

2005

# Unexpected estimates of variance components with a true model containing genetic competition effects

L. D. Van Vleck

*USDA-ARS, Roman L. Hruska U.S. Meat Animal Research Center*

J. P. Morota

*North Carolina State University*

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



Part of the [Genetics and Genomics Commons](#), and the [Meat Science Commons](#)

---

Van Vleck, L. D. and Morota, J. P., "Unexpected estimates of variance components with a true model containing genetic competition effects" (2005). *Faculty Papers and Publications in Animal Science*. 973.

<http://digitalcommons.unl.edu/animalscifacpub/973>

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.

# Unexpected estimates of variance components with a true model containing genetic competition effects<sup>1</sup>

L. D. Van Vleck\*<sup>2</sup> and J. P. Cassady†

\*ARS, USDA, Roman L. Hruska U.S. Meat Animal Research Center, Lincoln 68583-0908; and †North Carolina State University, Raleigh 27695-7627

**ABSTRACT:** Simulation of a model containing genetic competition effects was initiated to determine how well REML could untangle variances due to direct and competition genetic effects and pen effects. A two-generation data set was generated with six unrelated males that were each mated to five unrelated females to produce 300 progeny, from which 30 females (one per mating in previous generation) were mated to six unrelated males to produce 300 more progeny. Progeny were randomly assigned, six per pen, to 50 pens per generation. Parameters were  $V_g$ ,  $V_c$ ,  $C_{gc}$ ,  $V_p$ , and  $V_e$ , representing direct and competition genetic variance with covariance, and pen and residual variance. Eight statistical models were used to analyze each of 400 replicates of 16 sets of parameters. Both  $V_g$  and  $V_e$  were fixed at 16.0. Values of  $C_{gc}$  were -2.0, -1.0, 0.1, 1.0, and 2.0. Values of  $V_c$  were 1.0 and 4.0, and values of  $V_p$  were 0.1, 1.0, and 10.0. With the full model, average estimates resembled true parameters, except that  $V_p$  was consistently overestimated when small (0.1 and 1.0), which

in turn slightly changed other estimates. The most unexpected result was overestimation of  $V_p$  when  $V_c$  and  $C_{gc}$  were ignored. Overestimation depended on  $V_c$  and the number of competitors in common between records in a pen. Upward bias was somewhat greater when  $C_{gc}$  was positive than when it was negative. For example, with  $C_{gc} = 2$ ,  $V_c = 4$ , and  $V_p = 0.1$ , the mean estimate of  $V_p$  was 20.4 when  $C_{gc}$  and  $V_c$  were dropped from the model and 15.3 when  $C_{gc} = -2.0$ . When  $V_p$  was ignored, estimates of both  $C_{gc}$  and  $V_c$  increased in proportion with  $V_p$ . Also  $V_g$  increased more with greater  $V_p$ . Another unexpected result occurred when pen was considered fixed. Empirical sampling standard errors of estimates of  $C_{gc}$  and  $V_c$  were decreased generally by factors of 2 to 30. Based on these results, we conclude a high estimate of pen variance may indicate genetic competition effects are important, and ignoring either the pen or competition effects will bias estimates of other components.

Key Words: Genetic Parameters, Pen Effects, Restricted Maximum Likelihood, Simulation

©2005 American Society of Animal Science. All rights reserved.

J. Anim. Sci. 2005. 83:68–74

## Introduction

Competition among animals in the same pen or group may decrease performance. Federer (1955) briefly discussed competition effects specifically for plants and also more generally for animals. Griffing (1967) developed a theory for plants to account for direct and associative effects. Muir and Schinkel (2002) introduced the idea of predicting both direct genetic effects and competition (associative) genetic effects for animals. Muir (2004) reviewed previous analyses of competition effects with plants. He also contrasted whole pen

(group) selection with selection on an index of direct and competition genetic effects for both a simulation study and an experiment with quail. Muir (2004) also presented the mixed model equations for analyses of plant data with competition effects proportional to distances and for animals interacting in a pen.

The first goal of this simulation study was to determine whether REML procedures with relationships could untangle the covariance structure of direct and competition genetic variances (with covariance) and variance due to pen (contemporary management effects) from samples of relatively limited size and a relatively simple numerator relationship structure. The second goal was to determine the effect of dropping various effects from the statistical model, such as competition effects or pen effects. A third goal was to determine empirical sampling standard deviations for estimates of components of (co)variance and for corresponding genetic parameters. The fourth goal was to compare

<sup>1</sup>Published as Paper No. 14738, Journal Ser., Nebraska Agric. Res. Div., Univ. of Nebraska, Lincoln 68583-0908.

<sup>2</sup>Correspondence: A218 Animal Sciences (phone: 402-472-6010; fax: 402-472-6362; e-mail: lvanvleck@unlnotes.unl.edu).

Received September 3, 2004.

Accepted September 29, 2004.

estimates of other (co)variances with pens considered fixed or random.

