

Environmental Correlations in Pedigree Estimates of Breeding Value¹

G. M. THOMSON and A. E. FREEMAN

Department of Animal Science, Iowa State University, Ames 50010

Abstract

Effects of environmental correlations among the records of ancestors of bulls in artificial breeding were evaluated with data from the Holstein-Friesian Association of America. Formulas developed for the variances and covariances of means of records and progeny group means included the effects of distribution of records in both herds and year-seasons. Selection index procedures were used to estimate the breeding value of bulls from ancestor information.

Estimates of heritability were 0.338 and 0.358 for first-lactation mature-equivalent milk and records deviated from their adjusted herd-mate average. Repeatabilities were 0.622 for mature-equivalent milk and 0.501 for deviation milk. The environmental correlations among records of cows calving in the same herd and year-season, in the same herd but different year-seasons, and in the same year-season but different herds were 0.328, 0.226, and 0.028 for mature-equivalent records. Similar estimates for deviation records were 0.0, 0.0, and 0.003. Additional environmental correlations for half-sisters in the same herd were 0.005 and 0.102 for mature-equivalent and deviation records.

Correlations between computed indexes and the "predicted difference" of the bulls in artificial breeding were near zero for mature-equivalent records and near 0.25 for indexes of deviation records.

Introduction

The theoretical effects of environmental correlations have been demonstrated by Lush (5, 6). He shows that the limiting value of the

correlation between the breeding value of a bull and the average of single records of his daughters depends on heritability and the environmental correlation among the progeny. When there is no environmental correlation and heritability is not zero, the correlation has a limit of one, regardless of the magnitude of heritability. However, when an environmental correlation is present, the limit to the correlation is always less than one and may be considerably below one, the exact magnitude depending upon the magnitude of the parameters.

Plowman and McDaniel (7), in a discussion of methods to estimate the breeding value of bulls by the United States Department of Agriculture, show how a residual correlation (c^2) among half-sisters in the same herd affects the regression of future progeny on initial progeny. They illustrate the effect of number of daughters, number of herds, and the distribution of daughters among herds with 0.19 for heritability of single records and 0.14 for the residual correlation among half-sisters. These values are currently used by the U.S. Department of Agriculture.

The purpose of this study was to evaluate the effects of environmental correlations among the records of ancestors of bulls in artificial breeding.

Description of Data

From the Holstein-Friesian Association of America there were 158,336 records by 62,389 cows in 450 herds from 38 states. Starting dates of records were from March, 1952, through April, 1961. All records were standardized to 305-day, two-times-a-day milking, mature-equivalent by conversion factors of Kendrick (4) and the Holstein-Friesian Association (3).

Records were also deviated from their adjusted herd-mate average. The herd-mate average for any cow was the average of lactations started in the same herd and year-season by all paternally unrelated cows. Records of all cows, except the cow in question and her paternal half-sisters, calving in the same herd and within five months centered on the month the cow

Received for publication December 30, 1969.

¹ Journal Paper no. J-6433 of the Iowa Agriculture and Home Economics Experiment Station, Ames. Project 1053. This research was part of the North Central Regional Project NC-2.

calved, are in the herdmate average (a rolling five-month season.)

Experimental Procedure

Formulas developed for the variances and covariances of means of records and progeny group means included the effects of distribution of records in both herds and year-seasons. A lactation record was represented by the following model:

$$Y_{ijklm} = \mu + a_i + h_j + s_k + d_{jkl} + e_{ijklm}$$

where Y_{ijklm} is the m^{th} record of the l^{th} cow, by the k^{th} sire, made in the j^{th} herd and started in the i^{th} year-season; μ is the population mean; a_i is an effect common to all records started in the i^{th} year-season; h_j is an effect common to all records made in the j^{th} herd; s_k is an effect common to all records made by daughters of the k^{th} sire; d_{jkl} is an effect common to all records of the l^{th} cow, by the k^{th} sire, made in the j^{th} herd ($l = 1, 2 \dots n_{ijk}$); and e_{ijklm} is an effect peculiar to the m^{th} record of the l^{th} cow, by the k^{th} sire, made in the j^{th} herd, and started in the i^{th} year-season ($m = 1, 2 \dots n_{jkl}$).

