January 1982

Derivation of Henderson's Method of Incorporating Artificial Insemination Sire Evaluations into Intraherd Prediction of Breeding Values

L. Dale Van Vleck
University of Nebraska-Lincoln, dvan-vleck1@unl.edu

Derivation of Henderson's Method of Incorporating Artificial Insemination Sire Evaluations into Intraherd Prediction of Breeding Values

L. D. VAN VLECK
Department of Animal Science
Cornell University
Ithaca, NY 14853

ABSTRACT

A derivation is given for the method proposed by Henderson to incorporate evaluations of sires based on daughter records in other herds into the intraherd prediction of breeding values. The derivation consists of pretending the daughters in other herds have records which have been adjusted for the herd-year-season effects of their origin. These equations then are absorbed into the equation for the sire which is tied to both sets of daughters by elements of the inverse of the matrix of relationships.

DISCUSSION

Henderson (2) presented a procedure for what was called intraherd prediction of breeding and producing abilities using all relatives. A significant section of the paper, however, described incorporation of sire evaluations from records of artificially sired (AI) daughters in other herds into the intraherd predictions. The procedure essentially provides ties between herds with AI daughters, and also provides the same genetic base for evaluation of cows in such herds as the base for AI sire evaluation. Thus, the procedure provided a logical basis for effective comparison of the genetic values of cows in different herds. Bolgiano et al. (1) demonstrated that the error variance of prediction between genetic evaluations of cows in different herds is only slightly larger than that for comparisons of cows in the same herd. Their results were based on direct inverses of coefficients for simultaneous mixed model equations for cows in two herds tied together

by some AI sires in common.

Derivation of the procedure for incorporation of AI sire evaluations was not given in the paper (2). The purpose of this note is to present a simple derivation which is within the framework of mixed model equations. This derivation may not be the same envisioned by Henderson, but the results are identical. Quaas (4) obtained a different deviation which is more general. The derivation in this paper is useful for teaching purposes as well as for establishing the logical basis for the approximation.

Derivation will be by example for a sire having a daughter in the herd and q other daughters in other herds with only their first records included in the external sire evaluation. The derivation includes the technique, not described until a later paper by Henderson (3), of including breeding values of animals without records in the mixed model equations. The animal without records is the sire. Records of daughters in the AI proof will be assumed to be adjusted perfectly for fixed effects (e.g., the herd-year-season effects associated with their records). This incorrect assumption probably is a good approximation because most procedures for sire evaluation will have included adjustment for such effects. Equations for genetic values of cows with records in the herd are not tied directly to equations for cows in other herds as the only ties would be through the inverse of the numerator relationship matrix and those elements for paternal half sisters are zero. Thus, ties are indirect through the sire because the inverse element of the relationship matrix for a sire with any of his daughters is $-2/3$.

Let g_{sp} be the genetic value of a representative daughter of the p daughters of the sire in the herd, g_{si} ($i = 1, \ldots, q$) be genetic values of q daughters in other herds, and g_s be the genetic value of the sire. Because the model to predict genetic values also includes permanent

environmental effects, permanent environmental effects will be absorbed as described by Henderson (2).

The mixed model equations then become (ignoring equations for other cows in the herd and fixed effects for cows in the herd)

$$
\begin{bmatrix}
n_{sp}k/(n_{sp}+k)+4t/3 & -2t/3 & 0 & \cdots 0 \\
-2t/3 & (1+1/3+q/3)t & -2t/3 & \cdots -2t/3 \\
0 & -2t/3 & k/(k+1)+4t/3 & 0 \cdots 0 \\
\cdot & \cdot & 0 & \cdot \ \ \cdot \\
\cdot & \cdot & \cdot & \ \cdot \ \ \cdot \\
\cdot & \cdot & \cdot & \ \ \cdot \\
0 & -2t/3 & 0 & 0 \cdots k/(k+1)+4t/3
\end{bmatrix}
\begin{bmatrix}
\hat{g}_{sp} \\ \hat{g}_{s} \\ \hat{g}_{sl} \\ \cdot \\ \cdot \\ \cdot \\ \hat{g}_{sq}
\end{bmatrix}
=
\begin{bmatrix}
ky_{sp}/(n_{sp}+k) \\ 0 \\ ky_{sl}/(k+1) \\ \cdot \\ \cdot \\ \cdot \\ ky_{sq}/(k+1)
\end{bmatrix}
$$

Coefficients of t are elements of the inverse of the relationship matrix. The diagonal inverse element for the sire equation is partitioned into three parts: the 1 for the sire, the 1/3 for daughter sp, and the q/3 for the other q daughters. Henderson (2) defined $t = (1-r)/h^2$ and $k = (1-r)/(r-h^2)$ with r and h^2 repeatability and heritability, which are the ratios of the residual variance to the genetic and permanent environmental variances. The terms with k account for absorption of permanent environmental effects where daughter sp has n_{sp} records which sum to y_{sp} and the other daughters each have a single record represented as the y_{si}.

