University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Papers and Publications in Animal Science

Animal Science Department

January 1982

Monte Carlo Study of Genetic Groups in Sire Evaluation

T. R. Famula 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

Famula, T. R. and Van Vleck, L. Dale, "Monte Carlo Study of Genetic Groups in Sire Evaluation" (1982). *Faculty Papers and Publications in Animal Science*. 388. https://digitalcommons.unl.edu/animalscifacpub/388

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.

T. R. FAMULA¹ and L. D. VAN VLECK Department of Animal Science Cornell University Ithaca, NY 14853

ABSTRACT

Fixed genetic groups have been defined arbitrarily in linear models for sire evaluation. This simulation examined genetic groups under several selection strategies. Sires with highest estimated transmitting abilities were selected to have sons and additional female progeny. The process of evaluation and selection of sires and dams was continued for four generations. Final estimates of transmitting ability were compared to actual transmitting ability for several sire evaluation models on 300 simulated data sets. Genetic groups defined by generation increased the accuracy of estimating sire transmitting abilities over ignoring genetic groups when selected dams and sires of young bulls were mated. A joint selection model was presented which assists in interpreting results of simulation.

INTRODUCTION

Use and definition of genetic groups for sire evaluation has received attention in (6, 7, 8, 10). In a review of sire evaluation and genetic groups (2), effects of selection were examined in reference to the problem of defining genetic groups. One goal of this simulation study was to show that one answer to the question of defining genetic groups is not possible.

Genetic groups are used to represent different populations of sires where expected breeding values of sires may be different from population to population. However, except for a few examples where genetic groups are distinct (13, 16), definition of genetic groups often is based on arbitrary distinctions. Because group definitions are formulated with the visualization of a repeated sampling strategy, explicit group definitions cannot be derived. Thus, analytic evaluation of grouping definitions is not possible. An alternative is to simulate realistic sampling strategies and examine the accuracy of sire evaluation under several group definitions.

MATERIALS AND METHODS

The basic strategy of the simulation was to select sires with the highest estimated transmitting abilities (ETA) to have sons and additional female progeny. Sons became sires of the following generation. The process of evaluation and selection on ETA was continued for four generations. Final sire ETA's were compared with actual transmitting abilities (TA's) for several models of sire evaluation.

Data

Data were simulated for a population of dairy cattle under several sampling (selection) strategies. Simulation of the data began with the creation of eight sires (independently) from a base population with a mean transmitting ability of zero. Fifty female progeny records were simulated for each sire under the model:

$$y_{ij} = s_i + e_{ij} \qquad [1]$$

where y_{ij} is the jth $(1 \le j \le 50)$ progeny record of the ith $(1 \le i \le 8)$ sire, s_i is the random TA of the ith sire with expectation zero (for sires of the first generation) and variance $\sigma_s^2 = 1/16$, and e_{ij} is a random residual effect (independent of s_i) with expectation zero and variance $\sigma_e^2 =$ 15/16. In matrix notation model [1] is:

$$\mathbf{y} = \mathbf{Z}\mathbf{s} + \mathbf{e}$$
 [2]

where y is a vector of progeny records, Z is a known incidence matrix associated with s, the unobservable random vector of sire TA's, and eis an unobservable random vector of residuals. No herd or other fixed effects were included in simulation of progeny records.

Received April 21, 1981.

¹Department of Animal Science, University of California, Davis 95616.

Estimates of transmitting ability (i.e., predictions of s) for the eight sires were computed by best linear unbiased prediction (BLUP) (3) by solving:

$$(\mathbf{Z}'\mathbf{Z} + \mathbf{I}\boldsymbol{\gamma})\mathbf{\hat{s}} = \mathbf{Z}'\mathbf{y}$$
 [3]

where \hat{s} is BLUP of s and $\gamma = \sigma_e^2 / \sigma_s^2 = 15$. The four sires with the highest ETA's were selected to have 2 sons and 50 additional female progeny in each subsequent generation. Computation of the TA of the sons is described in the next section. Fifty female progeny records also were simulated for each of the eight young sires of the following generation. New progeny records for all sires were generated as in [1]. From all progeny records, a new set of ETA's were computed for both generations of sires. The process of evaluation and selection (on ETA) of sires for the next generation of bulls was repeated until there were four generations of eight sires per generation. Final predictions of TA were computed with all available records for all 32 sires and compared with the known TA's.

