

# MULTIBREED SIRE EVALUATION PROCEDURES WITHIN A COUNTRY

M. A. Elzo<sup>1</sup> and T. R. Famula<sup>2</sup>

University of California,  
Davis 95616<sup>3</sup>

## Summary

Frequently, exotic, hybrid and native sires are evaluated based on progeny records, using additive models for direct effects and assuming common genetic variances across breed groups. In some cases, however, additive maternal and nonadditive direct and maternal fixed and random effects may be important. Also, heterogeneous variances across genetic groups may exist. Hence, sire proofs from additive models will be biased and not of minimum variance. Sire-maternal grandsire models that account for direct and maternal, additive and nonadditive fixed and random effects as well as different additive and nonadditive variances across genetic groups are presented. Direct and maternal effects are treated as separate characters. Sire proofs for direct and maternal traits contain group (fixed) and sire (random) additive and nonadditive effects. Problems of confounding and multicollinearity may arise due to empty cells in the design. Ridge regression may be used to help remove problems of multicollinearity. Extension to multitrait evaluation is straightforward.

(Key Words: Sire Evaluation, Best Linear Unbiased Prediction, Additive Models, Non-additive Models.)

## Introduction

Importation of semen from improved breeds is often used in attempts to increase local

livestock productivity. Frequently this is done without systematic evaluation of the introduced stocks in the new environment. When straightbred and(or) crossbred sires have been evaluated, single trait additive models for direct effects, which assume equal additive variances across genetic groups, have been used (e.g., Mahadevan et al., 1970; Franklin et al., 1976; Leroy et al., 1978; Parekh and Sahu, 1978; Parekh and Pande, 1982). However, nonadditive direct and maternal effects have been found to be important for various traits in several studies (e.g., Touchberry, 1970; Taneja and Bhat, 1974; Donald et al., 1977; Alenda et al., 1980a,b; Massey and Benyshek, 1981; Peacock et al., 1981; Robison et al., 1981; Rincon et al., 1982). Also, heterogeneity of additive covariances across breed groups may exist, but this has not been adequately studied. Thus, the objective of this research is to present procedures to evaluate exotic, native and crossbred sires using data from native environments. These procedures will account for direct and maternal, additive and nonadditive effects, as well as unequal additive and nonadditive covariances across breed groups and breed group combinations, respectively.

## Development of Multibreed Sire-Maternal Grandsire Evaluation Models

Willham (1963) presented a genetic model accounting for direct and maternal effects. Based on Willham's model and using gametic modeling techniques (Quaas and Pollak, 1980), a maternal multibreed sire-maternal grandsire interaction (MSMI) model can be derived. This model (MSMI) contains environmental fixed effects, fixed additive genetic effects (sire groups, maternal granddam groups), fixed nonadditive effects (sire group  $\times$  dam group interactions) random sire effects, random nonadditive effects (sire  $\times$  dam group), interactions and residual effects. Both direct and maternal effects are included in the MSMI

<sup>1</sup>Present address: Dept. of Anim. Sci., Cornell Univ., Ithaca, NY 14853.

<sup>2</sup>The authors gratefully acknowledge A. P. Fenech for his statistical assistance, G. E. Bradford for many useful discussions and E. J. Pollak for editorial comments on this paper. The senior author also acknowledges the financial support of the Univ. of California, Davis, throughout his graduate studies.

<sup>3</sup>Dept. of Anim. Sci.

Received March 20, 1984.

Accepted December 4, 1984.

model. Effects are either present or absent in a given record (i.e., there is a 1 or a 0 in the appropriate incidence matrix). Before explaining the meaning of the term "group" in the context of this paper, two other terms must be defined. First is breed group subclass, which refers to all animals having a given breed composition (straightbred or crossbred). The second is time group subclass, which alludes to all animals belonging to a given time period, such as generation, year of entry into the stud, et cetera. Thus, group (sire group, maternal granddam group) is defined to be the mean additive genetic effect of all animals (sires, maternal granddams) within a breed group  $\times$  time group subclass. In matrix notation the MSMI model is:

$$y = X\alpha + Z^S Q^S g^S + Z^{mgd} Q^{mgd} g^{mgd} + Z^{im} Q^{im} \delta_1^{im} + Z^S a^S + Z^{im} \delta_2^{im} + e \quad (1)$$

$$E(y) = X\alpha + Z^S Q^S g^S + Z^{mgd} Q^{mgd} G^{mgd} + Z^{im} Q^{im} \delta_1^{im}$$

$$\text{var}(y) = \text{var}(Z^S a^S + Z^{im} \delta_2^{im} + e),$$

where  $y$  is the observation vector,  $\alpha$  is a vector of fixed environmental effects,  $a^S$  is a vector of sire additive values for direct and maternal effects, ordered within sire,  $g^S$  is a vector of direct and maternal sire group effects,  $\delta_2^{im}$  is a vector of sire  $\times$  group-of-dam interactions for both direct and maternal effects,  $\delta_1^{im}$  is a vector of direct and maternal group-of-sire  $\times$  group-of-dam interaction effects and  $e$  is a vector of residual effects. The  $X$ ,  $Z^S$ ,  $Z^{mgd}$  and  $Z^{im}$  are known incidence matrices relating observations to environmental fixed effects, direct and maternal sire effects, direct plus maternal maternal-granddam effects and direct and maternal sire  $\times$  dam group interaction effects, respectively. The  $Q^S$ ,  $Q^{mgd}$  and  $Q^{im}$  are known incidence matrices relating sires and maternal grandsires, maternal granddams and sire  $\times$  dam group interaction effects to their respective group effects. For the remainder of this paper, the terms sire interactions and sire group interactions will be used instead of sire  $\times$  group-of-dam interactions and group-of-sire  $\times$  group-of-dam interactions.

