

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Faculty Papers and Publications in Animal  
Science

Animal Science Department

---

January 1983

## Variances of Prediction Errors for Differences of Estimated Genetic Values of Cows in the Same and in Different Herds

D. C. Bolgiano  
*Cornell University*

R. L. Quaas  
*Cornell University*

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

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



Part of the [Animal Sciences Commons](#)

---

Bolgiano, D. C.; Quaas, R. L.; and Van Vleck, L. Dale, "Variances of Prediction Errors for Differences of Estimated Genetic Values of Cows in the Same and in Different Herds" (1983). *Faculty Papers and Publications in Animal Science*. 393.

<https://digitalcommons.unl.edu/animalscifacpub/393>

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.

# Variances of Prediction Errors for Differences of Estimated Genetic Values of Cows in the Same and in Different Herds

D. C. BOLGIANO,<sup>1</sup> R. L. QUAAS, and L. D. VAN VLECK

Department of Animal Science  
Cornell University  
Ithaca, NY 14853

## ABSTRACT

A modification of the evaluation method for cows in the northeast United States allowed simultaneous evaluation of cows in pairs of herds and computation of variances of prediction errors of differences of genetic potential between pairs of cows within and across herds. These variances were examined for 20 pairs of herds in Holstein, Jersey, and Brown Swiss breeds. Current evaluations of cows are nearly as accurate for comparing cows from different herds as they are for comparing cows from the same herd.

## INTRODUCTION

Cows in Dairy Herd Improvement Association (DHIA) herds in the northeast United States currently are evaluated by the Northeast A<sup>-1</sup> estimated transmitting ability (NEA<sup>-1</sup> ETA) procedure (2, 5). This method uses Henderson's mixed model equations (3) to obtain best linear unbiased predictions (BLUP) of the additive genetic values of all cows in a particular herd. The predicted genetic values then are halved to arrive at predictions of cows' transmitting abilities. The procedure uses all records on all cows in a herd and incorporates information from proofs of any artificial insemination (AI) sires having progeny in that herd, including the sires' predicted transmitting abilities and the number of daughter records used to evaluate the bulls (5, 6).

The NEA<sup>-1</sup> ETA evaluations aid the dairy-producer in choosing which cows to rebreed so

that they may have another lactation and also which cows to breed to produce replacement females. The evaluations also might be used by an AI organization for selecting dams of bulls. The AI stud wants to mate its best bulls to the best cows in the population, keeping the male offspring for progeny testing. Because an AI bull has the possibility for wide use, the best possible parents should be chosen for such a sire. The primary objective of this study was to examine the accuracy of the NEA<sup>-1</sup> ETA evaluations when they are used to compare genetic values of cows from different herds as against comparing genetic values of cows from within the same herd. If there is not a substantial loss of accuracy when cows from different herds are compared, then an AI stud may use these evaluations to select dams of bulls with more confidence.

For Henderson's mixed model equations and BLUP of random effects, a measure of accuracy associated with these predictions is the variance of prediction errors, i.e.,  $\text{var}(\hat{u} - u)$ , where  $u$  is a random effect in the model with mean 0 and variance  $\sigma_u^2$ . Variance of prediction errors is a desirable standard of accuracy, because it gives a measure of dispersion of the predictors about the true predictands and can be calculated from the elements of the inverse of the coefficient matrix of the mixed model equations (3, 4).

In our study, cows from pairs of herds were evaluated simultaneously by the NEA<sup>-1</sup> ETA procedure, and variances of prediction errors of differences of genetic value between cows from different herds were compared to variances of prediction errors of differences of genetic value between cows from the same herd.

## MODEL OF COW EVALUATION

The additive genetic model in the NEA<sup>-1</sup> ETA cow evaluation can be written

Received October 14, 1982.

