing ancestors were joined in a single base population, as frequently performed.

A second analysis of the same subset of data was done to evaluate effects of using standardized test day records, instead of actual records, for the calculation of variance components. Standardized records were calculated for the last day of the four different periods using simple linear interpolation between the two closest records around the fixed day. Following this strategy, a standardized test day record was calculated for all four periods. In order to be able to calculate standardized yields using linear interpolations, animals were required to have at least one test day record in or before period 4 and one in or after period 7. Variance components for standardized test day records were estimated using Model [1] but excluding the covariate. A total of 9648 lactations were used; 76% had records for all four traits analyzed, and the other 24% lacked at least one record in one of the periods. To enable a good comparison, variance components for test day records of actual data in this subset were estimated using Model [1].

RESULTS AND DISCUSSION

Data Description

The number of completed lactations (at least 305 d) represented 52.6% of the total number of lactations in the original data file, before editing. The number of samples was reduced considerably by the editing procedure (Table 2). The mean number of samples per lactation before editing was 14.2 ± 11.7 (SD); number of samples ranged from 1 to 90. After editing, this mean decreased to 7.31, which was similar to values recently reported for other countries (20, 26). The main causes of this reduction were additional samples in the same period, which constituted 35.5% of the test day records, and samples taken before d 5 or after d 305 in the lactation, which constituted 10.1% of the samples. A total of 24.6% of the samples pertained to first lactation cows. The mean milk yield per lactation, calculated from 40,318 finished and unfinished lactations with >250 d was 4427 ± 1685 (SD). This number is substantially lower than has been reported for Holstein cattle in the US and Germany (20, 22).

TABLE 2. Structure of the data file before and after editing.

	Data file	
Parameter	Original	Final
Farms included	230	222
Breed types	107	6
Cows	29,702	28,417
Lactations	62,405	57,891
Individual samples	886,253	423,366

The initial number of breed types was high (Table 2). However, the number of samples per breed type was very low in some cases. For the analysis, breed types were joined in six different groups: Holstein (56.61%), Jersey (19.95%), Brown Swiss (9.20%), *Bos taurus* × *Bos indicus* crosses (8.06%), *Bos taurus* × *Bos taurus* crosses (3.83%), and Guernsey (0.99%). Most common breed crosses involved Holstein or Jersey and *Bos indicus* breeds, such as Brahman.

Only a small fraction of cows had both parents identified (Table 3), which is common in countries with relatively new breed registration and milk recording organizations. This fact is demonstrated by the high proportion of non-AI sires, which was approximately 56.5% of the total number of sires. However, a total of 54.2% of the cows with known sires were daughters of AI sires.

The weighted means for daily milk yield (Table 4) were 14.0 and 16.9 kg for first and later parities, respectively. Standard deviations for daily milk yield were higher than previously reported (6, 20), partly because of the high number of breed types included and the wider range in management practices.

Variance-Covariance Matrix and Genetic Parameters

Estimates of heritability ranged from 0.15 to 0.23 and from 0.13 to 0.24 for traits of first and later parities, respectively (Table 5). The standard errors of these estimates ranged from 0.02 to 0.03 for test

TABLE 3. Number of individuals (cows and sires) in the relationship matrix according to the existence or nonexistence of identified parents.

	Identified parents									
Class	Both		Only dam		Only sire		None		Total	
	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)
Cows	3398	(11.5)[1]	922	(3.1)	12,173	(41.2)	11,924	(40.3)	28,417	(96.1)
Sires	0	(0.0)	0	(0.0)	425	(1.4)	736	(2.5)	1161	(3.9)
Total	3398	(11.5)	922	(3.1)	12,598	(42.6)	12,660	(42.8)	29,578	(100.0)

[1]Relative frequency given in parentheses.

days in the first parity and were always <0.01 for later parities. Heritabilities were slightly higher for test day records during midlactation than for those at the beginning or end of lactation. In general, heritabilities for daily milk yield are low compared with literature estimates. Higher estimates of heritability for test day records in first lactation cows using a sire model have been reported (8, 11, 13). Heritability in those studies ranged between 0.17 to 0.27 (8), 0.27 to 0.39 (11), and 0.10 to 0.37 (13). Heritability estimates for test day records using a multiple-trait approach with a sire model and canonical transformation of the data have been reported (8). The heritability estimates reported with this procedure ranged from 0.24 to 0.35, which were considerably higher than those found in the present analysis. A characteristic of all these studies was that lactations were required to have a minimum number of samples, and partial lactation records were sometimes removed. These two restrictions implied the elimination of a large quantity of data.