Selection of Dams of Young Bulls

Based on ETA, genetically superior sires are chosen to have male progeny which are sires of the following generation. To represent the common practice of contracting mates for these superior sires, a procedure to simulate these mates was developed. The procedure began by computing the average of all available records to represent the population mean. With this average two strategies for selection of dams of young sires were employed. In the first strategy, a record which was two standard deviations above the mean was computed [where $2\sigma_V$ = $2(\sigma_s^2 + \sigma_e^2)^5$] and was taken as the selected dam's own record, y_D. The expected TA of an individual making such a record can be computed from selection index theory (15). For females with record y_D, the expected TA is TA = $.5h^2y_D$. In the second strategy of dam selection, y_D was taken as the population mean instead of two standard deviations above the mean.

Given the expected transmitting ability of the dam (Dam TA) and the known TA of the selected sire (Sire TA), the actual TA of the young sire (Son TA) was computed as:

Son TA
$$\approx .5$$
(Sire TA + Dam TA) + $\phi \sqrt{1-.25} \sigma_s$
[4]

where ϕ is a standard normal deviate.

Selection of Sires of Young Bulls

Selection of sires of young bulls was based on ETA. Those four sires with superior estimated transmitting abilities were selected to have additional female progeny and two sons to be sires of the following generation. Selection decisions were based on either 1) within generation comparison of ETA's or 2) across generation comparison of ETA's. In either case, the top four sires were chosen to have two sons each. Across generation selection is more intense than selection within generation because all sires are compared with one another. For selection within generation, only the top four sires out of the eight from each generation were selected to have sons. As mentioned, each young bull received 50 female progeny records generated as in [1], and selected sires of young bulls also received 50 additional female progeny records each generation.

Use of Relationships

As sires are selected to have sons, the use of additive relationships in sire evaluation becomes important. The primary function of the use of relationships is to reduce the error variance of prediction (4). However, simultaneous use of genetic groups and additive relationships has been questioned (1, 6, 14). One goal of this simulation was to examine genetic groups with relationships. The availability and use of relationships proceeded two ways: 1) all relationships among the 32 sires were used in prediction of TA; 2) 25% of the young bulls did not have their sire's identification recorded (this generates bulls with no ties to the base population) but were sons of sires with superior ETA's. The fraction of young bulls with unidentified sires was not exact. The decision to record the identification of the sire of a young bull was random where the probability of having an unidentified sire was .25.

The generation of TA for young bulls was not influenced by identification of his sire. The TA of the sire of the young bull was used to generate the young bull's TA, but the sire identification was not recorded.

Genetic Groups

This simulation intended to examine the accuracy of sire evaluations under several definitions of genetic groups. Thus, ETA's were computed under four sire evaluation models: 1) BLUP without genetic groups; 2) BLUP with genetic groups by generation; 3) BLUP with genetic groups by generation where young sires without an identified sire were in a separate genetic group; and 4) BLUP with genetic groups by relationship (i.e., all sires within a genetic group were related with no relationships across genetic groups). Sires of the first generation first were evaluated by BLUP without genetic groups as in [3]. However, when additional generations of sires were included, sires were evaluated under the model (when genetic groups were used):

$$\mathbf{y} = \mathbf{Z}\mathbf{Q}\mathbf{g} + \mathbf{Z}\mathbf{s} + \mathbf{e}$$
 [5]

where y, Z, s, and e are as in [2] (where the order of these matrices changes as the number of sires and progeny records increases), and Q is a known incidence matrix of order $t \times r$ (t is the number of sires to be evaluated and r is the number of genetic groups) which defines the structure of genetic groups, and g is a vector of unknown fixed group effects of order r. With relationships included:

$$\operatorname{Var} \begin{pmatrix} s \\ e \end{pmatrix} = \begin{pmatrix} A1/\gamma & O \\ O & I \end{pmatrix} \sigma_{e}^{2}$$

where A is the numerator relationship matrix among sires. Accordingly the predictor of sire transmitting abilities is $\hat{\mathbf{u}} = \mathbf{Q}\hat{\mathbf{g}} + \hat{\mathbf{s}}$, where $\hat{\mathbf{u}}$ is a vector of estimated sire transmitting abilities and $\hat{\mathbf{g}}$ and $\hat{\mathbf{s}}$ are solutions to:

$$\begin{pmatrix} \mathbf{Q}'\mathbf{Z}'\mathbf{Z}\mathbf{Q} & \mathbf{Q}'\mathbf{Z}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{Z}\mathbf{Q} & \mathbf{Z}'\mathbf{Z}+\mathbf{A}^{-1}\gamma \end{pmatrix} \begin{pmatrix} \hat{\mathbf{g}} \\ \hat{\mathbf{g}} \end{pmatrix} = \begin{pmatrix} \mathbf{Q}'\mathbf{Z}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix}.$$
[6]

Under model [5], equation [6] has full rank and $Q\hat{g}$ is estimable. Furthermore, for BLUP

Journal of Dairy Science Vol. 65, No. 7, 1982

without genetic groups, \mathbf{Q} does not exist. Parenthetically, the use of a groups model as in [5] does not imply that a group effect was added in simulation of progeny records.

Summary of Simulation

There were two strategies for selection of dams of young bulls, two strategies for selection of sires of young bulls, three uses of relationships, and four models of sire evaluation for a total of 24 possible combinations (BLUP with genetic groups by relationship was not computed when 25% of the relationships were ignored and BLUP with unrelated sires in a separate genetic group was not computed when all relationships were used). Each of the 24 combinations was replicated on 300 sets of 32 sires.

RESULTS AND DISCUSSION

Interpretation of simulation results can be assisted by consideration of a joint selection model by Famula et al. (2). As in this simulation, the joint selection model considers selection of sires based on ETA which can be formulated as $L_1'y$ selection (after Henderson, 4) and selection on dams which directly influences sire transmitting abilities and can be presented as $L_2'u$ selection for two matrices L_1' and L_2' of full row rank. The mixed model equations associated with this selection model are (after model [5]).

$$\begin{pmatrix} \mathbf{Z}'\mathbf{Z}+\mathbf{A}^{-1}\gamma & \mathbf{O} & -\mathbf{L}_{2} \\ \mathbf{O} & \mathbf{L}_{1}'\mathbf{V}\mathbf{L}_{1} & \mathbf{L}_{1}'\mathbf{Z}\mathbf{A}\mathbf{L}_{2}(1/\gamma) \\ -\mathbf{L}_{2}' & \mathbf{L}_{2}'\mathbf{A}\mathbf{Z}'\mathbf{L}_{1}(1/\gamma) & \mathbf{L}_{2}'\mathbf{A}\mathbf{L}_{2}(1/\gamma) \end{pmatrix}$$
$$\begin{pmatrix} \mathbf{\hat{u}} \\ \mathbf{\hat{t}}_{1} \\ \mathbf{\hat{t}}_{2} \end{pmatrix} = \begin{pmatrix} \mathbf{Z}'\mathbf{y} \\ \mathbf{L}_{1}'\mathbf{y} \\ \mathbf{O} \end{pmatrix} [7]$$

where $V = Var(y) = (ZAZ'(1/\gamma) + I)\sigma_e^2$ (4). Quaas and Pollak (12) have shown a correspondence between the groups model in [5] and the selection model [7] for $L_2' = Q'A^{-1}$ and L_1' not existing. As demonstrated in Famula et al. (2), if $L_1'ZAL_2 = O$, proofs computed under the groups model of [5] (i.e., using the mixed model equations of [6]) are equivalent to proofs computed under the selection model [7]. The importance of this result is that if $L_2' = Q'A^{-1}$ (as in the equivalence drawn by Quaas and Pollak (12)), then $L'_1ZAL_2 = L'_1ZQ$, which is null if sires are selected within genetic groups.