The covariance matrix of the random effects of the MSMI model is:

$$\text{var} \begin{bmatrix} a^S \\ \delta_2^{im} \\ e \end{bmatrix} = \begin{bmatrix} G & 0 & 0 \\ 0 & D_2^{im} & 0 \\ 0 & 0 & R \end{bmatrix}$$

The  $G$  is an additive covariance matrix among male ancestors only, for both direct and maternal effects, which accounts for unequal covariances across sire groups. Assuming constant additive covariances within base breeds over time and absence of linkage, additive covariances for crossbred groups can be calculated as a weighted mean of the additive covariances of the base breeds represented in them (Elzo, 1983). The weights for the base breed additive covariances are the expected fractions of these base breeds in a crossbred group. This reduces the number of additive covariances required to be known (or estimated) to only the number of base breed additive covariances. Also,  $G$  can easily be computed using the recursive rules given in Appendix C, Elzo (1983).

The  $D_2^{im}$  is a covariance matrix of direct and maternal sire interaction effects, assumed to be diagonal. Nonadditive covariances among different sire interaction effects, which might exist, are ignored to facilitate computations. Under the assumption of equal direct and maternal interaction variances across sire groups, only one direct and one maternal interaction variance are required.

It will also be assumed that the vector of observations contains only one record per animal and that no environmental covariance among records of different animals exists. Thus, if covariances among records due to dam-dam and sire-dam relationships are ignored, the covariance matrix of the residual effects ( $R$ ) is a diagonal matrix. The expression of the residual variances will depend on the male ancestors ( $s$ ,  $mgs$ ) identified (see Appendix for explicit expressions). The value of these residual variances will depend on the expression used and on the value of the genetic covariances corresponding to the groups of the individual producing the record and its ancestors.

The MSMI model is not well suited for evaluating sires from several breed groups in open crossbreeding programs because additive and nonadditive direct and/or maternal effect differences among sire groups cannot be estimated separately. The expected value of sire-group solution differences for additive direct or maternal effects will always con-

tain a nonadditive direct or maternal component and, depending on the data set, various degrees of confounding between direct and maternal genetic effects may exist. This means that sires can be compared separately for additive and nonadditive effects within mating type only. Across-group comparisons for direct effects will include at least one nonadditive maternal component when sires from different genetic groups are compared for maternal effects. Also, the information that contributes to the estimate of sire group differences and the prediction of sire interaction effects comes only from records of the specific subclasses involved in the comparison. Hence, it would be desirable: a) to improve the utilization of the data to estimate additive and nonadditive group effects, as well as to predict nonadditive sire interaction effects, b) to obtain separate additive and nonadditive sire comparisons for direct and maternal effects and c) to be able to predict the outcome of all possible sire  $\times$  group-of-dam interactions, not only the ones corresponding to the filled subclasses in the design. Procedures that meet these conditions will enable breeders to make better mating plans for a given population.

Reparameterization and use of regression procedures will be used to solve these problems. The modifications of the MSMI model will be carried out in two steps: first, for the fixed and random interaction effects (model 1) and second, for the additive fixed effects (model 2).

#### Strategy for Nonadditive Effects

Nonadditive interaction effects will be explained through intra- and inter-locus interactions between alleles of the same and different breed origin, within a given sire group or sire interaction subclass. Sire group and sire interaction subclasses will now be defined to be the sire-time group  $\times$  dam-time group subclass (SGIS) and the sire  $\times$  dam-time group subclass (SIS), respectively. On the other hand, genetic group-of-sire  $\times$  genetic group-of-dam interactions (SGI) will be assumed to be proportional to the probabilities of occurrence of alleles from the same or from different base breeds, at one or two loci in a SGIS. Sire  $\times$  genetic group-of-dam interactions (SI) are defined as SGI, but measured within a SIS. This reparameterization of the nonadditive effects will yield one set of nonadditive parameters per sire group or sire  $\times$  dam-time group subclass.

However, if we wanted pooled sets of nonadditive parameters across dam-time groups, SGIS and SIS could be redefined to include all dam-time groups. Under these conditions, there will be only one set of nonadditive parameters associated with a sire group or with a sire.

There will be two sets of nonadditive parameters explaining fixed and random interactions: one for intra-locus and another for inter-locus (two loci) interaction effects (see Appendix). Higher-order interactions are assumed to be negligible, but this approach can be extended to account for them if necessary. The intra-locus interaction parameters contain the sum of all intra-locus interaction effects due to alleles of any two breeds, over all loci, within a sire group or a sire interaction subclass. The two-locus interaction parameters contain the sum of all two-locus interaction effects due to alleles of any two, three or four breeds, over all two-locus combinations, within a sire group or a sire interaction subclass. The individual fixed interaction parameters will not be estimable, but, depending on the model, differences between them will.

The resulting number of new interaction parameters will probably make it impossible to use the model in most open crossbreeding data sets. Problems of confounding and multicollinearity will arise due to empty subclasses in the design. Also, data may be insufficient to obtain accurate estimates of some nonadditive variances and any estimate at all for others. Therefore, simplifying assumptions relative to the nonadditive parameters defined above may be required for most data sets. Explicit expressions for fixed and random interaction effects under various simplifying assumptions as well as the complete model are given in Appendix A, Elzo (1983). The simplest alternative is to assume that: (i) all inter-locus interactions at two or more loci are negligible, (ii) all intra-locus intrabreed interactions are the same for all base breeds, (iii) all intra-locus interbreed interactions are the same for all interbreed combinations. Under these assumptions, the number of new interaction parameters reduces to two per trait (direct, maternal) and per sire group or sire interaction subclass. Several authors have made this assumption when evaluating breed groups (e.g., Touchberry, 1970; Parekh and Touchberry, 1974; Koger et al., 1975; Bhat et al., 1978; Peacock et al., 1981, 1982), while others have separated the intra-locus effects due to combination of alleles

of different breeds (Dillard et al., 1980; Robison et al., 1981).

