

Influence of Inbreeding Depression on Genetic (Co)Variance and Sire-by-Year Interaction Variance Estimates for Weaning Weight Direct-Maternal Genetic Evaluation¹

C. Lee² and E. J. Pollak

Department of Animal Science, Cornell University, Ithaca, NY 14853, USA

ABSTRACT: This study examined the effects of ignoring inbreeding depression on (co)variance components for weaning weight through the use of Monte Carlo simulation.

Weaning weight is of particular interest as a trait for which additive direct and maternal genetic components exist and there then is the potential for a direct-maternal genetic covariance. Ignoring inbreeding depression in the analytical model (.8 kg reduction of phenotypic value per

1% inbreeding) led to biased estimates of all genetic (co) variance components, all estimates being larger than the true values of the parameters. In particular, a negative bias in the direct-maternal genetic covariance was observed in analyses that ignored inbreeding depression. A small spurious sire-by-year interaction variance was also observed.

(Key Words: Variance Components, Genetic Correlation, Simulation)

INTRODUCTION

Inbreeding occurs over generations as individuals in a population of finite size inevitably become more related. For many characteristics, increased level of inbreeding is associated with decrease in performance (inbreeding depression). Levels of inbreeding in populations of cattle have been evaluated as well as levels of inbreeding depression for performance (Bowman et al., 1978; Hudson and Van Vleck, 1984; Short and Lawlor, 1992; Ward, 1993). In most genetic evaluations of beef cattle, inbreeding depression has been ignored, both when estimating the parameters used in evaluations and when estimating breeding values. In beef cattle, recent studies (Robinson, 1994; Mallinckrodt et al., 1995; Lee and Pollak, 1996c) have examined the genetic correlation between additive direct and maternal effects for weaning weight (WW). Biased estimates of the correlation resulted when certain effects were ignored in the analytical model. Ignoring sire-by-year interaction (SY) effects in the model

used to estimate (co)variance components for WW in beef cattle has been identified as a potential source of bias in genetic parameter estimation (Robinson, 1994). In particular, ignoring this effect can lead to a spurious negative genetic correlation between additive direct and maternal effects when, in fact, one is absent. In an analysis of Simmental data, a small SY variance (3% of the total variance) was detected that, when ignored, changed the estimate of the correlation from $-.14$ to $-.29$ (Lee and Pollak, 1996c).

The biological reason for existence of SY is not clear. The objective of this study was to determine the effect of inbreeding depression for WW performance on the estimation of the direct and maternal genetic (co)variances and the SY variance.

MATERIALS AND METHODS

Simulated data

Monte Carlo simulation was used to investigate the impact of ignoring inbreeding depression on (co)variance components, focusing on SY variance and the direct-maternal genetic covariance. Weaning weight was simulated with inbreeding depression but not with SY.

Records were simulated on a yearly basis. Each year, the calf crop was produced by mating animals selected on their contemporary group-adjusted observations. Simulated WW observations included contemporary group fixed

¹ This research was conducted using the resources of the Cornell Theory Center, which receives major funding from the Natl. Science Foundation (NSF) and New York State. Additional funding comes from the Advanced Research Projects Agency (ARPA), the Natl. Center for Research Resources at the Natl. Institutes of Health (NIH), and IBM Corporation.

² Address reprint requests to C. Lee, Institute of Environment and Life Science, Hallym Univ. Chuncheon, Kangwon-Do 200-702, Korea.

Received October 23, 1996; Accepted February 12, 1997

effects, a regression on level of inbreeding, additive direct and maternal genetic random effects, permanent environmental random effects, and residuals. Mendelian sampling for direct and maternal genetic effects was simulated with a variance adjusted for inbreeding. A random number generator based on the Box-Muller method was used to generate random Gaussian deviates (Press et al., 1992).

Ten herds were simulated with 100 dams per herd per year for 25 years, for a total of 25,000 records. There were 1,244 base animals, 1,200 dams and 44 sires.

Connections across herds and years were obtained by use of AI sires. For each calf crop, three AI bulls and two herd bulls were used per herd. Six of the AI bulls were kept from the previous year, and four new bulls were selected each year from bulls born in the previous calf population based on the largest phenotypic deviation (adjusted for contemporary group) across all herds. Selection for within-herd bulls occurred after selection of AI bulls. Any dam over 20 years old was culled.

The population included 1.4% sires, 22.1% dams, and 76.5% nonparents. Average numbers of progeny per parent were 141.5, 38.5, and 4.3 for AI sires, within-herd sires, and dams, respectively. A detailed description of the simulation and population structures is in Lee and Pollak (1996a).

Inbreeding depression.

Weaning weight performance data were simulated with inbreeding depression. Performance was reduced by .8 kg per 1% inbreeding for additive direct genetic effects but not for additive maternal genetic effects. The population included 55% inbred animals with the average inbreeding coefficient of .04. There were 25 replicates simulated.