## Materials and Methods

A model with direct genetic effects and the direct phenotype containing competition effects of pen mates is similar to the model for direct-maternal effects. In both models, the phenotype of one trait is embedded in the phenotypic measure of the other trait: maternal in direct and competition in direct. The differences are as follows: 1) several animals can contribute competition effects compared with only the dam contributing a maternal effect; 2) in the absence of inbreeding, the numerator relationship between a dam and her progeny with the record is one-half, whereas the relationship between the animal and its competitors can range from zero to one-half (if full sibs); 3) in the absence of inbreeding, relationships among competitors generally would also range from zero to one-half; and 4) phenotypic variance for the direct trait will depend on number of competitors, direct-competition genetic covariance, and relationships among pen mates.

These differences need to be considered for simulation of genetic values for direct and competition effects, as well as for statistical analyses of models with competition effects.

### Simulation Model

For this simulation, only one fixed factor,  $\mu$ , was considered; thus, the model for simulation of the direct phenotype ( $y_{ik}$ ) of animal  $i$  in pen  $k$  is:

$$y_{ik} = \mu + a_i + \Sigma c_j + p_k + e_i, \text{ where;}$$

$a_i$  is the direct genetic value of animal  $i$ ;  $\Sigma c_j$  is the sum of competition genetic values for penmates of animal  $i$  (in this simulation  $i$  had five penmates);  $p_k$  is an independent random pen effect; and  $e_i$  is an independent residual effect, which is actually the direct environmental effect associated with animal  $i$  plus the sum of competition environmental effects of the pen mates (all assumed to be uncorrelated). This model will be used later to explain some unexpected estimates of variance components when competition effects are ignored.

The simulation requires a vector,  $\mathbf{a}$ , for direct genetic values of animals with records; a vector,  $\mathbf{c}$ , with competition genetic values for animals with records; a vector,  $\mathbf{p}$ , for pen effects; and a vector,  $\mathbf{e}$ , of residual effects for each animal with a record. A simple design of two generations was chosen; for this design, the simulation and analysis is easier if  $\mathbf{a}$  and  $\mathbf{c}$  are augmented by genetic values for foundation sires and dams that do not have records.

What is needed for simulation of  $(\mathbf{a} \ \mathbf{c})'$  is that sampling be from a distribution with:

$$E\begin{pmatrix} \mathbf{a} \\ \mathbf{c} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ and } V\begin{pmatrix} \mathbf{a} \\ \mathbf{c} \end{pmatrix} = \begin{pmatrix} A\sigma_a^2 & A\sigma_{ac} \\ A\sigma_{ac} & A\sigma_c^2 \end{pmatrix} \text{ where}$$

$\mathbf{A}$  is the numerator relationship matrix augmented for foundation sires and dams,  $\sigma_a^2$  is the direct genetic variance,  $\sigma_c^2$  is the competition genetic variance, and  $\sigma_{ac}$  is the direct-competition genetic covariance.

The mating design for the first generation was to mate each of six unrelated males to five unrelated females. Each of the 30 matings produced 10 progeny. For the second generation, each of another six unrelated males was mated to five females with one female randomly chosen (the first one) from each litter of Generation 1. The first male was mated to females from the first five litters of generation one; the second male to females from the next five litters, etc. Thus, the relationship structure was always the same for each simulated data set, which allowed for simulation of  $\mathbf{a}$  and  $\mathbf{c}$  as follows. Total number of animals was  $6 + 6 + 30 + 300 + 300 = 642$ , with 600 having records. The numerator relationship matrix,  $\mathbf{A}$ , was calculated once for this design. For the simulation, let  $\mathbf{L}_A$  be the lower Cholesky factor of  $\mathbf{A}$  (see Van Vleck, 1994).

Let

$$\mathbf{V} = \begin{pmatrix} \sigma_a^2 & \sigma_{ac} \\ \sigma_{ac} & \sigma_c^2 \end{pmatrix} \text{ with}$$

$\mathbf{L}_V$  = lower Cholesky factor of  $\mathbf{V}$ . Let  $\mathbf{v}$  be a vector of order 1,284 ( $2 \times 642$ ) of randomly generated values from a pseudonormal distribution with a mean of zero and a variance of one. Then, calculate  $\begin{pmatrix} \mathbf{a} \\ \mathbf{c} \end{pmatrix} = (\mathbf{L}_A \otimes \mathbf{L}_A)\mathbf{v}$  with  $\otimes$  = the right direct product operator.