Table 1 presents residuals and mean square errors of prediction averaged for each sire in the 300 simulated data sets. Prior to discussion of the results, a discussion of the expected values of residuals and mean square errors is appropriate.

A problem with evaluating the use of genetic groups is that the model rarely can be defined exactly. This study was principally an examination of model building to determine which definition of genetic groups may be optimal for a given sampling design. Examining bias and error variance of prediction as a function of \mathbf{Q} , Z and A alone assumes the true model is known and does not consider y a result of nonrandom sampling.

The entries in the residuals (RES) column of Table 1 are computed as:

$$\begin{array}{ccc} & 32 & 300 \\ \text{RES} = & \Sigma & [\Sigma & (TA_{ij} - TA_{ij})/300] (1/32) \\ & i=1 & j=1 \end{array}$$

where TA_{ij} denotes transmitting ability of the ith sire (1 < i < 32) in the jth (1 < j < 300) data set and TA_{ij} is its estimate. In a model without selection, expectation of this residual would be zero (if the model used to predict TA was the true model). However, expectation of

TABLE 1. Average residuals (RES) and mean square errors (MSE) of prediction for sire evaluation averaged over all generations.

Sampling strategy ^a	RES (× 10) ^b	MSE (X 10) ^c	Sampling strategy	RES (X 10)	MSE (X 10)
Without genetic	groups			<u></u>	
WTL	.25	.12	CTL	.29	.13
WTS	.32	.13	CTS	.36	.14
WAL	.10	.11	CAL	.08	.12
WAS	.14	.12	CAS	.14	.13
With genetic gro	oups				
		groups defined	by generation		
Without a grou	up for unidentified si	res			
WTL	.07	.12	CTL	.11	.12
WTS	.08	.12	CTS	.09	.13
WAL	.08	.12	CAL	.10	.13
WAS	.09	.12	CAS	.07	.13
With a group f	or unidentified sires				
WTS	.08	.12	CTS	.09	.13
WAS	.09	.12	CAS	.10	.13
<u> </u>		groups defined l	by relationships ——		
WTL	.12	.12	CTL	.14	.13
WAL	.10	.12	CAL	.09	.13

^aThe triplets of letters are combinations from the following parts: Sire selection, W = within generation, C = across generation; dam selection, T = selected dams have records $2\sigma_y$ above average average, S = selected dams have average records; relationships used, L = all, S = 25% of sire identifications are ignored.

^bRES =
$$\sum_{i=1}^{32} \sum_{j=1}^{300} (TA_{ij} - TA_{ij})/300)$$
 (1/32), where TA_{ij} is the transmitting ability of the ith sire in the jth

data set and TA_{ii} is its estimate.

$$^{C}MSE = \sum_{i=1}^{32} (\sum_{j=1}^{300} (TA_{ij} - TA_{ij})^{2}/300) (1/32).$$

Journal of Dairy Science Vol. 65, No. 7, 1982

this residual is not zero for most sampling strategies because the assumption of random sampling for y has been violated.

Theoretical limits for the mean square error also can be determined if the model used to compute ETA's was the true model (without selection). Mean square error (MSE) was computed as

$$MSE = \sum_{i=1}^{32} \sum_{j=1}^{300} (TA_{ij} - TA_{ij})^2 / 300] (1/32).$$
[9]

In BLUP without genetic groups, the expected value of [9] can be written in matrix notation as:

$$E(MSE) = tr(Z'Z + A^{-1}\gamma)^{-1}\sigma_e^2/32 \quad [10]$$

where tr denotes the trace of a matrix. Under a no groups model with all relationships known E(MSE) = .011. With genetic groups:

$$\begin{pmatrix} \mathbf{Q}'\mathbf{Z}'\mathbf{Z}\mathbf{Q} & \mathbf{Z}'\mathbf{Z}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{Z}\mathbf{Q} & \mathbf{Z}'\mathbf{Z}+\mathbf{A}^{-1}\gamma \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{Q}' \\ \mathbf{I} \end{pmatrix} \sigma_{\mathbf{C}}^{2}/32 .$$
[11]

E(MSE) for all definitions of Q, with all relationships known is .014. E(MSE) is larger when 25% of the relationships are ignored. As the true model is not known and selection changes the distribution of y, entries in Table 1 can only approach the theoretical expectation.

