

January 1990

Genotype by Environment Interactions for Milk and Fat Production Across Regions of the United States

M. J. Carabano
Cornell University

K. M. Wade
Cornell University

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

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

 Part of the [Animal Sciences Commons](#)

Carabano, M. J.; Wade, K. M.; and Van Vleck, L. Dale, "Genotype by Environment Interactions for Milk and Fat Production Across Regions of the United States" (1990). *Faculty Papers and Publications in Animal Science*. 135.
<http://digitalcommons.unl.edu/animalscifacpub/135>

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.

Genotype by Environment Interactions for Milk and Fat Production Across Regions of the United States

M. J. CARABAÑO,¹ K. M. WADE, and L. D. VAN VLECK²

Department of Animal Science
Cornell University
Ithaca, NY 14853

ABSTRACT

(Co)variance components for regional production of milk, fat, and fat percentage were estimated via multiple-trait REML using DHIA data from California, New York, and Wisconsin. Data consisted of registered and grade first lactation yields of Holstein cattle from 1970 through 1984. Records were limited to daughters from sires common to the pairs of states analyzed. Averaged numbers of records from California, New York, and Wisconsin were 419,185, 460,296, and 449,866. Genetic correlations between New York and Wisconsin for milk, fat, and fat percentage (.99, .98, and .99) were larger than those between California and New York (.95, .95, and .98) and those between California and Wisconsin (.94, .93, and .98). Heritabilities in New York for milk, fat, and fat percentage (approximately .27, .26, and .48) were larger than those of Wisconsin (approximately .23, .22, and .43) and California (approximately .22, .21, and .34). Grade records were also investigated; numbers of records in California were reduced by about 25%, whereas records in both New York and Wisconsin were reduced by about 66%. Genetic correlations were virtually unchanged, but heritabilities for milk and fat in the three states were similar (approximately .17) and significantly smaller than those from combined registered and grade daughters. In conclu-

sion, there appeared to be no sire by region interaction.

(Key words: genotype by environment interaction, milk production, restricted maximum likelihood)

INTRODUCTION

Use of bulls across regions of the United States raises questions about the use of records from daughters located in subpopulations with high heritabilities to evaluate and select sires for milk traits in other subpopulations and also whether sires are expected to rank similarly in different regions within the country. Previous studies (17, 19) showed that heritabilities for milk yield, fat yield, and fat percentage do not differ significantly when samples from California (CA), New York (NY), and Wisconsin (WI) are compared. Those results suggest there would not be significant gain from using a larger percentage of daughters from any of these states in order to evaluate sires. In a study by Lytton and Legates (9), a high estimate of the genetic correlation among sires' breeding values and a low estimate of the sire by region interaction component of variance led them to conclude that no interaction between northern and southern regions of the US existed, but they suggested that additional confirmation was needed.

The primary purpose of this study was to investigate the possibility of a genotype by environment interaction for dairy production traits across different regions of the US. The three states chosen were CA, NY, and WI as these states might be considered representative of different environments in the US, representing three important producing regions differing in location, climate, mean production, and management. Consequently, traits were defined according to "state of production", and (co)variance components and genetic correlations across states were estimated.

Received March 7, 1989.

Accepted May 29, 1989.

¹Present address: Departamento de Genética, Instituto Nacional de Investigaciones Agrarias, Carretera de La Coruña, km 7, 28040 Madrid, Spain.

²Present address: Department of Animal Science, University of Nebraska, Lincoln 68583.

MATERIALS AND METHODS

Data

Data sets from CA, NY, and WI were obtained from the Animal Improvement Programs Laboratory (AIPL) of the USDA. Data contained 305-d, mature equivalent, first lactation records for milk production, fat production, and fat percentage of Holstein cows freshening from 1970 through 1984. Both registered and grade cattle were represented. The original size of the data set for each state was about 600,000 records. Each data set was edited for sires with 10 or less daughters and herd-year-season (HYS) with only one sire. Sires common to each pair of the three states being analyzed were then identified. Because the number of sires found in each case corresponded roughly to the computational limits, records in each analysis were limited to those of sires common to the pair of states in question. Thus, two overlapping data sets were used for each state. A summary of the data and edits is in Table 1. Statistics for the combination of registered and grade (whole) and grade only data sets are in Table 2.