The year-season, herd, sire, and cow effects are assumed to be randomly distributed with means equal to zero and variances σ_a^2 , σ_h^2 , σ_s^2 , and σ_d^2 . These effects are assumed independent. $\sigma_{e_{ijklm}, e_{ij'k'l'm}}$ is assumed equal to zero for all k not equal to k' and (or) j not equal to j' . For $k = k'$ and $j = j'$, $\sigma_{e_{ijklm}, e_{ij'k'l'm}}$ is assumed equal to $\sigma_{ee'}$, the environmental covariance between half-sisters calving in the same herd.

Where a cow makes all her records in one herd, the mean of the cow's records would be

$$\bar{Y}_{.jkl} = \mu + \frac{\sum a_i}{n_{jkl}} + h_j + s_k + d_{jkl} + \frac{\sum e_{ijklm}}{n_{jkl}},$$

with expected value

$$E[\bar{Y}_{.jkl}] = \mu + h_j + s_k + d_{jkl}$$

and variance

$$V[\bar{Y}_{.jkl}] = \frac{1}{n_{jkl}} \sigma_a^2 + \sigma_h^2 + \sigma_s^2 + \sigma_d^2 + \frac{1}{n_{jkl}} \sigma_e^2,$$

which is equal to

$$V[Y_{ijklm}] \left[\frac{1 + (n_{jkl} - 1)r}{n_{jkl}} \right],$$

where r (the repeatability of single records) is

equal to

$$\frac{\sigma_h^2 + \sigma_s^2 + \sigma_d^2}{\sigma_a^2 + \sigma_h^2 + \sigma_s^2 + \sigma_d^2 + \sigma_e^2}.$$

Only first lactations of daughters of sires were included. The average of first lactations of all daughters of a particular sire, say sire k' , is

$$\begin{aligned} \bar{Y}_{.k'.} &= \frac{1}{n_{.k'}} \sum_{ijl} Y_{ijk'l} \\ &= \mu + \frac{\sum a_i}{n_{.k'}} + \frac{\sum h_j}{n_{.k'}} + s_{k'} + \\ &\quad \frac{\sum d_{jk'l}}{n_{.k'}} + \frac{\sum e_{ijk'l'm}}{n_{.k'}}, \end{aligned}$$

which has the expected value

$$E[\bar{Y}_{.k'.}] = \mu + s_{k'}.$$

These records were assumed in the same herd for the variance of the mean of records of a cow. Also, a cow cannot have more than one record in any year-season. The distribution of records of half-sisters will not be so restricted for herds and year-seasons. Records of any two half-sisters will fall into one of the following classifications: 1) same herd and same year-season, 2) same herd but different year-seasons, 3) different herds but same year-season, or 4) different herds and different year-seasons. In addition to the sire component, σ_s^2 , half-sisters in the same year-season will have a year-season component, σ_a^2 , in common. For half-sisters in the same herd, a herd component, σ_h^2 , will be included in the covariance term. In addition, a $\sigma_{ee'}$ term will be included for half-sisters in the same herd. These situations are summarized in Table 1.

The component $\sigma_{ee'}$ would, by definition, contribute only to the covariance among half-sisters in the same herd. It would not appear in the covariance between half-sisters in different herds, whether or not they calve in the same year-season.

The variance of a paternal half-sister mean is

TABLE 1. Components of covariance for paternal half-sibs in the same and different herds and year-seasons.

