

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Faculty Papers and Publications in Animal  
Science

Animal Science Department

---

July 1988

## Effect of Relationships on Estimation of Variance Components with an Animal Model and Restricted Maximum Likelihood

M. C. Dong  
*Cornell University*

L. Dale Van Vleck  
*University of Nebraska-Lincoln, dvan-vleck1@unl.edu*

G. R. Wiggans  
*Agricultural Research Service, United States Department of Agriculture, Beltsville, MD*

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



Part of the [Animal Sciences Commons](#)

---

Dong, M. C.; Van Vleck, L. Dale; and Wiggans, G. R., "Effect of Relationships on Estimation of Variance Components with an Animal Model and Restricted Maximum Likelihood" (1988). *Faculty Papers and Publications in Animal Science*. 307.

<https://digitalcommons.unl.edu/animalscifacpub/307>

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.

# Effect of Relationships on Estimation of Variance Components with an Animal Model and Restricted Maximum Likelihood

M. C. DONG and L. D. VAN VLECK

Department of Animal Science  
Cornell University  
Ithaca, NY 14853

G. R. WIGGANS

Animal Improvement Programs Laboratory  
Agricultural Research Service  
United States Department of Agriculture  
Beltsville, MD 20705

## ABSTRACT

Milk, fat, and protein yields in first lactation of two samples of 12 herds consisting of 4020 and 4024 Holstein cows in New York were used to estimate genetic and phenotypic covariances with an animal model by REML. Relationships were calculated within herd in three ways to provide inverses of relationship matrices. Full relationships consisted of tracing relationships to parents (base animals) of animals with milk and fat records between 1970 and 1980. Full relationships less base animals consisted of tracing relationships only one generation back from when protein records became available in 1981. Sire-only relationships were compiled using only relationships through sires. Fractions of nonzero elements in inverses of within-herd relationship matrices were .72 for full, .62 for full-less-base, and .04 for sire-only relationships. Heritability estimates averaged over milk, fat, and protein were .35 with full, .33 with full-less-base, and .25 with sire-only relationship matrices. Estimates of genetic and phenotypic correlations were not affected by how relationships were computed. Comparable analyses of milk and fat yields from other sets of California, New York, and Wisconsin herds resulted in estimates of heritability that were slightly greater when only daughter and dam records were used than when all records in a herd were used.

## INTRODUCTION

The animal model was used by Swalve and Van Vleck (6) with REML to estimate variances and covariances for the first three lactations of New York Holsteins. The method required assuming all numerator relationships across herds were zero. Dong (1) used the same method for first lactation milk, survival, and calving interval. Because of computational necessity, he further assumed relationships within herd were due to sires alone. His estimate of heritability for first lactation milk yield was slightly less than the estimate of .33 by Swalve and Van Vleck (6). Dong (1) suggested that heritability estimates would be larger if more relationships were included with REML and an animal model.

Van Vleck and Hudson (9) showed for Henderson's method 3 (2) that including relationships among sires for a sire model would increase the estimate of heritability. The situation with REML and an animal model is more complex. A relationship matrix that is not the identity matrix would seem necessary to break confounding between additive genetic and environmental effects. In addition, REML incorporates inverse of the relationship matrix for computing quadratics for genetic effects and for expectation of quadratics. Lawlor (4) with REML and a sire model reported that estimates of heritability increased from .17 to .19 for milk yield and from .22 to .26 for fat yield when relationships among sires were utilized.

In a study of California, New York, and Wisconsin records, Kevin Wade (1987, personal communication) noticed that a smaller fraction of daughter and dam pairs was found in comparable data from California (14%) than from New York (36%) and Wisconsin (35%). This

---

Received March 10, 1988.

Accepted July 6, 1988.

observation suggested that fewer relationships are known among cows in California herds than in New York and Wisconsin. Effect of these differences on estimating heritability is not known.

The purpose of this study was to examine the hypothesis of Dong (1) that including more relationships would increase estimates of heritability.