The next step is to absorb equations for the other q daughters into the sire equation. This absorption does not affect any other equations. After considerable algebra, the equations after absorption are

$$
\begin{bmatrix}
n_{sp}k/(n_{sp}+k)+4t/3 & -2t/3 \\
-2t/3 & q(1-r)/(4-h^2)+(1+1/3)t
\end{bmatrix}
\begin{bmatrix}
g'_{sp} \\ \hat{g}_s
\end{bmatrix}
=
\begin{bmatrix}
ky_{sp}/(n_{sp}+k) \\ 2(1-r)y_{s.}/(4-h^2)
\end{bmatrix}
$$

where $y_{s.}$ is the sum of the records of the q daughters in the sire evaluation. The qt/3 in the diagonal of the sire equation now is included in $q(1-r)/(4-h^2)$. The other two parts of the inverse of the relationship matrix of 1 and 1/3 are shown separately. The equations are now as given by Henderson except for the right-hand-side of the sire equation. Since $y_{s.}$ is not known nor is q, Henderson advised substituting estimates of q and $y_{s.}$ based on the sire evaluation which is available: for q, the number of daughters which have equivalent accuracy or "repeatability" with one record per daughter. If the AI sire evaluation, $\hat{v}$, is equated to the algebraic equation for the corresponding selection index estimate of $1/2g_s$, an approximation for $y_{s.}$ can be found: $\hat{v} = y_{s.}/[q + (4-h^2)/h^2]$ so that

$\hat{v}[q + (4-h^2)/h^2]$ is substituted for y_s. Thus, the right-hand side for the sire equation becomes

$$
\frac{2(1-r)}{(4-h^2)}\left[\frac{4+(q-1)h^2}{h^2}\right]\hat{v},
$$

as given by Henderson (2).

If the sire had no daughters in the herd, the sire equation would not be tied to any other equations and $\hat{g}_s$ would be $2\hat{v}$. Since the t/3 for a daughter in the same herd would not be added to the diagonal, the coefficient of $\hat{g}_s$ is

$$
\frac{q(1-r)}{(4-h^2)} = \frac{(1-r)[4+(q-1)h^2]}{h^2(4-h^2)},
$$

so that the identity

$$\frac{(1-r)[4+(q-1)h^2]}{h^2(4-h^2)}\hat{g}_s =$$

$$\frac{2(1-r)[4+(q-1)h^2]}{(4-h^2)(h^2)}\hat{v}$$

is established; i.e., $\hat{g}_s = 2\hat{v}$.

If daughter records from all lactations were used in the AI sire evaluation, the absorption step for the approximation would be possible if the number of daughters with 1, 2, etc., records were known. The diagonals of the coefficients for daughters with n_i records would be $4t/3 + k/(n_i + k)$. A concise formula for the diagonal of the sire equations after absorption does not seem possible but could be calculated numerically. A major difficulty is the right-hand-side. The symbolic absorption can be done with the right-hand sides of the daughter equations represented as $ky_{si}/(n_i + k)$, where y_{si} is the sum of records of the ith daughter. The sire equation after absorption (ignoring daughters in the same herd) becomes if $w_{si} = k/(n_{si} + k)$:

$$[(1 + q/3)t - (4t^2/9)\sum_i(4t/3 + w_{si})^{-1}]\hat{g}_s =$$

$$(2t/3)\sum_i[w_{si}y_{si}/(4t/3 + w_{si})].$$

Because the right-hand-side is a weighted average of daughter records, a simple substitution based on the AI sire evaluation is not possible. Therefore, a reasonable approach seems to be the approximation outlined by Henderson (2) which translates the accuracy of prediction to the equivalent number of first lactation records and the right-hand-side to a sum which is equivalent to that many first lactation records which in turn can be approximated by a function of the AI sire evaluations derived from selection index principles.

REFERENCES

1 Bolgiano, D. C., R. L. Quaas, and L. D. Van Vleck. 1980. Accuracy of comparing cows in different herds. Page 16 in Genetics research. 1979-80 Rep. Eastern Artif. Insem. Coop., Inc. Dept. Anim. Sci., Cornell Univ., Ithaca, NY.

2 Henderson, C. R. 1975. Use of all relatives in intraherd prediction of breeding values and producing abilities. J. Dairy Sci. 58:1910.

3 Henderson, C. R. 1977. Best linear unbiased prediction of breeding values not in the model for records. J. Dairy Sci. 60:785.

4 Quaas, R. L. 1979. Incorporating AI sire proofs into intraherd BLUP of breeding values. Mimeo, Dept. Anim. Sci., Cornell Univ., Ithaca, NY.