METHODS

(Co)variance components for milk production, fat production, and fat percentage, within and among states, were estimated using a two-trait REML procedure. Therefore, analyses were performed on pairs of states rather than on all three states simultaneously. Genetic correlations that were estimated were restricted to the same trait across states (e.g., milk in CA with milk in NY, but not milk in CA with fat in NY) for all data and for the grade only data, a total of 18 analyses.

The model equation used for the analyses was:

$$y_{ijkl} = h_i + p_j + s_{k(j)} + e_{ijkl}$$

where:

y_{ijkl} = the vector of milk production (or fat production or fat percentage) in states 1 and 2, or 1 and 3, or 2 and 3;

h_i = the fixed effect due to HYS i of freshening; there were two seasons (December through April and May through November).

p_j = the fixed effect due to period j of entry to AI service by the sire; $j = 1, 2, 3$ (1 was 1970 through 1974, 2 was 1975 through 1979, and 3 was 1980 through 1984).

$s_{k(j)}$ = the random effect due to sire k within period j ; and

e_{ijkl} = the residual effect for the record of cow l in HYS i , daughter of sire k that entered service in period j .

In matrix notation, the model used for the analyses was:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} Z_1 & 0 \\ 0 & Z_2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \end{pmatrix},$$

where y_1 and y_2 are vectors of observations for a trait (milk production or fat production or fat percentage) in states 1 and 2 or 1 and 3 or 2 and 3; X_i and Z_i are known incidence matrices, associated with vectors of fixed (β_i) and random (u_i) effects; and e_i is the vector of residual effects. If g_{ii} is the sire genetic variance for state i , g_{ij} is the sire genetic covariance between states i and j , and r_{ii} is the residual variance for state i ; then for N_i being the total number of observations in state i , the assumptions for u and e are:

$$\begin{pmatrix} u \\ e \end{pmatrix} \sim \text{MVN} \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} G & 0 \\ 0 & R \end{pmatrix} \right\}$$

where:

G = $\text{var}(u) = G_0 \otimes A$,

for $G_0 = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$;

A = the numerator relationship matrix among sires; and

$$R = \text{var}(e) = \begin{pmatrix} r_{11} I_{N_1} & 0 \\ 0 & r_{22} I_{N_2} \end{pmatrix}.$$

where:

- C = the inverse of the coefficient matrix;
- C_{ij}^{22} = block of C corresponding to sires for states i and j;
- W = [X:Θ]; and
- q = number of sires.

To simplify computations, an equivalent model, described by Lawlor (8), was defined by use of the decomposition of $A = LL'$ as follows:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \Theta_1 & 0 \\ 0 & \Theta_2 \end{pmatrix} \begin{pmatrix} u_1^* \\ u_2^* \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

where $\Theta = ZQ$, for $Q = I_2 \otimes L$ and $u^* = Q^{-1}u$, for I_2 an identity matrix of order 2. Thus, $\text{var}(u^*) = G^* = G_0 \otimes I$, rather than $G_0 \otimes A$.

Restricted maximum likelihood estimates of the (co)variance components were obtained via an expectation maximization algorithm (3). The equations required for iteration to solution were:

$$\hat{u}_i^* \hat{u}_j^* + \text{tr}(C_{ij}^{22}) = \text{tr}(g_{ij} I_q)$$

$$\hat{e}_i' \hat{e}_i + \text{tr}(WCW')_{ii} = \text{tr}(r_{ii} I_{N_i})$$

These methods follow those described by Carabaño et al. (1).