The new sire group and sire interaction terms, in matrix notation, are  $Zsg^d Qsg^d \delta_1$  and  $Zsg^d \delta_2$ , respectively. The MSMI model modified this way will be called model 1. The  $\delta_1$  is a vector of alternating fixed direct and maternal interaction effects associated with a SGIS. The  $\delta_2$  is a vector of alternating random direct and maternal interaction effects associated with a SIS. The  $Qsg^d$  is a known incidence matrix relating elements of  $\delta_2$  to elements of  $\delta_1$ . The  $Zsg^d$  is a known incidence matrix relating observations to direct and maternal elements of  $\delta_2$  through  $t$ . The  $t$ 's are the probabilities of having alleles of particular breeds at one and two loci. The definitions of these probabilities (i.e.,  $t$ ) will depend on the assumptions made with respect to the heterogeneity of the fixed and random nonadditive effects in the various one- and two-locus allelic combinations present in a data set (see Appendix A, Elzo, 1983 for further details). The number of nonzero elements per row in  $Zsg^d$  will vary according to the male ancestors identified for an individual with a record. For instance, consider the case of two breeds in a data set: Simmental (S) and Hereford (H). There are three intra-locus direct, three intra-locus maternal, three inter-locus (two loci) direct and three inter-locus maternal nonadditive parameters according to formula (4) in the Appendix. The values of these 12 elements in a row of  $Zsg^d$  for a progeny of .75S .25H sire and a .5S .5H dam, and whose maternal grandsire is 100% S and whose maternal granddam is 100% H, are: .375, 0, .5, 1.0, .125, 0, .140625, 0, .84375, 1.0, .015625 and 0, which correspond to  $\delta_{W11}^D$ ,  $\delta_{W11}^M$ ,  $\delta_{W12}^D$ ,  $\delta_{W12}^M$ ,  $\delta_{W22}^D$ ,  $\delta_{W22}^M$ ,  $\delta_{B11}^D$ ,  $\delta_{B11}^M$ ,  $\delta_{B12}^D$ ,  $\delta_{B12}^M$ ,  $\delta_{B22}^D$  and  $\delta_{B22}^M$ , respectively, in the notation of formula (4) in the Appendix. Also, the superscripts D and M denote direct and maternal nonadditive effects, respectively, and the subscripts 11, 12, 22 of W and B refer to nonadditive effects among Simmental alleles only, Simmental and Hereford alleles and Hereford alleles only, respectively.

The covariance matrix of  $\delta_2$ , i.e.,  $D_2$ , is assumed to be diagonal, and it contains one set of direct and one set of maternal nonadditive variances per SIS. As for the MSMI model, nonadditive covariances among SIS for either

direct or maternal nonadditive effects are ignored, to ease computations. Assuming constant nonadditive variances across sire interaction subclasses, only one set of direct and one of maternal sire interaction variances are required per data set.

Model 1 has the advantage over the MSMI model that any potential breed group-of-sire  $\times$  breed group-of-dam interaction or sire  $\times$  breed group-of-dam interaction can be estimated or predicted as a weighted mean of its fixed or its random nonadditive parameters. The weights are the probabilities of occurrence of these nonadditive effects in a particular mating. Also, information from all progeny breed groups within a SGIS or within a SIS is used to estimate fixed or to predict random nonadditive parameters. Thus, when the nonadditivity assumptions are true and the number of nonadditive parameters in model 1 is less than or equal to that of the MSMI model, estimable linear contrasts in model 1 should be at least as precise as those of the MSMI model.

#### Additive Grouping Strategy

The additive grouping system used in the MSMI model and in model 1 is based on time  $\times$  breed group subclasses. However, this system has the inconvenience of having to deal with an always-increasing number of breed groups over time. One solution is to limit the number of breed groups by setting up breed composition ranges. For instance, in a two-breed situation, a breed composition range of 10% could be set, such that breed group 1 would contain all animals with up to 10% alleles from breed A, and the remainder from breed B, breed group 2 would have animals with more than 10% and up to 20% of breed A alleles and the remainder from breed B, et cetera, up to the 10th breed group, whose breed composition is the opposite to breed group 1. Nevertheless, this system still uses information from only a particular breed group subclass to estimate the corresponding breed group effect. Another system is to express breed groups within generations as a function of the base breeds. Let model 1 modified this way be model 2. All breed groups within a generation will furnish information to estimate these generation  $\times$  breed additive group effects. Consider, for instance, two base breeds and two sires from two different breed subclasses within a generation. If sire groups are defined as in the MSMI model, each sire will

contribute to a different sire group. If breed composition ranges are used to merge sire groups, each sire will help estimate one sire group only if it falls within one breed composition range. However, when breed groups are expressed as a function of the base breeds, records from both sires will contribute to estimate all base breeds in proportion to their breed composition. The number of new additive parameters will always be less than or equal to the number of the old ones. It can be formally proven (A. P. Fenech, personal communication) that the variance of estimable linear combinations of additive groups in model 2 is at least as small as that of model 1. Another advantage of model 2 is that any potential breed group can be estimated. This type of regression approach to fixed additive effects has been used to evaluate straightbred and crossbred groups in several beef and dairy crossbreeding experiments (Touchberry, 1970; Parekh and Touchberry, 1974; Koger et al., 1975; Dillard et al., 1980; Peacock et al., 1981, 1982; Robison et al., 1981). However, none of the above research had sire evaluation in a multibreed population as an objective.

