

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Papers and Publications in Animal
Science

Animal Science Department

June 1988

Genetic (Co)Variances for Milk, Fat, and Protein Yield in Holsteins Using an Animal Model

L. Dale Van Vleck

University of Nebraska-Lincoln, dvan-vleck1@unl.edu

M. C. Dong

Cornell University

Follow this and additional works at: <https://digitalcommons.unl.edu/animalscifacpub>



Part of the [Animal Sciences Commons](#)

Van Vleck, L. Dale and Dong, M. C., "Genetic (Co)Variances for Milk, Fat, and Protein Yield in Holsteins Using an Animal Model" (1988). *Faculty Papers and Publications in Animal Science*. 308.

<https://digitalcommons.unl.edu/animalscifacpub/308>

This Article is brought to you for free and open access by the Animal Science Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Papers and Publications in Animal Science by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Genetic (Co)Variances for Milk, Fat, and Protein Yield in Holsteins Using an Animal Model

L. D. VAN VLECK and M. C. DONG
Department of Animal Science
Cornell University
Ithaca, NY 14853

ABSTRACT

First lactation milk, fat, and protein yields for first lactations of 8044 Holstein cows in New York from 24 herds were used to estimate genetic and phenotypic covariances with an animal model by restricted maximum likelihood. Numerator relationships within herd, including those from sires, were utilized, although relationships across herds were ignored. Each analysis was terminated after 300 rounds of iteration. Average milk production (twice daily milking, 305-d lactation, mature equivalent) was 8630 kg. Estimates were obtained for each individual herd and groups of 3, 6, and 12 herds. Average estimates from separate analyses of the 24 herds were nearly identical to those from combined analysis. Heritability estimates from combined analysis were .36, .35, and .33 for milk, fat, and protein yield with standard errors of approximately .03. Genetic correlations were .72 for milk and fat, .88 for milk and protein, and .77 for fat and protein. Phenotypic correlations were .81, .91, and .82. Heritability estimates for individual herds were quite variable with standard errors of about .15. The largest environmental standard deviation for a herd was about twice the smallest. Estimates from only one analysis failed to converge—a group of three herds. All single herd analyses reached convergence as did the other 7 trios of herds and all sets of 6 and 12 herds.

INTRODUCTION

When inexpensive protein testing was introduced in the early 1980's, reports of heritability for protein yield and correlations with milk and fat yield were rare and based on relatively small and difficult to obtain data sets (4, 16, 27). Review papers (1, 10, 14, 23) outlined potential benefits and cost effectiveness of protein testing and selection. When multiple-trait (milk, fat, and protein) sire evaluation was introduced (9) in the Northeast, estimates of required genetic and phenotypic covariances were needed. Estimates obtained from small data sets available from paid testing for protein (15, 17) were marked by unrealistic estimates of the correlation between milk and fat and relatively small or inconsistent heritability estimates for milk, fat and protein. In fact, the genetic covariance matrix reported by Manfredi et al. (17) was negative definite. Lawlor's (15) estimates obtained from records based on the first 15 mo of widescale protein testing (July 1981 to September 1982) with a relatively large data set (37,233 first lactation records of daughters of 702 sires) using REML procedures, but not relationships among the sires, also featured very small estimates of heritability (.20, .20, and .15 for milk, fat, and protein) and a relatively small estimate of the genetic correlation between milk and fat yield (.51). Because the data resulted mostly from daughters of proved bulls, heritability estimates might have been reduced because of selection. Hargrove et al. (11), using differences from modified contemporary averages and method 1 of Henderson (12), reported estimates similar to those of Lawlor (15). Cue et al. (5), also using a sire model and REML with relationships among sires considered to be zero and a Canadian Holstein data set smaller than that of Lawlor, reported estimates remarkably similar to those that are reported in this paper except for a smaller estimate of heritability for protein yield.

Received March 10, 1988.
Accepted June 9, 1988.

Swalve and Van Vleck (21) used REML with a multiple-trait animal model to estimate the covariance structure among first three milk lactation records. The animal model for simultaneous bull and cow evaluation (20, 25, 26) will require estimates of variances appropriate to an animal model.

The purposes of this study were 1) to obtain estimates of covariances needed for genetic evaluations for protein, fat, and milk with an animal model, 2) to examine groups of data sets to obtain some idea of the sampling variances of covariance matrices estimated from relatively small samples with an animal model, and 3) to determine the range of estimates obtained from individual herds, which is one possible approach to evaluation with heterogeneous genetic and environmental variances.