Estimates of phenotypic correlations in the present analysis ranged from 0.58 to 0.90 and from 0.56 to 0.91 for test day records in first and later parities, respectively (Table 5). The estimates of genetic correlations ranged from 0.49 to 1.0 and 0.55 to 1.0 for test days in first and later parities, respectively. Standard errors for genetic correlations were between 0.02 and 0.07 for test days in first parity and always lower than 0.01 in later parities. The estimates of the correlations among daily milk yield from different periods were inversely related to the relative distance within the lactation (i.e., correlation decreased as the interval increased). However, this trend was not always consistent, especially for first parity traits (Table 5). Genetic correlations presented in the literature ranged from 0.39 to 0.95 (8), 0.73 to 0.99 (11), and 0.43 to 0.95 (15). Differences between the results reported in literature and those found in the present analysis may be due to the differences in the model used, the quantity of data available, and the differences in the definition of test day records.

TABLE 4. Total number of records (n), arithmetic mean, and standard deviation of test day records[1] of cows in first and later parities.

	First parity				Later parities		
Variable	n	$\bar{X}$	SD	Variable	n	$\bar{X}$	SD
y_1	6410	14.8	4.8	y_{12}	20,225	19.3	6.8
y_2	8236	16.1	5.4	y_{13}	26,104	20.7	7.5
y_3	11,561	16.0	5.8	y_{14}	36,686	20.5	7.7
y_4	11,441	15.4	5.8	y_{15}	36,067	19.4	7.5
y_5	10,917	14.6	5.7	y_{16}	34,372	18.1	7.1
y_6	10,661	13.9	5.5	y_{17}	33,420	16.9	6.7
y_7	10,123	13.3	5.3	y_{18}	31,849	15.7	6.3
y_8	9714	12.8	5.1	y_{19}	30,475	14.6	5.9
y_9	9088	12.3	4.9	y_{20}	28,737	13.4	5.5
y_{10}	8083	11.7	4.8	y_{21}	24,929	12.2	5.1
y_{11}	6174	11.2	4.6	y_{22}	18,094	11.4	4.8
Total	102,408	14.0			320,958	16.9	

[1]Traits y_1 and y_{12} are test day records between d 4 and 16 of first and later parities, respectively. Traits y_2 and y_{13} are test day records between d 15 and 31 of first and later parities, respectively. Traits y_3 to y_{11} and y_{14} to y_{22} are test day records between d 30 and 306 of first and later parities, divided into 30.4-d periods.

A previous result has been reported using a similar methodology (26). Estimates of heritability were between 0.14 and 0.23 for first parity traits. Phenotypic and genetic correlations among test day records ranged from 0.20 to 0.63 and 0.50 to 1.0, respectively. Those results are generally in agreement with our analysis; however, the differences between phenotypic and genetic correlations in the mentioned study (26) were much higher, and the trend in genetic correlations was more consistent. In our analysis, mainly for first parity traits, the trend in the correlation was not always consistent. As an example, the genetic correlation between y_1 and y_9 was lower than the respective value between y_1 and y_{10}. A possible reason for this result is the use of unstandardized data without restrictions on the number of samples per lactation. In addition, the size of our data file was relatively small, and, consequently, the estimates of standard errors for genetic correlations of first parity traits in our analysis were sometimes close to 0.07. As was expected from the larger number of records, the results obtained in our analysis in later parities were more consistent with estimates in the literature.