The results in Table 1 show the use of genetic groups can increase the accuracy of sire evaluation under particular sampling strategies. For example, the accuracy of sire evaluation increases when genetic groups by generation are used, as compared to no genetic groups, when there has been intense selection of dams. This result is expected given equations [7]. The correspondence between the groups model of [5] and the $L'_2 u$ selection model shows genetic groups can account for dam and sire selection if sires are selected on within group comparisons (i.e., $L'_1 ZQ = O$) of progeny means or previously computed proofs. When groups are defined by generation, within generation selection is also within group selection. As a result there is

a larger residual (bias) in grouping by relationship than in grouping by generation for within generation sire selection. However, the increased accuracy is not maintained for across generation selection, and this, once again, can be expected under the joint selection model equations [7]. Note, if sire selection was based on within family comparisons, grouping by relationship would increase the accuracy of sire evaluation over grouping by generation. Table 1 also shows an increase in mean square error of prediction when 25% of the relationships are ignored compared to when all relationships are known as expected (5). When selection of dams is absent (a case where groups are not necessary by equations [7]), actual mean square error approaches the theoretical value. Residuals (bias) also increase when relationships are ignored as can be explained with equation [7]. Quaas and Pollak (12) showed a correspondence between groups model and Henderson's (4) L'u selection model such that in equations [7], $L'_2 = Q'A^{-1}$. However, if relationships are ignored, the A^{-1} used to construct L_2 is not the inverse of the true relationship matrix. Thus, $L_1^{\prime}ZAL_2 \neq 0$ even for within group sire selection, and if this term is not included in computation of sire proofs (which it is not, under the groups model), there is a potential for bias in the resulting sire proofs. This point, perhaps, is not recognized in many sire evaluation models, i.e., that the A^{-1} used in the computation of the sire proofs is not the inverse of the true relationship matrix. If the true A^{-1} is not used, the groups model is not adequate to account for dam of sire selection as can be seen by comparing the equations for the modified groups model of (12) with equations [7].

Table 2 presents the average group solutions for sampling strategies with grouping by generation. Note that group effects are estimable in the models. If relationships among sires are ignored, the average group solutions increase. This result has been explained by Pollak and Quaas (9), who showed that group solutions are functions of the "average genetic superiority of the selected ancestors to their generation mean." Thus, as relationships are ignored, additional terms appear in the group solutions because they are no longer corrected for as in using the complete relationship matrix. Group solutions under across generation selection of sires are larger than under within generation

	Group (Generation)					
Sampling strategy ^a	I	II	III	IV		
WTL	01	.11	.19	.24		
WTS	01	.15	.22	.27		
WAL	01	01	02	.01		
WAS	01	.01	.03	.05		
CTL	01	.11	.17	.23		
CTS	.01	.15	.22	.28		
CAL	01	01	01	.03		
CAS	.01	.03	.04	.08		

TABLE 2. Average group solutions for models with genetic groups defined by generation.

^aThe triplets of letters are combinations from the following parts: Sire selection, W = within generation, C = across generation; dam selection, T = selected dams have records $2\sigma_y$ above average average, S = selected dams have average records; relationships used, L = all, S = 25% of sire identifications are ignored.

selection. The increase in selection intensity amplifies the "average genetic superiority of the selected ancestors" (9). However, when there is no selection of dams of young bulls and all relationships are used, the average group solutions exhibit no definite trend. As Tables 1 and 3 indicate, genetic groups are not necessary when dams are not selected and all relationships are known.