Model 2 in matrix notation is:

$$y = X\alpha + Z^S Q^S P^S b^S + Z^{mgd} Q^{mgd} P^{mgd} b^{mgd} + Z^{sgd} Q^{sgd} \delta_1 + Z^S a^S + Z^{sgd} \delta_2 + e \quad (2)$$

$$E(y) = X\alpha + Z^S Q^S P^S b^S + Z^{mgd} Q^{mgd} P^{mgd} b^{mgd} + Z^{sgd} Q^{sgd} \delta_1$$

$$\text{var}(y) = \text{var}(Z^S a^S + Z^{sgd} \delta_2 + e),$$

where  $b^S$  is a vector of alternating direct and maternal time  $\times$  breed of sire additive group effects,  $b^{mgd}$  is a vector of direct plus maternal time  $\times$  breed of mgd group effects,  $P^S$  is a known incidence matrix relating elements of  $g^S$  to elements of  $b^S$  through the  $p^S$  and  $P^{mgd}$  is a known incidence matrix relating elements of  $g^{mgd}$  to elements of  $b^{mgd}$  through the  $p^{mgd}$ . The  $p^S$  and  $p^{mgd}$  are the expected fractions of alleles of the  $i^{\text{th}}$  base breed in the breed group of the sire and the maternal granddam, respectively. For instance, in the Simmental-Hereford example presented above, the values for the  $p^S$  are (.75, .25), (1.0, 0) and (1.0, 0) and for  $p^{mgd}$  are: (0, 1.0), which correspond to the direct sire breed effects (first pair), direct maternal grandsire effects (second pair), mater-

nal maternal-grandsire effects (third pair) and maternal granddam effects (fourth pair), relative to the Simmental and Hereford breeds, respectively, for the appropriate sire, maternal grandsire- and maternal granddam-time groups. The remaining terms of (2) are as previously defined.

The inverses of the matrices of residual effects ( $R^{-1}$ ), nonadditive sire effects ( $D_2^{-1}$ ) and additive sire effects ( $G^{-1}$ ) are needed to write the mixed model equations (MME) for model 2. The  $R^{-1}$  and  $D_2^{-1}$  are simply the reciprocals of their diagonal elements. Because the direct and maternal additive variances are allowed to differ among breed groups,  $G^{-1}$  will not be equal to  $4Vg^{-1} * A^{-1}$  (Henderson, 1976b), where  $Vg$  = additive covariance matrix of direct and maternal effects and  $A$  = numerator relationship matrix among sires. As an extension of Henderson (1975a,b; 1976a,c) and Quaas (1975), Elzo (1983) derived rules to write  $G^{-1}$  directly for multiple traits and heterogeneous covariance matrices across breed groups. Application of these rules for the case of two traits yields  $G^{-1}$  for model 2. As an illustration, assume, in the Simmental-Hereford example above, that the 100% Simmental sire is a base animal within unknown ancestors and also the sire of the .75S .25H bull, which has unknown maternal granddam. Let the elements of the additive covariances matrices for direct and maternal effects (ordered by column) be (4, 2, 2, 3) for Simmental and (9, 4, 4, 6) for Hereford, for a given trait. By formula C.10, Elzo (1983), the additive covariance matrix for the .75S .25H bull is .75 times the Simmental covariance matrix, plus .25 times the Hereford covariance matrix, i.e., (5.25, 2.5, 2.5, 3.75). The inverses of the Simmental and .75S .25H breed groups are (.375, -.25, -.25, .5) and (.27907, -.18605, -.18605, .39070), approximately. The inverse of the additive covariance matrix for this example is a  $4 \times 4$  symmetric matrix, whose upper triangular entries, calculated using the recursive rules given in Appendix B, Elzo (1983), are

$$\begin{aligned} g^{11} &= .375 + .25(.27907), \\ g^{12} &= -.25 + .25(-.18605) = g^{21}, \\ g^{13} &= -.5(.27907), \\ g^{14} &= -.5(-.18605) = g^{23}, \\ g^{22} &= .5 + .25(.39070), \\ g^{24} &= -.5(.39070), \\ g^{33} &= .27907, \\ g^{34} &= -.18605, \\ g^{44} &= .39070. \end{aligned}$$

So, the full  $G^{-1}$  covariance matrix is

$$\hat{s} = L \hat{d}, \quad (3)$$

$$\begin{bmatrix} .445 & -.297 & -.140 & .093 \\ -.297 & .598 & .093 & -.195 \\ -.140 & .093 & .279 & -.186 \\ .093 & -.195 & -.186 & .391 \end{bmatrix},$$

where

$$L = [Q^S P^S : T^S Q^S : I : T^S],$$

$$\hat{d} = [\hat{b}^S : \hat{\delta}_1 : \hat{a}^S : \hat{\delta}_2']'$$

$\hat{s}$  = BLUP of  $s$ ,

$\hat{b}^S$  = generalized least-squares (GLS) estimator of  $b^S$ ,

$\hat{\delta}_1$  = GLS estimator of  $\delta_1$ ,

$\hat{a}$  = BLUP of  $a$ ,

$\hat{\delta}_2$  = BLUP of  $\delta_2$ ,

$T^S$  = known matrix relating direct and maternal sire by group-of-dam interaction effect to elements of  $\delta_2$  through  $\{t^d\}$  and  $\{t^m\}$ .

where all entries were approximated to three digits.

The contributions of progeny records to the left hand side (LHS) and right hand side (RHS) of the MME can be found using the expression  $c_i \times r_i^{-1} \times c_i'$  and  $c_i \times r_i^{-1} \times y_i$ , respectively, where  $c_i$  and  $c_i'$  are the coefficients of the  $i$ th and  $i$ 'th parameters of model 2 for a record,  $r_i^{-1}$  is the inverse of its residual variance and  $y_i$  is the value of the progeny record. The  $c_i$  and  $r_i^{-1}$  are given in table 1, where all terms are as previously defined.

*Sire Best Linear Unbiased Prediction (BLUP) for Direct and Maternal Traits.* The vector of sire proofs for direct and maternal traits is:

The remaining matrices are as defined previously. Additive and nonadditive sire group effects are taken as deviations from base groups within sire time groups. The estimators and predictors come from solving the MME for model 2.