The Cholesky factor of  $\mathbf{A}$  needs to be calculated only once. The Cholesky factor of  $\mathbf{V}$  needs to be calculated only once for each set of parameters.

$$\begin{aligned} \text{Note that } V\begin{pmatrix} \mathbf{a} \\ \mathbf{c} \end{pmatrix} &= E[(\mathbf{L}_A \otimes \mathbf{L}_V)\mathbf{v}\mathbf{v}'(\mathbf{L}_A \otimes \mathbf{L}_V)'] \\ &= (\mathbf{L}_A \otimes \mathbf{L}_V) E[\mathbf{v}\mathbf{v}'](\mathbf{L}_A \otimes \mathbf{L}_V)' \\ &= (\mathbf{L}_A \otimes \mathbf{L}_V) (\mathbf{L}_A \otimes \mathbf{L}_V)' \\ &= (\mathbf{L}_A \mathbf{L}_A') \otimes (\mathbf{L}_V \mathbf{L}_V') \\ &= \mathbf{A} \otimes \mathbf{V} \text{ as needed} \end{aligned}$$

Uncorrelated pen and residual effects were also generated from a pseudonormal distribution,  $[N(0,1)]$ , with each pseudonormal value multiplied by the appropriate standard deviation, either  $\sigma_p$  or  $\sigma_e$ .

Table 1 shows 16 combinations of parameters used for simulation. The pattern was to fix  $\sigma_a^2$  and  $\sigma_e^2$  and to vary the other three (co)variances. Rather than attempt a simulation with zero variances or a zero covariance, a small value was used in an attempt to avoid convergence to boundaries of the parameter space. Instead of zero, 0.1 was used for small values of  $\sigma_{ac}$  and  $\sigma_p^2$ . No

**Table 1.** Combinations of parameters (1 to 16) used for simulation of direct, competition, pen, and residual effects<sup>a</sup>

Parameter set	$\sigma_a^2$	$\sigma_{ac}$	$\sigma_c^2$	$\sigma_p^2$	$\sigma_e^2$
1	16	-2.0	4	10.0	16
2	16	2.0	4	10.0	16
3	16	-2.0	4	1.0	16
4	16	2.0	4	1.0	16
5	16	-1.0	1	10.0	16
6	16	1.0	1	10.0	16
7	16	0.1	4	10.0	16
8	16	0.1	4	1.0	16
9	16	0.1	1	10.0	16
10	16	0.1	1	1.0	16
11	16	-1.0	1	1.0	16
12	16	1.0	1	1.0	16
13	16	-2.0	4	0.1	16
14	16	2.0	4	0.1	16
15	16	0.1	4	0.1	16
16	16	0.1	1	0.1	16

<sup>a</sup> $\sigma_a^2$  = direct genetic variance,  $\sigma_{ac}$  = direct-competition genetic covariance,  $\sigma_c^2$  = competition genetic variance,  $\sigma_p^2$  = variance due to pen effects, and  $\sigma_e^2$  = residual variance.

attempt was made to simulate a broad range of direct heritabilities. Thus, as with any simulation, the results may not apply to other combinations of parameters.

For each set of variance parameters, 400 replicates were generated and analyzed. Table 2 summarizes the eight statistical models used. The model in matrix notation corresponding to the full model used in calculating the mixed model equations, which in turn were used in calculating the restricted likelihood to obtain derivative-free REML estimates, was:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Za} + \mathbf{Wc} + \mathbf{Sp} + \mathbf{e}$$

where  $\mathbf{1}$  is a vector of 600 ones,  $\mu$  is a constant (100),  $\mathbf{Z}$  is a matrix of order  $600 \times 642$  augmented for the 42 sires and dams without records, which associates direct

**Table 2.** Assumptions for eight statistical models used for REML analyses of simulated data<sup>a,b</sup>

Analysis model	$\sigma_a^2$	$\sigma_{ac}$	$\sigma_c^2$	$\sigma_p^2$	$\sigma_e^2$
1	✓	✓	✓	✓	✓
2	✓	0	✓	✓	✓
3	✓	0	0	✓	✓
4	✓	✓	✓	0	✓
5	✓	0	0	0	✓
6	✓	✓	✓	Fix	✓
7	✓	0	✓	Fix	✓
8	✓	0	0	Fix	✓

<sup>a</sup>✓ indicates included, 0 indicates not included, and Fix indicates treated as a fixed factor.

<sup>b</sup> $\sigma_a^2$  = direct genetic variance,  $\sigma_{ac}$  = direct-competition genetic covariance,  $\sigma_c^2$  = competition genetic variance,  $\sigma_p^2$  = variance due to pen effects, and  $\sigma_e^2$  = residual variance.

genetic effects with records (will have a single 1 in column  $i$  of the  $i$ th row corresponding to a record of animal  $i$ ),  $\mathbf{W}$  is an augmented matrix of order  $600 \times 642$ , which associates competition genetic effects with records (will have five ones in the  $i$ th row corresponding to columns of pen mates associated with the record of animal  $i$ ),  $\mathbf{S}$  is a matrix of order  $600 \times 100$ , which associates records to pens. Within each of the two generations, six progeny were allocated randomly to a pen.

