Method R Estimates of Additive Genetic, Dominance Genetic, and Permanent Environmental Fraction of Variance for Yield and Health Traits of Holsteins

C. P. Van Tassell,* I. Misztal,† and L. Varona†
*Agricultural Research Service,
USDA, Beltsville, MD 20705
†Department of Animal and Dairy Science,
University of Georgia, Athens 30602

ABSTRACT

Fractions of variance accounted for by additive genetic, dominance genetic, and permanent environmental effects for milk, fat, and protein yields; somatic cell score; and productive life were estimated from Holstein data used for national genetic evaluations. Contemporary group assignments were determined using the national procedure. Data included 1,973,317 milk and fat records for 812,659 cows, 1,019,421 protein records for 462,067 cows, 468,374 lactation average somatic cell score (SCS) records for 232,909 cows, and 735,256 cows with productive-life records. Variance components were estimated with the JAADOM program, which uses iteration on data and second-order Jacobi iteration for obtaining solutions to the mixed-model equations and Method R for estimation of variance components. Ten different random data subsets were used to estimate parameters for each trait. Estimated additive genetic, dominance genetic, and permanent environmental fractions of variance were 0.34, 0.05, and 0.10 for milk yield; 0.34, 0.05, and 0.11 for fat yield; 0.31, 0.05, and 0.10 for protein yield; and 0.17, 0.01, and 0.16 for lactation average SCS. Estimated additive genetic and dominance genetic fractions of variance were 0.12 and 0.06 for productive life. Mean empirical standard errors of additive genetic, dominance genetic, and permanent environmental variance fractions were 0.003, 0.006, and 0.006.
(**Key words:** heritability, dominance variance, milk yield, somatic cell score, productive life)

Abbreviation key: LSCS = lactation average $\log_2$-transformed SCC; **PL** = productive life.

INTRODUCTION

Although animal breeders have traditionally assumed that additive genetic variance accounts for most or all of total genetic variance, the total genetic value of an animal may include additive, dominance, and epistatic effects (3):

$$g = a + d + aa + dd + ad + \dots ,$$

where g is the total genetic value of an animal, a is the additive genetic effect, d is the dominance genetic effect, and aa, dd, and ad are epistatic genetic effects. If nonadditive genetic variance exists, then the genetic merit of a mating combination (i.e., the specific combining ability) has a contribution associated with interactions of genes at the same locus (dominance) or at multiple loci (epistasis). Because of the structure of the dairy cattle population historically, mean dominance relationships have been relatively low. With the increased application of embryo transfer and increased fraction of full sibs in the population, however, mean dominance relationships are likely to increase.

Use of nonadditive genetic effects in animal evaluation and mating decisions requires the calculation of inverses of matrices of additive and dominance relationships between individuals, as well as the knowledge of genetic variances. Given those relationships, the covariance of genetic values for two noninbred individuals can be written as

$$\mathrm{cov}(x,y) = \sum_{i,j} a_{xy}^{i} d_{xy}^{j} \sigma_{ij}^{2},$$

where a_{xy} (d_{xy}) is the additive (dominance) relationship between x and y, σ_{ij}^2 is the appropriate genetic variance (e.g., σ_{10}^2, σ_{01}^2, and σ_{20}^2 are additive, dominance, and additive by additive genetic variances), and $i + j \geq 1$ (3). Variances with values of i or j greater than 2 likely would not be easily estimated or applied in mating programs.

Henderson (6) described a method to rapidly obtain the inverse of the numerator relationship matrix. By using a similar general technique, Hoeschele and VanRaden (8) derived algorithms to obtain inverses of dominance relationship matrices. Prior to the avail-

---

Received November 12, 1999.
Accepted February 25, 2000.
Corresponding author: C. P. Van Tassell; e-mail: curtvt@aipl.arsusda.gov.

ability of the algorithm to calculate the inverse of a dominance relationship matrix directly, analyses used relatively small data sets, because direct calculation of inverses were required or relationships were ignored to facilitate simplified models (e.g., 23). More recent analyses sometimes were restricted to data only from three-quarter or full sibs to maximize dominance information with minimum computing requirements (26).