MATERIALS AND METHODS

Data

Records on first lactation (twice daily milking, 305-d lactation mature equivalent) milk, fat, and protein yields of Holstein cows were obtained from the Northeast Dairy Records Processing Laboratory. These records were for cows freshening after free protein testing began in 1981. Cows in the herds before that time were utilized to create relationship matrices. Herds were chosen for analysis that had approximately 300 to 400 cows with milk, fat, and protein records. These herds had greater production (average of 1070 kg milk) than the entire data set. Characteristics of the 24 herds are in Table 1.

Methods

The REML estimation procedure with the expectation-maximization algorithm (REML-EM) used by Swalve and Van Vleck (21) as derived from Henderson (13) was used with an animal model. The model for a record contained a fixed herd-year-season effect (two seasons per year), the additive genetic value of the animal, and an environmental effect. Covariances between genetic and environmental effects were assumed to be zero. All numerator relationships within a herd (including sires) were utilized. The approximation involved with the procedure (21) is that numerator relationships across herds are assumed to be zero.

Because each cow was required to have all three yield records, and because only one random effect other than the residual effect was in the model, a canonical transformation (2, 22), as outlined by Lawlor (15) for a sire model, was used to reduce calculation to single-trait analyses. The portion of the inverse of the relationship matrix due to sires and cows without protein records was absorbed into the portion for cows with all three traits measured. The numerator relationship matrix among those cows was then obtained by inverting the result of the absorption. To reduce the number of calculations in each round of iteration, another transformation was employed involving the Cholesky decomposition of the relationship matrix as described by Lawlor (15) referenced by personal communication to R. L. Quaas (1983) to diagonalize the portion of the mixed model equations corresponding to the additive genetic values. Householder matrices (19) were then employed to tridiagonalize the mixed model equations, which, except for a constant times the identity matrix, do not change from round to round. The simplicity of the resulting tridiagonal form allows rapid calculation of solutions and inverse elements needed for the REML-EM procedure. Each analysis was stopped after 300 rounds of iteration when convergence was assumed. Convergence was monitored by the norm of the difference in estimates from round to round. To equalize significant digits, milk, fat, and protein records (lb) were scaled by dividing by 3200 lb, 120 lb, and 100 lb. Converged estimates in the tables are shown in scaled form with the conversion to (kg)² shown in the footnotes.

Starting values for all analyses were those obtained after 24 rounds for the data set that included the first 12 herds. Heritabilities were .318, .372, and .315 for milk, fat, and protein. Genetic correlations for milk-fat, milk-protein, and fat-protein were .778, .905, and .834; corresponding phenotypic correlations were .805, .929, and .827.

RESULTS AND DISCUSSION

Two data sets (herds 1 to 12 and herds 13 to 24) were analyzed first. The two data sets gave such similar results that later individual herds, groups of three, groups of six, and the complete set (herds 1 to 24) were analyzed separately

TABLE 1. Characteristics of 24 herds used to estimate genetic and environmental covariance matrices and REML estimates of environmental variances.

Herd	Number		Mean (kg)			Environmental variance (kg) ²		
	Sires	Cows ¹	Milk	Fat	Protein	Milk ²	Fat ³	Protein ⁴
1	77	303	8801	310	284	.97	.73	.70
2	105	305	8971	336	296	.93	.49	.90
3	84	303	8493	290	267	.42	.44	.52
4	125	392	8604	314	273	.83	.61	.86
5	163	348	7465	261	241	.45	.35	.40
6	80	362	8220	300	267	.49	.40	.48
7	133	386	7696	279	250	.84	.78	.78
8	94	334	9201	344	297	.70	.74	.62
9	73	344	8585	315	275	.50	.27	.43
10	127	305	8557	312	288	.61	.50	.60
11	109	330	8255	286	263	1.16	.64	1.10
12	49	308	8701	311	275	.50	.70	.45
13	93	348	8858	319	282	1.11	.99	1.19
14	176	353	8557	303	279	1.28	.75	1.34
15	69	332	10,048	364	328	.67	.62	.46
16	109	310	8722	331	290	.66	.90	.69
17	82	295	9931	350	315	.96	.71	1.29
18	128	397	8272	296	266	.38	.34	.32
19	126	331	9488	340	307	1.20	.74	1.62
20	97	349	8431	325	281	.60	.79	1.06
21	78	308	7189	252	228	.63	.57	.61
22	108	304	7868	300	258	.67	.63	.77
23	116	366	8734	322	280	.81	.78	1.07
24	66	331	9790	367	315	.64	.82	.77
Avg.	103	335	8630	313	279	.75	.64	.79

¹ With first lactation milk, fat, and protein records after July 1, 1981.