Variances during the lactation were clearly heterogeneous. For test day yields in first parity (Figure 1), the estimate of total variance for y_1 and y_{11} were 11.5 and 8.9 kg^2, respectively. Estimates of residual variance for the same traits were 8.8 and 7.8 kg^2, respectively. For test day yields in later parities (Figure 2), the estimate of total variance for y_{12} was 19.3 kg^2, which then increased to around 20.5 kg^2 for y_{13} and y_{14} and decreased to a value of 13.3 kg^2 for y_{22}. The residual variance decreased progressively from 14.8 kg^2 for y_{12} to 8.0 kg^2 for y_{22}. Estimates of residual variance have been reported in the range of 3.5 to 9.5 kg^2 (5) and 5.16 to 8.31 kg^2 (8). The higher values found in the present analysis were partly due to the lower estimates of heritability and to the use of samples from later lactations.

The estimates of repeatability, defined as the ratio of genetic plus permanent environmental variance, divided by the total phenotypic variance, increased from 0.23 for y_{12} to 0.42 for y_{20} and then decreased to 0.35 for y_{22}. Estimates of repeatability of 0.52, 0.71, and 0.66 for the first, second, and third trimester of lactation have been given previously (6). However, in that report, the permanent environmental effect was defined for repeated observations within the same trimester in the lactation, and, thus, the repeatability

TABLE 5. Heritabilities (diagonal) and genetic (below diagonal) and phenotypic (above diagonal) correlations of test day records in first and later parities.[1]

	First parity										
	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}
y_1	0.23	0.82	0.78	0.73	0.70	0.68	0.66	0.65	0.62	0.60	0.58
y_2	0.81	0.15	0.89	0.85	0.82	0.79	0.77	0.75	0.73	0.70	0.65
y_3	0.85	0.96	0.20	0.89	0.87	0.83	0.82	0.80	0.78	0.74	0.70
y_4	0.75	0.77	0.88	0.21	0.90	0.87	0.85	0.83	0.81	0.77	0.72
y_5	0.71	0.72	0.83	0.92	0.17	0.90	0.87	0.85	0.82	0.78	0.74
y_6	0.60	0.54	0.82	0.89	0.90	0.15	0.90	0.87	0.85	0.81	0.76
y_7	0.63	0.54	0.79	0.87	0.97	0.99	0.20	0.90	0.88	0.83	0.78
y_8	0.72	0.75	0.89	0.86	0.94	0.96	0.96	0.23	0.90	0.85	0.80
y_9	0.49	0.76	0.78	0.85	0.98	0.96	0.97	0.97	0.19	0.89	0.84
y_{10}	0.65	0.75	0.74	0.88	0.84	0.94	0.84	0.90	0.96	0.23	0.87
y_{11}	0.62	0.77	0.77	0.71	0.78	0.87	0.77	0.60	0.81	0.97	0.16
	Later parities										
	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}	y_{21}	y_{22}
y_{12}	0.13	0.85	0.83	0.79	0.77	0.75	0.72	0.70	0.66	0.62	0.56
y_{13}	0.93	0.16	0.90	0.87	0.84	0.81	0.79	0.77	0.72	0.66	0.60
y_{14}	0.96	1.00	0.21	0.90	0.87	0.85	0.82	0.80	0.75	0.70	0.63
y_{15}	0.89	0.95	0.99	0.20	0.91	0.88	0.85	0.82	0.78	0.72	0.65
y_{16}	0.85	0.93	0.96	1.00	0.22	0.91	0.88	0.84	0.80	0.74	0.66
y_{17}	0.79	0.88	0.93	0.98	1.00	0.22	0.90	0.87	0.82	0.76	0.69
y_{18}	0.75	0.82	0.90	0.94	0.98	0.99	0.24	0.90	0.85	0.79	0.71
y_{19}	0.71	0.81	0.86	0.93	0.95	0.97	1.00	0.24	0.88	0.82	0.73
y_{20}	0.68	0.78	0.82	0.87	0.90	0.94	0.97	0.99	0.23	0.87	0.78
y_{21}	0.61	0.69	0.76	0.82	0.84	0.87	0.94	0.97	0.99	0.20	0.84
y_{22}	0.55	0.62	0.67	0.75	0.79	0.80	0.89	0.91	0.94	0.98	0.18

