

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 and Fat Yield in California, New York, and Wisconsin for an Animal Model by Restricted Maximum Likelihood

L. Dale Van Vleck

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

M. C. Dong

Cornell University

G. R. Wiggans

Agricultural Research Service, USDA Beltsville, MD

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



Part of the [Animal Sciences Commons](#)

Van Vleck, L. Dale; Dong, M. C.; and Wiggans, G. R., "Genetic (Co)Variances for Milk and Fat Yield in California, New York, and Wisconsin for an Animal Model by Restricted Maximum Likelihood" (1988). *Faculty Papers and Publications in Animal Science*. 286.
<https://digitalcommons.unl.edu/animalscifacpub/286>

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 and Fat Yield in California, New York, and Wisconsin for an Animal Model by Restricted Maximum Likelihood

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

G. R. WIGGANS
Animal Improvement Programs Lab
Agricultural Research Service, USDA
Beltsville, MD 20705

ABSTRACT

Two samples of data from approximately 4000 Holstein cows were analyzed by REML with a multivariate (milk and fat yields) animal model for first lactations started in 1970 through 1977 and in 1978 through 1985 and also for first lactation records started in low, middle, and high production herds in 1978 through 1985 separately for California, New York, and Wisconsin herds. Heritability estimates for milk yield (similar for fat yield) ignoring category of herd production were .29 for 1978 and later and .34 for 1977 and earlier. Estimates for low, middle, and high production herds were .23, .29, and .36. Estimates of genetic correlation between milk and fat averaged .62 for 1978 and later and .73 for 1977 and earlier; for 1978 and later these estimates were greater for California than for New York and Wisconsin, .70 vs. .58 overall. Phenotypic and environmental variances associated with low production herds were less than for middle and high production herds, but differences between middle and high herds varied by state and sample. The pattern of estimates of genetic and environmental parameters suggests that differences between California, New York, and Wisconsin are not very important for genetic evaluation procedures. Production of herds does seem important.

INTRODUCTION

Estimates of heritability of milk yield from both daughter on dam regression and paternal half-sib correlations have been found to increase with production of the herd (2, 3, 11, 13, 16). Earlier studies showed a similar pattern as recent estimates, although most were based on less data. Heritability estimates from countries with low production are generally smaller than estimates from countries with high production (11). Many recent results are from analyses of records from northeastern United States. High average production in California suggests that heritability for herds in California might be greater than where average production is less. In that case, large California herds might have some advantage for sampling bulls. The Midwest and Northeast, however, contain a large number of cows. The purpose of this project was to compare heritability and variation at different herd production in California, New York, and Wisconsin using REML with an animal model. The three states were chosen to represent three regions with high concentrations of cows and different management practices.

MATERIALS AND METHODS

Data

First lactation milk and fat records of Holstein cows freshening from 1970 through 1985 in California, New York, and Wisconsin were furnished by the Animal Improvement Programs Laboratory of the USDA. To reduce the time period and range of yearly production within a herd, only production records started in 1978 and after were included in one set of analyses. Pedigree information from 1970 through 1977 records also was used to account

Received March 21, 1988.
Accepted June 23, 1988.

for numerator relationships among animals that had records in the analyses. Another data set was created from records started before 1978 to examine changes with time as average production was less before 1978 than later (see bottom of Table 3). For the 1977 and earlier analyses, relationships were computed from cows in data set (1970 through 1977) and parents of those cows. For the 1978 and later analyses, relationships were computed from cows in data set (1978 through 1985) and from ancestors of those cows appearing from 1970. Management practices may also have changed through time, although such information is not available. Records were eliminated for inconsistent dates, missing information, and milk records less than 2268 kg or greater than 18,144 kg and fat records less than 68 kg or greater than 680 kg.

Herds were chosen from the overall data to contain 300 to 400 cows with first lactation records in the time period to be analyzed. The goal was to have about 12 herds in each sample for analysis. Two samples were chosen from each state and time period. Characteristics of the data sets are in Table 1. The same procedure was followed when samples were taken according to herd production category except that considerably more than 12 herds were sometimes needed to obtain samples of about 4000 cows (Table 2).