With  $\mathbf{G}_0 = \begin{pmatrix} \sigma_a^2 & \sigma_{ac} \\ \sigma_{ac} & \sigma_c^2 \end{pmatrix}$  and  $\sigma_e^2 \mathbf{G}_0^{-1} = \begin{pmatrix} g^{aa} & g^{ac} \\ g^{ac} & g^{cc} \end{pmatrix}$  and with  $\lambda = \sigma_e^2/\sigma_p^2$ , the mixed model equations, multiplied by  $\sigma_e^2$ , are

$$\begin{pmatrix} \mathbf{X}\mathbf{X} & \mathbf{X}\mathbf{Z} & \mathbf{X}\mathbf{W} & \mathbf{X}\mathbf{S} \\ \mathbf{Z}\mathbf{X} & \mathbf{Z}\mathbf{Z} + \mathbf{A}^{-1}g^{aa} & \mathbf{Z}\mathbf{W} + \mathbf{A}^{-1}g^{ac} & \mathbf{Z}\mathbf{S} \\ \mathbf{W}\mathbf{X} & \mathbf{W}\mathbf{Z} + \mathbf{A}^{-1}g^{ac} & \mathbf{W}\mathbf{W} + \mathbf{A}^{-1}g^{cc} & \mathbf{W}\mathbf{S} \\ \mathbf{S}\mathbf{X} & \mathbf{S}\mathbf{Z} & \mathbf{S}\mathbf{W} & \mathbf{S}\mathbf{S} + \mathbf{I}\lambda \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{c}} \\ \hat{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}\mathbf{y} \\ \mathbf{Z}\mathbf{y} \\ \mathbf{W}\mathbf{y} \\ \mathbf{S}\mathbf{y} \end{pmatrix}$$

These equations are also the basis for derivative-free REML. To obtain natural logarithms of the likelihoods given the data and to obtain estimates for each replicate to accommodate the statistical models shown in Table 2, equations were modified appropriately: 1) when the direct-genetic competition correlation was dropped,  $\sigma_{ac}$  was fixed at zero; 2) when the competition effect was dropped ( $\sigma_c^2 = 0$  and  $\sigma_{ac} = 0$ ), parts with  $\mathbf{W}$  were dropped out; 3) when the pen effect was dropped ( $\sigma_p^2 = 0$ ), parts with  $\mathbf{S}$  were dropped out; 4) when both competition and pen effects were dropped, both 2) and 3) applied; and 5) when statistical models 6 to 8 were used, pen effects were treated as fixed ( $\lambda = 0$ ).

Derivative-free REML estimates were obtained for each of the eight statistical models with the MTDFREML programs (Boldman et al., 1995) modified to include competition effects (Van Vleck and Cassady, 2004). After REML estimates were obtained for each of the 400 replicates for a parameter set, means and empirical standard deviations were calculated for estimates of parameters included in the statistical model used, including the genetic correlation, direct and competition heritabilities, and fraction of variance due to pen effects. For calculation of the latter three parameters, phenotypic variance was calculated with estimates of variance components as  $\hat{\sigma}_a^2 + \hat{\sigma}_p^2 + \hat{\sigma}_e^2 + 5\hat{\sigma}_c^2$ , which ignores relationships among animals in a pen that may change from pen to pen and that generally would have little effect on the phenotypic variance.

## Results and Discussion

Means of estimates of variances and covariance with eight statistical models are given in Tables 3 through 6 for Parameter Sets 1 to 16. Rather than discuss each table separately, an attempt will be made to demonstrate patterns. Patterns may be similar for several tables, but all tables will be shown here. Estimates of

**Table 3.** True parameters and means of estimates from 400 replications with eight statistical models for Parameter Sets 1 to 4<sup>a</sup>

Analysis	$\sigma_a^2$	$\sigma_{ac}$	$\sigma_c^2$	$\sigma_p^2$	$\sigma_e^2$
True Set 1	16.0	-2.0	4.0	10.0	16.0
1	15.9	-2.1	4.0	10.0	16.0
2	17.6	-	4.4	6.5	17.0
3	22.6	-	-	25.5	16.9
4	17.4	0.1	6.5	-	16.3
5	18.7	-	-	-	43.2
6	15.3	-2.3	4.2	Fix	16.0
7	19.5	-	4.6	Fix	16.0
8	24.1	-	-	Fix	16.0
True Set 2	16.0	2.0	4.0	10.0	16.0
1	15.4	1.7	3.9	10.6	16.1
2	13.9	-	3.5	13.8	15.4
3	16.2	-	-	30.7	16.0
4	18.2	4.4	6.5	-	16.2
5	19.0	-	-	-	48.2
6	16.0	2.2	4.3	Fix	16.1
7	11.2	-	4.8	Fix	16.1
8	16.0	-	-	Fix	16.1
True Set 3	16.0	-2.0	4.0	1.0	16.0
1	16.3	-2.5	3.8	2.3	15.4
2	18.0	-	3.5	0.6	17.1
3	22.7	-	-	16.5	16.9
4	16.4	-2.0	4.3	-	15.6
5	19.2	-	-	-	34.2
6	16.0	-2.3	4.2	Fix	15.5
7	20.2	-	4.6	Fix	15.5
8	24.8	-	-	Fix	15.5
True Set 4	16.0	2.0	4.0	1.0	16.0
1	15.4	1.5	3.5	3.0	16.1
2	14.3	-	3.4	4.8	15.2
3	16.2	-	-	21.1	16.1
4	16.1	2.3	4.2	-	16.2
5	19.0	-	-	-	34.1
6	15.8	2.1	4.4	Fix	16.1
7	11.1	-	4.8	Fix	16.1
8	15.9	-	-	Fix	16.1