RESULTS AND DISCUSSION

Estimates for genetic correlations are in Table 3. All were greater than .90, which seems to indicate little evidence of any important genotype by environment interactions for production traits across the three states, because "if the genetic correlation is high, then performance in two different environments represents very nearly the same characters, determined by very nearly the same set of genes" [Falconer (7)]. If these states are representative, then there should be no significant reranking of bulls across these different regions of the US. Lytton and Legates (9) concluded similarly with regard to northern and southern regions of the US.

TABLE 1. Structure of data sets.

	Original data sets					
	California (CA) (576,006 records)		New York (NY) (668,678 records)		Wisconsin (WI) (632,309 records)	
	Whole data sets (edited)					
State with records ¹	CA ^a	CA ^b	NY ^a	NY ^c	WI ^b	WI ^c
Paired state	NY	WI	CA	WI	CA	NY
Number of sires common to both states	1235	1416	1235	1175	1416	1175
Number of records	398,158	440,211	494,049	426,543	494,582	405,150
Number of herd-year-seasons	17,553	17,953	72,354	65,729	78,946	70,092
	Grade data sets (edited)					
State with records ¹	CA ^a	CA ^b	NY ^a	NY ^c	WI ^b	WI ^c
Paired state	NY	WI	CA	WI	CA	NY
Number of sires common to both states	854	1144	854	719	1144	719
Number of records	286,493	338,503	190,789	127,437	207,522	111,489
Number of herd-year-seasons	13,741	14,764	32,271	24,652	37,208	22,459

¹Columns with the same superscript were involved in the same analysis.

Lack of knowledge concerning the distribution of the obtained estimates of genetic correlation allows neither any accurate calculation of standard errors nor an objective statistical model to test whether or not the genetic correlations are significantly different from 1. However, the authors trust that the standard errors associated with the estimates of genetic correlation are relatively small, given that the average number of daughters per sire (>300) and the average total number of observations (>300,000) were relatively large. Lack of evidence for genotype by environment interaction in the US from other studies (5, 9, 15) also supports this claim. In an effort to clarify this, approximate standard errors for the estimates of genetic correlation were calculated after Robertson (12), based on approximate standard errors for the heritabilities after Swiger et al. (14). The standard errors for the estimates of genetic correlation were all small, the largest being .003. It should be noted that these approximations are not particularly good (12) when the genetic correlation is near 1, as is the case here.

An obvious trend in rankings of the correlations within each production trait was observed. Those for NY and WI were largest, followed by CA and NY, with CA and WI smallest. These trends were also found for grade data with smaller, but probably not significantly smaller, correlations. If it is reasonable to consider CA, NY, and WI as high, medium, and low in milk production (despite the fact that NY and WI are quite close), then these correlations agree with work done by DeVeer and Van Vleck (5). They found the highest correlation (.97) between medium and low production herds from Northeast data during 1970 and 1971.

The largest correlations were found for fat percentage; correlations were similar across pairs of states and also across analyses (whole versus grade only). These correlations may indicate that proportion of fat in milk is essentially the same trait across states, even though actual means for the three states were different. Although no significant reranking of bulls is expected across different regions, differences among bulls may be larger or smaller when evaluations are compared in various states. Such variation in differences is an interaction also, but it was not considered in this study.

Because variance components were avail-

able, estimates of heritability were calculated for each trait. The two estimates within state for each trait were in close agreement, probably because of the large number of cows and sires common to the two samples. Consequently, averages of estimates from both samples are in Table 4. Heritability estimates for milk and fat production were about .24 for all states when using whole data and about .17 for grade data. Differences between the two sets were less for fat percentage; whole data averaged about .42 whereas grade data averaged around .37. New York data had the highest heritability estimates for all three traits.