Herds	Year-seasons	Components of covariance
Same	Same	$\sigma_s^2 + \sigma_a^2 + \sigma_h^2 + \sigma_{ee'}$
Same	Different	$\sigma_s^2 + \sigma_h^2 + \sigma_{ee'}$
Different	Same	$\sigma_s^2 + \sigma_a^2$
Different	Different	σ_s^2

$$V(Y_{ijklm}) \left[\frac{4 + (n_{..k'} - 1)h^2 + 4 \sum_{ij} \frac{n_{ijk'}(n_{ijk'} - 1)}{n_{..k'}} (s + p^2)}{4n_{..k'}} \right. \\ \left. + \frac{4 \sum_j (\sum_{i \neq i'} \frac{(n_{ijk'}) (n_{ij'k'})}{n_{..k'}}) (t + p^2)}{4n_{..k'}} \right. \\ \left. + \frac{4 \sum_i (\sum_{j \neq j'} \frac{(n_{ijk'}) (n_{ij'k'})}{n_{..k'}}) (w)}{4n_{..k'}} \right]$$

where

$$h^2 = \frac{4\sigma_s^2}{(\sigma_a^2 + \sigma_h^2 + \sigma_s^2 + \sigma_d^2 + \sigma_e^2)}, \\ s = \frac{(\sigma_a^2 + \sigma_h^2)}{(\sigma_a^2 + \sigma_h^2 + \sigma_s^2 + \sigma_d^2 + \sigma_e^2)}, \\ t = \frac{\sigma_h^2}{(\sigma_a^2 + \sigma_h^2 + \sigma_s^2 + \sigma_d^2 + \sigma_e^2)}, \\ w = \frac{\sigma_a^2}{(\sigma_a^2 + \sigma_h^2 + \sigma_s^2 + \sigma_d^2 + \sigma_e^2)}, \text{ and} \\ p^2 = \frac{\sigma_{ee'}}{(\sigma_a^2 + \sigma_h^2 + \sigma_s^2 + \sigma_d^2 + \sigma_e^2)}.$$

If all half-sisters are in different herds and different year-seasons, this reduces to

$$V[\bar{Y}_{..k'..}] = V[Y_{ijklm}] \left[\frac{4 + (n_{..k'} - 1)h^2}{4n_{..k'}} \right].$$

The variance will be the same for maternal half-sisters if the daughters are sired by different bulls and maternal effects are nonexistent.

Covariances between means of records of a cow, between progeny group means, and between a mean of records and a progeny group mean equal

$$V[Y_{ijklm}] (a_{12}h^2 + \frac{n_{(i=i', j=j')}}{n_1 n_2} s + \frac{n_{(i \neq i', j=j')}}{n_1 n_2} t + \frac{n_{(i=i', j \neq j')}}{n_1 n_2} w),$$

where

- a_{12} = the numerator of Wright's (8) coefficient of relationship,
 - n_1 = the number of records (or progeny) in Group 1,
 - n_2 = the number of records (or progeny) in Group 2,
 - $n_{(i=i', j=j')}$ = the number of cases that the records in Group 1 and Group 2 occur in the same herd and year-season,
 - $n_{(i \neq i', j=j')}$ = the number of cases that the records occur in the same herd but different year-seasons,
 - $n_{(i=i', j \neq j')}$ = the number of cases that the records occur in the same year-season, but in different herds and
- h^2 , s , t , and w are as defined previously.

If all records were started in different herds and different year-seasons, the covariances would reduce to:

$$V(Y_{ijklm}) (a_{12}h^2).$$

Heritability was estimated from between- and within-sire analyses of variance of both mature-equivalent and deviated data. Sire groups were excluded when they had daughters in less than five herds. Only one daughter per herd and only first lactations were included. A record was a first lactation if the cow was from 20 to 35 months old at first calving. Cows calving at less than 20 months and not calving in the range of 20 to 35 months were excluded from the analyses. Repeatability estimates were obtained from between- and within-cow analyses of variance similar to that for heritability.