Average mean square error of prediction of TA for fourth generation bulls is in Table 3. From this table, the accuracy of sire evaluation can be assessed for sires with few progeny and no ties to the base population. Mean square error of prediction (MSE) for individual bulls of the fourth generation was computed as:

$$MSE = \sum_{i=25}^{32} \sum_{j=1}^{300} (TA_{ij} - TA_{ij})^2 \delta_{ij} / \sum \sum_{i j} \delta_{ij}$$
[12]

for young bulls with identified sires and:

$$32 \quad 300 \\ MSE = \sum_{i=25}^{300} \sum_{j=1}^{300} (TA_{ij} - TA_{ij})^{2} \\ (1-\delta_{ij}) / \sum_{ij} (1-\delta_{ij}) \\ ij$$
 [13]

for young bulls with unidentifed sires where $\delta_{ij} = 1$ if the ith (25 < i < 32) young bull in the jth (1 < j < 300) data set has an identified sire, and

O otherwise. Sires of the fourth generation are identified by numbers 25 through 32.

Theoretical limits for MSE of individual sires can be computed if the model used to compute ETA's is the true model (without selection). For young bulls with a tie to the base population E(MSE) = .014 under all definitions of **Q**. For young bulls with unidentified sires E(MSE) =.014 under a no-groups model and under other definitions of **Q**, E(MSE) = .015. Because the true model is not known and selection changes the distribution of **y**, entries in Table 3 do not approach theoretical estimates.

As in Table 3, there is essentially no difference in accuracy of evaluation for young bulls with identified sires regardless of the use of genetic groups. The importance of genetic groups arises in evaluation of young bulls without identified sires when there has been selection of dams of sires. In the absence of dam selection, the use of genetic groups does not increase accuracy of evaluating bulls (as also in Table 1).

The results of this simulation suggest that the definition of genetic groups will remain arbitrary. However, guidelines under assumed selection strategies can be developed. For example, although grouping by generation is not possible in actual populations, genetic groups defined by birthdate or some other function of time would be useful in accounting for sire selection based on comparisons of sires within years. This conclusion is particularly relevant if all relationships are not known,

Sampling strategy ^a	MSE of young bulls with identified sires (X 10) ^b	MSE of young bulls with unidentified sires (× 10) ^c	
Without genetic grou	ps		
WTS	.16	.20	
WAS	.14	.16	
CTS	.16	.25	
CAS	.15	.15	
With genetic groups	defined by generation		
Without a group for	r unidentified sires		
WTS	.16	.16	
WAS	.15	.15	
CTS	.15	.16	
CAS	.14	.16	
With a group for un	identified sires		
WTS	.15	.15	
WAS	.15	.15	
CTS	.15	.15	
CAS	.15	.16	

TABLE 3. Average mean square errors (MSE) of prediction for fourth generation bulls.

^aThe triplets of letters are combinations from the following parts: Sire selection, W = within generation, C = across generation; dam selection, T = selected dams have records $2\sigma_y$ above average, S = selected dams have average records; relationships used, L = all, S = 25% of sire identifications are ignored.

 $b_{MSE = \Sigma}^{32 300} (\Sigma (TA_{ij} - TA_{ij})^2 (\delta_{ij})) (\frac{1}{\Sigma \Sigma \delta_{ij}}),$ i=25 j=1 ($\sum_{i=1}^{2} (\Delta_{ij})^2 (\delta_{ij})$) ($\sum_{i=1}^{2} (\Delta_{ij})$),

where TA_{ij} is the transmitting ability of the ith sire of the fourth generation in the jth data set, TA_{ij} is its estimate, and $\delta_{ij} = 1$ if the ith sire in the jth data set has an identified sire and zero otherwise.

 $\label{eq:MSE} \begin{array}{c} {}^{32} \quad {}^{300} \\ {}^{c} \text{MSE} = \sum \left(\sum \limits_{j=1}^{\infty} \left(\text{TA}_{ij} - \text{TA}_{ij} \right)^2 \left(1 - \delta_{ij} \right) \left(\frac{1}{\sum \sum \limits_{i=1}^{\infty} \left(1 - \delta_{ij} \right)} \right). \end{array}$

which makes grouping by relationship impractical. However, if sire selection decisions are based on comparisons within family, then grouping by relationship is optimal (i.e., $L_1'ZAL_2 = O$ from equation [7]). In the absence of dam selection, genetic groups may be dropped from the model. Thus, choice of a definition for genetic groups is dependent upon visualization of a selection strategy that fits the data. In most practical situations, a definition of groups which mimics generations appears to be most appropriate.