Analyses by herd production category were done only for records of cows that first freshened in 1978 or after. Categories of production were determined from herd average for milk yield; records in the herd average associated with each record include average mature equivalent yield of cows with completed lactations that calved in the 12 mo preceding fresh date. Averages associated with individual records were reviewed, and the largest for any lactation started in the herd after 1977 was selected as the indicator of production category for all records from that herd. The same limits were used for the three states: low production herds had largest herd average of less than 7711 kg, middle production herds had largest herd average between 8392 and 9299 kg, and high production herds had largest herd average greater than 9979 kg. Gaps between categories were set deliberately to separate low and high groups from the middle group. California had few herds and cows in the low category; nearly all were included in the analysis with first lactation herd sizes ranging from 44 to 357. The count for herd size included all cows in the herd for the time period studied. Herd sizes for low herds ranged from 174 to 290 in New York and 170 to 399 in Wisconsin. Similarly, nearly all herds categorized as high in Wisconsin were needed for those analyses with herd sizes ranging from 75 to 233. Sizes of New York

TABLE 1. Numbers of cows with records and herds in each sample used to estimate genetic and environmental covariances for milk and fat yield. The fraction of nonzeros in the inverse of the relationship matrices indicate the fraction of ties caused by relationships within herd.

State and sample	1978 ⁺		1977 ⁻		Fraction nonzero off-diagonals in inverse of relationship matrix (within herd)	
	No. cows	No. herds	No. cows	No. herds	1978 ⁺	1977 ⁻
California						
Set 1	4168	12	4183	12	.34	.25
Set 2	4208	12	4051	13	.26	.39
New York						
Set 1	4065	12	3768	13	.80	.81
Set 2	3973	12	4629	15	.80	.76
Wisconsin						
Set 1	4675	12	4428	19	.86	.59
Set 2	4288	12	4162	18	.61	.54

¹ 1978⁺ = Records started 1978 through 1985, 1977⁻ = 1970 through 1977.

TABLE 2. Numbers of cows and herds in each sample used to estimate genetic and environmental covariances for milk and fat yield. The fraction of nonzeros in the inverse of the relationship matrices indicate the fraction of ties caused by relationships within herd.

State and sample	Herd production category					
	Low		Middle		High	
	No. cows	No. herds	No. cows	No. herds	No. cows	No. herds
California						
Set 1	2260	16	4145	12	4014	13
Set 2	1356	14	3945	12	3107	13
New York						
Set 1	3630	16	4344	13	3534	18
Set 2	3700	16	4303	15	3246	19
Wisconsin						
Set 1	3791	17	3794	14	2374	16
Set 2	3690	17	3909	14	2077	17
Fraction nonzero offdiagonals in inverse of relationship matrix (within herd)						
California						
Set 1		.16		.39		.29
Set 2		.09		.36		.36
New York						
Set 1		.82		.87		.94
Set 2		.81		.90		.88
Wisconsin						
Set 1		.73		.69		.83
Set 2		.67		.58		.57

herds in the high category ranged from 143 to 540. The range of sizes for middle category herds was about 300 to 400 for all three states.

Methods

Restricted maximum likelihood estimates (14) were obtained with a multivariate (milk and fat) animal model that included fixed herd-year-season effects (two seasons per year: February through June and July through January), animal genetic value, and random environmental value. Covariances between residual genetic and environmental effects and among environmental effects of relatives were assumed to be zero. The procedure is outlined in Henderson (9) and detailed in Swalve and Van Vleck (15) and Van Vleck and Dong (17). All numerator relationships within a herd (including sires and base animals without records) were used to determine the rela-

tionship matrix among animals with records. The inverse of that relationship matrix was used in the REML procedure to make the procedure computationally possible. If relationships are assumed to be zero with the animal model, genetic and environmental variances cannot be separated. Relationships across herds were assumed to be zero even if bulls were used in more than one herd in the sample. Such relationships do not seem to have much effect on estimates of variances and covariances (Robert P. Yerex, 1988, personal communication) because of the sparsity of across-herd relationships. Canonical transformation, factorization of the relationship matrix, and tridiagonalization were used to speed up the computations as described by Lawlor (10). Number of rounds of iteration was 300 per analysis for overall samples. Little change in estimates occurred after 50 rounds, although at 50 rounds some drift in estimates seemed to be

occurring. For the analyses by category of herd production, 500 rounds of iteration were done because drift was more noticeable in some of these analyses. Later inspection revealed no practically important changes after 300 rounds. Only samples for low production herds in California showed changes as much as .02 for estimates of heritability and genetic correlations from round 300 to round 500. Changes were in the third decimal place for most other analyses.