TABLE 1. EXPRESSIONS FOR  $c_i$  AND  $r_i^{-1}$  USED TO CALCULATE THE CONTRIBUTIONS OF AN INDIVIDUAL RECORD TO THE MME FOR MODEL 2

Parameter	Sire, mgs identified	Only sire identified	Only mgs identified
$\mu$	1	1	1
$b_i^{ds}$	$p_i^s$	$p_i^s$	0
$b_i^{dmgs}$	$.5 p_i^{mgs}$	0	$.5 p_i^{mgs}$
$b_i^{mmgs}$	$p_i^{mgs}$	0	$p_i^{mgs}$
$b_i^{mgd}$	$p_i^{mgd}$	0	$p_i^{mgd}$
$\delta_{1i}^d$	$t_{1i}^d$	$t_{1i}^d$	0
$\delta_{1i}^m$	$t_{1i}^m$	0	$t_{1i}^m$
$a_i^{ds}$	1	1	0
$a_i^{dmgs}$	.5	0	.5
$a_i^{mmgs}$	1	0	1
$\delta_{2i}^d$	$t_{2i}^d$	$t_{2i}^d$	0
$\delta_{2i}^m$	$t_{2i}^m$	0	$t_{2i}^m$
$r_i^{-1}$	IE1A <sup>a</sup>	IE2A <sup>b</sup>	IE3A <sup>c</sup>

<sup>a</sup>IE1A = inverse of equation 1 in the Appendix.

<sup>b</sup>IE2A = inverse of equation 2 in the Appendix.

<sup>c</sup>IE3A = inverse of equation 3 in the Appendix.

The error variance of prediction (EVP) is:

$$\text{var}(\hat{s} - s) = L(\text{LHS})^{-1}L', \quad (4)$$

where  $(\text{LHS})^{-1}$  = generalized inverse of the left hand side of the MME for model 2.

A complete example illustrating models 1 and 2 is presented in the Appendix.

#### Applications and Discussion

The models presented above were designed to account for additive and nonadditive, intra- and inter-locus, direct and maternal genetic effects in open crossbreeding mating schemes. Examples of such systems are: (a) new breed formation, (b) isolated or recurrent rounds of semen importation and (c) stratified crossbreeding schemes. In new breed formation, the final expected breed composition of the resulting breeds would be determined through the evaluation of animals using field data. For instance, this approach could be used in the formation of the Kamaduk cattle breed of India (Bhat, 1974). Isolated or recurrent rounds of semen importation intend to upgrade the performance of local stocks. The situations described by Wilkins et al. (1979) in Bolivia and by Madalena (1981) in Brazil for milk production in cattle fall into this category. This system also includes grading up, which would be the case when successive rounds of exotic semen belong to sires of only one breed, e.g., the Simmental breed in the United States. Stratified crossbreeding schemes involve at least two levels of production conditions. The top level would produce and evaluate the bulls to be used by the second level. Osman and Russell (1974) suggested a stratified crossbreeding system for dairy cattle in the Sudan. Finally, models 1 and 2 might be useful in systematic crossbreeding schemes when sires are mated to dams of several breed groups.

The parameters used in the present research to explain two-locus interactions are intended to measure any type of epistatic interaction included in the specified configurations. This includes recombination losses as defined by Dickerson (1973) and Kinghorn (1980). They do not account for linkage, however, because they are defined on the assumption of independence among loci.

Model 2 has several advantages. The first is its flexibility to be adapted to data sets of different sizes. The size and complexity of the equations can be reduced (or expanded) through

assumptions on the nonadditive parameters and on the additive and nonadditive covariance matrices. Second, it uses more information to estimate additive parameters than model 1, and to estimate and predict additive and nonadditive parameters than the MSMI model. Thus, if the nonadditive assumptions are true for a given data set, sire proofs from model 2 will be at least as precise as those from the MSMI model and model 1. Third, it permits the design of flexible crossbred mating plans. This flexibility refers to the emphasis given to the additive and to the nonadditive parts of a sire's proof, when choosing a sire. If a breeder's goal were to achieve long-term genetic improvement, then more emphasis should be given to the additive part of a sire's proof. On the other hand, if the goal were to maximize the expected immediate returns (e.g., commercial producers), then sires with high predicted total genetic value (i.e., additive plus nonadditive) should be chosen. Also, because bulls of different breed composition may have proofs of similar value, breeders will be able to choose them according to the breed composition they would like to have in their herds.

However, model 2 has disadvantages as well. First, given a specific form of model 2 for a data set, a minimum number and type of equations (i.e., matings) is required for comparisons between any pair of sires to be feasible. Empty cells in the design (i.e., mating types not performed) can cause problems of confounding and multicollinearity among the fixed parameters. This means that some sire comparisons will be biased. One way to approach this problem is to use the Ridge Regression Method (Hoerl and Kennard, 1970a,b), which in its simplest form consists of adding  $kI$ ,  $k > 0$ , to the fixed effects submatrix of the MME. The constant  $k$  will be chosen such that the mean square errors of the ridge estimators of the fixed effects, although biased, will be smaller than that of the unbiased ones. Second, it may be computationally costly. Hence, a compromise between accuracy of prediction and cost of computation will have to be reached in many instances.

The final form of model 2 for a data set will be determined by the two points raised above as well as the feasibility of estimating the necessary additive and nonadditive covariances. Estimation of these covariances is to be done before using model 2. The estimation of the necessary additive and nonadditive covariances

could be done by means of the usual variation component estimation procedures (e.g., Henderson's methods 2 and 3, Henderson, 1953; restricted maximum likelihood, Corbeil and Searle, 1976; minimum norm quadratic unbiased estimation, Searle, 1979) using a model similar to the one suggested by Elzo (1983). In addition, under the assumptions of independence of loci and constant covariances over time, additive covariances for crossbred groups can be calculated as a linear combination of the additive covariances of the breeds present in a crossbred group. Thus, only base breed additive covariances need to be estimated, which greatly reduces the task of estimating additive covariances. On the other hand, the number of nonadditive variances to be estimated will be determined not only by the number of base breeds involved, but also by the assumptions made with respect to the nonadditive effects. Thus, their number can be made as small as two per trait, under the assumptions mentioned previously. If later on, data allowing estimation of more covariances became available and(or) more subclasses are filled, model comparisons and reevaluation of sires could be made.

Interactions at more than two loci as well as grandmaternal additive and nonadditive effects were not included in models 1 and 2 due to the additional complications they would cause the prospects that field data would be inadequate for their estimation and(or) prediction.