<sup>a</sup> $\sigma_a^2$  = direct genetic variance,  $\sigma_{ac}$  = direct-competition genetic covariance,  $\sigma_c^2$  = competition genetic variance,  $\sigma_p^2$  = variance due to pen effects, and  $\sigma_e^2$  = residual variance.

genetic correlations and heritabilities will not be discussed, although patterns for estimates of those parameters will follow patterns for estimates of the (co)variances.

With the full model, the REML algorithm with relationships for the two generations generally did well for these parameter sets in partitioning total variance into variances due to direct and competition genetic effects, their covariance, and pen and residual variances. The full model, with pen effects treated as fixed, also did well in partitioning the other variances and covariances. Exceptions were made for the full model when the true pen variance was small (1.0) and very small (0.1). In those cases, means of estimates of pen variance were consistently larger than the true variance. The magnitude of overestimation seemed to depend on the magnitude and sign of  $\sigma_{ac}$  and the magnitude of  $\sigma_c^2$  with greater overestimation with larger  $\sigma_c^2$ , which corresponded with

**Table 4.** True parameters and means of estimates from 400 replications with eight statistical models for Parameter Sets 5 to 8<sup>a</sup>

Analysis	$\sigma_a^2$	$\sigma_{ac}$	$\sigma_c^2$	$\sigma_p^2$	$\sigma_e^2$
True Set 5	16.0	-1.0	1.0	10.0	16.0
1	15.8	-1.0	1.0	9.9	16.1
2	16.9	-	1.1	8.5	16.6
3	18.0	-	-	13.1	16.7
4	18.6	1.7	3.2	-	16.6
5	16.4	-	-	-	30.4
6	15.7	-1.1	1.1	Fix	16.1
7	17.8	-	1.2	Fix	16.1
8	18.9	-	-	Fix	16.1
True Set 6	16.0	1.0	1.0	10.0	16.0
1	16.0	1.1	1.1	9.8	16.0
2	14.9	-	0.9	11.5	15.6
3	15.5	-	-	15.8	15.8
4	19.9	4.0	3.1	-	16.1
5	17.3	-	-	-	30.0
6	16.0	1.1	1.1	Fix	16.0
7	13.8	-	1.2	Fix	16.0
8	15.0	-	-	Fix	16.0
True Set 7	16.0	0.1	4.0	10.0	16.0
1	15.8	0.1	4.0	9.9	16.1
2	15.8	-	4.0	10.0	16.1
3	19.3	-	-	28.1	16.5
4	17.9	2.4	6.5	-	16.2
5	18.9	-	-	-	43.1
6	15.5	0.1	4.3	Fix	16.1
7	15.0	-	4.6	Fix	16.1
8	19.6	-	-	Fix	16.1
True Set 8	16.0	0.1	4.0	1.0	16.0
1	15.8	-0.2	3.6	2.8	16.0
2	16.0	-	3.6	2.6	16.1
3	19.3	-	-	18.9	16.4
4	16.3	0.5	4.2	-	16.1
5	19.2	-	-	-	34.0
6	15.5	0.1	4.4	Fix	16.0
7	15.1	-	4.6	Fix	16.0
8	19.7	-	-	Fix	16.0

<sup>a</sup> $\sigma_a^2$  = direct genetic variance,  $\sigma_{ac}$  = direct-competition genetic covariance,  $\sigma_c^2$  = competition genetic variance,  $\sigma_p^2$  = variance due to pen effects, and  $\sigma_e^2$  = residual variance.

some decrease in estimates of  $\sigma_c^2$ . The overestimation of  $\sigma_p^2$  may be due partially to negative estimates not being allowed for REML, which will bias upward estimates of variances at or near zero.

When  $\sigma_{ac}$  was fixed at zero, estimates of  $\sigma_c^2$  and  $\sigma_a^2$  tended to increase if the true covariance was negative and decrease if the true covariance was positive.