² Multiply by (1451.5)².

³ Multiply by (54.43)².

⁴ Multiply by (45.36)².

to explore convergence problems and sampling variances with smaller data sets.

Estimates for the sample of 24 herds are shown in the next to last row of Table 2. Averages of separate estimates from the 24 herds are shown in the last row of Table 2. Estimates in the two rows are in general agreement. These estimates are considerably different from previous estimates from Northeast records by Manfredi et al. (17) and Lawlor (15) who found much smaller estimates of heritability, especially for protein yield. Estimates of the genetic correlations between milk and protein (.88 and .85) are also somewhat greater than the previous two studies (.81 and

.79) and between milk and fat (.72 and .70) are considerably larger than the .43 and .51 reported by Manfredi et al (17) and Lawlor (15). Heritabilities and correlations are remarkably similar to those reported by Cue et al. (5) for Holstein data except that their estimate of heritability for protein yield was only .25. The animal model may be less affected by selection than a sire model using records of daughters of selected bulls without a full relationship matrix. The Canadian results may be due to less selection on bulls. Both genetic and phenotypic correlations between milk and protein in this and other reports are greater than between milk and fat and between fat and protein. Cor-

TABLE 2. Estimates of heritabilities, environmental variances, and genetic and phenotypic correlations from different numbers of herd.

Herds	Heritability			Correlation						Environmental variance (kg) ²				
	M ¹	F ²	P ³	Genetic			Phenotypic			M ⁴	F ⁵	P ⁶		
				M ₁ F	M ₁ P	F ₁ P	M ₁ F	M ₁ P	F ₁ P					
1 to 3	.36	.48	.36	.80	.94	.84	.82	.82	.94	.85	.809	.523	.729	
4 to 6	.49	.49	.48	.76	.93	.79	.80	.76	.95	.83	.505	.423	.478	
7 to 9	.30	.30	.40	.77	.74	.66	.80	.78	.89	.78	.715	.626	.654	
10 to 12	.26	.29	.27	.52	.79	.52	.78	.52	.92	.82	.824	.653	.775	
13 to 15	.32	.36	.37	.83	.84	.84	.82	.82	.91	.83	.982	.775	1.003	
16 to 18	.34	.32	.33	.67	.88	.72	.77	.77	.88	.75	.709	.646	.807	
19 to 21	.22	.30	.15	.55	.77	.66	.80	.80	.88	.80	.853	.722	1.114	
22 to 24				Did not converge										
1 to 6	.43	.49	.42	.78	.93	.81	.81	.81	.95	.84	.637	.466	.586	
7 to 12	.29	.30	.35	.65	.75	.67	.79	.79	.90	.80	.760	.635	.703	
13 to 18	.34	.34	.34	.75	.88	.80	.80	.80	.90	.80	.835	.712	.929	
19 to 24	.39	.31	.28	.69	.89	.76	.82	.82	.90	.82	.798	.758	1.035	
1 to 12	.36	.39	.37	.72	.86	.75	.80	.80	.92	.82	.703	.554	.658	
13 to 24	.36	.33	.31	.72	.88	.78	.81	.81	.90	.81	.820	.735	.978	
1 to 24	.36	.35	.33	.72	.88	.77	.81	.81	.91	.82	.761	.646	.827	
Avg. of 24	.36	.35	.36	.70	.85	.76	.80	.80	.91	.82	.751	.638	.793	

¹ Milk.

² Fat.

³ Protein.

⁴ Multiply by (1451.5)².

⁵ Multiply by (54.43)².

⁶ Multiply by (45.36)².

TABLE 3. Estimates of heritabilities and genetic and phenotypic correlations for each of 24 herds.

