

A Computationally Feasible Test Day Model with Separate First and Later Lactation Genetic Effects

G.R. Wiggans¹ and M.E. Goddard

Animal and Genetics Breeding Unit, University of New England, Armidale, NSW, Australia

¹Animal Improvement Programs Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350, USA

Abstract

In a two step analysis test day effects were estimated within herd with adjustment for across herd effects and then the adjusted data were analysed across herd. Genetic effects were defined for each of 10 months in milk within first and later lactation and for milk, fat and protein giving 60 traits. The rank of the genetic (co)variance matrix (G) was reduced to 6 such that G retained the information to evaluate the selection objective. A repeatability model allowed for multiple lactations with each lactation, conceptually, expressing all 60 traits, but missing observations for 30 or more. A canonical transformation was applied to create uncorrelated traits. Missing values were replaced by their expectations at each round. Because of the rank reduction, only 6 canonical traits were solved for. This system should make evaluations more stable by removing biases due to genetic differences in persistency and rate of maturity.

Introduction

Milk yield in dairy cattle is typically estimated from periodic measurement of milk volumes and analysis of milk samples for fat and protein percentages. These measurements have historically been on a monthly interval, but the need to reduce the cost of milk recording has made less frequent measurements common. For genetic evaluations, the test day yields have been combined into a lactation measure (Wiggans, 1985). Greater precision in accounting for environmental influences is possible by analysing the test day yields directly (Ptak and Schaeffer, 1993). The Australian system has estimated test day effects and removed them before creating a lactation measure (Jones and Goddard, 1990). Recently, Johnson (1996) developed a method of combining test days into a 270 day lactation yield that weights the

individual tests according to the correlations among them. The correlation declines with increasing days between tests.

Each test day may be treated as a separate trait which allows for genetic differences in the shape of the lactation curve, but substantially increases the computational requirements. Jamrozik et al. (1995) have proposed that a curve be fitted for each cow. This approach reduces the magnitude of the problem and is the basis for test day models being developed in Canada.

A repeatability model is typically used in dairy evaluation. It assumes that each lactation is an expression of the same genetic trait. Estimated genetic correlations in the range of .88 to 1.02 have been found for lactations (Meyer, 1985) and .58 to .88 for test day yields (Rekaya et al., 1995) indicating some divergence from this

assumption. A multi-trait model with lactations as separate traits allows for correlations of less than 1 between pairs of lactations.

A multi-trait analysis increases the accuracy of the evaluation by using information from correlated traits. However, when all traits are measured on each individual, the heritability of all traits is the same and the genetic correlations are equal to the phenotypic correlations, there is no gain in accuracy. If recording of milk volumes is more frequent than analysis for component percentages, a multi-trait analysis allows information from milk volumes to contribute to the accuracy of evaluation of fat and protein yields. With inline milk recording equipment, continuous milk volumes and two component samples per lactation might become a common milk recording option.

Where there are no missing data, a canonical transformation has been used to convert a multi-trait analysis into single trait analyses which substantially reduces computing requirements. Ducrocq and Bebes (1993) extended the application of the canonical transformation to cases where not all traits were measured by replacing missing values with their expectations during each iteration.

Users expect stability in genetic evaluations as information accumulates. If daughters of a bull tend to be more persistent than their contemporaries, that bull's evaluation will increase as his daughters progress through their lactations. Similarly, if a bull's daughters reach mature yield at a younger age than contemporaries, his evaluation will decline as his daughters add later lactations. A similar volatility can occur when a bull is returned to service and has a large number of new daughters. Estimation of separate genetic effects within and across lactation will dampen these swings because the correlations of less than 1 limit the effect information on one trait has on the evaluation of another.

The goal of this project was to develop a computationally feasible test day model that would account for genetic differences within and across lactation.

Data

Methods were developed using test day records of milk, fat and protein yields from 92 herds in Victoria, Australia. Calving dates were from February 1983 through January 1995. Only tests where days in milk was in the interval 5 to 305 days were included. Test day yields were adjusted for age at test day and stage of lactation before analysis with the factors routinely used in the Australian Dairy Herd Improvement Scheme. Details of the data are in Table 1.

Model

The multi-trait model for test day yield was:

$$Y_{ijklmno} = htd_{ijkop} + hys_{ijlno} + bv_{imno} + pe_{imno} + e_{ijklmno}$$

where

y was age, stage of lactation adjusted yield of product i (milk, fat, or protein) in herd j on test day k in year-season l from cow m in month of milking n of parity group o (first or later) and lactation within parity group p. Within later parity group the 4 classes of p were 2, 3, 4, and 5 and greater. Seasons were February through July and August through January. Months of milking were intervals of 30 days starting 5 days after calving. The model effects were htd for herd test day, hys for herd year season, bv for breeding value, pe for permanent environment, and e for residual.