The most unexpected result (to the authors) was the large overestimation of pen variance when competition effects were dropped from the model. Overestimation was greater when the true direct-competition covariance was positive. When the true direct-competition covariance was negative,  $\sigma_a^2$  was also overestimated, but not nearly as much as  $\sigma_p^2$ . Examination of the sire model may help to explain the overestimation of  $\sigma_p^2$ , although the analogy is not perfect. With the intraclass correlation model, the assumption is usually that the



**Table 5.** True parameters and means of estimates from 400 replications with eight statistical models for Parameter Sets 9 to 12<sup>a</sup>

Analysis	$\sigma_a^2$	$\sigma_{ac}$	$\sigma_c^2$	$\sigma_p^2$	$\sigma_e^2$
True Set 9	16.0	0.1	1.0	10.0	16.0
1	16.1	0.1	1.1	9.6	16.0
2	16.0	–	1.0	9.9	16.0
3	16.8	–	–	14.4	16.1
4	19.3	2.8	3.2	–	16.2
5	17.1	–	–	–	29.9
6	16.1	0.1	1.1	Fix	16.0
7	15.8	–	1.2	Fix	16.0
8	16.9	–	–	Fix	16.0
True Set 10	16.0	0.1	1.0	1.0	16.0
1	16.1	–0.1	0.9	1.4	15.8
2	16.1	–	0.9	1.4	15.8
3	16.9	–	–	5.3	16.1
4	16.4	0.4	1.2	–	16.0
5	17.1	–	–	–	21.1
6	16.1	0.1	1.1	Fix	15.9
7	15.8	–	1.2	Fix	15.9
8	17.0	–	–	Fix	15.9
True Set 11	16.0	–0.1	1.0	1.0	16.0
1	15.8	–1.1	0.9	1.4	16.1
2	16.3	–	0.8	0.7	17.0
3	17.1	–	–	3.9	17.3
4	15.8	–0.6	1.1	–	16.6
5	16.2	–	–	–	16.3
6	15.3	–1.1	1.1	Fix	16.3
7	17.4	–	1.2	Fix	16.3
8	18.6	–	–	Fix	16.3
True Set 12	16.0	1.0	1.0	1.0	16.0
1	15.5	0.8	0.9	1.7	16.1
2	15.1	–	0.9	2.4	15.5
3	15.5	–	–	6.4	15.9
4	16.0	1.3	1.2	–	16.3
5	16.5	–	–	–	21.5
6	15.9	1.1	1.1	Fix	16.1
7	13.6	–	1.2	Fix	16.1
8	14.8	–	–	Fix	16.1

<sup>a</sup> $\sigma_a^2$  = direct genetic variance,  $\sigma_{ac}$  = direct-competition genetic covariance,  $\sigma_c^2$  = competition genetic variance,  $\sigma_p^2$  = variance due to pen effects, and  $\sigma_e^2$  = residual variance.

variance component for the class effect is equivalent to the covariance between any pair of records in a level of the class (e.g., a record of one progeny with a record of another progeny of the same sire). With pens being the class of effects, pen variance would be the same as the covariance between records of any pair of animals in the same pen. If the relationship matrix accounts for covariances due to direct genetic effects, then competition effects, as well as the pen effect itself, will be left in the records of pen mates.

For this situation with six penmates (five competitors for each animal), let  $y_1$  and  $y_2$  be a representative pair of records from the same pen after adjustment for fixed effects:

$$y_1 = a_1 + c_2 + c_3 + c_4 + c_5 + c_6 + p + e_1$$

$$y_2 = a_2 + c_1 + c_3 + c_4 + c_5 + c_6 + p + e_2$$

**Table 6.** True parameters and means of estimates from 400 replications with eight statistical models for Parameter Sets 13 to 16

Analysis	$\sigma_a^2$	$\sigma_{ac}$	$\sigma_c^2$	$\sigma_p^2$	$\sigma_e^2$
True Set 13	16.0	–2.0	4.0	0.1	16.0
1	16.5	–2.5	3.6	2.3	15.8
2	18.0	–	3.4	4.3	15.0
3	22.7	–	–	20.4	15.9
4	16.6	–2.1	4.1	–	15.9
5	19.0	–	–	–	33.4
6	16.0	–2.3	4.3	Fix	15.9
7	20.3	–	4.8	Fix	15.9
8	24.9	–	–	Fix	15.9
True Set 14	16.0	2.0	1.0	1.0	16.0
1	15.5	1.4	0.9	1.7	16.1
2	14.5	–	0.9	2.4	15.5
3	16.4	–	–	6.4	15.9
4	16.0	2.0	1.2	–	16.3
5	18.8	–	–	–	21.5
6	16.2	2.2	1.1	Fix	16.1
7	11.4	–	1.2	Fix	16.1
8	16.2	–	–	Fix	16.1
True Set 15	16.0	0.1	4.0	0.1	16.0
1	15.7	–0.5	3.5	2.4	15.7
2	16.2	–	3.5	1.9	16.0
3	19.5	–	–	18.1	16.4
4	16.2	0.1	4.1	–	15.8
5	18.9	–	–	–	33.5
6	16.0	0.1	4.4	Fix	15.8
7	15.6	–	4.6	Fix	15.8
8	20.2	–	–	Fix	15.8
True Set 16	16.0	0.1	1.0	0.1	16.0
1	16.2	–0.3	0.9	1.1	15.6
2	16.3	–	0.8	0.8	15.7
3	17.0	–	–	4.4	16.1
4	16.4	0.1	1.0	–	15.8
5	17.0	–	–	–	20.3
6	16.4	0.1	1.1	Fix	15.8
7	16.1	–	1.2	Fix	15.8
8	17.2	–	–	Fix	15.8

<sup>a</sup> $\sigma_a^2$  = direct genetic variance,  $\sigma_{ac}$  = direct-competition genetic covariance,  $\sigma_c^2$  = competition genetic variance,  $\sigma_p^2$  = variance due to pen effects, and  $\sigma_e^2$  = residual variance.