Estimates of dominance variance for milk, fat, and protein yields have been estimated by many authors (9, 12, 13, 23, 26), and estimates range from 0.03 to 0.24, although most estimates are <0.10. When measured as a fraction of additive genetic variance, dominance fraction of variance estimates ranged from 0.04 to 0.75 with most estimates <0.20. Estimates for lactation average $\log_2$-transformed SCC (**LSCS**) were quite small with only 1.3% of the phenotypic variance attributed to dominance corresponding to 8% of additive genetic variance (12). Previous estimates of dominance variance were not available for productive life (**PL**), although estimates for traits related to herd life were available. Variance fraction estimates for conformation traits ranged from 0.02 to 0.16 when measured as a fraction of phenotypic variance and from 0.08 to 0.99 when measured as a fraction of additive genetic variance (15, 22). Dominance variance estimates for fertility traits ranged from 0.5 to 3.4% of phenotypic variance and from 25 to 350% of additive genetic variance (7).

Inbreeding depression for milk, fat, and protein yields have been estimated by many authors (13, 20, 21, 33). Estimates of inbreeding depression for milk, fat, and protein yield ranged from 23 to 30 kg, 0.90 to 1.2 kg, and 0.78 to 1.2 kg loss for each 1% increase in inbreeding. Previous estimates of inbreeding for SCS were near zero and opposite in direction (12, 21). Only one previous study looked at the effect of inbreeding on PL (21) and estimated inbreeding loss as 0.43 mo of PL for each percentage increase of inbreeding.

Because of advances in computing hardware and algorithms, estimation of nonadditive genetic variances with large data sets and calculation of nonadditive predicted genetic values on a regular basis are feasible (14). Procedures developed by Misztal (14) allow dominance genetic solutions to be calculated for parental combinations present in the data at a cost of less than twice that of calculating solutions for additive genetic effects. Varona and Misztal (28) described an algorithm to predict parental dominance combinations for matings not represented in the data set, which might allow for development of mate allocation algorithms with increased response to selection (5, 28).

The objectives of this study were to estimate fractions of genetic variance associated with additive genetic, dominance genetic, and permanent environmental effects and to estimate the magnitude of inbreeding depression for production traits in Holstein dairy cattle.

MATERIALS AND METHODS

Data

Data were obtained from the Animal Improvement Programs Laboratory database, which includes historical yield and pedigree information for cows enrolled in DHIA programs throughout the US. Records were extended to 305 d. Further adjustments were made so that yields were to a mature-equivalent and twice-daily milking basis. Data for LSCS (18) and PL (25) were also obtained. Records were adjusted with expansion factors to stabilize phenotypic variance and were adjusted to account for heterogeneous variance. Contemporary group assignments were determined using the same procedure that is implemented in the national genetic evaluation system (32).

The complete data set included lactation records of all cows born after 1980. Lactation records were discarded when sequential parity records were not present, starting with first lactation (e.g., if first and third lactations were present, only the first lactation was used). These data were reduced by removing single record contemporary groups. Herds with fewer than 11 animals across time were discarded. Full sibs were identified across herds. Data were selected to maximize the number of full sibs in the analysis and to use as large a data set as possible in an attempt to keep the data representative of the US population. Data were included from herds with sufficiently large numbers of full sibs so that the overall fraction of full-sib animals was at least 20%, when full sibs were determined using all herds. Because all herds were not used, the actual fraction of full sibs was likely lower. Based on this cutoff, data from herds with at least 15% full sibs were used for milk, fat, and protein as well as PL. A cutoff of 14% was used for LSCS.

