

EFFECT OF USING THE ADDITIVE GENETIC RELATIONSHIP MATRIX ON SIRES' PROOFS

M. H. Sadek¹, A. A. Halawa² and A. A. Ashmawy¹

- 1) Department of Animal Production, Faculty of Agriculture, University of Ain Shams, Cairo, Egypt
- 2) Animal Production Research Institute, Ministry of Agriculture, Cairo, Egypt

SUMMARY

Two analyses were performed using a mixed model that included herd-year-seasons and groups of sires as fixed effects and sires within groups as a random effect. In one model the additive genetic relationship inverse matrix (A^{-1}) was included and in the other the A^{-1} was not included. Data represent 1812 305-day first lactation records, daughters of 79 sires, collected over the period from 1960 up to 1989. Two seasons of freshening were defined, November to April and May to October with a total of 94 herd-year-seasons (HYS). The 305-day milk overall mean was $1,967 \pm 445$ kg.

Transmitting ability (TA) estimates disregarding the A^{-1} ranged from 182 kg to -143 kg with an average difference of 4.1 kg among sires while these estimates stretched from 197 kg to -143 kg with an average difference among sires of 4.3 kg when the A^{-1} was included. The product moment correlation among sires TA' estimates was 0.96 with and without including the A^{-1} . Rank correlations was 0.98 for the whole list of sires. It was 0.62 for the top ten sires and 0.99 for the lowest 10 sires. Slight changes in variance components and heritability estimates were revealed as a result of including the A^{-1} .

Keywords: Transmitting ability, rank and relationship matrix

INTRODUCTION

The concept of sire evaluation means estimating a value for each sire representing his transmitting ability (TA). The rank of these sires' proofs permits the optimum choice of sires. The accuracy of sire evaluation and consequently sires' rank depend to a large magnitude on the factors that are included in the statistical model used in estimating these values. The resemblance between relatives has long been known as one of the most important phenomena in population genetics theory. Including the inverse of the relationship matrix (A^{-1}) in the statistical model Henderson (1975a), as a measure of resemblance among sires, provides relationship ties among contemporary groups. The lower the heritability of the trait, the more important this aspect of including the A^{-1} becomes (Kennedy and Moxley 1975; Jensen 1980; Wood *et al.* 1991)

Henderson (1975 a) discovered that relationships could readily be included explicitly in the best linear unbiased prediction (BLUP) models because the A^{-1} is easier to compute than the relationship matrix itself. He presented a rapid method for computing the A^{-1} . Henderson (1975b) reported that the potential advantages in using relationships among sires in addition to, some female ancestors in sire evaluation are: (1) increases accuracy of prediction, particularly for sires with few or no progeny; (2) requires fewer groups to account for genetic trend and for genetic differences among populations; (3) evaluates sires earlier through the use of records on their dams, dams' paternal sisters, and sires' paternal sisters.

There is general agreement that the inclusion of genetic groups can account for differences among sires beyond those which are strictly sire differences. Keown (1974) reported that the reason for grouping is to put sires into homogeneous genetic groups and the intent is to have as much diversity between groups and as little diversity within groups as possible. As expressed by Powel and Freeman (1974) " Virtually any grouping is preferable to ignoring groups of sires in general". There are different methods of grouping: grouping by year of birth, by stud, by region, or by pedigree index. Robinson (1986) reported that no criterion to define group effects in models for estimating breeding values

is universally accepted. The criteria for grouping is a topic of some debate. However, if complete relationship matrix is available, groups are not needed.

Therefore, the objective of this study was to evaluate the effect of including the A^{-1} on sires' proofs and their rank using the BLUP method.

MATERIAL AND METHODS

Data were 1812 305-day first lactation records collected over the period from 1960 up to 1989 from two herds of Friesian cattle raised at Sakha and El-Karada experimental stations, located in the northwest of the Nile Delta in Kafr El-Sheikh governorate. These herds belong to Animal Production Research Institute, Ministry of Agriculture and Land reclamation, Egypt.

A total of 94 herd-year-seasons (HYS) were formed by specifying two seasons of freshening, November to April and May to October.

All records for cows whose sires and/or dams had unknown identification numbers and those records for cows whose sire had less than 10 daughters were discarded. Searle (1964) reported that the efficiency of a sire's merit is dependent on the number of daughters available and on heritability of the trait concerned. Henderson (1973) reported that the decreased number of daughters per sire would reduce the accuracy of predicted sire breeding value. Abubakar *et al.*, (1986b) suggested that 10 daughters per sire be considered minimum for progeny testing purposes in tropical areas by using BLUP procedures.

