3rd World Congress on Genetics Applied to
Livestock Production

Animal Science Department

1986

ESTIMATION OF GENETIC RELATIONSHIPS AMONG MILK RECORDS FOR FIRST THREE LACTATIONS USING REML FOR AN ANIMAL MODEL

H. H. Swalve
Institut für Tierzucht

L. D. Van Vleck
Cornell University

ESTIMATION OF GENETIC RELATIONSHIPS
AMONG MILK RECORDS FOR FIRST THREE
LACTATIONS USING REML FOR AN ANIMAL MODEL

H. H. SWALVE
Institut für Tierzucht
Göttingen, West Germany

L. D. VAN VLECK
Department of Animal Science
Cornell University
Ithaca, New York

SUMMARY

In dairy cattle breeding, genetic relationships among lactation records are of special interest because most selection operates on first lactations. This selection also complicates estimation of genetic parameters. Techniques which give estimates unbiased by selection should be used. Estimation was done using EM-type REML for an animal model neglecting relationships across herds. Records were from 3,070 Holstein cows which had the first lactation recorded. Estimates after 17 rounds of iteration for heritabilities and genetic correlations were: $h_1^2 = .33$, $h_2^2 = .32$, $h_3^2 = .33$, $r_{12} = .88$, $r_{13} = .83$, $r_{23} = .86$. Within herd-year-season phenotypic standard deviations were 1,223 kg, 1,323 kg, and 1,265 kg.

INTRODUCTION

Animal breeding plans and methods for evaluating breeding values require knowledge of the heritabilities and phenotypic and genetic correlations of the traits considered. In dairy cattle, relationships among traits of main economic importance, i.e., milk, fat and protein, and also among lactations are of interest. Often performance in later lactations is assumed due to the same genes which influence performance in first lactation. The validity of this assumption is difficult to evaluate because selection has occurred based on knowledge of part of the data. In this case, estimators of variances and covariances by methods such as Henderson's Method 3 are likely to be biased (Rothschild et al., 1979; Meyer and Thompson, 1984; Sorensen and Kennedy, 1984) and methods such as minimum variance quadratic unbiased estimation (MIVQUE), maximum likelihood (ML) and restricted maximum likelihood (REML) should be used (Henderson, 1984a).

MATERIALS AND METHODS

General Model

If all available data are used, estimators unbiased from selection can be expected from ML procedures (Meyer and Thompson, 1984; Sorensen and Kennedy, 1984). An animal model with complete relationship matrix meets this requirement. The model for this study is

$$y = Xb + Za + e \qquad [1]$$

where
 y is a vector of observations,
 b is a vector of fixed effects (herd-year-seasons),
 a is a vector of random effects which represent additive genetic values of individual animals,
 e is the vector of residuals, and
 X and Z are known incidence matrices.
Expectations and variances are:

$$
E \begin{Bmatrix} a \\ e \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \quad , \quad Var \begin{Bmatrix} a \\ e \end{Bmatrix} = \begin{Bmatrix} G & 0 \\ 0 & R \end{Bmatrix}
$$

and the mixed model equations are:

$$
\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z+G^{-1} \end{bmatrix} \begin{Bmatrix} \hat{b} \\ \hat{a} \end{Bmatrix} = \begin{Bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{Bmatrix} \qquad [2]
$$

Data are ordered lactations within animals and missing lactations are included as zero rows/columns so that R is block diagonal where each block corresponds to the combination of traits recorded for a specific animal. Each block is derived from R_O, the variance-covariance matrix of the residuals by crossing out rows/columns associated with missing records. $G^{-1} = A^{-1} * G_O^{-1}$, where * denotes the direct product operation, A is the numerator relationship matrix and G_O is the variance-covariance matrix among additive genetic effects for the three traits.

The Method

The method applied was REML using the EM algorithm (Henderson, 1984a,b). A property is that solutions for variance and covariance components are within the parameter space. Iteration is on the following two sets of equations for $i,j = 1,...,3$:

estimation of G_O;
$$
tr\ A^{-1}G = \hat{a}_i'A^{-1}\hat{a}_j + tr\ A^{-1}C_{ij} \qquad [3]
$$
estimation of R_O;
$$
tr\ Q_{ij}R = \hat{e}'Q_{ij}\hat{e} + tr\ Q_{ij}\ WCW' \qquad [4]
$$

where C is a generalized inverse of the coefficient matrix in [2], C_{ij} represents the block pertaining to traits i and j from the lower right submatrix of C (pretending in [3] that animals are ordered within traits), $W = [X:Z]$, $\hat{e} = y - \hat{b} - \hat{a}$ and $\hat{b}$ and $\hat{a}$ are solutions to [2]. Derivation of Q_{ij}, the matrix for estimating the ij[th] element of R_O was as described in Henderson (1984a,b). The main computational problem in iterating on [3] and [4] is the need for the inverse of the coefficient matrix of [2]. The number of equations in [2] is the number of classes for fixed effects plus the number of animals times the number of traits. To simplify computations, transformation of [2] in some cases results in unrelated residuals (Quaas, 1984). The mixed model equations [2] then simplify to those for single trait analysis. This method, however, requires equal incidence matrices for all traits which is not fulfilled for this case of subsequent lactations. Other assumptions had to be made. By ignoring relationships across herds it

was possible to solve [2] herd by herd. After obtaining A_+^{-1}, the inverse of the numerator relationship matrix including base animals without records, equations for all base animals were absorbed into equations for cows actually having records by using rules for inversion of partitioned matrices.