#### MATERIALS AND METHODS

Data were those that Van Vleck and Dong (7) used to estimate genetic and environmental variances and covariances among first lactation milk, fat, and protein yields of New York Holsteins. Their data included two samples of 12 herds (Table 1). Analyses were based on REML (3, 4, 5) with all relationships traced back to base animals (parents without records but with daughters with milk and fat records after 1970). Only first lactation milk, fat, and protein records started after July 1981 were included in analyses. This set of analyses is denoted as with full relationships. The same records were used with the same REML procedure but with relationships computed in two other ways. First, relationships were traced from cows with milk, fat, and protein records only to first ancestors without protein records. Because of the limited time period for the full data set and for convenience, these analyses were denoted as with full-less-base relationships. Second, relationships among cows with milk, fat, and protein records were computed only from sires and denoted as with sire-only relationships. Inbreeding was not considered in any of the analyses.

Another set of analyses was done with California, New York, and Wisconsin first lactation milk and fat yields of Holsteins to examine effect of using records from daughter and dam pairs. These data were the same as those from which Van Vleck et al. (8) estimated genetic parameters with REML and an animal model. Two samples, each with 12 herds and about 3600 cows, were chosen from each state. These analyses were denoted as with all relationships. Other herds were used for analyses that included records from daughter and dam pairs. To make number of cows comparable per sample, more herds (Table 2) were included in these analyses. Other relationships in the data in addition to daughter and dam relationships were utilized. These analyses were denoted as with paired relationships.

#### RESULTS

Characteristics of inverses of within-herd relationship matrices for milk, fat, and protein analyses are shown in Table 3. These characteristics were fractions of off-diagonal elements that were negative or positive as well as ratio of sum of diagonal elements to number of diagonals. If relationships were all 0, the ratio would be 1. Fractions of nonzero, off-diagonal elements were slightly larger with full than with full-less-base relationships (.70 and .73 vs. .59 and .64). Ratios also were slightly larger with full than with full-less-base relationships. With sire-only relationships, the only nonzero elements, except for diagonals, are negative and were a fraction of only .03 and .04 for the two data sets. Ratios were much smaller for sire-

TABLE 1. Mean yields and number of cows and sires per herd for two samples of 12 New York Holstein herds with first lactation milk, fat, and protein records.

Sample	Mean yield			Mean no. per herd	
	Milk	Fat	Protein	Cows	Sires
	(kg)				
1	8435	304	272	335	102
2	8811	322	286	335	104

TABLE 2. Mean yields and number of cows and sires per herd for four samples from California, New York, and Wisconsin.

	Mean yield		Mean no. per herd		No. of herds
	Milk	Fat	Cows	Sires	
	(kg)				
California					
All <sup>1</sup>					
Sample 1	8548	305	347	79	12
Sample 2	8575	313	351	86	12
Paired <sup>2</sup>					
Sample 3	8989	320	325	125	13
Sample 4	8882	317	332	150	13
New York					
All <sup>1</sup>					
Sample 1	8023	294	339	97	12
Sample 2	8463	313	328	109	12
Paired <sup>2</sup>					
Sample 3	8482	308	304	107	14
Sample 4	8089	292	310	109	13
Wisconsin					
All <sup>1</sup>					
Sample 1	8073	299	390	148	12
Sample 2	8080	302	357	137	12
Paired <sup>2</sup>					
Sample 3	7990	294	227	96	17
Sample 4	8235	306	202	106	17

<sup>1</sup> First lactation records of all cows from random sample of herds with desired number of first lactation records.

<sup>2</sup> Only first lactation records of daughter and dam pairs included in another random sample of herds.

only as compared with the other two types of relationships.