RESULTS

Estimates from samples over all herd production categories are in Table 3 for the two time periods and three states. Standard errors for heritability estimates (17) are probably about .03, larger than that for genetic

correlations and smaller for phenotypic correlations.

A somewhat surprising result is that estimates of heritabilities and genetic and phenotypic correlations are consistently larger for the early time period than for the last time period in spite of yield increases. The only exception is for milk yield in Wisconsin. This trend agrees with paternal half-sib analyses of New York data (13). Daughter on dam regressions from New York data (2, 16), however, showed no trend over time. Those analyses contained considerably larger data sets, although the procedures did not account for relationships other than parent-progeny and paternal half-sibs. Heritability estimates from New York data including protein yield by the same method, however, were about .36 for records made from 1981 through 1985 (17). Cue et al. (1) also

TABLE 3. Estimates¹ from analyses by restricted maximum likelihood of milk and fat yield for first lactation cows in California, New York, and Wisconsin for records started 1978 through 1985 (1978⁺) and started 1970 through 1977 (1977⁻).

Statistic	California	New York	Wisconsin	Mean
Heritability				
Milk				
1978 ⁺	.26	.28	.34	.29
1977 ⁻	.39	.32	.31	.34
Fat				
1978 ⁺	.29	.27	.30	.29
1977 ⁻	.43	.33	.34	.37
Correlation				
Genetic				
1978 ⁺	.70	.54	.62	.62
1977 ⁻	.82	.70	.67	.73
Phenotypic				
1978 ⁺	.82	.77	.78	.79
1977 ⁻	.84	.82	.81	.82
Phenotypic variance ²				
Milk, kg ²				
1978 ⁺	.8358	.9870	.8562	.8930
1977 ⁻	.9442	.6942	.6425	.7603
Fat, kg ²				
1978 ⁺	.7111	.7842	.7340	.7431
1977 ⁻	.7549	.5789	.5921	.6420
Mean milk, kg				
1978 ⁺	8562	8245	8077	
1977 ⁻	7751	7462	7018	

¹ Average of two samples.

² For milk, multiply by (1451.5)²; for fat, multiply by (54.43)².

found estimates of about .35 from Canadian data for a sire model with relationships assumed to be zero. Within the limits of sampling variance (records of about 8000 cows in each estimate), differences in estimates among states or milk and fat are not apparent.

Although the samples may be too small to be definitive, the observed decrease from early to late time period in genetic correlation is sizeable and seems fairly consistent over the three states. Genetic correlations between milk and fat from California data appear considerably larger than from data of the other states. Based on only two samples for each estimate, the differences across states might be attributed to sampling variance. Estimates from production category analyses, however, follow the same pattern, resulting in the average of estimates from eight samples of records from California probably being significantly greater than from 16 samples of New York and Wisconsin records. New York (17) and Canadian (1) studies, however, with protein data found genetic correlations of .70 and .73 between milk and fat yield for late 1970's and early 1980's data. Northeast data have generally produced smaller estimates of genetic correlations: .40, .43, .38 (7, 10, 12), although Wilcox et al. (20) found estimates of about .70 for 1960's data. Majjala and Hanna (11) suggested the average of genetic correlation estimates to be .81.

Phenotypic correlations follow the trend shown by the genetic correlations, but the magnitudes of differences are smaller. The New York study (17) of 1981 through 1985 records with protein reported a phenotypic correlation of .80 between milk and fat yield.

What is indicated by apparently larger genetic correlations in California as compared with those in New York and Wisconsin? Is the difference due to effects of management practices on genetic expression? In the production analyses (Table 4), genetic correlations seem to become smaller as production increases (except for one California estimate). Average milk production for the California samples that ignored herd yield, however, was greater, although not substantially so, than for New York and Wisconsin data.