<sup>1</sup>Department of Biostatistics, University of Washington, Seattle 98195.

$$y = X\beta + (Z \ 0) \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + Zp + e \quad [1]$$

where

$y$  is a vector of all 305-day, mature equivalent (ME) lactation records on all cows in a herd,

$X$  is a design matrix relating fixed effects in vector  $\beta$  to the records,

$\beta$  is a vector of fixed year-season effects,

$(Z \ 0)$  is a design matrix relating random effects in vectors

$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$  to lactation records,

$a_1$  is a vector of random additive genetic effects of all cows in the herd with records,

$a_2$  is a vector of random additive genetic effects of all AI sires with daughters in the herd,

$p$  is a vector of random nonadditive genetic and permanent environmental effects of all cows with records, and

$e$  is a vector of random residual terms associated with the lactation records.

The following conditions concern the model

$$E(y) = X\beta \quad E(a'_1 \ a'_2) = 0 \quad E(p) = 0 \quad E(e) = 0$$

$$\text{Var} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} h^2 \sigma^2$$

$$\text{Var}(p) = I(r - h^2) \sigma^2$$

$$\text{Cov}(a'_1 \ a'_2, p) = 0 \quad \text{Cov}(a_1 \ a'_2, e) = 0$$

$$\text{Cov}(p', e) = 0$$

$$\text{Var}(e) = I(1 - r) \sigma^2$$

where

$A_{ij}$  represents a submatrix of the numerator relationship matrix for the cows and sires represented in  $a'_1$  and  $a'_2$ ,

$h^2$  is heritability in the narrow sense = .25,

$r$  is repeatability = .40, and

$\sigma^2$  is the phenotypic variance.

The resulting mixed model equations are

$$\begin{bmatrix} X'X & X'Z & 0 & X'Z \\ Z'X & Z'Z + A^{11}t & A^{12}t & Z'Z \\ 0 & A^{21}t & D + A^{22}t & 0 \\ Z'X & Z'Z & 0 & Z'Z + Ik \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{p} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \\ Q'\hat{p} \\ Z'y \end{bmatrix} \quad [2]$$

where  $t = (1 - r)/h^2$ ,  $k = (1 - r)/(r - h^2)$ ,  $D$  is a diagonal matrix of the additions made to the diagonal elements of the sire equations, and  $Q$  is a diagonal matrix of the adjustments made to AI sire's breeding values ( $\hat{p}$ ) when they are incorporated as described in equation [6] (1, 5). The  $A_{ij}$  is a submatrix of the inverse of the numerator relationship matrix for the cows and sires in the herd. The solution of this set of mixed model equations (MME) yields BLUP of the additive genetic values of all cows in the herd with records.

These procedures can be extended to handle simultaneous evaluation of any two herds. Now  $\beta$  will be a vector of fixed herd-year-season (HYS) effects representing all year-seasons from both herds in which cows had records. That no cow has a daughter in the other herd and that no cow has records in two herds are assumed. Additive genetic values of cows to be evaluated from both herds will be in  $a_1$ . The records are sorted so that all cows from one herd will appear first in  $a_1$  and cows from the other herd will follow. Genetic values of all sires used in either one or both herds will be in  $a_2$ . Finally,  $p$  will represent the nonadditive genetic plus permanent environmental effects for cows in both herds. The order of animals in  $p$  is the same as in  $a_1$ . Because  $Z'Z + Ik$  is a diagonal matrix, equations for  $\hat{p}$  can be absorbed into equations for  $\hat{\beta}$  and  $\hat{a}_1$ . This reduces the set of equations to be solved to

$$\begin{bmatrix} \mathbf{X}'\mathbf{M}\mathbf{X} & \mathbf{X}'\mathbf{M}\mathbf{Z} & \mathbf{0} \\ \mathbf{Z}'\mathbf{M}\mathbf{X} & \mathbf{Z}'\mathbf{M}\mathbf{Z} + \mathbf{A}^{11}\mathbf{t} & \mathbf{A}^{12}\mathbf{t} \\ \mathbf{0} & \mathbf{A}^{21}\mathbf{t} & \mathbf{D} + \mathbf{A}^{22}\mathbf{t} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{a}}_1 \\ \hat{\mathbf{a}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{M}\mathbf{y} \\ \mathbf{Z}'\mathbf{M}\mathbf{y} \\ \mathbf{Q}'\hat{\mathbf{v}} \end{bmatrix} \quad [3]$$