To limit the number of extra equations for dominance effects, a single level of recursion of rules described by Hoeschele and VanRaden (8) was used to build the inverse of the parental dominance relationship matrix. That limitation resulted in inclusion of full sibs, three-quarter sibs, and some cousins in dominance relationships. Some dominance information was lost because some dominance connections through grandparents were lost. But, because dominance relationships decrease rapidly as the generations to connected animals increase, the effect of these lost relationships might have been small.

Model

The model used for data analysis was

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Zb}\Delta + \mathbf{Za} + \mathbf{ZWf} + \mathbf{e},$$

where $\mathbf{y}$ is a vector of observations, β is a vector of fixed effects, $\mathbf{b}$ is a vector of inbreeding coefficients, Δ is coefficient of inbreeding depression, $\mathbf{a}$ is a vector of random additive animal effects, $\mathbf{f}$ is a vector of random parental animal effects, $\mathbf{e}$ is a vector of random residual effects, and $\mathbf{X}$, $\mathbf{Z}$, and $\mathbf{W}$ are appropriately dimensioned incidence matrices. Standard assumptions were made about parameter means and variances:

$$\mathrm{E}\begin{bmatrix}\mathbf{y}\\ \mathbf{a}\\ \mathbf{f}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{X}\beta\\ \mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\end{bmatrix};\ \mathrm{Var}\begin{bmatrix}\mathbf{a}\\ \mathbf{f}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{A}\sigma_a^2 & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & .25\mathbf{F}\sigma_d^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_e^2\end{bmatrix},$$

where $\mathbf{A}$ is the additive numerator relationship matrix, and $\mathbf{F}$ is the matrix of dominance relationships among parent classes.

Parameter Estimation

Variance components were estimated with the JAA-DOM program, which uses iteration on data and second-order Jacobi iteration for obtaining solutions to the mixed-model equations and Method R (16) for variance component estimation as described by Misztal (14). Method R was used for parameter estimation, because other methods were not practical for such large data sets (14). Even with Method R, computing requirements were substantial, with 18 to 35 h required to obtain solutions for each random subset of data.

Method R requires R values. These values are regressions of predicted random effects calculated using "complete" data on predicted random effects calculated with random subsets of the same data. All R values will be 1 if the parameters are appropriate for the population. The main advantage of using Method R is that large data sets can be used for parameter estimation because the procedure is based on repeated solutions of standard mixed-model equations. Additionally, standard errors and confidence intervals for variance fractions can be estimated using the Method R sample estimates (11). The difficulty of this method when applied to multiple random effects is that constrained optimization is needed to obtain parameters that result in all regressions of 1. The convergence criterion for Method R was R = 1 ± 0.0002. Ten random subsets of the complete data set were used for each trait. Estimated sampling variances of the parameter estimates were approximated as the variance of estimates obtained from the 10 subsets. Once mean estimates of the variance fractions were obtained, inbreeding depression for each trait was estimated by solving the complete set of mixed-model equations.

Table 1. Number of records, cows with records, management groups, total animals included in additive relationship matrix, and number of parent classes in dominance relationship matrix for milk and fat, protein, lactation average SCS (LSCS), and productive life (PL).

	Milk and fat	Protein	LSCS	PL
Records	1,973,317	1,019,421	468,374	735,256
Cows with records	812,659	462,067	232,909	735,256
Management groups	305,567	139,639	64,155	123,408
Total animals	1,456,664	1,244,828	611,820	1,360,506
Dominance classes	2,523,587	1,430,278	714,518	2,296,343

RESULTS AND DISCUSSION

Data set characteristics are given in Table 1. Nearly 2 million lactation records for more than 810,000 cows were used to estimate parameters for milk and fat yields. Nearly 1 million lactation records for over 400,000 cows were used for protein yield. Because of less complete data recording, the number of records available for LSCS was considerably smaller than for the yield traits, with over 460,000 records on more than 230,000 cows. Approximately 735,000 cows had observations for PL.