The environmental correlations, s , t , and w , were estimated by the intraclass correlations from analyses of variance between- and within-herd-year-season, herd, and year-season. All available records were included in the analyses for the correlation (s) among cows in the same herd and same year-season. In the analyses for the correlation (t) among cows in the same herd but different year-seasons, and the correlation (w) among cows in the same year-season but different herds, one cow's record was selected at random from each herd-year-season.

In data selected in this manner (i.e., one record from each herd-year-season), the herd component of variance in a between- and within-herd analysis is expected to contain only the differences between herds. Similarly, the year-season component in a between- and within-year season analysis is expected to contain only the differences between year-seasons.

Table 2 is a summary of genetic and environmental correlations. The correlation among mature-equivalent records of cows freshening in a year-season was 0.328, close to the correlation of 0.332 reported by Bereskin (2). He used 33,139 Holstein records from the Iowa Central Processing Center made from May, 1957, through September, 1960. Bereskin used fixed

TABLE 2. Estimates of genetic and environmental correlations and their standard errors.

Correlation	Number of classes	Number of cows	Mature-equivalent records	Deviation records
h^2	215	1,911	0.338 ± 0.108	0.358 ± 0.108
r	39,873	65,951	0.622 ± 0.004	0.501 ± 0.005
s	35,137	158,336	0.328 ± 0.003	0.0 ^a
t	450	35,137	0.226 ± 0.012	0.0 ^a
w	96	35,137	0.028 ± 0.004	0.003 ± 0.001
$\frac{1}{4} h^2 + t + p^2$	386	2,544	0.335 ± 0.023	0.176 ± 0.019
$\frac{1}{4} h^2$	386	2,050	0.104 ± 0.020	0.074 ± 0.018
p^2			0.005	0.102

^a Negative component—assumed to be zero.

seasons in his analysis. In this study five-month rolling seasons were used, and a season defined as one month.

The results of the analyses of deviation records indicate that deviating records was effective in removing herd and year-season effects. Only the correlation among cows in the same year-season but different herds was above zero, and this only slightly so.

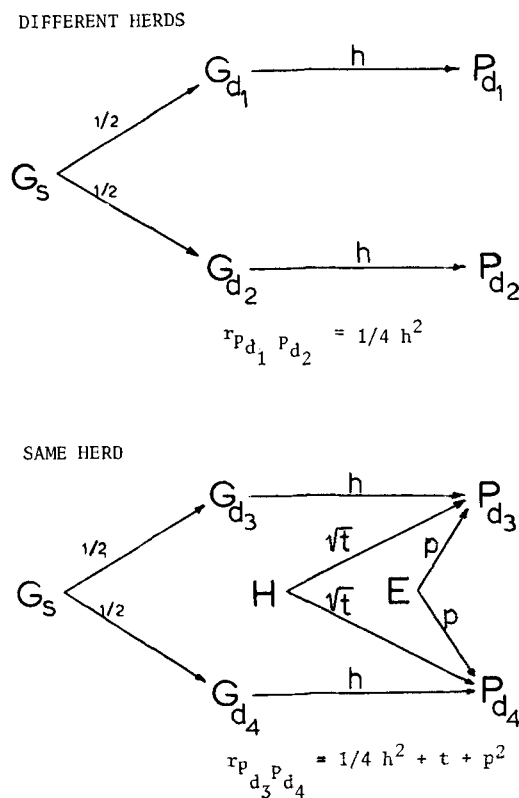


FIG. 1. Path diagram illustrating the correlation between half-sibs in the same and different herds.

It is possible that paternal half-sisters in the same herd are treated differently from other cows in the herd (i.e., if they were all housed together in the barn or given extra care). In addition to their being half-sisters and being in the same herd, the correlation among these half-sisters would be higher than expected because of correlated environment. Figure 1 shows this possibility in a path diagram. The correlation between P_{d1} and P_{d2} would be one-fourth heritability whereas the correlation between P_{d3} and P_{d4} would also include t , the correlation among cows calving in the same herd, and the additional correlation, p^2 , due to common environmental conditions.