If competition effects are ignored and the animals are unrelated:

$$\text{COV}(y_1, y_2) = 4 \sigma_c^2 + \sigma_p^2 + 2 \sigma_{ac}$$

Thus the “expected” bias would be  $4 \sigma_c^2 + 2 \sigma_{ac}$ . If this equation is used to model the expected estimate of  $\sigma_p^2$  when  $\sigma_c^2$  and  $\sigma_{ac}$  are ignored, general agreement can be seen with means of the estimates. For example, for Parameter Sets 1 to 4,

1)  $4(4) + 10 + 2(-2) = 22$  vs. the mean estimate of 25.5

2)  $4(4) + 10 + 2(2) = 30$  vs. the mean estimate of 30.6

3)  $4(4) + 1 + 2(-2) = 13$  vs. the mean estimate of 16.5

4)  $4(4) + 1 + 2(2) = 21$  vs. the mean estimate of 21.1

The analogy may become more tenuous when only  $\sigma_{ac}$  is ignored. If both  $\sigma_a^2$  and  $\sigma_c^2$  are accounted for:

$$\text{COV}(y_1, y_2) = \sigma_p^2 + 2\sigma_{ac}$$

Now, the bias will be  $2\sigma_{ac}$ . Comparisons with means for Parameter Sets 1, 2, 3, 4, 13, and 14 are as follows:

- 1)  $10 + 2(-2) = 6.0$  vs. the mean estimate of 6.5
- 3)  $1 + 2(-2) = -3.0$  vs. the mean estimate of 0.6
- 13)  $0.1 + 2(-2) = -3.9$  vs. the mean estimate of 0.5
- 2)  $10 + 2(2) = 14.0$  vs. the mean estimate of 13.8
- 4)  $1 + 2(2) = 5.0$  vs. the mean estimate of 3.4
- 14)  $0.1 + 2(2) = 4.1$  vs. the mean estimate of 4.3

The covariance between records of pen members seems to explain much of the bias when the genetic covariance is ignored. Parameter Sets 3 and 13 lead to a negative expectation of pen variance, which is out of the parameter space for REML, and in those two cases, the estimates of pen variance were small. In the other cases, agreement is quite good between the estimate of pen variance and the theoretical covariance between a pair of records in the pen after adjustment for competition genetic values. Although the true situation is more complex, the simple expectations between records of pairs of animals in the same pen do explain most of the bias.

If pen effects were ignored and true pen variance was relatively large ( $\sigma_p^2 = 10$ ), the estimate of residual variance was not affected much, but the other three components were greatly inflated except for Model 5, which also ignored  $\sigma_c^2$  and  $\sigma_{ac}$ . With the true covariance being negative or near zero, estimates became positive or were substantially greater than zero when pen effects were ignored.

If both pen and competition effects were ignored, estimates of direct genetic variance increased with the increase associated more with the magnitude of the competition variance than with the magnitude of the pen variance. Most of the ignored variance went to increase estimates of residual variance by approximately  $4\sigma_c^2 + \sigma_p^2$ .

A few minor surprises showed up when  $\sigma_{ac}$  or both  $\sigma_{ac}$  and  $\sigma_c^2$  were dropped from the model with pens considered to be fixed. In both these cases, estimates of residual variance were more similar to the true residual variance than when pens were considered to be random. With  $\sigma_{ac}$  ignored in the statistical analysis and having a true negative value, estimates of direct and competition genetic variances increased and the increase was more

**Table 7.** Empirical standard deviations of estimates of direct-competition genetic covariance,  $\sigma_{ac}$ , and competition genetic variance,  $\sigma_c^2$ , with pen random or pen fixed from 400 replicates for each of 16 parameter sets

Parameter set	Pen random		Pen fixed	
	$\sigma_{ac}$	$\sigma_c^2$	$\sigma_{ac}$	$\sigma_c^2$
1	2.1	1.7	0.4	0.7
2	2.3	1.9	0.4	0.6
3	1.7	1.1	0.3	0.7
4	1.9	1.3	0.4	0.6
5	1.4	0.7	0.2	0.1
6	1.6	0.8	0.2	0.2
7	2.3	1.9	0.03	0.8
8	1.8	1.2	0.03	0.8
9	1.5	0.8	0.03	0.2
10	1.0	0.5	0.03	0.2
11	1.1	0.5	0.2	0.1
12	1.0	0.5	0.2	0.2
13	1.7	1.1	0.4	0.8
14	2.0	1.2	0.4	0.6
15	1.7	1.2	0.03	0.9
16	1.0	0.4	0.03	0.2

when pen effects were considered fixed effects than when pens were considered to be random effects. With a positive true genetic covariance, estimates of direct genetic variance decreased and the decrease was more than when pens were considered as fixed effects. Increases in the estimates of variance of competition genetic effects were slight.