[1]Traits y_1 and y_{12} are test day records between d 4 and 16 of first and later parities, respectively. Traits y_2 and y_{13} are test day records between d 15 and 31 of first and later parities, respectively. Traits y_3 to y_{11} and y_{14} to y_{22} are test day records between d 30 and 306 of first and later parities, divided into 30.4-d periods.

TABLE 6. Variance-covariance components estimates of residual and genetic effect in traits y_4, y_5, y_6, and y_7[1] obtained by using actual data with genetic groups, actual data with a base population, and standardized data with genetic groups.

	Actual data with genetic groups[2]				Actual data with the base population				Standardized data with genetic groups			
	y_4	y_5	y_6	y_7	y_4	y_5	y_6	y_7	y_4	y_5	y_6	y_7
Residual	8.02	8.04	7.74	7.44	7.78	7.54	7.42	6.62	6.52	6.40	6.23	6.09
Genetic	1.96	1.18	1.20	1.41	2.48	1.90	1.68	2.43	2.11	1.94	1.87	1.84
Total	9.98	9.22	8.94	8.85	10.26	9.44	9.10	9.05	8.63	8.34	8.10	7.93

[1]Traits y_4, y_5, y_6, and y_7 comprised test days records taken in the periods between d 30 to 61, 60 to 91, 90 to 121, and 120 to 151 in the lactation, respectively.

[2]Genetic groups (n = 208) were defined according to estimated birth year, selection path, and breed type.

should be compared with the phenotypic correlation between test day yields (Table 5) and not the repeatability across lactations as in the present analysis.

Genetic Grouping Versus a Single Base Population

Few animals in the data file had both parents identified (Table 3). In addition, the data spanned a relatively long period, and pedigree could not be traced for all animals. A large proportion of the sires was imported. Assignment of missing ancestors to a single base population instead of to genetic groups resulted in a decrease in the residual variance and an increase in the genetic variance (Table 6). This trend was the same for all four traits studied (y_4 to y_7). As a consequence, heritabilities were also higher (Table 7). The effect of the use of a conditional model to account for selected base animals has been previously addressed (23). Genetic grouping to account for effects of selection is more likely to have a major effect on the estimates of variance components when the number of animals lacking pedigree information is high (25), as is the case in our analysis. When genetic groups are assigned, missing ancestors are allocated to different groups, which resulted in decreased additive variance; this decrease is a reflection of genetic differences between breeds and genetic trend within breeds.

Standardized Versus Actual Data

Some discrepancies have been found among studies in heritability estimates for test day records. These differences may be due to the use of actual data in some studies instead of standardized records. Standardized records are obtained by the test interval method, which uses interpolation or extrapolation on observed test day records to obtain yields at fixed points in time.

The consequences of using standardized test day records for the calculation of variance components in a TDM have been quantified (Table 6). Standardization of milk yield to fixed days in lactation resulted in a decrease in the residual variance and an increase in the genetic variance compared with the results that were obtained when actual data were used. Standardization also reduced the total variance, and, consequently, the estimates of heritability increased (Table 7). Standardization of test day records can be compared with averaging actual records flanking the time for which the yield is to be calculated. Residual correlations between subsequent yields are substantially <1. Consequently, error variance on average is lower than that on a single observation. The effect on the genetic variance is small when the genetic correlation is high, as is the case here (Table 5). The