An examination of the maternal grandsire model of sire evaluation of Quaas et al. (11) was not possible in this simulation. However, application of their model to data sets with intense selection of dams of young bulls should reduce the need for grouping if all sires can be tied by relationships to the base population. Famula et al. (2) showed that selection on the female side of a sire's pedigree can be accounted for either by genetic groups or by inclusion of dam information in the mixed model equations. Quaas et al. (11) demonstrated that information on female relatives can be included in sire evaluation by approximating some numerator relationship coefficients. Thus, the use of a maternal grandsire model for sire evaluation can reduce the need for genetic groups for individuals with ties to the base population, provided maternal grandsire information is available.

CONCLUSION

In development of this simulation, selection strategies were chosen to reflect realistic practices. To account for effects of nonrandom mating of sires and dams, genetic groups should be defined such that sire selection is based on comparisons within groups. However, when all bulls can be tied by relationships to the base population and there has been no selection of dams, the need for genetic groups vanishes. The results of this study suggest that a unique optimal definition of genetic groups does not exist and that any successful definition of genetic groups will be dependent on visualization of the selection scheme underlying the data.

ACKNOWLEDGMENTS

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

REFERENCES

- 1 Dempfle, L. 1978. BLUP unter Verwendung der Verwandtschaftsmatrix. Pages 76-85 In Zuchwertschatzung beim Rind. Habilitationsschrift, Munchen-Weihenstephen.
- 2 Famula, T. R., E. J. Pollak, and L. D. Van Vleck. 1982. The use of genetic groups in dairy sire evaluation under a selection model. J. Dairy Sci. (Submitted).
- 3 Henderson, C. R. 1973. Sire evaluation and genetic trends. Page 10 *in* Proc. Anim. Breeding Genet. Symp. in Honor of Dr. Jay L. Lush, ASAS and ADSA, Champaign, IL.
- 4 Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics 31:423.
- 5 Henderson, C. R. 1976. A simple method for computing the inverse of a numerator relationship

matrix used in prediction of breeding values. Biometrics 32:69.

- 6 Jensen, E. L. 1980. Bull groups and relationships among sires in best linear unbiased prediction sire evaluation models. J. Dairy Sci. 63:2111.
- 7 Kennedy, B. W. 1981. Bias and mean square error in mixed model sire evaluation when genetic groups are ignored or included. J. Dairy Sci. 64:689.
- 8 Kennedy, B. W., and J. E. Moxley. 1975. Comparison of genetic group and relationship methods for mixed model sire evaluation. J. Dairy Sci. 58:1507.
- 9 Pollak, E. J., and R. L. Quaas. 1981. Definition of group effects in sire evaluation models. J. Dairy Sci. (Suppl. 1) 64:72. (Abstr.)
- 10 Pollak, E. J., G. R. Ufford, and S. J. Gross. 1977. Comparison of alternative models for within-herd genetic evaluation of beef cattle. J. Anim. Sci. 45:1010.
- 11 Quaas, R. L., R. W. Everett, and A. E. McClintock. 1979. Maternal grandsire model for dairy sire evaluation. J. Dairy Sci, 62:1648.
- 12 Quaas, R. L., and E. J. Pollak. 1981. Modified equations for sire models with groups. J. Dairy Sci. 64:1868.
- 13 Stolzman, E. M. 1979. Testing different strains of Friesian cattle in Poland. Inst. Zootech., Cracow, Poland. 5th Annu. Rep.
- 14 Thompson, R. 1979. Sire evaluation. Biometrics 35:339.
- 15 Van Vleck, L. D. 1979. Notes on the theory and application of selection principles for the genetic improvement of animals. Dept. Anim. Sci., Cornell Univ., Ithaca, NY.
- 16 Willham, R. L. 1979. Evaluation and direction of beef sire evaluation program. J. Anim. Sci. 49:592.