Herd	Heritabilities			Correlations					
	M ¹	F ²	P ³	Genetic			Phenotypic		
				M,F	M,P	F,P	M,F	F,P	F,P
1	.22	.26	.37	.94	.93	.89	.85	.94	.88
2	.47	.64	.44	.79	.97	.86	.83	.97	.87
3	.49	.29	.29	.87	.76	.77	.90	.90	.78
4	.36	.38	.23	.70	.93	.75	.79	.96	.81
5	.28	.32	.32	.49	.84	.58	.79	.94	.84
6	.29	.27	.28	.56	.84	.67	.77	.92	.83
7	.20	.20	.23	.83	.84	.88	.82	.92	.84
8	.36	.18	.42	.51	.82	.57	.74	.92	.78
9	.48	.68	.65	.92	.77	.69	.83	.84	.74
10	.27	.25	.35	.10	.62	.51	.67	.86	.74
11	.24	.48	.26	.81	.90	.90	.85	.94	.86
12	.48	.18	.42	.66	.95	.71	.77	.95	.82
13	.32	.28	.28	.97	.98	.98	.88	.95	.89
14	.10	.33	.12	.51	.52	.56	.84	.94	.86
15	.45	.46	.72	.77	.82	.78	.73	.85	.74
16	.48	.26	.46	.65	.94	.80	.82	.96	.85
17	.34	.41	.32	.71	.80	.66	.72	.81	.62
18	.63	.65	.67	.87	.96	.90	.83	.96	.86
19	.08	.38	.06	.21	.55	.52	.73	.88	.72
20	.51	.36	.28	.76	.74	.80	.85	.84	.84
21	.17	.13	.15	.88	.92	.88	.84	.94	.90
22	.39	.29	.31	.88	.92	.98	.87	.93	.91
23	.51	.40	.35	.80	.97	.82	.84	.92	.83
24	.64	.37	.59	.82	.92	.87	.84	.90	.82
Avg.	.36	.35	.36	.70	.85	.76	.80	.91	.82

¹ Milk.² Fat.³ Protein.

relations between milk and protein are more similar from study to study than correlations between milk and fat (4, 11, 15, 16, 27).

The greater heritabilities in this study may be due to the tendency shown in many studies [e.g., (7)] for heritability to increase with herd production, which is much greater in these data than in data used in earlier studies.

The two initial data sets (herds 1 to 12 and 13 to 24) yielded similar estimates of heritabilities and correlations, although the phenotypic and environmental variances were considerably greater in the second data set, possibly associated with a mean production level difference of 375 kg. Even data sets of only 6 herds (Table 2) resulted in somewhat similar estimates, especially for phenotypic correlations.

Estimates from samples containing 3 herds were more variable (Table 2). In one case, for herds 22 to 24, estimates for the covariance matrices did not converge, even though convergence was obtained when those herds were combined with 3 other herds (19 to 21) and when analyzed as separate herds (Table 3).

Individual herds yield quite variable results, although all separate analyses led to convergence. Estimates of phenotypic correlations were least variable over all herds, especially between milk and protein. The ranges suggest a standard deviation for separate herd estimates of heritability of about .15 for herds of size 300 to 400. The pooled estimates from 24 herds would then have standard error of about .03. These herds were relatively large. Estimates from smaller herds would be expected to be even more variable. Genetic variances therefore would seem to be difficult to estimate for individual herds and for herds within a limited period of time.

Environmental variances from analyses for individual herds are shown in Table 1 and suggest considerable differences in variance from herd to herd with environmental standard deviations being different from smallest to largest by a factor of about 2 for milk or fat yield and of >2 for protein yield. The importance (3, 24) of differences of this magnitude will depend on the distribution of variances and, in the case of sire proving, on the distribution of daughters across herds.

CONCLUSIONS

Estimates of heritability are similar for milk, fat, and protein and are considerably different from those used in genetic evaluations. Estimates of heritability for milk are similar to those obtained from several other samples of similar sizes using REML with an animal model (8, 21). These estimates are consistently greater than those obtained earlier when production was less and particularly when sire models were used [e.g., (16)]. Heritability estimates from daughter-dam regression with similar production are about the same as found in this study [e.g., (6)]. Estimates from a sire model by Cue et al. (5) and Monardes and Hayes (18) are also in agreement with results reported here. If these estimates represent variance-covariances from an unselected population, then consideration must be given to using these estimates rather than those now used in genetic evaluation programs.

The method of pooling data from a sample of herds appears to lead to estimates with relatively small standard errors for herd sizes of 300 to 400 and as few as 24 herds. Problems of convergence can occur as in this study when small samples are used, which may create problems if genetic evaluation is done jointly with estimation of variance components.

ACKNOWLEDGMENTS

Computing support by the National Cornell Supercomputing Facility was essential for this project. This research was supported in part by grants from Eastern AI Cooperative, Inc., Ithaca, NY, and US-Spain Joint Committee for Scientific and Technological Cooperation.