Mean estimates of fractions of variance for milk, fat, and protein yield are shown in Table 2. Heritability estimates averaged 0.33 across traits for the yield traits. Those estimates are in the middle of the range of estimates obtained by Van Tassell et al. (27), in which heritabilities were estimated using first lactation records on cows in subsets, based on herd-year variance. Standard errors for all parameters were small (<0.015), although estimates were relatively large compared with parameter estimates in some cases, particularly for LSCS for which the dominance variance estimate was near zero. For yield traits, dominance estimates were small (approximately 0.05), however dominance variance accounted for 15 to 17% of additive genetic variance. These estimates were similar to estimates by Tempelman and Burnside (23) for milk yield. However, these estimates of dominance fraction were higher than those previously estimated for milk, fat, and protein yields by Miglior et al. (13) and milk yield by VanRaden et al. (26). Estimates for milk and fat yields were smaller than those by Hoeschele and Vollema (9) and were much smaller

Table 2. Mean estimates of fraction of variance ($\bar{x}$) and estimates of standard errors (SE) of additive genetic variance (h^2), dominance variance (d^2), permanent environmental variance (p^2), cow effect fraction of variance (c^2), and residual fraction of variance (e^2), ratio of dominance and additive genetic variances (d^2/h^2), and inbreeding depression (INBRD) for milk, fat, protein, lactation average SCS (LSCS), and productive life (PL).

	Milk		Fat		Protein		LSCS		PL	
	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE
h^2	0.342	0.003	0.338	0.003	0.312	0.004	0.166	0.003	0.118	0.002
d^2	0.053	0.010	0.050	0.005	0.052	0.004	0.010	0.007	0.057	0.011
p^2	0.104	0.010	0.111	0.006	0.098	0.005	0.160	0.007	NA[1]	NA
c^2	0.499	0.001	0.499	0.002	0.462	0.002	0.335	0.003	0.175	0.011
e^2	0.501	0.001	0.501	0.002	0.538	0.002	0.665	0.003	0.825	0.011
d^2/h^2	0.15	—	0.15	—	0.17	—	0.06	—	0.48	—
INBRD	–31.7 kg	—	–1.2 kg	—	–1.0 kg	—	0.0037 LSCS	—	–0.242 mo	—

[1] NA = not applicable.

than fat yield estimates by Tempelman and Burnside (23). Inbreeding depression estimates indicate that production losses were 32 kg of milk, 1.2 kg of fat, and 1.0 of kg protein for each percentage increase in inbreeding. These losses are quite similar to those reported by Smith et al. (21) and Wiggans et al. (33). The estimates were larger than those estimated by Short et al. (20), whose data included only records for registered cows, and Miglior et al. (13), whose estimates were based on Canadian Holsteins.

Estimates of variance fractions for LSCS are also shown in Table 2. Fractions of variance due to additive genetic and permanent environmental effects were very similar, with estimates of 16.6 and 16.0%, respectively. Estimates of heritability were higher than in many studies (e.g., 1, 2, 4, 19), although similar to those by Miglior et al. (12) and Welper and Freeman (31). Estimated dominance fraction was the lowest of all traits examined when measured as a fraction of phenotypic variance (1.0%) and as a fraction of additive genetic variance (6.1%). The estimated fractions attributed to additive and dominance genetic effects agree with those obtained by Miglior et al. (12), who used data from Canadian Holstein cows. Estimated inbreeding depression was quite small, with an estimated change of 0.0037 units of LSCS for each 1% of inbreeding. This estimate is smaller than that obtained by Miglior et al. (12) and opposite to that obtained by Smith et al. (21).

Estimates of variance fraction for PL are also shown in Table 2. Additive genetic effects were estimated to account for 11.8% of the phenotypic variance. This estimate was higher than most previous estimates (e.g., 10, 24, 25). The estimated dominance fraction was the highest of all traits considered in this study when measured either as a fraction of phenotypic variance (5.7%) or as a fraction of additive genetic variance (48.3%). Dominance fraction was also higher than any of a large number of linear traits considered by Misztal et al. (15), which are genetically associated with productive life (17, 29, 30). The dominance fraction was also higher than that estimated for fertility traits considered by Hoeschele and Vollema (9), which is also likely related to length of life. Estimated inbreeding depression indicate that PL losses were 0.24 mo of PL for each percentage increase in inbreeding. These losses are approximately half those reported by Smith et al. (21) but still suggest extreme consequences of inbreeding.