The Data

The original data set consisted of mature equivalent lactation milk yields from 294,401 New York Holstein cows, first freshening between 1977 and 1984 and passing usual edits. All cows had at least a recorded first lactation, later lactations were accepted if the preceding lactations were recorded. From this data set a computationally manageable data set containing 3,070 cows (3,070 first, 1,943 second, 1,080 third lactations in 15 herds) was selected. Herds were selected if they contained a high frequency of related animals and the number of cows was in the range of 170 to 230. The average number of equations including fixed effects was 650 per herd.

RESULTS AND DISCUSSION

Table 1 shows prior values and estimates of additive genetic and residual variances and covariances from round 0 to round 17. All records were scaled by dividing by the overall standard deviation of first lactation milk yield to avoid round-off problems during the computations.

Starting with round 4 all estimates monotonically approached a point of convergence which was not reached by round 17. However, changes are very small. Estimates for R_0 are increasing, those for G_0 are decreasing; heritabilities are decreasing. A similar pattern was found with ML estimation by Rothschild et al. (1979). Estimated parameters obtained from round 17 solutions are given in Table 2. In general, the estimates are in

Table 2. Estimated heritabilities (on the diagonal), additive genetic (above the diagonal) and phenotypic (below the diagonal) correlations for milk yield for three lactations after 17 rounds of iteration.

		Lactation	
Lactation	1	2	3
1	.328	.877	.828
2	.587	.320	.860
3	.530	.648	.334

agreement with estimates found by other authors for different Holstein populations. However, heritabilities are higher especially for second and third lactations than those commonly found, which may be due to using an animal model which should be a better approximation of the true model as compared to sire models, which may be affected more by selection. Genetic correlations are slightly lower than in Meyer (1983), Meyer (1984) and Rothschild and Henderson (1979), but in agreement with Tong et al. (1979) and Meyer (1985) who all used ML procedures.

CONCLUSIONS

Estimation of genetic variances and covariances for first three lactations using REML for an animal model is computationally feasible. The results show relatively high heritabilities of equal size for all three lactations. Genetic correlations are not greatly different from unity. Phenotypic correlations are somewhat higher than usually reported for repeatability. The expected effects of selection and evaluations based on first lactations only or evaluations based on a constant repeatability therefore seem to be justified using these estimates. Further investigation to find computing techniques which would allow processing of larger data sets seems necessary.

REFERENCES

HENDERSON, C. R. 1984a. Applications of linear models in animal breeding. University of Guelph, Ontario, Canada.

HENDERSON, C. R. 1984b. Estimation of variances and covariances under multiple trait models. J. Dairy Sci. 67:1581-1589.

MEYER, K. 1983. Maximum likelihood procedures for estimating genetic parameters for later lactations of dairy cattle. J. Dairy Sci. 66:1988-1997.

MEYER, K. 1984. Estimation of genetic parameters for milk and fat yield for the first three lactations in British Friesian cows. Anim. Prod. 38:313-322.

MEYER, K. 1985. Genetic parameters for dairy production of Australian black and white cows. Livest. Prod. Sci. 12:205-219.

MEYER, K. and THOMPSON, R. 1984. Bias in variance and covariance component estimation due to selection on a correlated trait. Z. Tierzuchtg. Zuchtungsbiol. 101:33-50.

QUAAS, R. L. 1984. Linear prediction. In BLUP School Handbook. AGBU, University of New England, Australia.

ROTHSCHILD, M. F. and HENDERSON, C. R. 1979. Maximum likelihood estimates of parameters of first and second lactation milk records. J. Dairy Sci. 62:990-995.

ROTHSCHILD, M. F., HENDERSON, C. R. and QUAAS, R. L. 1979. Effects of selection on variances and covariances of simulated first and second lactations. J. Dairy Sci. 62:996-1002.

SORENSEN, D. A. and KENNEDY, B. W. 1984. Estimation of genetic variances from unselected and selected populations. J. Anim. Sci. 59:1213-1223.

TONG, A. K. W., KENNEDY, B. W. and MOXLEY, J. E. 1979. Heritabilities and genetic correlations for first three lactations from records subject to culling. J. Dairy Sci. 62:1784-1790.

Table 1. Starting values and estimates[a] for residual (R_o) and additive genetic (G_o) variances and covariances for milk yield for the first three lactations for 17 rounds of iteration.

Estimates	Starting values	Round										
		0	1	3	5	7	9	11	13	15	17
R_o:												
r_{11}	.456	.475	.485	.493	.496	.497	.499	.500	.501	.502	.503
r_{12}	.197	.213	.223	.231	.235	.237	.239	.241	.242	.244	.245
r_{13}	.152	.167	.176	.184	.188	.191	.193	.194	.196	.197	.198
r_{22}	.551	.566	.575	.582	.586	.588	.590	.591	.593	.594	.595
r_{23}	.282	.284	.286	.291	.295	.298	.300	.302	.303	.305	.307
r_{33}	.494	.506	.513	.519	.523	.525	.527	.528	.530	.531	.532
G_o:												
g_{11}	.250	.254	.256	.256	.254	.253	.251	.249	.248	.247	.245
g_{12}	.239	.243	.244	.244	.242	.240	.237	.235	.233	.232	.230
g_{13}	.221	.225	.226	.225	.224	.221	.219	.217	.215	.214	.212
g_{22}	.288	.292	.293	.293	.291	.289	.287	.285	.283	.281	.280
g_{23}	.249	.252	.253	.252	.250	.247	.245	.242	.240	.237	.235
g_{33}	.277	.280	.282	.281	.279	.277	.275	.273	.271	.269	.267

[a] Estimates are standardized; multiply by 2,000,000 to convert to kg^2.