Estimates of heritability for the whole data set, although smaller than those obtained in some studies (2, 4, 6, 13, 16, 17, 19), agree with those of Lawlor (8) and DeVeer and Van Vleck (5) and are similar to those quoted by Majjala and Hanna (10). Some studies (11, 18) found a decrease in heritability over time since 1977. Smaller estimates found for grade only data are difficult to explain. Differences in heritability estimates were found between grade and registered records by Wade and Van Vleck (19) from daughter on dam regression, but these differences were smaller. Other studies using daughter on dam regression (13) and paternal half-sib correlation (18) found little or no differences in estimates between data for grade and registered cows. Because the present study was primarily intended to investigate ge-

TABLE 2. Summary of means¹ and standard deviations¹ for combined registered and grade (whole) and grade only (grade) Holstein data in California, New York, and Wisconsin from 1970 through 1984.

	Whole data		Grade data	
	Mean	SD	Mean	SD
	California			
Milk, kg	8632	1651	8576	1616
Fat, kg	309	57	307	55
Fat, %	3.59	.37	3.59	.37
	New York			
Milk, kg	7584	1592	7642	1595
Fat, kg	275	57	276	57
Fat, %	3.65	.40	3.63	.41
	Wisconsin			
Milk, kg	7560	1478	7580	1464
Fat, kg	281	54	281	53
Fat, %	3.74	.41	3.72	.41

¹Averages from pairs of analyses.

TABLE 3. Estimates of genetic correlations between pairs of states from Holstein records in California (CA), New York (NY), and Wisconsin (WI) from 1970 through 1984 using the whole and grade only data sets.

Pair	Whole			Grade		
	Milk	Fat	Fat %	Milk	Fat	Fat %
CA and NY	.954	.947	.983	.935	.914	.981
CA and WI	.939	.929	.980	.903	.901	.979
NY and WI	.987	.984	.992	.980	.975	.990

TABLE 4. Estimates¹ of heritability for production from Holstein records in California (CA), New York (NY), and Wisconsin (WI) from 1970 through 1984 using the whole and grade only data sets.

Item	Whole			Grade		
	Milk	Fat	Fat %	Milk	Fat	Fat %
CA	.22	.22	.34	.18	.17	.32
NY	.27	.26	.48	.17	.17	.41
WI	.23	.22	.44	.17	.17	.38
Overall ²	.24	.23	.42	.17	.17	.37

¹Average of two estimates from paired states.

²Unweighted average.

notype by environment interaction, a breakdown of the data into grade and registered records was not considered necessary initially. However, after small estimates of heritability were found in grade only data, one registered only data set was analyzed (milk in NY and WI). The same procedure and edits applied for

this last analysis. A summary of both the data and results is in Table 5. The correlation was large (.99), as expected, and heritabilities were larger than those obtained from the whole analysis (.32 for NY and .29 for WI). If one assumes that the other heritabilities from data of registered cows follow a similar fashion, then

TABLE 5. Summary of means (\pm SD), estimates of sire (σ_s^2) and residual (σ_r^2) variance and between-state sire covariance (σ_{s_i, s_j}), and estimates of heritability and genetic correlation from Holstein records of registered cows in New York (NY) and Wisconsin (WI) from 1970 through 1984.

	Registered New York data	Registered Wisconsin data
Number of records	208,468	208,993
Number of sires		855
Mean		
Milk, kg	7593 \pm 1602	7615 \pm 1494
Fat, kg	277 \pm 58	283 \pm 55
Fat, %	3.67 \pm .41	3.75 \pm .41
Variance components for milk, kg ²		
σ_s^2	127,902	105,681
σ_r^2	1,478,220	1,368,177
Covariance component for milk, kg ²		
$\sigma_{s_{NY-WI}}$	115,225	
Heritability (milk)	.32	.29
Genetic correlation (milk for NY and WI)	.991	

TABLE 6. Estimates of sire ($\hat{\sigma}_s^2$) and residual ($\hat{\sigma}_r^2$) variances¹ and between state sire covariances ($\hat{\sigma}_{s_i s_j}$) from Holstein records in California (CA), New York (NY) and Wisconsin (WI) from 1970 through 1984 using the combination of registered and grade (whole) and grade only data sets.