where  $\mathbf{M} = \mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{Z} + \mathbf{I}\mathbf{k})^{-1}\mathbf{Z}'$ . Solutions to these equations are best linear unbiased estimates (BLUE) of  $\beta$  and BLUP of  $\mathbf{a}_1$  and  $\mathbf{a}_2$  (1). Equations in [3] can be set up directly from a list of records sorted by freshening date within cow within herd.

Our interest centered on the inverse of the coefficient matrix of the mixed model equations in [3] as variances of prediction errors of linear functions of random effects are functions of elements of this inverse. For this model the coefficient matrix has full rank.

Let

$$\begin{bmatrix} \mathbf{X}'\mathbf{M}\mathbf{X} & \mathbf{X}'\mathbf{M}\mathbf{Z} & \mathbf{0} \\ \mathbf{Z}'\mathbf{M}\mathbf{X} & \mathbf{Z}'\mathbf{M}\mathbf{Z} + \mathbf{A}^{11}\mathbf{t} & \mathbf{A}^{12}\mathbf{t} \\ \mathbf{0} & \mathbf{A}^{21}\mathbf{t} & \mathbf{D} + \mathbf{A}^{22}\mathbf{t} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \mathbf{C}_{23} \\ \mathbf{C}_{31} & \mathbf{C}_{32} & \mathbf{C}_{33} \end{bmatrix} \quad [4]$$

then the variance of prediction errors of a linear function of cows' genetic values,  $\mathbf{m}'\mathbf{a}_1$ , is equal to  $\mathbf{m}'\mathbf{C}_{22}\mathbf{m}\sigma_e^2$  (3, 4), where  $\sigma_e^2 = (1 - r)\sigma^2$ . To examine the variance of prediction errors of the difference of genetic value between any two cows, let  $\mathbf{m}'$  be a vector with all elements null except  $\mathbf{m}_i = 1$  and  $\mathbf{m}_j = -1$ . Then the variance of prediction errors of  $\mathbf{m}'\mathbf{a}_1 = (\mathbf{a}_i - \mathbf{a}_j)$ , the difference of additive genetic values of the  $i$ th and  $j$ th cows, is

$$\begin{aligned} & \mathbf{m}'\mathbf{C}_{22}\mathbf{m}\sigma_e^2 \\ & = (\mathbf{c}_{22,ii} + \mathbf{c}_{22,jj} - 2\mathbf{c}_{22,ij})\sigma_e^2 \end{aligned} \quad [5]$$

where  $\mathbf{c}_{22,ii}$  represents the  $i$ th diagonal element of  $\mathbf{C}_{22}$ ;  $\mathbf{c}_{22,jj}$  represents the  $j$ th diagonal element of  $\mathbf{C}_{22}$ ; and  $\mathbf{c}_{22,ij}$  represents the  $ij$ th off-diagonal element. Diagonal elements of  $\mathbf{C}_{22}\sigma_e^2$  represent variances of prediction errors of individual elements of  $\mathbf{a}_1$ , and off-diagonal elements represent covariances (3, 4). All expressions for variances of prediction errors involve the scalar  $\sigma_e^2$ , the residual variance. Therefore, in comparing variances of different functions, only the magnitude of functions of elements of the inverse need be compared.

### COMPUTATIONAL PROCEDURES

Twenty pairs of herds were selected randomly from each of the Holstein, Jersey, and Brown Swiss breeds. Herd size was limited to 100 cows because of limited computer storage. For each pair of herds, a set of mixed model equations of the form in [3] was constructed (5). A list of AI sires used in one or both herds was compiled simultaneously. Estimating transmitting abilities and numbers of daughter records for all proven AI sires used in either herd were obtained from an AI sire file. This information was used to calculate diagonal elements and right-hand sides of the sire equations according to Henderson's procedure (5):

$$\text{sire's diagonal element} = n(1 - r)/(4 - h^2)$$

$$\begin{aligned} & \text{sire's right-hand side} \quad [6] \\ & = 2(1 - r)[4 + (n - 1)h^2]v/[h^2(4 - h^2)] \end{aligned}$$

where  $n$  is the number of daughter records and  $v$  is the sire's estimated transmitting ability. If the sire was not a proven AI sire, these elements remained zero (1, 5).