For each cow, there were 60 bv defined, 3 products $\times$ 10 months of milking $\times$ 2 parity groups. For each herd year season, a hys effect was defined for each of the 60 traits. The htd effect was separate for each product and parity group, but included all months of milking. The pe effect allowed for lactations beyond 2 to be included.

The genetic (G) and residual (R) (co)variance matrices were based on estimates from Meyer and Goddard (Accuracy of selection for dairy production using individual test-day records and a "canonical" index. In preparation), who used matrices of order 30. The 60 $\times$ 60 matrices were constructed by duplicating the first parity block to the later parity position. For G, the off diagonal blocks were assumed to be 90% of the diagonal blocks, and for phenotypic variance, 80%. The phenotypic variance after subtracting G was partitioned 60% for R and 40% for var(pe). The covariances between months of milking across trait and parity had not been specifically estimated so were approximated from lactation values. Further work may provide better estimates of G and R. Table 2 summarises these matrices by giving heritabilities and correlations on a lactation basis (sum of 10 months of milking) and Table 3 on a month of milking basis for first lactation.

Reduced rank G

To reduce the computational requirements, the genetic covariance matrix G was approximated by G* which has only 6 nonzero eigenvalues. The approximation was done so the (co)variances of important linear combinations of test day yields such as total yield of milk, fat, and protein were unchanged (Goddard, M.E. Multitrait BLUP using a genetic covariance matrix of reduced rank. In preparation).

Canonical transformation

The residual covariance matrix R and the G* were used in creating a transformation matrix to construct a canonical transformation of the 60 traits. This generated 60 transformed traits which were environmentally and genetically uncorrelated. However, because the rank of G* was 6, only 6 canonical transformations had any genetic variance. Therefore, the 60 trait analysis was replaced by 6 single trait analyses on the transformed scale.

Two-stage analysis

To simplify calculations, the herd test day effect was solved for separately within herd. First, hys, bv and pe effects were subtracted from test day yields. Then htd effects were estimated and subtracted from y.

The second stage was a multi-trait animal model with missing values replaced by their expectations in each round following Ducrocq and Bebes (1993). Each lactation could contribute a maximum of 30 test day yields for the first or later parity traits; the other 30 or more were missing. Lactations with fewer than 10 tests would have more missing values. It was not necessary to return to the original scale to replace the missing values by their expectations.

Solutions from the second stage were back transformed and used in stage 1 to improve estimates of the htd effects. This feedback is primarily useful in accounting for culling over the lactation which could cause only the genetically superior cows to be included in late season test days.

Several functions of the solutions were defined. The lactation yield bv was $30 \sum bv_i$. The bv for first and second half of lactation were similarly defined.

Results

In solving the BLUP equations, the first stage of analysis was carried out 4 times. The first time, 150 iterations of the second stage were required for the convergence criterion (sum of squared differences/sum of squared solutions) to fall below 1×10^{-6} for all 6 traits. For the second, 14; third, 10; and fourth, 9 iterations were required to reach the same value.

Table 4 displays functions of solutions for the 58 bulls with 30 or more (39 maximum) daughters. Only results for protein are presented. The means of all functions of solutions are near zero. The maximum values have larger absolute values than the minimums indicating some skewness to the right in these distributions, a likely result from selection. The standard deviations decline with increase in months of milking for both first and later lactations. The statistics are nearly identical for first and later lactations. This probably results from the partitions of the (co)variance matrices being the same for each lactation. The first half of lactation minus the second half is a measure of persistency. The variance of this measure is about 9% of the variance bv for the first or last half of lactation. A measure of rate of maturity is the first lactation bv minus the later lactation bv. The variance of this measure is about 5% of that of the lactation bv, less than for persistency.

Discussion

This study shows that a multi-trait animal model using test day data is practical. The canonical transformation with missing data and the reduced rank of the genetic (co)variance matrix allow multi-trait analysis to be applied to large field data sets. Multi-trait solutions allow detection of differences in rate of maturity and persistency, but perhaps more importantly, should contribute to increased stability of bv as data accumulate.

Acknowledgements

This project was supported by a grant from the Australian Dairy Research and Development Corporation. Bruce Tier of the Animal Genetics and Breeding Unit provided advice on computational techniques. Data were provided by Kevin Beard and Les Jones of the Australian Dairy Herd Improvement Scheme. The senior author received salary support from the United States Department of Agriculture while on sabbatical leave at the University of New England.