Records were eliminated if: (1) first parity was before 24 months or after 36 months of age; (2) number of days in milk were less than 180 days; (3) a HYS contains daughters of only one sire. It was important to check that each HYS had progenies of at least two sires to avoid the complete confounding of sires and HYSs or having disconnected sub sets of data. Therefore, the minimum number of sires per HYS was two sires and the maximum was 14 sires; the minimum number of daughters per HYS was two daughters and the maximum was 99 daughters.

Data editing left 1,812 valid 305-day first lactation records daughters of 79 sires, each sire had at least 10 daughters, dispersed in 94 HYS.

To account for genetic trend, five genetic groups were arbitrarily defined according to the year of birth of sires. A genetic group consisted of those sires whose birth date were within 4 to 6 years.

Age at first calving had been incorporated in the models of the analyses as a fixed effect to adjust for age differences at calving within first parity. Ignoring certain fixed effects actually present in the model leads to biased estimators, Henderson (1975 c).

Two analyses were performed using Harvey (1990) program to obtain BLUP of random effects, best linear unbiased estimates (BLUE) of fixed effects and restricted maximum likelihood estimates (REML) of the variance components. One was based on a population of unrelated sires, i.e., without including the A^{-1} , and the other based on a population of related sires, i.e., with including the A^{-1} .

The following model, in matrix notation, was used:

$$y = X_1b_1 + X_2b_2 + ZQg + Zs + e \quad [1]$$

where:

y denotes a vector of observations of milk yield;
 b_1 , b_2 and g denote unknown vectors of herd-year-season, age at first calving and group of sires fixed effects, respectively;

s denote unknown vector of sire random effect;
 X_1 , X_2 , Q and Z denote known incidence matrices for fixed and random effects;

e denote nonobservable random vector of errors.

$$E(y) = X_1b_1 + X_2b_2 + ZQg; \quad E(s) = E(e) = 0;$$

$V(e) = I\sigma_e^2$ and $V(s) = Ik^{-1}$; in the case of including the A^{-1} $V(s) = Ak^{-1}$; (I) is an identity matrix; (A) is the relationship matrix, the numerator of Wright (1922) coefficient of relationship; and $k = \sigma_e^2/\sigma_s^2$, where σ_s^2 is the component of variance due to sires and σ_e^2 is the error variance component.

The mixed model equations, Henderson (1975 c) and Quaas and Pollak (1981), for [1] are symbolized in equations [2.a] in the case of not including the A^{-1} and [2.b] in the case of including A^{-1} .

$$\begin{bmatrix} X_1'X_1 & X_1'X_2 & X_1'ZQ & X_1'Z \\ X_2'X_1 & X_2'X_2 & X_2'ZQ & X_2'Z \\ Q'Z'X_1 & Q'Z'X_2 & Q'Z'ZQ & Q'Z'Z \\ Z'X_1 & Z'X_2 & Z'ZQ & Z'Z + Ik \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{g} \\ \hat{s} \end{bmatrix} = \begin{bmatrix} X_1'y \\ X_2'y \\ Q'Z'y \\ Z'y \end{bmatrix} \quad [2.a]$$

$$\begin{bmatrix} X_1'X_1 & X_1'X_2 & X_1'ZQ & X_1'Z \\ X_2'X_1 & X_2'X_2 & X_2'ZQ & X_2'Z \\ Q'Z'X_1 & Q'Z'X_2 & Q'Z'ZQ & Q'Z'Z \\ Z'X_1 & Z'X_2 & Z'ZQ & Z'Z + \lambda^{-1}k \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{g} \\ \hat{s} \end{bmatrix} = \begin{bmatrix} X_1'y \\ X_2'y \\ Q'Z'y \\ Z'y \end{bmatrix} \quad [2.b]$$