Item	Whole						Grade					
	Milk (kg ²)		Fat (kg ²)		Fat (% ²)		Milk (kg ²)		Fat (kg ²)		Fat (% ²)	
	$\hat{\sigma}_s^2$	$\hat{\sigma}_r^2$	$\hat{\sigma}_s^2$	$\hat{\sigma}_r^2$	$\hat{\sigma}_s^2$	$\hat{\sigma}_r^2$	$\hat{\sigma}_s^2$	$\hat{\sigma}_r^2$	$\hat{\sigma}_s^2$	$\hat{\sigma}_r^2$	$\hat{\sigma}_s^2$	$\hat{\sigma}_r^2$
State	Variance components within state											
CA	109,476	1,880,412	123	2168	.0096	.1032	83,784	1,857,128	95	2130	.0088	.1027
NY	106,584	1,478,571	120	1777	.0171	.1252	67,552	1,511,281	78	1793	.0144	.1275
WI	83,107	1,357,213	99	1717	.0161	.1326	56,745	1,356,815	71	1690	.0138	.1337
Trail	Between-state covariances											
Milk, kg ²	$\hat{\sigma}_{CA-NY}$	$\hat{\sigma}_{CA-WI}$	$\hat{\sigma}_{NY-WI}$	$\hat{\sigma}_{CA-NY}$	$\hat{\sigma}_{CA-WI}$	$\hat{\sigma}_{NY-WI}$	$\hat{\sigma}_{CA-NY}$	$\hat{\sigma}_{CA-WI}$	$\hat{\sigma}_{NY-WI}$	$\hat{\sigma}_{CA-NY}$	$\hat{\sigma}_{CA-WI}$	$\hat{\sigma}_{NY-WI}$
Fat, kg ²	103,104	86,122	96,338	71,009	61,916	60,341	79	75	71	.0112	.0138	.0138
Fat, % ²	116	99	110	79	75	71	.0107	.0112	.0112	.0112	.0112	.0138
Test, % ²	.0123	.0120	.0171	.0171	.0171	.0171	.0107	.0107	.0107	.0107	.0107	.0138

¹Average of two estimates from paired states.

this result seems to indicate that estimates from the whole data were an average of the estimates from registered and grade data, at least in the case of NY and WI. Also, these estimates then would be in good agreement with previously cited studies (2, 4, 6, 13, 16, 17, 19). The questions as to what the breakdown of the data is in the other studies and whether or not this difference between registered and grade is a real phenomenon are not answered.

Two studies (17, 19) merit special mention since they dealt with the same three states over roughly the same time period as that studied here. Both studies reported larger heritabilities for milk, fat, and fat percentage than those in this study. There are two main reasons offered for the smaller estimates found here. First, records in the data sets were restricted to daughters of sires with 10 or more daughters. About 62% of sires in NY and WI had 50 or more daughters in each of those two states, and this proportion was closer to 70% for California data. Second, daughters were limited to bulls common to at least two of the three states analyzed. Both of these facts, especially the latter, imply that the bulls were a selected group, which would be expected to reduce the genetic variance and, hence, heritability. The estimates of sire and residual variance are in Table 6. The estimates of genetic variance found in this study (four times the estimate of the sire component) do not agree with those obtained by Van Vleck et al. (17) using an animal model. This difference supports the notion that selection on bulls contributed to reduced heritability estimates in this study. There is also the chance pedigree information available for grade animals is less accurate; when the grade only subsets were investigated, numbers of records were only reduced by about 25% in CA but by about 66% in both NY and WI. If the situation in CA is indicative of misidentification, this may lead to smaller estimates of heritability for grade data, as has been discussed in a previous study (19).