Table 4 lists characteristics of the California, New York, and Wisconsin samples. As noted by Van Vleck et al. (8), fraction of nonzero off-diagonal elements in the inverses of within-herd relationship matrices was much less with all relationships for California (.34 and .26) than for New York (.79 and .79) or Wisconsin (.85 and .61). Ratios of traces to number of diagonals also were smaller with all relationships for California (1.33 and 1.34) than for New York (1.55 and 1.61) or Wisconsin (1.63 and 1.57). With relationships from paired data, characteristics of inverses are more similar for California, New York, and Wisconsin.

Estimates of heritabilities, genetic and phenotypic correlations, and phenotypic variances for milk, fat, and protein analyses of New York Holstein data are in Table 5. Phenotypic variances and phenotypic and genetic correlations were similar for the three ways of accounting for relationships. Heritability estimates with sire-only relationships were considerably smaller than with full and full-less-base relationships in agreement with the hypothesis of Dong (1). All six estimates of heritability (milk, fat, and protein for two samples) were slightly larger for full than for full-less-base relationships.

Table 6 compares California, New York, and Wisconsin analyses of milk and fat yields with

TABLE 3. Characteristics of elements of inverses of within herd relationship matrices for two samples of New York herds with milk, fat, and protein records.

Sample	Full relationship			Full-less-base relationship			Sire-only relationship		
	Fraction off-diagonals		Ratio <sup>1</sup>	Fraction off-diagonals		Ratio	Fraction off-diagonals		Ratio
	Negative	Positive		Negative	Positive		Negative	Positive	
1	.39	.31	1.47	.36	.23	1.46	.03	.00 <sup>2</sup>	1.19
2	.41	.32	1.51	.40	.24	1.50	.04	.00	1.20

<sup>1</sup> Ratio of sum of diagonal elements to number of diagonals.

<sup>2</sup> Only diagonals positive.

all and paired relationships. In this case, different herds and cows in the all and paired analyses make interpretation more difficult, especially for variances. Average yields were not greatly different for herds with all and with paired data. No trend was evident for estimates of genetic or phenotypic correlations. Heritability estimates also were not greatly different, although five of the six averages were larger for paired than for all relationships. Estimate of heritability for California for fat yield was the exception despite the California analyses having the greatest differences in

TABLE 4. Characteristics of inverses of relationship matrices for four samples from California, New York, and Wisconsin.

	Elements of inverse of relationship matrix		
	Fraction off-diagonals		Ratio <sup>1</sup>
	Negative	Positive	
California			
All <sup>2</sup>			
Sample 1	.22	.12	1.33
Sample 2	.17	.09	1.34
Paired <sup>3</sup>			
Sample 3	.32	.29	1.56
Sample 4	.41	.40	1.48
New York			
All <sup>2</sup>			
Sample 1	.43	.36	1.55
Sample 2	.41	.38	1.61
Paired <sup>3</sup>			
Sample 3	.41	.39	1.72
Sample 4	.38	.37	1.70
Wisconsin			
All <sup>2</sup>			
Sample 1	.44	.41	1.63
Sample 2	.32	.29	1.57
Paired <sup>3</sup>			
Sample 3	.44	.43	1.63
Sample 4	.39	.38	1.53

<sup>1</sup> Ratio of sum of diagonal elements to number of diagonals.

<sup>2</sup> First lactation records of all cows from random sample of herds with desired number of first lactation records.

<sup>3</sup> Only first lactation records of daughter and dam pairs included in another random sample of herds.

inverses of relationship matrices. Large samples seem to be needed to show whether selecting data to include only daughter and dam pairs increases heritability estimates.

The small fraction of daughter and dam pairs in California data may indicate lack of identification or purchase of replacements from other herds. Misidentification would be expected to reduce estimates of heritability. Greater milk yield (Table 2), however, usually is associated with larger estimates of heritability (8). Whether these factors affected estimates from analyses of California records is not known.