The mixed model equations [2.a] and [2.b] with HYS equations absorbed can be rewritten as in equations [3.a] and [3.b]

$$\begin{bmatrix} X_2'MX_2 & X_2'MZQ & X_2'MZ \\ Q'Z'MX_2 & Q'Z'MZQ & Q'Z'MZ \\ Z'MX_2 & Z'MZQ & Z'MZ + Ik \end{bmatrix} \begin{bmatrix} \hat{\beta}_2 \\ \hat{g} \\ \hat{s} \end{bmatrix} = \begin{bmatrix} X_2'y \\ Q'Z'y \\ Z'y \end{bmatrix} \quad [3.a]$$

$$\begin{bmatrix} X_2'MX_2 & X_2'MZQ & X_2'MZ \\ Q'Z'MX_2 & Q'Z'MZQ & Q'Z'MZ \\ Z'MX_2 & Z'MZQ & Z'MZ + \lambda^{-1}k \end{bmatrix} \begin{bmatrix} \hat{\beta}_2 \\ \hat{g} \\ \hat{s} \end{bmatrix} = \begin{bmatrix} X_2'My \\ Q'Z'MY \\ Z'My \end{bmatrix} \quad [3.b]$$

where :

$$M = I - X_1 (X_1' X_1)^{-1} X_1'$$

Equations [3.a] and [3.b] were solved to obtain solutions for \hat{g} , and \hat{s} . For each sire evaluated, $u = Q\hat{g} + \hat{s}$, Henderson (1973) and Quaas and Pollak (1981). That is, a sire proof is an estimate of the sum of the sire effect plus that sire's group solution.

Equations [3.a] and [3.b] were solved using Harvey (1990) program to obtain the variance components for sires (σ_s^2) and error (σ_e^2) as minimum norm quadratic estimators (MINQUE). Iterative MINQUE estimators equal to REML estimators, Searle (1979). Therefore, σ_s^2 and σ_e^2 were obtained as REML estimators. Heritability estimate was computed as:

$$h^2 = 4\sigma_s^2 / (\sigma_s^2 + \sigma_e^2).$$

Where:

h^2 = heritability, σ_s^2 = Sires variance, and σ_e^2 = error variance.

RESULTS AND DISCUSSION

The 305-day milk overall mean was $1,967 \pm 445$ kg. The TA estimates of 79 sires predicted when A^{-1} was omitted ranged from 182 kg to -143 kg with an average difference of 4.11 kg among sires. The TA estimates when the A^{-1} was included stretched from 197 kg to -143 kg with an average difference among sires of 4.31 kg. Schaeffer et al. (1973) evaluated 63 Holstein sires and found that the TA estimates ranged from 456 kg to -520 kg with an average difference of 15.5 kg among sires. Abubakar et al. (1986 a) evaluated 60 Holstein sires under tropical conditions and found that the predicted sires values ranged from 358 kg to -340 kg with an average difference of 12 kg among sires.

Out of all sires, 41 sires showed positive TA estimates and 38 sires showed negative TA estimates. This distribution did not change by using the A^{-1} .

The difference between sires TA estimates with and without including the A^{-1} , showed no specific trend and many deviations were large. The greatest absolute change was 47.86 kg for a sire with 14 daughters. Table 1 shows the distribution of the absolute difference among TA estimates obtained with and without the A^{-1} . There were 59 sires, represent about 74.7 %, showed an absolute difference of less than 10 kg and 20 sires, represent about 25.3 %, showed an absolute difference of 10 kg or more.

As a result of using the relationship matrix, there were 44 sires out of 79 sires, showed positive difference which did indicate that the TA estimates increased by using the A^{-1} . However, there were 35 sires showed negative differences.

Table 1. Distribution of the absolute difference among TA estimates with and without including the A^{-1} .

Absolute Difference, kg	No. Sires	%
< 10	59	74.7
10 - 19	13	16.5
20 - 29	3	3.8
30 - 39	2	2.5
40 - 49	2	2.5

The product moment correlation (SAS, 1988) was 0.96 among sires' TA estimates with and without including the A^{-1} . Rank correlations are not a good indicator of ranking changes. Even with 0.98 rank correlations, using the A^{-1} changed the rank of 65 sires out of 79 sires. Rank correlations of a whole list of sires are not a sensitive measure of ranking changes in the top and bottom portion of the list. For any stud the top and bottom ranking sires, say top and bottom 5 %, are very important and have a large impact on the genetic improvement. Culling decisions are based on the bottom portion, usually more than the bottom 5 % would be culled, and decisions for intensively using sires are usually on the top portion. In this study, there was not enough number of sires, only 79 sires, the top ten sires and the bottom ten sires were used to show the effect of using the relationship matrix on sires' rank. The rank correlations were 0.62 for the top ten sires and 0.99 for the lowest 10 sires.