References

- Ducrocq, V. and Bebes, B. 1993. Solution of multiple trait animal models with missing data on some traits. *Journal of Animal Breeding and Genetics* 110, 81-92.
- Jamrozik, J., Schaeffer, L.R. and Dekkers, J.C.M. 1995. Genetic evaluation of dairy cattle using test day and random regression model. *Journal of Dairy Science* 78 (Suppl. 1), 249.
- Johnson, D.L. 1996. Estimation of lactation yield from repeated measures of test-day yields. *Proceedings of the New Zealand Society of Animal Production* 56: [in press].
- Jones, L.P. and Goddard, M.E. 1990. Five years experience with the animal model for dairy evaluations in Australia. *Proceedings of 4th World Congress on Genetics Applied to Livestock Production XIII*, 382-385.
- Meyer, K. 1985. Genetic parameters for dairy production of Australian Black and White cows. *Livestock Production Science* 12, 205-219.
- Ptak, E. and Schaeffer, L.R. 1993. Use of test day yields for genetic evaluation of dairy sires and cows. *Livestock Production Science* 34, 23-34.
- Rekaya, R., Bèjar, R., Carabaño, M.J. and Alenda, R. 1995. Genetic parameters for test day measurements in Spanish

Holstein-Friesian. Proceedings of the Open Session of the INTERBULL Annual Meeting, Prague, Czech Republic, Bulletin 11 (page numbering within article 1-8).

Wiggans, G.R. 1985. Procedures for calculating lactation records. Natl. Coop. Dairy Herd Impr. Progr. Handbook. Fact Sheet G-1, 10 pp. Ext. Serv., USDA, Washington, DC.

Table 1. Characteristics of data

		Counts by Parity		
		Parity	Count	%
Herds	92			
Cows	23,052			
Years of calving	1983-1994	1	18,649	25.5
		2	14,189	19.4
Lactations/cow	3.2	3	11,337	15.5
		4	8,829	12.1
Tests/lactation	6.8	5	20,053	27.4

Table 2. Heritabilities (diagonal), genetic correlations (above diagonal) and phenotypic correlations (below diagonal) for sum of 10 test days for milk, fat and protein yields in first and second lactation

First			Second		
M	F	P	M	F	P
.31	.77	.91	.91	.68	.82
.90	.20	.72	.68	.92	.64
.95	.91	.24	.82	.64	.91
.80	.72	.76	.31	.77	.91
.72	.80	.73	.90	.20	.72
.76	.73	.80	.95	.91	.24

Table 3. Heritabilities (diagonal), genetic correlations (above diagonal) and phenotypic correlations (below diagonal) for first lactation yield for 10 months in milk

Months in Milk									
1	2	3	4	5	6	7	8	9	10
Milk									
.14	.97	.92	.88	.84	.76	.72	.61	.58	.50
.59	.18	.99	.97	.94	.87	.84	.72	.63	.48
.51	.63	.19	.99	.98	.93	.91	.79	.67	.48
.46	.57	.63	.20	.99	.97	.95	.85	.72	.53
.43	.51	.56	.63	.20	.99	.98	.90	.78	.58
.40	.47	.56	.56	.63	.23	1.00	.95	.83	.63
.34	.42	.47	.51	.56	.62	.22	.96	.86	.65
.25	.33	.38	.42	.45	.50	.57	.18	.95	.80
.25	.29	.33	.37	.40	.43	.49	.55	.19	.94
.20	.26	.28	.29	.34	.36	.38	.48	.61	.14
Protein									
.09	.94	.84	.77	.73	.70	.65	.69	.74	.74
.44	.08	.97	.94	.91	.88	.83	.80	.82	.75
.38	.49	.14	.99	.96	.92	.87	.80	.79	.68
.33	.43	.50	.13	.99	.97	.93	.86	.84	.71
.32	.38	.43	.51	.12	.99	.97	.91	.89	.76
.32	.37	.40	.45	.51	.16	.99	.95	.93	.81
.28	.35	.39	.43	.46	.53	.18	.97	.94	.83
.21	.28	.33	.39	.41	.45	.53	.14	.99	.92
.21	.26	.30	.31	.34	.36	.43	.52	.12	.96
.21	.26	.26	.26	.31	.33	.37	.45	.54	.09

Table 4. Standard deviation, minimum (Min) and maximum (Max) of combinations of breeding values for protein of 58 bulls with 30 to 39 daughters

Period of Milking	Mean	SD	Min	Max
First lactation				
Test day 1	-.006	.036	-.080	.093
Test day 5	-.005	.034	-.072	.092
Test day 10	-.003	.029	-.064	.075
First half lact.	-.887	5.511	-12.120	14.850
Last half lact.	-.759	5.280	-11.940	13.020
First-last	-.127	1.656	-4.230	3.930
Total	-1.646	10.665	-23.130	27.870
Later lactations				
Test day 1	-.006	.037	-.085	.088
Test day 5	-.006	.035	-.073	.080
Test day 10	-.004	.030	-.063	.078
Total	-1.866	10.996	-23.220	26.310
First-later				
First-later	.220	2.490	-5.220	6.960