CONCLUSIONS

Estimates of genetic correlation reported here indicate no important genotype by environment interaction for milk yield, fat yield, or fat percentage among the states examined. If these states can be considered representative of

their regions, the conclusion is that bulls with daughters in one region would not be expected to be significantly reranked on records of daughters in another region of the US.

Heritability estimates from grade data were smaller than estimates from the whole data set, and indications are that estimates from grade data are smaller than those from records of registered cows. Estimates for the whole data set are somewhat smaller than sometimes reported, possibly because of bias from selection on bulls. Heritability estimates across states are so similar that there is no indication that selection of bull dams should be made only in some regions. There are indications of differing variances across states, and as has been suggested (19), perhaps heterogeneity of variance should be accounted for in current evaluation schemes.

ACKNOWLEDGMENTS

The authors wish to thank G. R. Wiggans at AIPL, USDA for supplying the data used in this study. Partial support was given by the US-Spain Joint Committee for Scientific and Technological Cooperation. Computing support was from the Dairy Records Processing Laboratory, Ithaca, NY, and the Cornell National Supercomputer Facility, which is funded in part by the National Science Foundation, New York State, and by the IBM Corporation.

REFERENCES

- 1 Carabaño, M. J., L. D. Van Vleck, and G. R. Wiggans. 1989. Estimation of genetic parameters for milk and fat yields of dairy cattle in Spain and the United States. *J. Dairy Sci.* 72:3013.
- 2 Cue, R. L., H. G. Monardes, and J. F. Hayes. 1987. Correlations between production traits in first lactation Holstein cows. *J. Dairy Sci.* 70:2132.
- 3 Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. Maximum likelihood from incomplete data via an EM algorithm. *J. R. Stat. Ser. B* 39:1.
- 4 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.
- 5 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.
- 6 Dong, M. C., L. D. Van Vleck, and G. R. Wiggans. 1988. Effects of relationships on estimation of variance components with an animal model and restricted maximum likelihood. *J. Dairy Sci.* 71:3047.
- 7 Falconer, D. S. 1981. Introduction to quantitative genetics: 2nd ed. Longman Inc., New York, NY.

- 8 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.
- 9 Lytton, V. H., and J. E. Legates. 1966. Sire by region interaction for production traits in dairy cattle. *J. Dairy Sci.* 49:874.
- 10 Maijala, K., and M. Hanna. 1974. Reliable phenotypic and genetic parameters in dairy cattle. Page 541 in *Proc. Int. Congr. Quant. Genet. Appl. Livest. Prod.*, Madrid, Spain.
- 11 Mirande, S. L., and L. D. Van Vleck. 1985. Trends in genetic and phenotypic variances for milk production. *J. Dairy Sci.* 68:2278.
- 12 Robertson, A. 1959. The sampling variance of the genetic correlation coefficient. *Biometrics* 15:469.
- 13 Schneider, J. C., and L. D. Van Vleck. 1986. Heritability estimates for first lactation milk yield of registered and nonregistered Holstein cows. *J. Dairy Sci.* 69:1652.
- 14 Swiger, L. A., W. R. Harvey, D. O. Everson, and K. E. Gregory. 1964. The variance of intraclass correlation involving groups with one observation. *Biometrics* 20: 818.
- 15 Van Vleck, L. D. 1963. Genotype and environment in sire evaluation. *J. Dairy Sci.* 46:983.
- 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., 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.
- 18 Van Vleck, L. D., J. C. Schneider, and C. P. Van Tassel. 1988. Comparison of heritability and variation for milk yield of registered and nonregistered Holstein cows. *J. Dairy Sci.* 71:2271.
- 19 Wade, K. M., and L. D. Van Vleck. 1989. Genetic parameters for production traits of Holsteins in California, New York and Wisconsin. *J. Dairy Sci.* 72:1280.