Table 2 presents the TA estimates and the rank of the top ten sire with and without including the A^{-1} . The first two sires have the same rank with or without including the A^{-1} . The remaining top sires ranked differently when the A^{-1} was included. Two sires were not in the top ten sires according to their TA estimates without the A^{-1} . They ranked the eleventh and the

fourteenth. However, after including the A^{-1} they ranked the seventh and the sixth, respectively.

Table 2. TA (BLUP) estimates of the top ten sires with and without using the A^{-1} in kg.

Sire ID	No. of Daughters	With A^{-1}		Without A^{-1}		Difference. ^a
		BLUP	Rank	BLUP	Rank	
51663	38	197	1	182	1	15
55360	20	171	2	170	2	1
51955	14	154	3	106	4	48
29080	21	129	4	122	3	7
29522	65	103	5	96	5	7
60311	10	97	6	59	14	38
51589	13	93	7	63	11	30
58125	12	90	8	77	9	13
57958	22	88	9	87	6	1
51208	19	87	10	81	7	6

^a Difference = BLUP with A^{-1} - BLUP without A^{-1} .

Some of the top ten sires showed large difference with using the A^{-1} in both the TA estimates and the rank. The maximum absolute difference among TA estimates with and without using the A^{-1} for the top ten sires was 48 kg of milk for one sire and the rank for that sire changed from 4 to 3.

Table 3 shows variance components and heritability estimates with and without including the A^{-1} . Using the A^{-1} increased the total variance by 149.7 kg². There are vast differences in variance components estimates in the literature, Schaeffer *et al.* (1975) and Abubakar *et al.* (1986a). These large differences may be attributed to the enormous differences of the records used and/or different method of estimations. Also, there is a scale effect on variances. Larger means are associated with larger variances. In general, very little changes in heritability and variance components estimates were revealed as a result of using the relationship matrix. Using the relationship matrix caused a little increase in sire variance and heritability estimates and little decrease in error variance. Kennedy and Moxley (1975); and Jensen (1980) reported that use of relationships

reduces the variance of error of prediction and they added, in general, inclusion of the additive genetic relationship matrix increases accuracy of genetic evaluations of the sires. Henderson (1975 b) reported that using relationships among sires increases accuracy of the prediction, particularly for sires with few or no progeny.

Table 3. Variance components estimates for sires (σ_s^2) and error (σ_e^2) and heritability (h^2) for milk with and without using the A^{-1} .

	σ_s^2	σ_e^2	h^2
Without (A^{-1})	8,044.3	189,293.4	.16
With (A^{-1})	8,282.0	189,205.4	.17

In conclusion, relationship ties do the same things as reference sires having progeny in many different contemporary groups. It also gives predictions that include the paternal half-sib information that is available. This concept is equivalent to the selection index theory approach which combines sources of information into one predicted value. Wood *et al.*, (1991) reported that the accuracy of evaluation of moderately to highly heritable traits can be improved by using a single-trait mixed model with relationships. To maximize genetic progress the best sires have to be selected. Increasing the accuracy of selection could be accomplished in part by choosing the right model for the evaluation. Using the additive genetic relationship matrix in the evaluation model resulted in large differences in sires' TA estimates and consequently sires ranking. The importance of the relationship matrix increases as the number of daughters by sire decreases (Suzuki and Mitsumoto, 1986).

REFERENCES

- Abubakar, B.Y., R.E. McDowell and L.D. Van Vleck, 1986a. Genetic evaluation of Holsteins in Colombia. J. Dairy Sci. 69:1081.