When pens were considered to be fixed effects and competition genetic effects were dropped from the model (no variance or covariance in model), with a true negative genetic covariance, the estimate of direct genetic variance was inflated more than when pens were considered as random effects. With a positive true covariance, estimates of direct genetic variance were similar whether pens were considered random or fixed.

A next step would be to compare predictions of breeding values for statistical analyses that account for or ignore competition effects when competition effects are in the simulation model. Such a study, while limited to parameters used in the simulation, would indicate the importance of considering competition effects in selection for both direct and competition breeding values.

The results from this study are conditioned on the model used for the simulation. Other more complicated models for competitive interactions among animals, as well as methods of analysis, may lead to different conclusions.

The second major unexpected result came from a comparison of the empirical standard deviations (which would correspond to standard errors for estimates with a single data set) for full models with pens as fixed or random effects. Standard deviations for estimates of genetic and residual variances were similar whether pens were considered fixed or random. For estimates of genetic variance, standard deviations ranged from

4.8 to 6.1 with pens random and from 3.9 to 6.4 with pens fixed. For estimates of residual variance, standard deviations ranged from 2.8 to 3.8 with pens as random and also with pens as fixed effects.

The surprise was that standard deviations for estimates of the competition genetic variance and especially for estimates of genetic covariance were smaller when pens were considered to be fixed rather than random effects. For all sets of parameters as shown in Table 7, empirical standard deviations were smaller when pens were treated as fixed effects.

For estimates of genetic covariance, standard deviations were from 5 to 63 times smaller when pens were considered fixed. Standard deviations were similar when pens were considered random, no matter the size of the true genetic covariance (range of 1.0 to 2.3), but with pens considered fixed, standard deviations scaled more to the magnitude of the true covariance (0.03 for  $\sigma_{ac} = 0.1$ ; approximately 0.4 for  $\sigma_{ac} = -2.0$  or 2.0; and approximately 0.2 for  $\sigma_{ac} = -1.0$  or 1.0).

For estimates of competition genetic variance, the disparity in empirical standard deviations was not as great as with estimates of the genetic covariance. The standard deviations of estimates of competition genetic variance ranged from 0.4 to 1.9, with pens considered to be random effects, and from 0.1 to 0.9 with pens considered to be fixed.

Larger standard deviations with pens considered to be random may reflect the difficulty the REML algorithm has, even with the required relationship matrix, of partitioning the pen and competition variances and the direct-competition genetic covariance. That difficulty seems to be much less when pens are considered fixed effects. Similar standard deviations for direct genetic and residual variances indicate that those components of variance are partitioned similarly, whether pens are considered fixed or random.

### Implications

Variance due to direct and competition genetic effects and pen effects can be partitioned. When effects were

dropped from the model, changes in estimates for components left in the model were generally small, except when competition effects were ignored, in which case, estimates of pen variance increased greatly because of the covariance between records of pen mates due to competition effects. Thus, a large variance due to pen effects from an analysis not including competition effects may indicate that competition effects should be included in the analysis. Not including pen effects in the model may bias estimation of direct-competition genetic covariance needed to calculate indexes of overall genetic value and expected responses due to selection. A result with possible implications for other models is that treating pen effects as fixed greatly decreased standard errors of estimates of genetic covariance between the direct and embedded competition effects and estimates of genetic variance for competition effects.

### Literature Cited

- Boldman, K. G., L. A. Kriese, L. D. Van Vleck, C. P. Van Tassell, and S. D. Kachman. 1995. A manual for use of MTDFREML. A set of programs to obtain estimates of variance and covariances [Draft]. USDA-ARS, Clay Center, NE.
- Federer, W. T. 1955. Chapter 3 in *Experimental Design: Theory and Application*. The MacMillan Co., New York.
- Griffing, B. 1967. Selection in reference to biological groups. I. Individual and group selection applied to populations of unordered groups. *Aust. J. Biol. Sci.* 10:127–139.
- Muir, W. M. 2004. Incorporation of competitive effects in breeding programs. *Genetics* (In press).
- Muir, W. M., and A. Schinckel. 2002. Incorporation of competitive effects in breeding programs to improve productivity and animal well being. CD-ROM Communication No. 14–07 in *Proc. 7th World Cong. Genet. Appl. Livest. Prod.*, Montpellier France.
- Van Vleck, L. D. 1994. Algorithms for simulation of animal models with multiple traits and with maternal and non-additive genetic effects. *Braz. J. Genet.* 17:53–57.
- Van Vleck, L. D., and J. P. Cassady. 2004. Modification of MTDFREML to estimate variance due to genetic competition effects. *J. Anim. Sci.* 82(Suppl.2):38. (Abstr).