The sire component in a between- and within-sire analysis would contain the additional herd and common half-sister environmental covariance if all the daughters of a sire were in the same herd. To evaluate this additional correlation, heritability was estimated under two conditions. First, only one daughter was included from any one herd and, secondly, only daughters in a single herd were included in analyses of variance. The difference between the intra-class correlations should measure $t + p^2$. Reducing this by t from a previous analysis where half-sisters were not considered, estimated the correlation from common environment over and above the common herd effect.

Estimates of Breeding Values of Sires

Sires in artificial breeding whose dam had at least one record in the data were chosen. The progeny performance of these bulls was from the U.S. Department of Agriculture, Dairy Herd Improvement Association Sire Summary Lists. The value used was the Predicted Difference, which estimates half of the sire's breeding value. Sires with five or more daughters were included. Pedigrees of each bull were obtained from the Holstein-Friesian Herd Books.

Records of ancestors of each sire were then assembled from the available data. Information from the bull's dam, his paternal and maternal half-sisters, his grandams, and the paternal and maternal half-sisters of his sire and dam was used. A record was accepted if it started no later than five months after the bull's birth. Thus, information available when the bull was approximately 15 months old was used. A total of 176 sires with varying amounts of pedigree information was available.

Selection index procedures were used to make pedigree estimates of breeding value for both mature-equivalent and deviation records. The index equations were set up with variances and covariances of the forms shown above. Estimates were made by substituting the correlations in Table 2 into the index equations.

Estimates were made 1) with environmental correlations, and 2) with these correlations zero. The pedigree information was used three different ways, using 1) all information available, 2) only the paternal side of the pedigree, and 3) only the maternal side of the pedigree. The dam's records were excluded from both of her parents' progeny groups. Full sisters were each treated as two half-sisters, one paternal and one maternal. All production information on dams and grandams was used, but progeny averages included only the first available record of each daughter.

Averages of indexes and correlations between breeding value and index are presented in Table 3.

Environmental correlations in the selection index lowered the correlation between the index and the breeding value of the bull for both mature-equivalent milk and deviation milk. The effect is not as severe for deviation milk as it is for mature-equivalent milk. This is because

deviating records was effective in removing herd-year-season effects. The environmental correlations applied to deviation records were, therefore, small or zero. The correlations were lower for mature-equivalent records than for deviation records, due to the difference in the estimates of heritability and repeatability.

Product-moment correlations between the different indexes based on the same information are presented in Table 4. Indexes from deviation records with and without environmental correlations were more highly correlated than the indexes from mature-equivalent records. This was expected, since environmental correlations were low in deviated records. Correlations between indexes from mature-equivalent records and indexes from deviation records were low, whether or not environmental correlations were taken into account.

Barr (1), with Canadian Record of Performance data, used pedigree information to estimate breeding values of bulls. His data contained more complete pedigree information than in this study. He obtained an average r_{GI} of about 0.55, which is equivalent to information about eight daughters of the bull. The average value of r_{GI} obtained in this study for complete pedigree information would be equivalent to information about two or three daughters of the bull.

Correlations between the estimates of breeding value in this study and estimates of breeding value from progeny performance (Predicted Difference) are presented in Table 5. Two correlations are presented. The product-moment correlation measures the degree or intensity of association of the two indexes. The rank correlation measures the similarity in the way the two indexes rank the bulls.

The most obvious conclusion from Table 5

TABLE 3. Average breeding value and r_{GI} estimates.

Index	Number of bulls	$r_e = 0^a$		$r_e > 0^b$	
		Index (kg)	r_{GI}	Index (kg)	r_{GI}
Mature-equivalent records					
All	176	5,201	0.414	2,794	0.343
Paternal	170	3,266	0.309	2,034	0.246
Maternal	150	2,401	0.295	1,367	0.274
Deviation records					
All	176	312	0.425	274	0.399
Paternal	170	145	0.313	113	0.283
Maternal	150	202	0.308	194	0.302

^a Environmental correlations assumed to be zero.