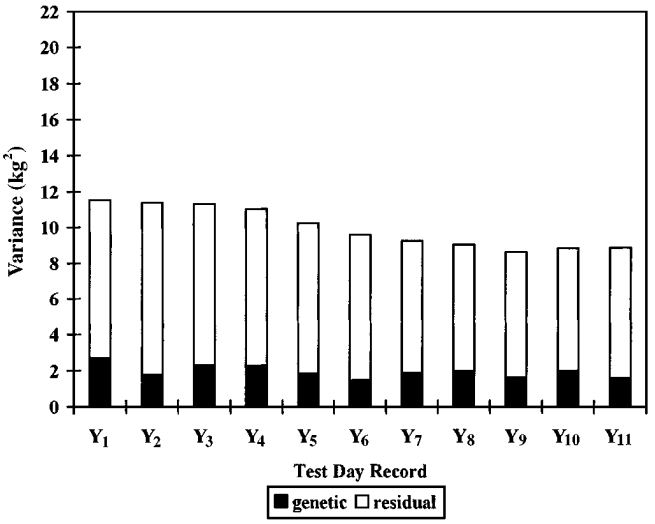

Figure 1. Estimates of genetic and residual variance for test day records of first parity. Trait y_1 comprised test day records between d 4 and 16, trait y_2 comprised test day records between d 15 and 31, and traits y_3 to y_{11} comprised test day records between d 30 and 306, divided into 30.4-d periods.

TABLE 7. Heritabilities and genetic correlations for traits y_4, y_5, y_6, and y_7[1] obtained by using actual data with genetic groups, actual data with a base population, and standardized data with genetic groups.

	Actual data with genetic groups[2]				Actual data with the base population				Standardized data with genetic groups			
	y_4	y_5	y_6	y_7	y_4	y_5	y_6	y_7	y_4	y_5	y_6	y_7
y_4	0.20	0.89	0.81	0.78	0.24	0.95	0.93	0.88	0.24	0.99	0.88	0.87
SE	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.00)	(0.01)	(0.01)
y_5		0.13	0.91	0.89		0.20	0.94	0.95		0.23	0.94	0.90
SE		(0.01)	(0.01)	(0.01)		(0.01)	(0.01)	(0.01)		(0.01)	(0.01)	(0.01)
y_6			0.13	0.94			0.18	0.98			0.23	0.94
SE			(0.01)	(0.01)			(0.01)	(0.01)			(0.00)	(0.00)
y_7				0.16				0.27				0.23
SE				(0.01)				(0.01)				(0.00)

[1]Traits y_4, y_5, y_6 and y_7 comprise test days records taken in the periods between d 30 to 61, 60 to 91, 90 to 121, and 120 to 151 in the lactation, respectively.

[2]Genetic groups (n = 208) were defined according to estimated birth year, selection path, and breed type.

results in Tables 6 and 7 are in agreement with these expectations.

The reasons are less obvious for the increase in additive genetic variance caused by standardization. The high correlations between yields might have played a role. The correlations for the standardized data were clearly higher than those for actual data. In the calculation of standardized yields in subsequent months, one actual record contributes to two standardized records (i.e., the preceding and the following), which introduces an extra source of covariance. When the interval between actual test records increased, the size of the additional covariance increased.

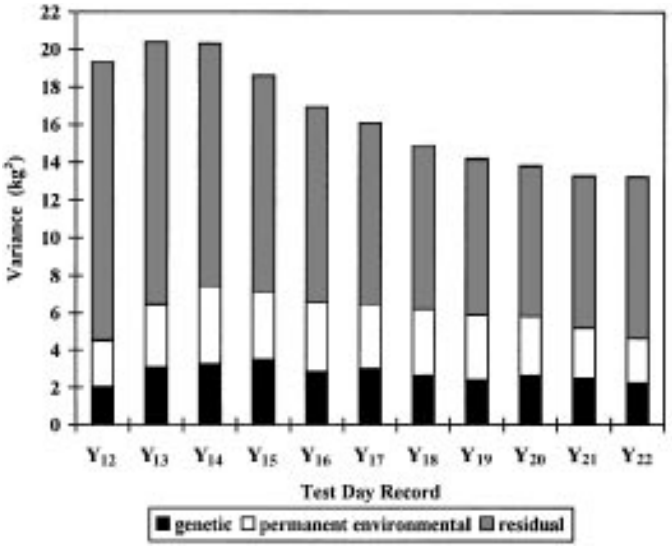

Figure 2. Estimates of genetic, permanent environmental, and residual variance for test day records of later parities. Trait y_{11} comprised test day records between d 4 and 16, trait y_{12} comprised test day records between d 15 and 31, and traits y_{13} to y_{22} are test day records between d 30 and 306, divided into 30.4-d periods.