The inverse of the numerator relationship matrix ( $\mathbf{A}^{-1}$ ) was computed by Henderson's rapid method (5). Elements of  $\mathbf{A}^{-1}$  multiplied by  $\mathbf{t}$  were added directly to the corresponding elements of the coefficient matrix of the MME already stored in core. The addition of  $\mathbf{A}^{-1}\mathbf{t}$  to the left-hand sides completed construction of the MME. The inverse of the coefficient matrix was computed and stored in core. Variances of prediction errors of differences in genetic values between pairs of cows within a herd or from the two different herds were calculated as linear functions of elements of this inverse.

## RESULTS AND DISCUSSION

The primary objective was to examine the accuracy of the NEA<sup>-1</sup> ETA cow evaluations for comparing genetic values of cows from different herds and within a single herd. The standard of accuracy was the variance of prediction errors of the difference of genetic value between pairs of cows, from hereon referred to as the variance of prediction errors. Pairs of cows compared were classified as within herd, i.e., both cows from the same herd, or across herds, i.e., each cow from a different herd. An average of variance of prediction errors was calculated for each classification in each breed and the difference between classification averages within each breed was tested for significance.

Each of the breeds, Holstein, Brown Swiss, and Jersey, were evaluated separately, but the same procedures were applied to all. Within each breed, 20 pairs of herds were selected randomly, and from within each pair of herds 20 pairs of cows were chosen randomly. Each pair of cows was classified within herd or across herds, and the variance of prediction errors was computed for each pair. If the selection process resulted in either classification having less than three pairs of cows, a new random sample of animals was drawn from that pair of herds. Once a suitable sample had been drawn, the average variance of prediction errors within each classification was calculated.

Let  $X_{ij}$  represent the average variance of prediction errors for the  $i$ th classification and  $j$ th pair of herds. We assumed that

$$X_{ij} \sim N(\mu_i, \sigma_{ij}^2) \quad i = 1, 2; j = 1, 2, \dots, 20$$

with

$$X_{ij} = \mu_i + e_{ij}$$

and

$$E(X_{ij}) = \mu_i, E(e_{ij} e_{i'j'}) = \sigma_{ij}^2 \\ \text{if } i = i' \text{ and } j = j' \text{ and } 0 \text{ otherwise}$$

where  $\sigma_{ij}^2 = \sigma^2/n_{ij}$ . Tables 1 to 3 present the  $X_{ij}$ 's for each breed and the corresponding  $n_{ij}$ 's, i.e., the number of observations in each average. For each breed, a weighted average of the  $X_{ij}$ 's was calculated within each classification. The

average variance of prediction errors for each herd pair ( $X_{ij}$ ) was weighted by the number of observations it contained so that within each breed  $\bar{X}w_i$  was calculated, where

$$\bar{X}w_i = \sum_{j=1}^{20} n_{ij} X_{ij} / n_i, \quad i = 1, 2$$

with  $E(\bar{X}w_i) = \mu_i$  and  $\text{var}(\bar{X}w_i) = \sigma^2/n_i$ . Under the null hypothesis,  $H_0: \mu_1 = \mu_2$ ; with the alternative hypothesis,  $H_a: \mu_1 \neq \mu_2$ ; differences between classifications within a breed were tested with the statistic

$$t = \bar{X}w_1 - \bar{X}w_2 / \sqrt{\hat{\sigma}^2 [(1/n_1) + (1/n_2)]}$$

This requires an estimate of the residual variance,  $\hat{\sigma}^2$ , which was derived:

$$\begin{aligned} \text{Var}(X_{ij} - \bar{X}w_i) &= \text{Var}(X_{ij}) + \text{Var}(\bar{X}w_i) - 2\text{Cov}(X_{ij}, \bar{X}w_i) \\ &= \sigma^2/n_{ij} + \sigma^2/n_i - 2E[(X_{ij} - \mu_i)(\bar{X}w_i - \mu_i)] \\ &= \sigma^2/n_{ij} + \sigma^2/n_i - 2E(X_{ij}\bar{X}w_i) + 2\mu_i^2 \\ &= \sigma^2/n_{ij} + \sigma^2/n_i - 2(\mu_i^2 + \sigma^2/n_i) + 2\mu_i^2 \\ &= \sigma^2 [(1/n_{ij}) - (1/n_i)] \\ &= k\sigma^2. \end{aligned}$$

Letting  $Z_{ij} = (1/\sqrt{k})(X_{ij} - \bar{X}w_i)$ , then  $\text{Var}(Z_{ij}) = (1/k)\text{Var}(X_{ij} - \bar{X}w_i) = \sigma^2$ . Therefore, a pooled estimate of the residual variance could be calculated as

$$S_p^2 = \sum_{i=1}^2 \sum_{j=1}^n Z_{ij}^2 / 2(n-1)$$

Tables 1 to 3 include the weighted averages for each classification at the bottom. Table 4 gives the weighted averages for each classification in each breed, the pooled estimates of the residual variance associated with these weighted averages for each breed and the resulting test statistic. All tests were conducted with a Type I error of .05; the critical value of  $t$  was 2.025.

The test statistic failed to exceed the critical value in all three cases so the null hypothesis  $\mu_1 = \mu_2$  was not rejected for any of the breeds. The weighted average for across classifications was greater than that for within classifications in all three breeds, but none of the differences was significant. This indicated that the average

TABLE 1. Average variance<sup>1</sup> of prediction errors of difference of additive genetic value between pairs of Holstein cows within and across herds, including the number of observations in each average.

Within herds	Number of observations included	Across herds	Number of observations included
.479586	8	.513556	12
.474871	11	.458537	9
.451340	10	.471693	10
.403496	13	.483792	7
.460812	8	.455959	12
.403324	9	.430134	11
.557065	8	.543029	12
.491841	10	.520080	10
.487824	11	.509348	9
.510149	7	.491450	13
.487776	12	.536773	8
.509681	8	.557852	12
.522173	12	.545185	8
.487766	12	.488777	8
.465036	9	.451789	11
.477285	12	.475003	8
.511110	10	.538471	10
.441121	8	.430362	12
.569931	10	.525123	10
.571804	11	.593673	9
.487636 = Average		.499780 = Average	

<sup>1</sup> Decimal fractions represent the average variance of prediction errors divided by  $\sigma_e^2$ .

TABLE 2. Average variance<sup>1</sup> of prediction errors of difference of additive genetic value between pairs of Jersey cows within and across herds, including the number of observations in each average.

Within herds	Number of observations included	Across herds	Number of observations included
.456449	12	.483280	8
.429218	10	.462190	10
.420467	14	.482780	6
.498647	13	.507936	7
.571269	10	.555643	10
.522779	9	.559611	11
.416580	17	.547079	3
.537394	11	.677559	9
.387160	9	.456290	11
.461760	13	.449086	7
.543002	12	.560076	8
.479774	12	.491524	8
.458409	7	.455931	13
.408051	11	.440543	9
.459260	16	.471815	4
.456840	11	.499314	9
.604314	9	.569966	11
.473843	9	.441315	11
.461730	17	.476168	3
.549142	11	.564661	9
.473001 = Average		.508234 = Average	

<sup>1</sup> Decimal fractions represent the average variance of prediction errors divided by  $\sigma_e^2$ .

variance of prediction errors was similar for comparing genetic values of cows from within a herd or from different herds. In other words, the results suggest that  $NEA^{-1}$  ETA evaluations may be used with nearly equal confidence for comparing cows from the same herd or differ-

ent herds in the Holstein, Jersey, and Brown Swiss breeds.