Inclusion of dominance effects in genetic evaluations for yield traits could increase accuracy of prediction of additive genetic effects, especially if use of MOET increases the incidence of large full-sib family groups in the dairy population. This inclusion may be especially beneficial for increasing PL, for which a large fraction of the genetic variance was estimated to be dominance variance.

CONCLUSIONS

Fractions of variance accounted for by additive genetic, dominance genetic, and permanent environmental effects for milk, fat, and protein yields were estimated from Holstein data used for national genetic evaluations. Estimation of dominance variance was desired because computational methods have been developed that may allow inclusion of dominance genetic effects in large-scale evaluations. Variance components were estimated using Method R. Dominance variance accounted for 15 to 17% as much of the phenotypic variance as additive genetic variance for yield traits, 6% for LSCS, and 48% for PL. Because of the magnitude of the estimates obtained and possible changes in population structure associated with reproductive technologies, further research is needed to determine the potential impact on predictions of additive genetic effects of including dominance effects in a genetic evaluation system.

REFERENCES

1 Banos, G., and G. E. Shook. 1990. Genotype by environment interaction and genetic correlations among parities for somatic cell count and milk yield. J. Dairy Sci. 73:2563–2573.

2 Boettcher, P. J., L. B. Hansen, P. M. VanRaden, and C. A. Ernst. 1992. Genetic evaluation of Holstein bulls for somatic cells in milk of daughters. J. Dairy Sci. 75:1127–1137.

3 Cockerham, C. C. 1954. An extension of the concept of partitioning heredity variance for analysis of covariances among relatives when epistasis is present. Genetics 39:859–882.

4 Da, Y., M. Grossman, I. Misztal, and G. R. Wiggans. 1992. Estimation of genetic parameters for somatic cell score in Holsteins. J. Dairy Sci. 75:2265–2271.

5 De Stefano, A. L., and I. Hoeschele. 1992. Utilization of dominance through mate allocation strategies. J. Dairy Sci. 75:1680–1690.

6 Henderson, C. R. 1976. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32:69–83.

7 Hoeschele, I. 1991. Additive and nonadditive genetic variance in female fertility of Holsteins. J. Dairy Sci. 74:1743–1752.

8 Hoeschele, I., and P. M. VanRaden. 1991. Rapid inversion of dominance relationship matrices for noninbred populations by including sire by dam subclass effects. J. Dairy Sci. 74:557–569.

9 Hoeschele, I., and A. R. Vollema. 1993. Estimation of variance components with dominance and inbreeding in dairy cattle. J. Anim. Breed. Genet. 110:93–104.

10 Jairath, L. K., J. F. Hayes, and R. I. Cue. 1994. Multitrait restricted maximum likelihood estimates of genetic and phenotypic parameters of lifetime performance traits for Canadian Holsteins. J. Dairy Sci. 77:303–312.

11 Mallinckrodt, C. H., B. L. Golden, and A. Reverter. 1997. Approximate confidence intervals for heritability from Method R estimates. J. Anim. Sci. 75:2041–2046.

12 Miglior, F., E. B. Burnside, and J. C. Dekkers. 1995. Nonadditive genetic effects and inbreeding depression for somatic cell counts of Holstein cattle. J. Dairy Sci. 78:1168–1173.

13 Miglior, F., E. B. Burnside, and B. W. Kennedy. 1995. Production traits of Holstein cattle: estimation of nonadditive genetic variance components and inbreeding depression. J. Dairy Sci. 78:1174–1180.

14 Misztal, I. 1997. Estimation of variance components with large-scale dominance models. J. Dairy Sci. 80:965–974.

15 Misztal, I., T. J. Lawlor, and N. Gengler. 1997. Relationships among estimates of inbreeding depression, dominance and additive variance for linear traits in Holsteins. Gen. Sel. Evol. 29:319–326.