Models

For (co)variance components estimation, a mixed model was used:

$$Y = X\beta + W\rho + Z_d u_d + Z_m u_m + Z_p u_p + Z_{sy} u_{sy} + e,$$

where Y is a vector of WW observations, β is a vector of contemporary group fixed effects, ρ is covariate for inbreeding depression effects, u_d is an unknown random vector of additive direct genetic effects, u_m is an unknown random vector of additive maternal genetic effects, u_p is an unknown random vector of permanent environmental effects, u_{sy} is an unknown random vector of SY effects, and e is an unknown random vector of residuals. The X , Z_d , Z_m , Z_p , and Z_{sy} are known incidence matrices relating observations to β , u_d , u_m , u_p , and u_{sy} , respectively. The W

is known vector including inbreeding of calves. The first and second moments of random components in the model are $[u_d' u_m' u_p' u_{sy}' e' y'] \sim N [0' 0' 0' 0' 0' (X\beta + W\rho)']$, Σ , where

$$\Sigma = \begin{bmatrix} A\sigma_d^2 & A\sigma_{dm} & 0 & 0 & 0 & AZ_d'\sigma_d^2 \\ A\sigma_{dm} & A\sigma_m^2 & 0 & 0 & 0 & AZ_m'\sigma_m^2 \\ 0 & 0 & I\sigma_p^2 & 0 & 0 & Z_p'\sigma_p^2 \\ 0 & 0 & 0 & I\sigma_{sy}^2 & 0 & Z_{sy}'\sigma_{sy}^2 \\ 0 & 0 & 0 & 0 & I\sigma_e^2 & I\sigma_e^2 \\ Z_d A\sigma_d^2 & Z_m A\sigma_m^2 & Z_p \sigma_p^2 & Z_{sy} \sigma_{sy}^2 & I\sigma_e^2 & V \end{bmatrix}$$

A is the matrix of numerator relationships among all animals, σ_d^2 is the additive direct genetic variance, σ_{dm} is the additive direct and maternal genetic covariance, σ_m^2 is the additive maternal genetic variance, σ_p^2 is the permanent environmental variance, σ_{sy}^2 is the SY variance, σ_e^2 is the residual variance, and $V = Z_d A Z_d' \sigma_d^2 + Z_m A Z_m' \sigma_m^2 + (Z_d A Z_m' + Z_m A Z_d') \sigma_{dm} + Z_p Z_p' \sigma_p^2 + Z_{sy} Z_{sy}' \sigma_{sy}^2 + I\sigma_e^2$.

Analyses were done with different analytical models that included all or a subset of the effects in the above model. Four models were used; the full model presented above and models ignoring SY and/or inbreeding depression.

Parameter estimation

Restricted maximum likelihood (REML) estimates of the (co)variance components in the analytical models were obtained using multiple trait derivative-free REML (Boldman et al., 1993). The program uses the procedures from Meyer (1989) and incorporates routines for handling sparse matrices (George et al., 1980). The program uses the downhill simplex method to minimize negative log likelihood function. Stopping criterion was simplex variance being less than 10^{-10} . Cold restarts with the converged (co)variance component estimates as new starting values were applied to every estimation.

RESULTS AND DISCUSSION

Influence of ignoring inbreeding depression

Estimates of parameters from four analytical models are shown in table 1. Estimates of (co)variance components and covariate for inbreeding depression using the true model were not different ($p > .05$) from their corresponding input values. In the analysis with the model excluding inbreeding depression covariate, all genetic parameters were inflated ($p < .01$) compared to those

estimated from the true model. The largest differences were observed for the direct genetic effect and for the direct-maternal covariance (the direct-maternal covariance became more negative). Both the residual and permanent environmental variances were biased upward, as was the phenotypic variance. In this paper, *bias* means that parameter estimate significantly differed from value to be compared (input value in simulation or estimate obtained from true model), not that the expected value of the estimate differed from its parameter. This is due to the 'bias' property of REML estimate.

Lee and Pollak (1996b) demonstrated that misidentification of sires leads to a spurious SY variance. Of interest here is whether ignoring inbreeding depression does as well, which might further explain the presence of an SY variance in the analysis of Simmental data. An analytical model ignoring inbreeding but including an SY effect was used (table 1). All genetic parameters were biased in a manner similar to ignoring inbreeding depression without fitting an SY effect. A small SY variance was estimated and was statistically different from

zero ($p < .01$). Although the estimated variance attributed to SY differed significantly from the input value, it accounted for a small proportion (.63%) of phenotypic variance. The small average variance estimated using REML might be attributed to the propriety of REML estimates being within the parameter space (non-negativity of REML estimates for variance components). When using zero variance in simulation, variance estimates will be zero or positive. The presence of a small SY variance estimate can be explained by this property of REML estimates. That is, if the maximum of a likelihood is located on the boundary of the parameter space, then the derivative of the likelihood is unlikely to equal zero at the maximum. Eight replicates had the SY variance on the boundary in this study.