### CONCLUSIONS

The current  $NEA^{-1}$  ETA cow evaluations provide predictions of genetic values and pro-

TABLE 3. Average variance<sup>1</sup> of prediction errors of difference of additive genetic value between pairs of Brown Swiss cows within and across herds, including the number of observations in each average.

Within herds	Number of observations included	Across herds	Number of observations included
.435719	11	.424221	9
.440884	9	.462373	11
.430259	17	.487400	3
.424408	12	.543792	8
.527112	13	.510017	7
.513490	14	.550252	6
.471496	12	.453102	8
.488955	13	.492666	7
.478349	14	.537968	6
.481468	11	.520436	9
.490800	11	.515501	9
.469159	10	.522557	10
.470307	10	.543820	10
.414187	9	.432918	11
.527009	14	.546030	6
.416491	14	.427707	6
.455355	14	.544287	6
.525221	9	.571502	11
.496419	11	.502180	9
.477052	10	.471183	10
.471759 = Average		.501611 = Average	

<sup>1</sup> Decimal fractions represent the average variance of prediction errors divided by  $\sigma_e^2$ .

TABLE 4. Summary and test statistics for each breed.

Breed	Breed average variances <sup>1</sup> of prediction errors		Pooled estimate of residual variance	t Statistic ( $\alpha = .05$ , $df = 38$ )
	Within herd	Across herd		
Holstein	.487636	.499780	.024393	-.778
Jersey	.473001	.508234	.042362	-1.688
Brown Swiss	.471759	.501611	.023714	-1.903

<sup>1</sup> Decimal fractions represent the average variance of prediction errors divided by  $\sigma_e^2$ .

ducing abilities for all cows in a herd based on all records. Inclusion of AI sire evaluations in the procedure tie the cow evaluations to the mean of the same base population from which the sire evaluations are deviated. Results of this study indicate that comparisons of cows' genetic values based on their NEA<sup>-1</sup> ETA evaluations are nearly as accurate for cows from different herds as for cows from the same herd. Therefore, an AI stud that wants to choose the best cow from a group of cows from different herds can do so with nearly the same confidence as a dairy producer who must choose a cow from the same size group in his own herd.

Although the ideal way to evaluate dairy cows would be to evaluate all animals in a particular breed and regional subpopulation simultaneously, this is not yet computationally feasible. Such a method would allow direct comparison of all animals. Currently the NEA<sup>-1</sup> ETA evaluations provide a working alternative that may be used to make both interherd and intraherd comparisons with nearly the same confidence.

#### ACKNOWLEDGMENTS

This research was supported in part by a grant from Eastern Artificial Insemination Cooperative, Inc., Ithaca, NY.

#### REFERENCES

- 1 Bolgiano, D. C. 1981. An examination of the accuracy of cow evaluations for interherd versus intraherd comparisons. M.S. Thesis, Cornell Univ., Ithaca, NY.
- 2 Everett, R. W., C. R. Henderson, and C. F. Hanson. 1977. The Northeast cow ETA report. Anim. Sci. Mimeo Ser. #19. Dep. of Anim. Sci., Cornell Univ., Ithaca, NY.
- 3 Henderson, C. R. 1973. Sire evaluation and genetic trends. *In* Proc. Anim. Breeding and Genet. Symp. in Honor of Dr. Jay L. Lush. Am. Soc. Anim. Sci. and Am. Dairy Sci. Assoc., Champaign, IL.
- 4 Henderson, C. R. 1974. General flexibility of linear model techniques for sire evaluation. *J. Dairy Sci.* 57:963.
- 5 Henderson, C. R. 1975. Use of all relatives in intraherd prediction of breeding values and producing abilities. *J. Dairy Sci.* 58:1910.
- 6 Van Vleck, L. D. 1982. Derivation of Henderson's method of incorporating artificial insemination sire evaluations into intraherd prediction of breeding values. *J. Dairy Sci.* 58:1910.