16 Reverter, A., B. L. Golden, R. M. Bourdon, and J. S. Brinks. 1994. Method R variance component procedure: application on the simple breeding value model. J. Anim. Sci. 72:2247–2253.

17 Rogers, G. W., G. Banos, and U. Sander-Nielsen. 1999. Genetic correlations among protein yield, productive life, and type traits from the United States and diseases other than mastitis from Denmark and Sweden. J. Dairy Sci. 82:1331–1338.

18 Schutz, M. M. 1994. Genetic evaluation of somatic cell scores for United States dairy cattle. J. Dairy Sci. 77:2113–2129.

19 Schutz, M. M., P. M. VanRaden, and G. R. Wiggans. 1994. Genetic variation in lactation means of somatic cell scores for six breeds of dairy cattle. J. Dairy Sci. 77:284–293.

20 Short, T. H., T. J. Lawlor, and R. W. Everett. 1992. Inbreeding in U. S. Holsteins and its effect on yield and type traits. J. Dairy Sci. 75(Suppl. 1):154. (Abstr.).

21 Smith, L. A., B. G. Cassell, and R. E. Pearson. 1998. The effects of inbreeding on the lifetime performance of dairy cattle. J. Dairy Sci. 81:2729–2737.

22 Tempelman, R. J., and E. B. Burnside. 1990. Additive and nonadditive genetic variation for conformation traits in Canadian Holsteins. J. Dairy Sci. 73:2214–2220.

23 Tempelman, R. J., and E. B. Burnside. 1990. Additive and nonadditive genetic variation for production traits in Canadian Holsteins. J. Dairy Sci. 73:2206–2213.

24 Van Doormaal, B. J., L. R. Schaeffer, and B. W. Kennedy. 1985. Estimation of genetic parameters for stayability in Canadian Holsteins. J. Dairy Sci. 68:1763–1769.

25 VanRaden, P. M., and E.J.H. Klaaskate. 1993. Genetic evaluation of length of productive life including predicted longevity of live cows. J. Dairy Sci. 76:2758–2764.

26 VanRaden, P. M., T. J. Lawlor, T. H. Short, and I. Hoeschele. 1992. Use of reproductive technology to estimate variances and predict effects of gene interactions. J. Dairy Sci. 75:2892–2901.

27 Van Tassell, C. P., G. R. Wiggans, and H. D. Norman. 1999. Method R estimates of heritability for milk, fat, and protein yields of United States dairy cattle. J. Dairy Sci. 82:2231–2237.

28 Varona, L., and I. Misztal. 1999. Prediction of parental dominance combinations for planned matings, methodology, and simulation results. J. Dairy Sci. 82:2186–2191.

29 Weigel, D. J., B. G. Cassell, and R. E. Pearson. 1997. Prediction of transmitting abilities for productive life and lifetime profitability from production, somatic cell count, and type traits in milk markets for fluid milk and cheese. J. Dairy Sci. 80:1398–1405.

30 Weigel, K. A., T. J. Lawlor, Jr., P. M. VanRaden, and G. R. Wiggans. 1998. Use of linear type and production data to supplement early predicted transmitting abilities for productive life. J. Dairy Sci. 81:2040–2044.

31 Welper, R. D., and A. E. Freeman. 1992. Genetic parameters for yield traits of Holsteins, including lactose and somatic cell score. J. Dairy Sci. 75:1342–1348.

32 Wiggans, G. R., I. Misztal, and L. D. Van Vleck. 1988. Implementation of an animal model for genetic evaluation of dairy cattle in the United States. J. Dairy Sci. 71:54–69.

33 Wiggans, G. R., P. M. VanRaden, and J. Zuurbier. 1995. Calculation and use of inbreeding coefficients for genetic evaluation of United States dairy cattle. J. Dairy Sci. 78:1584–1590.