Estimates of (co)variance components with the model including both inbreeding depression covariate and SY, an overparameterized model, were not different ($p > .05$) from estimates using the true model. Again, SY variance was small but significantly different from zero because of the reason addressed above (table 1).

Table 1. Estimates^a of (co)variance components and regression for inbreeding coefficient from the data simulated with inbreeding depression effects using models with and without sire-by-year interaction and/or inbreeding depression

	Additive direct variance	Additive direct and maternal covariance	Additive maternal variance	Permanent environmental variance	SY variance	Residual variance	Phenotypic variance	Covariate for inbreeding depression
Input value	219	-51	116	74	0	382	740	-.8
F	216.4 ± 4.3	-53.7 ± 2.3	120.1 ± 2.5	72.5 ± 1.2	—	384.7 ± 2.4	739.9 ± 1.5	-.788 ± .012
	ns	ns	ns	ns		ns	ns	ns
NONE	7.3 ± .6	-9.9 ± .3	1.3 ± .3	1.9 ± .2	—	2.4 ± .3	3.0 ± .3	—
	**	**	**	**		**	**	
SY	10.6 ± 2.5	-12.5 ± 1.6	3.5 ± 1.1	1.7 ± .3	.63 ± .18	.5 ± 1.2	4.4 ± .6	—
	**	**	**	**		ns	*	
SY+F	3.4 ± 2.2	-3.0 ± 1.8	3.3 ± 1.8	-.5 ± .3	.56 ± .17	-2.1 ± 1.1	1.8 ± .8	.000 ± .000
	ns	ns	ns	ns		ns	*	ns

^aMean ± SE and significance (ns $p > .05$, * $p < .05$, ** $p < .01$). The figures for the models (NONE, SY, and SY+F) are deviations from estimates with true model (F). Standard errors were empirically obtained from 25 replicates. $H_0: \mu_F = \text{input value for true model}$, and $E_0: \mu_{SY+F} (\mu_{SY}, \text{ or } \mu_{NONE}) - \mu_F = 0$ for the other models.

IMPLICATIONS

In this study, ignoring inbreeding depression effects in the analysis of the simulated WW data inflated direct and maternal genetic (co)variance components. The findings implied that inbreeding depression should be considered in the genetic evaluation for populations where inbreeding is present. Inaccurate parameter estimates caused by ignoring inbreeding depression would lead to reduction in

genetic progress because of reduction of correlation between estimated and true genetic merit.

REFERENCES

- Boldman, K. G., L. A. Kriese, L. D. Van Vleck, and S. D. Kachman. 1993. A manual for use of MTDFREML. A set of programs to obtain estimates of variances and covariances (DRAFT). U. S. Department of Agriculture, Agricultural Research Service.

- Bowman, J. C., E. A. Butler, and E. Tuncel. 1978. Coefficients of inbreeding and degree of relationship for the British Friesian herd. *Anim. Prod.* 27:269.
- George, A., J. Liu, and E. Ng. 1980. User guide for SPARSPAK: Waterloo sparse linear equations package. CS-78-30, Dept. of Computer Sci., Univ. of Waterloo, Waterloo, ON, Can.
- Hudson, G. F. S. and L. D. Van Vleck. 1984. Effects of inbreeding on milk and fat production, stayability, and calving interval of registered Ayrshire cattle in the Northeastern United States. *J. Dairy Sci.* 67:171.
- Lee, C., and E. J. Pollak. 1996a. Influence of partitioning data by sex on genetic variance and covariance components for weaning in beef cattle. *J. Anim. Sci.* 75:61.
- Lee, C., and E. J. Pollak. 1996b. Influence of sire misidentification on sirexyear interaction variance for weaning weight in beef cattle. *J. Anim. Sci.* (accepted).
- Lee, C., and E. J. Pollak. 1996c. Relationship between sirexyear interactions and direct-maternal genetic correlation for weaning weight of Simmental cattle. *J. Anim. Sci.* 75:68.
- Mallinckrodt, C. H., B. L. Golden, and R. M. Bourdon. 1995. The effect of selective reporting on estimates of weaning weight parameters in beef cattle. *J. Anim. Sci.* 73:1264.
- Meyer, K. 1989. Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free algorithm. *Genet. Sel. Evol.* 21:317.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. Numerical recipes in FORTRAN (2nd. ed.). Cambridge Univ. Press, New York.
- Robinson, D. L. 1994. Models which might explain negative correlations between direct and maternal genetic effects. *Proc 5th World Congr Genet Appl Livest Prod*, Guelph, 18:378.
- Short, T. H. and T. J. Lawlor. 1992. Inbreeding in U. S. Holsteins and its effect on yield and type traits. *J. Dairy Sci.* 75 (Suppl. 1):154.
- Ward, B. 1993. The level of inbreeding and its effect on weight traits in Simmental cattle. Ph. D. Dissertation, Cornell University, Ithaca, New York.