Models 1 and 2 become sire-maternal grand-sire models for crossbreeding data if maternal effects are unimportant, and sire models if maternal grandsires are ignored.

#### Appendix

*Residual Variances.* The expression of the residual variances according to the male ancestors (s, mgs) identified in individual i is:

$$\sigma_e^2 \approx \sum_{j=1}^{14} w_j$$

when both s and mgs are identified, (1)

$$\sigma_e^2 \approx w_1 + w_2 + w_4 + w_6 + w_8 + w_{10} + w_{14}$$

when only s is identified, (2)

$$\sigma_e^2 \approx w_1 + w_3 + w_5 + w_6 + w_7 + w_8 + w_9 + w_{10} + w_{11} + w_{13} + w_{14} \quad (3)$$

when only mgs is identified, where the "≈" sign means that relationships through male ancestors only are accounted for and:

- $w_1$  = additive variance for direct effects in breed group i,
- $w_2$  =  $-1/4$  additive variance for direct effects in breed group s,
- $w_3$  =  $-1/16$  additive variance for direct effects in breed group mgs,
- $w_4$  =  $-1/16$  additive covariance for direct effects between the sire and the mgs of s,
- $w_5$  =  $-1/64$  additive covariance for direct effects between the sire and the mgs of mgs.
- $w_6$  = additive variance for maternal effects in breed group d,
- $w_7$  =  $-1/4$  additive variance for maternal effects in breed group mgs,
- $w_8$  =  $1/4$  additive covariance for maternal effects between the sire and the mgs of d,
- $w_9$  =  $-1/16$  additive covariance for maternal effects between the sire and the mgs of mgs,
- $w_{10}$  = additive covariance between direct and maternal effects in breed group d,
- $w_{11}$  =  $-1/4$  additive covariance between direct and maternal effects in breed group mgs,
- $w_{12}$  =  $-1/2$  additive covariance between direct and maternal effects of the sire and the mgs of i,
- $w_{13}$  =  $-1/16$  additive covariance between direct and maternal effects of the sire and the mgs of mgs,
- $w_{14}$  = sum of the variance of direct sire group × dam interaction effects, variance of sire × dam interaction effects, variance of the mgs groups × mgd interaction effects, variance of the mgs × mgd interaction effects and the variance of the random direct and maternal environmental effects.

*Expression of an Interaction Parameter in Models 1 and 2.* Assuming nonadditive effects at more than two loci are negligible, the complete expression of interaction parameters (fixed or random) in models 1 and 2 is:



$$\sum_{\substack{i,j \\ j>i}}^I t_{W_{ij}} \delta_{W_{ij}} + \sum_{\substack{i,j \\ j>i}}^I t_{B_{ij}} \delta_{B_{ij}} + \sum_{\substack{i,j,k \\ k>j>i}}^I t_{B_{ijk}} \delta_{B_{ijk}} + \sum_{\substack{i,j,k,l \\ l>k>j>i}}^I t_{B_{ijkl}} \delta_{B_{ijkl}} \tag{4}$$

where

- I = number of straightbreds in the data set,  $1 < i, j, k, l < I$ ,
- $t_{W_r}$  = probability of having alleles of breeds r at one random locus, where  $r = i, j$ ,
- $t_{B_s}$  = probability of having alleles of breeds s at two random loci, where  $s = i, j; i, j, k; i, j, k, l$ ,
- $\delta_{W_r}$  = sum of all one locus interaction effects due to alleles belonging to breeds r over all loci,
- $\delta_{B_s}$  = sum of all two-locus interaction effects due to alleles belonging to breeds s over all two locus combinations.

Criteria for simplifying expression (4) above include ignoring unimportant interactions and assuming equal interaction effects for some allelic combinations at one and(or) two loci.

*Numerical Example for Models 1 and 2.* Consider the case of two breeds, e.g., Simmental (S) and Hereford (H) as in the main text, and the evaluation of two unrelated bulls from the same generation (i.e., same genetic group), one .75S .25H and another 100% S. Suppose that each bull has 10 progeny from .5S .5H dams and 10 progeny from straightbred Hereford dams. Let the means for hypothetical trait be 20 and 21 for the progeny groups

resulting from the matings of the .75S .25H bull and the .5S .5H dams and the 100% H dams, respectively, and 23 and 24 for the corresponding progeny groups from the 100% S sire when mated to dams of the same two breed groups, respectively. All maternal granddams are 100% H; hence maternal granddam effects are confounded with the overall mean ( $\alpha$ ). Thus, assuming that only intra-locus interbreed interaction effects are important for this trait, model 2 will consider seven effects: Simmental breed ( $b_1^S$ ), Hereford breed ( $b_2^S$ ), Simmental-Hereford intra-locus fixed effect ( $\delta_1$ ), .5S .5H sire effect ( $a_1^S$ ), 100% H sire effect ( $a_2^S$ ), .5S .5H sire interaction effect ( $\delta_{21}$ ) and 100% H sire interaction effect ( $\delta_{22}$ ).

Let  $G = \begin{bmatrix} 5.25 & 0 \\ 0 & 9 \end{bmatrix}$  and  $D_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ .

Also, let the environmental variance ( $\sigma_e^2 = 10$ ) be the same for all four progeny groups. So, by formula (2) in the Appendix, the computational rule to calculate additive covariances for crossbred groups and the Simmental and Hereford additive covariance matrices given in the main text, R turns out to be

$R = \text{diag} [20.1875 \ 24.1 \ 20.5 \ 23.0]$ .