### CONCLUSIONS

Estimates of heritability from REML with an animal model were considerably smaller if relationships were from sires only as compared with from more complete relationships. Full relationships with REML from ancestors of

about two generations resulted in slightly larger estimates of heritability as compared with relationships from about one generation. C. R. Henderson (1988, personal communication) has proven that ignoring relationships that exist will result in a reduction in REML estimates of genetic variance. Results of this study were in agreement. Estimates of genetic and phenotypic correlations were not affected by completeness of relationships associated with REML for an animal model.

### ACKNOWLEDGMENTS

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. Computing support provided by the Cornell National Supercomputing Facility made this project possible.

TABLE 5. Average estimates of parameters from two samples of first lactation milk, fat, and protein yields from New York herds from restricted maximum likelihood for an animal model with relationships computed three ways.

Parameter	Relationships		
	Full	Full-less-base	Sire-only
Heritability			
Milk	.358	.330	.251
Fat	.358	.342	.264
Protein	.340	.314	.224
Correlation			
Genetic			
Milk, fat	.724	.726	.701
Milk, protein	.872	.884	.884
Fat, protein	.766	.777	.775
Phenotypic			
Milk, fat	.806	.806	.806
Milk, protein	.914	.914	.910
Fat, protein	.816	.816	.814
Phenotypic variance, <sup>1</sup> kg <sup>2</sup>			
Milk <sup>2</sup>	1.1861	1.1833	1.1690
Fat <sup>3</sup>	1.0037	1.0021	.9929
Protein <sup>4</sup>	1.2391	1.2363	1.2214

<sup>1</sup> Scaling factors were used to reduce rounding errors in computations by making scaled variances of milk, fat, and protein approximately unity.

<sup>2</sup> Multiply by (1451.5)<sup>2</sup>.

<sup>3</sup> Multiply by (54.43)<sup>2</sup>.

<sup>4</sup> Multiply by (45.36)<sup>2</sup>.

TABLE 6. Average estimates of parameters from two samples of California, New York, and Wisconsin first lactation milk and fat yields from restricted maximum likelihood with records from all cows in herds and from daughter and dam pairs only.

Item	Heritability		Correlation		Phenotypic variance	
	Milk	Fat	Genetic	Phenotypic	Milk <sup>1</sup>	Fat <sup>2</sup>
California						
All <sup>3</sup>	.262	.291	.704	.816	.8360	.7085
Paired <sup>4</sup>	.296	.244	.630	.804	.9315	.7644
New York						
All	.276	.274	.536	.768	.9840	.7853
Paired	.300	.285	.600	.780	.9756	.8080
Wisconsin						
All	.342	.300	.616	.776	.8557	.7298
Paired	.378	.311	.606	.752	.8094	.6854

<sup>1</sup> Multiply by (1451.5 kg)<sup>2</sup>.

<sup>2</sup> Multiply by (54.43 kg)<sup>2</sup>.

<sup>3</sup> First lactation records of cows from two random samples of herds with desired number of first lactation records.

<sup>4</sup> Only first lactation records of daughter and dam pairs included from two other random samples of herds.

#### REFERENCES

- 1 Dong, M. C. 1987. Estimation of genetic (co)variances for production, survival and reproduction. Ph.D. thesis, Cornell Univ., Ithaca, NY.
- 2 Henderson, C. R. 1953. Estimation of variance and covariance components. *Biometrics* 9:226.
- 3 Henderson, C. R. 1984. Application of linear models in animal breeding. Univ. of Guelph Press, Guelph, Ont., Can.
- 4 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.
- 5 Patterson, H. D., and R. Thompson. 1971. Recovery of interblock information when block sizes are unequal. *Biometrika* 54:545.
- 6 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.
- 7 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.
- 8 Van Vleck, L. D., M. C. Dong, and G. R. Wiggans. 1988. Genetic (co)variances for milk and fat yield in California, New York, and Wisconsin for an animal model with restricted maximum likelihood. *J. Dairy Sci.* 71:3053.
- 9 Van Vleck, L. D., and G.F.S. Hudson. 1982. Relationships among sires in estimating genetic variance. *J. Dairy Sci.* 65:1663.