Except for California samples, phenotypic variances were greater in the 1978 and later records than the pre-1978 records as has been reported [e.g., by Mirande and Van Vleck (13)

for residual variances from paternal half-sib analyses having much larger degrees of freedom for the residual components]. The California result is partly due to the greater genetic variance in the early time period. In California, the environmental variances are less for the pre-1978 period than in the later period for scaled fat yield (.4273 vs. .5042) and for scaled milk yield (.5750 vs. .6168). When averaged over all three states, phenotypic standard deviations for both milk and fat yield were about 8% greater for records in 1978 and later compared with the pre-1978 records.

Phenotypic variances seem to be somewhat smaller for Wisconsin and New York than for California for pre-1978 records. For the last period, New York seems to have the largest phenotypic variance. The scaled standard errors of the scaled variances are probably about $[2(1^4/8000)]^{.5} = .016$ for the average of two samples of 4000 cows. Mean production, however, was greatest in California and smallest in Wisconsin, although differences were not great.

Samples by Production Categories

Estimates by state and by category of herd average are in Table 4. Heritability estimates for milk and fat follow the usual pattern [e.g., (2, 3, 16)] for the three states in that heritability becomes substantially greater from low to middle to high herd production categories. The pattern and average estimates essentially are the same for milk and fat yield.

With exception of California, genetic correlation estimates decreased as production increased. The average absolute difference between the two estimates for each production category and state combination was relatively large, .070. Phenotypic correlations decreased slightly with each increase in category of herd average for the three states. The average absolute difference between the two samples was only .022.

Estimates for both genetic and phenotypic correlations were greater for California than for New York and Wisconsin for all three production categories as well as for the analyses for samples from all herds.

With the exception of New York middle and high categories of production, phenotypic variance increased as category of production

increased. The overall pattern based on relatively small samples is consistent with other estimates of phenotypic variance [e.g., (3, 13, 19)]. Environmental variances, however, were slightly greater in middle than in high category herds for California and New York records. The average absolute difference between the two samples for environment variance was .0513 for milk and .0549 for fat yield.

In a parallel study using daughter and dam pairs, only about 40% as many pairs were found in the California data (13.7%) as in New York (36.0%) and Wisconsin (34.7%) data (Kevin Wade, 1987, personal communication). This discrepancy suggested that reliability of pedigree identification may be different in California from that in New York and Wisconsin or that more replacements may be purchased in

TABLE 4. Estimates¹ from analyses by restricted maximum likelihood of milk and fat yield by herd production category (low, middle, high) for first lactation cows in California, New York, and Wisconsin for records started in 1978 and after.

Statistic	California	New York	Wisconsin	Mean
Heritability				
Milk				
Low	.17	.26	.27	.23
Middle	.26	.30	.31	.29
High	.36	.37	.34	.36
Fat				
Low	.16	.24	.29	.23
Middle	.24	.27	.33	.28
High	.35	.32	.42	.36
Correlation				
Genetic				
Low	.70	.63	.65	.66
Middle	.70	.58	.63	.64
High	.77	.54	.55	.62
Phenotypic				
Low	.84	.82	.80	.82
Middle	.83	.78	.76	.79
High	.79	.73	.70	.74
Phenotypic and environmental variance²				
Milk kg²				
Low	.6990 (.5809) ³	.8118 (.6048)	.7207 (.5232)	.7438 (.5693)
Middle	.9053 (.6690)	1.0685 (.7437)	.8384 (.5802)	.9374 (.6643)
High	1.0286 (.6552)	1.0188 (.6378)	.9741 (.6478)	.9619 (.6469)
Fat, kg²				
Low	.5973 (.5047)	.6768 (.5171)	.6489 (.4614)	.6410 (.4944)
Middle	.7776 (.5894)	.8658 (.6329)	.7299 (.4890)	.7911 (.5704)
High	.8703 (.5662)	.8584 (.5854)	.9022 (.5224)	.8770 (.5580)
Mean milk (kg)				
Low	7165	7107	7099	
Middle	8499	8326	8174	
High	9675	9422	9423	

¹ Average of two samples.

² For milk, multiply by (1451.5)²; for fat, multiply by (54.43)².

³ Environmental variance in parentheses.

California than in New York and Wisconsin. Heritability estimates from REML and an animal model decreased when only relationships due to sires were used compared with when all relationships were used (4, 5). Van Vleck and Hudson (18) showed that with method 3 (8), incorporating the numerator relationship matrix in a sire model increased heritability estimates from about .31 to .32. That the sire component of variance (and heritability) would increase can be shown theoretically with method 3. With REML the effect of ignoring certain relationships is more complex because relationships enter into estimates of breeding value, quadratics for estimating (co)variance components, and expectations (or prediction error variance) of the quadratics. Nevertheless, inverses of the relationship matrices were examined.