The MME for model 2 are shown in table 2. The vector of solutions from these MME is [22.54, 20.23, .25, .00, .06, -.94, .94]. The matrix L, needed to obtain the sire proofs for total genetic value for both bulls when mated to both types of dams, for model 2 ( $L_2$ ), is

$$L_2 = \begin{bmatrix} .75 & .25 & .5 & 1 & 0 & .5 & 0 \\ .75 & .25 & .75 & 1 & 0 & .75 & 0 \\ 1 & 0 & .5 & 0 & 1 & 0 & .5 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

TABLE 2. MIXED MODEL EQUATIONS FOR MODEL 2 (EXAMPLE)

14.3462 <sup>a</sup>	1.7068	10.9785	6.8272	9.2259	4.1916	6.7869	$b_1^S$	=	363.6294
	.5689	1.3972	2.2757	.0	1.3972	.0	$b_2^S$		49.0288
		9.1397	5.5881	6.7869	3.5725	5.5673	$\delta_1$		280.2873
			9.2934	.0	5.5888	.0	$a_1^S$		196.1153
				9.3370	.0	6.7869	$a_2^S$		216.5424
					4.0724	.0	$\delta_{21}$		119.8419
						6.0673	$\delta_{22}$		160.4454
Symmetric									

<sup>a</sup>All numbers were approximated to four digits.

So, the four sire proofs and their corresponding EVP [calculated using formulae (3) and (4) in the main text] are:

$$\begin{bmatrix} \hat{s}_{11} \\ \hat{s}_{12} \\ \hat{s}_{21} \\ \hat{s}_{22} \end{bmatrix} = \begin{bmatrix} 21.62 \\ 21.45 \\ 23.19 \\ 23.78 \end{bmatrix} \text{ and } \begin{bmatrix} \text{EVP}(\hat{s}_{11}) \\ \text{EVP}(\hat{s}_{12}) \\ \text{EVP}(\hat{s}_{21}) \\ \text{EVP}(\hat{s}_{22}) \end{bmatrix} = \begin{bmatrix} .15 \\ .17 \\ .19 \\ .21 \end{bmatrix}.$$

The first subscript of the  $s$  represent sire number and the second the dam group number.

The MME for model 1 are the same as those for model 2, except for the first two equations, where all numbers are different. Also,  $g_1^S, g_2^S$  are substituted for  $b_1^S, b_2^S$ , to indicate that records of progeny from the .75S .25H sire will contribute only to  $g_1^S$  and records of progeny from the 100% S sire only to  $g_2^S$ . The numbers to be substituted for rows 1 and 2 in the LHS for model 2 MME are: 9.1029, 0.0, 5.5888, 9.1029, 0.0, 5.5888, 0.0 and 9.2259, 6.7868, 0.0, 9.2259, 0.0 and 6.7868, respectively. The corresponding substitutions for the RHS of model 2 MME are 196.1153 and 216.5430. The vector of solutions from model 1 MME is: [21.96, 22.54, .25, .00, .06, -.94, .94]. The matrix  $L$  for model 1 ( $L_1$ ) is also similar to the one for model 2 ( $L_2$ ), but its first 2 columns are: 1, 1, 0, 0 and 0, 0, 1, 1, respectively. The remaining columns of  $L_1$  are the same as those from  $L_2$ . Thus, the vector of sire proofs and their corresponding EVP for model 1, are:

$$\begin{bmatrix} \hat{s}_{11} \\ \hat{s}_{12} \\ \hat{s}_{21} \\ \hat{s}_{22} \end{bmatrix} = \begin{bmatrix} 21.62 \\ 21.45 \\ 23.19 \\ 23.78 \end{bmatrix} \text{ and } \begin{bmatrix} \text{EVP}(\hat{s}_{11}) \\ \text{EVP}(\hat{s}_{12}) \\ \text{EVP}(\hat{s}_{21}) \\ \text{EVP}(\hat{s}_{22}) \end{bmatrix} = \begin{bmatrix} .15 \\ .17 \\ .19 \\ .21 \end{bmatrix}.$$

So, in this particular example sire proofs and their EVP are the same using models 1 and 2. Finally, as an illustration of the ability of these models to predict sire performance for matings not made (to help plan matings), the value of the 100% S sire's proof and its EVP, if this sire were mated to .4S .6H dams, will be calculated. Here,  $L = [1 \ 0 \ .6 \ 0 \ 1 \ 0 \ .6]$  for model 2. So,  $\hat{s}_{23} = 23.31$  and  $\text{EVP}(\hat{s}_{23}) = .14$ . If model 1 is used,  $L = [0 \ 1 \ .6 \ 0 \ 1 \ 0 \ .6]$  and  $\hat{s}_{23}$  and  $\text{EVP}(\hat{s}_{23})$  have the same value as for model 2.