Computational Aspects

Available computer resources were not sufficient to solve the model when all traits were included simultaneously, as previously stated, and, for this reason, traits were grouped. A multiple-trait model including four first parity test days and using the Model [1] required an average 19.6 h of CPU (central processing unit) time (HP-9000/735 workstation; Hewlett Packard Co., Palo Alto, CA): 2.78 h (14.2%) setting the mixed model equations, 0.23 h (1.2%) inverting the equations, and 16.6 h (84.6%) iterating and finding the final solutions. The model allowed for missing observations, which complicated the application of canonical transformations that could have been used to reduce the computing time. A technique has been suggested to circumvent this problem (1) based on the substitution of the missing values by their expectations. However, although such a technique would have reduced the computing time, it would not be expected to affect the results.

CONCLUSIONS

One of the frequently mentioned advantages of the TDM is its ability to account for the heterogeneity of genetic and environmental variances during the lactation (8, 21, 22). In our analysis, heterogeneity of variance is clearly demonstrated (Figures 1 and 2). Previous research (27) has shown that the highest response to selection could be obtained by using only milk yield during the second trimester of the lactation because the consequences of lower genetic correlations are compensated by a shorter generation inter-

val and higher heritability. However, those estimates of heritability were obtained using standardized milk yield records. Results shown in Tables 6 and 7 indicate that the use of standardized milk yield may inflate the actual value of these genetic parameters.

There is still a further question to be answered about the adequacy of a multiple-trait approach (8, 26) or a repeatability model (12, 17). The increased computational burden for estimating breeding values using a multiple-trait approach may be reason to use a repeatability model instead. However, the heterogeneity of variances during the lactation and the patterns in genetic and phenotypic correlations suggest that a multiple-trait approach is more accurate than the repeatability model. A relatively new methodology based on the use of covariance functions (7) has been suggested that could increase the flexibility of the multiple-trait approach and could allow the inclusion of all observations. Rather than applying models with many traits, the variance-covariance structure of repeatedly measured traits over time is modeled using a covariate function.

Conditions in Costa Rica are ideal for the application of the TDM. Genetic and phenotypic parameters obtained in this paper can be used to develop management tools to be implemented in on-farm management programs and in the design of a breeding scheme for local dairy cattle.

ACKNOWLEDGMENTS

The authors thank the farmers and staff at Universidad Nacional de Costa Rica for collecting the data. B. Vargas gratefully acknowledges the financial support from Netherlands Organization for International Cooperation in Higher Education and the Wageningen Agricultural University, Wageningen, The Netherlands.