^b Environmental correlations included in the index.

TABLE 4. Correlations between estimates of breeding value.

Index		Mature equivalent		Deviation records	
		$r_e = 0^a$	$r_e > 0^b$	$r_e = 0$	$r_e > 0$
All information					
Mature equivalent	$r_e = 0$	1.00			
	$r_e > 0$	.70	1.00		
Deviation records	$r_e = 0$	.42	.32	1.00	
	$r_e > 0$	.41	.34	.98	1.00
Paternal information					
Mature equivalent	$r_e = 0$	1.00			
	$r_e > 0$	.81	1.00		
Deviation records	$r_e = 0$	.43	.23	1.00	
	$r_e > 0$	.39	.27	.97	1.00
Maternal information					
Mature equivalent	$r_e = 0$	1.00			
	$r_e > 0$	.54	1.00		
Deviation records	$r_e = 0$	.40	.45	1.00	
	$r_e > 0$	.39	.47	.99	1.00

^a Environmental correlations assumed to be zero.^b Environmental correlations included in the index.

is that deviation records are useful, but there is little or no resemblance between pedigree evaluation of mature-equivalent records and progeny performance. Almost all the correlations are near zero. The indexes on the maternal side of the pedigree alone are somewhat higher. For only the paternal half of the pedigree the correlations are slightly negative.

This indicates a negative correlation between the paternal and maternal sides of the pedigree,

particularly on mature-equivalent records. The negative correlations would most probably exist between the dam and sire. This may occur if negative assortive mating was common in these herds. This breeding system would seem unlikely for milk production, per se, but may be practiced on traits correlated with milk production. Thus, some increase in milk yield would be sacrificed to raise a low test or correct a structural defect.

TABLE 5. Correlations between estimated breeding value and the Predicted Difference of bulls in artificial breeding.

Index	$r_e = 0^a$		$r_e > 0^b$	
	r^c	r_s^d	r	r_s
Mature-equivalent records				
All information	0.06	0.06	-0.07	-0.07
Paternal	-0.03	-0.03	-0.09	-0.09
Maternal	0.22	0.22	0.11	0.12
Deviation records				
All information	0.24	0.15	0.24	0.15
Paternal	0.07	0.04	0.05	-0.01
Maternal	0.28	0.22	0.28	0.21

^a Environmental correlations assumed to be zero.^b Environmental correlations included in the index.^c Product-moment correlation coefficient.^d Spearman's rank correlation coefficient.

A negative environmental correlation between the paternal and maternal ancestors may give the same result. This could occur if the dam was given special treatment and the paternal ancestors were either in the same herd and not given special environment or were in different herds.

References

- (1) Barr, G. R. 1962. Selecting young bulls on differences between relatives and their contemporaries. Ph.D. thesis, Iowa State University, Ames.
- (2) Bereskin, B. 1963. Effects of genetic and environmental variance on dairy sire evaluation. Ph.D. thesis, Iowa State University, Ames.
- (3) Holstein-Friesian Association of America. *ca.* 1955. Factors for extending incomplete records. (Mimeo.) Brattleboro, Vermont.
- (4) Kendrick, J. F. 1955. Standardizing DHIA records in proving sires. USDA, ARS, Publ. ARS-52-1.
- (5) Lush, J. L. 1931. The number of daughters necessary to prove a sire. *J. Dairy Sci.*, 14: 209.
- (6) Lush, J. L. 1935. Progeny test and individual performance as indicators of an animal's breeding value. *J. Dairy Sci.*, 18: 1.
- (7) Plowman, R. D., and B. T. McDaniel. 1968. Changes in USDA sire summary procedures. *J. Dairy Sci.*, 51: 306.
- (8) Wright, S. 1922. Coefficients of inbreeding and relationship. *American Naturalist*, 56: 330.