#### Literature Cited

- Alenda, R., T. G. Martin, J. F. Lasley and M. R. Ellersieck. 1980a. Estimation of genetic and maternal effects in crossbred cattle of Angus, Charolais and Hereford parentage. I. Birth and weaning weights. *J. Anim. Sci.* 50:226.
- Alenda, R., T. G. Martin, J. F. Lasley and M. R. Ellersieck. 1980b. Estimation of genetic and maternal effects in crossbred cattle of Angus, Charolais and Hereford parentage. II. Postweaning growth, ribeye area and fat cover. *J. Anim. Sci.* 50:235.
- Bhat, P. N. 1974. Crossbreeding cattle in India. Results, aims and objectives of the All-India Coordinated Research Project on Dairy Cattle Breeding. *Proc. 1st World Congress on Genet. Appl. to Livestock Prod.* 2:237.
- Bhat, P. N., V. K. Taneja and R. C. Garg. 1978. Effects of crossbreeding on reproduction and production traits. *Indian J. Anim. Sci.* 48:71.
- Corbeil, R. R. and S. R. Searle. 1976. Restricted maximum likelihood (REML) estimation of variance components in the mixed model. *Technometrics* 18:31.
- Dickerson, G. 1973. Inbreeding and heterosis in animals. *Proc. of the Anim. Breeding and Genetics Symposium in Honor of Dr. J. L. Lush.* American Soc. of Anim. Sci. and Amer. Dairy Sci. Assoc., Champaign, IL p 54.
- Dillard, E. U., O. Rodriguez and O. W. Robison. 1980. Estimation of additive and non-additive direct and maternal genetic effects from crossbreeding beef cattle. *J. Anim. Sci.* 50:653.
- Donald, H. P., D. Gibson and W. S. Russell. 1977. Estimation of heterosis in crossbred dairy cattle. *Anim. Prod.* 25:193.
- Elzo, M. A. 1983. Multibreed sire evaluation within and across countries. Ph.D. Dissertation. Univ. of California, Davis.
- Franklin, I. R., R. H. Hayman and R. W. Hewetson. 1976. *Bos indicus* and *Bos taurus* crossbred dairy cattle in Australia. IV. Progeny testing and expected rate of genetic improvement. *Australian J. Agr. Res.* 27:309.
- Henderson, C. R. 1953. Estimation of variance and covariance components. *Biometrics* 9:226.
- Henderson, C. R. 1975a. Rapid method for computing the inverse of a relationship matrix. *J. Dairy Sci.* 58:1727.
- Henderson, C. R. 1975b. Inverse of a matrix relationship due to sires and maternal grand sires. *J. Dairy Sci.* 58:1917.
- Henderson, C. R. 1976a. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* 32:69.
- Henderson, C. R. 1976b. Multiple trait sire evaluation using the relationship matrix. *J. Dairy Sci.* 59:769.
- Henderson, C. R. 1976c. Inverse of a matrix of relationships due to sires and maternal grand sires in an inbred population. *J. Dairy Sci.* 59:1585.
- Hoerl, E. A. and R. W. Kennard. 1970a. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* 12:55.
- Hoerl, E. A. and R. W. Kennard. 1970b. Ridge regression: Application to nonorthogonal problems. *Technometrics* 12:69.
- Kinghorn, B. 1980. The expression of "Recombination Loss" in quantitative traits. *Z. Tierz. Züch-*

- tungsbiol. 97:138.
- Koger, M., F. M. Peacock, W. G. Kirk and J. R. Crockett. 1975. Heterosis effects on weaning performance of Brahman-Shorthorn calves. *J. Anim. Sci.* 40:826.
- Leroy, P., R. Hauser and A. Francois. 1978. L'évaluation (par la méthode B.L.U.P.) des taureaux Holstein-Friesian utilisés en race Pie-Noire de Herve. *Ann. Med. Vet.* 122:501.
- Madalena, F. E. 1981. Crossbreeding strategies for dairy cattle in Brazil. *Wld. Anim. Rev.* 38:23.
- Mahadevan, P., K. E. Wellington and K. L. Roache. 1970. An evaluation of Jamaica Hope bulls. *J. Agr. Res. (Camb.)* 74:473.
- Massey, M. E. and L. L. Benyshek. 1981. Interactions involving sires, breed of dam and age of dam for performance characteristics in Limousin cattle. *J. Anim. Sci.* 53:940.
- Osman, A. H. and W. S. Russell. 1974. Comparative performance of different grades of European-Zebu crossbred cattle at Ghurashi Dairy Farm, Sudan. *Trop. Agr.* 51:549.
- Parekh, H.K.B. and A. B. Pande. 1982. Genetic evaluation of exotic sires under different environment and their accuracy. *Proc. 2nd World Congress on Genet. Appl. to Livestock Prod.* VII:170.
- Parekh, H.K.B. and J. P. Sahu. 1978. Studies on some economic traits of Tharparkar X Holstein half-bred females. *Indian J. Anim. Sci.* 48:853.
- Parekh, H.K.B. and R. W. Touchberry. 1974. Comparative performance of purebreds and crossbreds in respect to reproductive traits: 1st parity records of four generations crossbreeding experiment. *Proc. 1st World Congress on Genet. Appl. to Livestock Prod.* 3:601.
- Peacock, F. M., M. Koger, T. A. Olson and J. R. Crockett. 1981. Additive genetic and heterosis effects in crosses among cattle breeds of British, European and Zebu origin. *J. Anim. Sci.* 52:1007.
- Peacock, F. M., M. Koger, Z. A. Palmer, W. Carpenter and T. A. Olson. 1982. Additive breed and heterosis effects for individual and maternal influences on feedlot gain and carcass traits of Angus, Brahman, Charolais and crossbred steers. *J. Anim. Sci.* 55:797.
- Quaas, R. L. 1975. From Mendel's laws to the A inverse. *Dept. Anim. Sci. Cornell Univ., Ithaca, NY (Mimeo.)* p 16.
- Quaas, R. L. and E. J. Pollak. 1980. Mixed model methodology for farm and ranch beef cattle testing programs. *J. Anim. Sci.* 51:1277.
- Rincon, E. J., E. C. Schermerhorn, R. E. McDowell and B. T. McDaniel. 1982. Estimation of genetic effects on milk yield and constituent traits in crossbred dairy cattle. *J. Dairy Sci.* 65:848.
- Robison, O. W., B. T. McDaniel and E. J. Rincon. 1981. Estimation of direct, maternal additive and heterotic effects from crossbreeding experiments in animals. *J. Anim. Sci.* 52:44.
- Searle, S. R. 1979. Notes on variance component estimation: A detailed account of maximum likelihood and kindred methodology. BU-673-M, Biometrics Unit, Cornell Univ., Ithaca, NY.
- Taneja, V. K. and P. N. Bhat. 1974. Estimation of additive and nonadditive effects in Sahiwal X Friesian crossbreds. *Proc. 1st World Congress on Genet. Appl. to Livestock Prod.* 3:611.
- Touchberry, R. W. 1970. A comparison of the general merits of purebred and crossbred dairy cattle resulting from twenty years (four generations) of crossbreeding. *Proc. 19th Annu. Natl. Breeders' Roundtable, Kansas City, MO* p 18.
- Wilkins, J. V., G. Pereyra, A. Ali and S. Ayola. 1979. Milk production in the tropical lowlands of Bolivia. *Wld. Anim. Rev.* 32:25.
- Willham, R. L. 1963. The covariance between relatives for characters composed of components contributed by related individuals. *Biometrics* 19:18.