The right half of Table 1 and lower half of Table 2 show the fraction of nonzero elements in inverses of the within herd relationship matrices. These fractions are considerably different for the California samples as compared to New York and Wisconsin, especially for samples for the low category of herd average. Whether this difference suggests that true heritability in California is larger than estimated in this study is unknown. In a study to be reported later (6) that included matched daughters and dams, the fractions of nonzero elements in the inverse of relationship matrices were comparable in all three states and similar to those for New York and Wisconsin. From records started in 1978 and later, heritability estimates were slightly larger for milk yield in all three states and for each state except California for fat yield. A different set of herds was analyzed so that the estimates may be affected by the herds chosen for the samples.

CONCLUSIONS

Heritability for milk and fat yield seems similar in California, New York, and Wisconsin herds. The pattern that heritability for milk and fat yield is smallest in low production herds and largest in high production herds is the same for California, New York, and Wisconsin herds. Differences in heritability would make testing of daughters in high production herds more effective for evaluation of young bulls than in low production herds if genetic correlations

between genetic expression in different production categories are as large ($.90 >$) as reported by DeVeer and Van Vleck (3).

Genetic correlation between milk and fat yield seems to have decreased from the 1970 through 1977 data sets to the 1978 through 1985 data sets. Genetic correlation also seems to have decreased slightly as herd production category increased in the same time period. Genetic correlations, however, from the California samples were consistently greater than those from the New York and Wisconsin samples.

The pattern of phenotypic and environmental variances is confusing except that variances were smallest in herds in the low production category.

The significance of differences in relationship matrices for California as compared with New York and Wisconsin herds needs to be studied. If the reason is inaccurate identification, then use of such herds may lead to inaccurate predictions of breeding value from either animal or sire models, depending on whether sires or dams or both are misidentified. Purchased animals with missing identification would not allow correction for mates in predicting sire breeding values from animal or maternal grandsire models.

ACKNOWLEDGMENTS

Computing support of 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, the National Association of Animal Breeders, and US-Spain Joint Committee for Scientific and Technological Cooperation.

REFERENCES

- 1 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.
- 2 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.
- 3 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.
- 4 Dong, M. C. 1987. Estimation of genetic (co)variances for production, survival and reproduction. Ph.D. thesis. Cornell Univ., Ithaca, NY.

- 5 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.)
- 6 Dong, M. C., L. D. Van Vleck, and G. R. Wiggans. 1988. Effect of relationships on estimation of variance components with an animal model and restricted maximum likelihood. *J. Dairy Sci.* 71:3047.
- 7 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.
- 8 Henderson, C. R. 1953. Estimation of variance and covariance components. *Biometrics* 9:226.
- 9 Henderson, C. R. 1984. Application of linear models in animal breeding. Univ. Guelph Press, Guelph, Ont., Can.
- 10 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.
- 11 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.
- 12 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 concentraions. *J. Dairy Sci.* 67:2028.
- 13 Mirande, S. L., and L. D. Van Vleck. 1985. Trends in genetic and phenotypic variances for milk production. *J. Dairy Sci.* 68:2278.
- 14 Patterson, H. D., and R. Thompson. 1971. Recovery of interblock information when block sizes are unequal. *Biometrika* 54:545.
- 15 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.
- 16 Van Vleck, L. D., L. R. Cox, and S. L. Mirande. 1985. Heritability estimates of milk production from daughter on dam regression by year and management level. *J. Dairy Sci.* 68:2964.
- 17 Van Vleck, L. D., and M. C. Dong. 1988. Genetic (co)variances for milk, fat, and protein yield in Holsteins using an animal model. *J. Dairy Sci.* 71:3040.
- 18 Van Vleck, L. D., and G.F.S. Hudson. 1982. Relationships among sires in estimating genetic variance. *J. Dairy Sci.* 65:1663.
- 19 Vinson, W. E. 1987. Potential bias in genetic evaluations from differences in variation within herds. *J. Dairy Sci.* 70:2450.
- 20 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.