REFERENCES

- 1 Anderson, R. D., R. W. Everett, and L. D. Van Vleck. 1978. Economic analysis of protein testing for selection. *J. Dairy Sci.* 61:102.
- 2 Arnason, T. 1982. Prediction of breeding values for multiple traits in small nonrandom mating (horse) populations. *Acta Agric. Scand.* 32:171.
- 3 Brotherstone, S., and W. G. Hill. 1986. Heterogeneity of variance amongst herds for milk production. *Anim. Prod.* 42:297.
- 4 Butcher, K. R., F. D. Sargent, and J. E. Legates. 1967. Estimates of genetic parameters for milk constituents and yields. *J. Dairy Sci.* 50:185.
- 5 Cue, R. I., H. G. Monardes, and J. F. Hayes. 1987. Correlations between production traits in first lactation Holstein cows. *J. Dairy Sci.* 70:2132.

- 6 DeVeer, J. C., and L. D. Van Vleck. 1986. Comparison of heritability estimates from daughter on dam regression with three models to account for production level of dam. *J. Dairy Sci.* 69:2890.
- 7 DeVeer, J. C., and L. D. Van Vleck. 1987. Genetic parameters for first lactation milk yield at three levels of herd production. *J. Dairy Sci.* 70:1434.
- 8 Dong, M. C., and L. D. Van Vleck. 1987. Estimation of genetic (co)variances for production, survival and reproduction. *J. Dairy Sci.* 70(Suppl. 1):232. (Abstr.)
- 9 Everett, R. W., R. L. Quaas, and E. J. Pollak. 1983. Multiple trait Northeast AI Sire Comparison. Mimeo Ser. #74, Dep. Anim. Sci., Cornell Univ., Ithaca, NY.
- 10 Gaunt, S. N. 1973. Genetic and environmental changes possible in milk composition. *J. Dairy Sci.* 56:270.
- 11 Hargrove, G. L., D. A. Mbah, and J. L. Rosenberger. 1981. Genetic and environmental influences on milk and milk component production. *J. Dairy Sci.* 64:1593.
- 12 Henderson, C. R. 1953. Estimation of variance and covariance components. *Biometrics* 9:226.
- 13 Henderson, C. R. 1984. Application of linear models in animal breeding. Univ. Guelph Press, Ont., Can.
- 14 Laben, R. C. 1963. Factors responsible for variation in milk composition. *J. Dairy Sci.* 46:1293.
- 15 Lawlor, T. J. 1984. Estimation of genetic and phenotypic parameters of milk, fat and protein yields of Holstein cattle under selection. Ph.D. Thesis, Cornell Univ., Ithaca, NY.
- 16 Majjala, K., and M. Hanna. 1974. Reliable phenotypic and genetic parameters in dairy cattle. Page 541 in Vol. I, Proc. Int. Congr. Quant. Genet. Appl. Livest. Prod., Madrid, Spain.
- 17 Manfredi, E. J., R. W. Everett, and S. R. Searle. 1984. Phenotypic and genetic statistics of components of milk and two measures of somatic cell concentration. *J. Dairy Sci.* 67:2028.
- 18 Monardes, H. G., and J. F. Hayes. 1985. Genetic and phenotypic relationships between lactation cell counts and milk yield and composition of Holstein cows. *J. Dairy Sci.* 68:1250.
- 19 Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. V. Vetterling. 1986. Numerical recipes. Cambridge Univ. Press, New York, NY.
- 20 Robinson, G. K. 1986. Group effects and computing strategies for models for estimating breeding values. *J. Dairy Sci.* 69:3106.
- 21 Swalve, H., and L. D. Van Vleck. 1987. Estimation of genetic (co)variances for milk yield in first three lactations using an animal model and restricted maximum likelihood. *J. Dairy Sci.* 70:842.
- 22 Thompson, R. 1977. Estimation of quantitative genetic parameters. Page 639 in Proc. Int. Conf. Quant. Genet., Iowa State Univ., Ames.
- 23 Van Vleck, L. D. 1978. Breeding for increased protein content in milk. *J. Dairy Sci.* 61:815.
- 24 Vinson, W. E. 1987. Potential bias in genetic evaluations from differences in variation within herds. *J. Dairy Sci.* 70:2450.
- 25 Westell, R. A., and L. D. Van Vleck. 1987. Simultaneous genetic evaluation of sires and cows for a large population of dairy cattle. *J. Dairy Sci.* 70:1006.
- 26 Wiggans, G. R., and I. Misztal. 1987. Supercomputer for animal model evaluation of Ayrshire milk yield. *J. Dairy Sci.* 70:1906.
- 27 Wilcox, C. J., S. N. Gaunt, and B. R. Farthing. 1971. Genetic interrelationships of milk composition and yield. South. Coop. Ser. Bull. 155, Univ. Florida, Gainesville.