REFERENCES

1 Ducrocq, V., and B. Babes. 1993. Solution of multiple trait animal models with missing data on some traits. J. Anim. Breed. Genet. 110:81–92.
2 Dwinger, R. H., E. Cappella, E. Pérez, M. Baaijen, and E. Muller. 1994. Application of a computerized herd management and production control program in Costa Rica. Trop. Agric. (Trinidad) 71:74–76.
3 Groeneveld, E. 1996. REML-VCE User's Guide, Version 3.2. Inst. Anim. Husb. Anim. Ethol. Mariansee, Germany.
4 Holdridge, L. R. 1967. Life zone ecology. Int. Am. Inst. Agric. Sci. (IICA), San José, Costa Rica.
5 Jamrozik, J., L. R. Schaeffer, and J.C.M. Dekkers. 1996. Random regression models for production traits in Canadian Holsteins. Pages 124–134 *in* Bull. no. 14. INTERBULL Annu. Mtg., INTERBULL, Veldhoven, The Netherlands. Int. Bull. Eval. Serv., Uppsala, Sweden.
6 Kettunen, A., and E. A. Mäntysaari. 1996. Estimation of genetic parameters for test day milk production at different stages of lactation in first parity dairy cows. Page 51 *in* Proc. 47th Annu. Mtg., Eur. Assoc. Anim. Prod. Wageningen Pers, Wageningen, The Netherlands.
7 Kirpatrick, M., W. G. Hill, and R. Thompson. 1994. Estimating the covariance structure of traits during growth and ageing, illustrated with lactation in dairy cattle. Genet. Res. 64:57–69.
8 Meyer, K., H. U. Graser, and K. Hammond. 1989. Estimates of genetic parameters for first lactation test day production of Australian Black and White cows. Livest. Prod. Sci. 21:177–199.
9 Noordhuizen, J.P.T.M., and J. Buurman. 1984. Veterinary automated management and production control program for dairy farms (VAMPP). The application of MUMPS for data processing. Vet. Q. 6:62–77.
10 Pander, B. L., and W. G. Hill. 1993. Genetic evaluation of lactation yield from test day records on incomplete lactation. Livest. Prod. Sci. 37:23–36.
11 Pander, B. L., W. G. Hill, and R. Thompson. 1992. Genetic parameters of test day records of British Holstein Friesian heifers. Livest. Prod. Sci. 55:11–21.
12 Ptak, E., and L. R. Schaeffer. 1993. Use of test day yields for genetic evaluation of dairy sires and cows. Livest. Prod. Sci. 34: 23–34.
13 Reents, R., J.C.M. Dekkers, and L. R. Schaeffer. 1994. Genetic parameters of test day somatic cell counts and production traits. Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, ON XVII:120–123.
14 Reents, R., J.C.M. Dekkers, and L. R. Schaeffer. 1995. Genetic evaluation for somatic cell score with a test day for multiple lactations. J. Dairy Sci. 78:2858–2870.
15 Reents, R., J. Jamrozik, L. R. Schaeffer, and J.C.M. Dekkers. 1995. Estimation of genetic parameters for test day records of somatic cell score. J. Dairy Sci. 78:2847–2857.
16 SAS® User's Guide: Statistics, Version 6.0 Edition. 1990. SAS Inst., Inc., Cary, NC.
17 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., Guelph, ON XVIII: 443–446.
18 Schaeffer, L. R., and J. Jamrozik. 1996. Multiple-trait prediction of lactation yields for dairy cows. J. Dairy Sci. 79: 2044–2055.
19 Soto, H., and A. Aragón. 1994. Genetic analysis of milk yield of Holstein cattle in Costa Rica. Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, ON XX:352–354.
20 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.
21 Swalve, H. H. 1995. Test day models in the analysis of dairy production data—a review. Arch. Tierz. Dummerstorf 38: 591–612.
22 Swalve, H. H. 1995. The effect of test day models on the estimation of genetic parameters and breeding values for dairy yield traits. J. Dairy Sci. 78:929–938.
23 Van der Werf, J.H.J. 1990. A note on the use of conditional models to estimate additive genetic variance in selected populations. Proc. 4th World Congr. Genet. Appl. Livest. Prod., Edinburgh, Scotland XIII:476–479.
24 Vargas, B., and C. Solano. 1995. Tendencias genéticas y ambientales en producción de leche en ganado lechero de Costa Rica. Rev. Asoc. Lat. Prod. Anim. 3:165–176.
25 Westell, R. A., L. R. Quaas, and L. D. Van Vleck. 1988. Genetic groups in an animal model. J. Dairy Sci. 71:1310–1318.
26 Wiggans, G. R., and M. E. Goddard. 1996. A computationally feasible test day model with separate first and later lactation genetic effects. Pages 19–21 *in* Proc. 56th New Zealand Soc. Anim. Prod. Annu. Mtg., Hamilton, New Zealand.
27 Wilmink, J.B.M. 1988. Effects of incomplete records on genetic parameters among cumulative yields in first lactation and on extension of part lactations. Livest. Prod. Sci. 18:19–34.