- Abubakar, B.Y., R.E. McDowell, K.E. Wellington and Van Vleck, L.D. 1986b. Estimating genetic values for milk production in tropics. *J. Dairy Sci.* 69:1087.
- Butcher, K.R. and J.E. Legates, 1976. Estimating Son's progeny test from his pedigree information. *J. Dairy Sci.* 59:137.
- Harvey, W.R., 1990. User's guide for LSMLMW and MIXMDL PC-2 version, mixed model least-squares and maximum likelihood computer program. Ohio State University, Columbus.
- Henderson, C.R., 1973. Sire evaluation and genetic trends. In proceedings of the Animal Breeding and Genetics Symposium in Honor of Dr. Jay L. Lush, 10-41. American Society of Animal Science and American Dairy Science Association, Champaign, Illinois.
- Henderson, C.R., 1975a. Rapid method for computing the inverse of relationship matrix. *J. Dairy Sci.* 58:1727.
- Henderson, C.R., 1975b. Use of relationships among sires to increase accuracy of sire evaluation. *J. Dairy Sci.* 58:1731.
- Henderson, C.R., 1975c. Comparison of alternative Sire evaluation methods. *J. Anim. Sci.* 41:760.
- Jensen, E.L., 1980. Bull groups and relationships among sires in best linear unbiased prediction sire evaluation models. *J. Dairy Sci.* 63:2111.
- Kennedy, B.W. and J.E. Moxley, 1975. Comparison of genetic group and relationship methods for mixed model evaluation. *J. Dairy Sci.* 58:1507.
- Keown, J.F., 1974. Comparison of mixed model methods of sire evaluation. *J. Dairy Sci.* 57:245.
- Powell, R.L. and A.F. Freeman, 1974. Estimators of sires merit. *J. Dairy Sci.* 57:1228.
- Quaas, R.L. and E.J. Pollak, 1981. Modified equations for sire models with groups. *J. Dairy Sci.* 64:1868.
- Robinson, G.K., 1986. Group effects and computing strategies for models for estimating breeding values. *J. Dairy Sci.* 69:3106.
- SAS, 1988. SAS procedures guide. Release 6.03 Edition, SAS Institute Inc., Cary, NC USA. p 130.

- Schaeffer, L.R., R.W. Everett and C.R. Henderson, 1973. Lactation records adjusted for days open in sire evaluation. *J. Dairy Sci.* 56:602.
- Schaeffer, L.R., M.G. Freeman, and E.B. Burnside, 1975. Evaluation of Ontario Holstein dairy sires for milk and fat production. *J. Dairy Sci.* 58:109.
- Searle, S.R. 1964. Progeny-testes of sire and son. *J. Dairy Sci.* 47:414.
- Searle, S. R., 1979. Notes on variance component estimation: A detailed account of maximum likelihood and kindred methodology. Paper BU-673M. Biometrix Unit, Cornell University, Ithaca, N.Y. USA. p 117.
- Suzuki, M. and T. Mitsumoto, 1986. A comparison of three models for BLUP sire evaluation of dairy cattle in Hokkaido. *Japanese Journal of Zootechnical science.* 57:718-725.
- Wood, C.M., L.L. Christian and M.F. Rothschild, 1991. Use of an animal model in situations of limited subclasses numbers and high degrees of relationships. *J. Anim. Sci.* 69:1420.
- Wright, S., 1922. Coefficients of inbreeding and relationship. *American Nat.* 56,330-8.

اثر استخدام مصفوفة علاقة القرابة التجمعية على تقويم الطلائق

محمد حسين صادق^١ - عرفة عطية حلاوة^٢ - عبد الحليم عثماوى^١
 ١- قسم الانتاج الحيوانى - كلية الزراعة - جامعة عين شمس ٢ - معهد بحوث
 الانتاج الحيوانى - وزارة الزراعة

تم اجراء تحليلين احصائيين احدهما باستخدام مقلوب مصفوفة علاقة القرابة التجمعية بين الطلائق والاخر بدونه . استخدم فى التحليلين نموذج رياضى مختلط اشتمل على: (قطيع- سنة - موسم) ، العمر عند اول ولادة، مجاميع الطلائق كعوامل ثابتة ، واثر الطلائق كعامل عشوائى ، اشتملت البيانات المستخدمة على ١٨١٢ سجل اول موسم حليب انتاج ٣٠٥ يوم بنات ٧٩ طلوقة. تم جمع هذه البيانات خلال الفترة من عام ١٩٦٠ الى عام ١٩٨٩. تم تحديد موسمين للولادة الاول من نوفمبر الى ابريل والثانى من مايو الى اكتوبر وبالتالي اشتملت البيانات المستخدمة على ٩٤ قطيع - سنة - موسم .

بلغ متوسط انتاج الـ ٣٠٥ يوم لأول موسم حليب ١٩٦٧ ± ٤٤٥ كجم تراوحت قيمة المقدرة التمريرية للطلائق من ١٨٢ كجم الى -١٤٣ كجم بمتوسط فرق ٤,١ كجم وذلك فى حالة عدم استخدام مقلوب مصفوفة القرابة بينما تراوحت هذه القيمة بين ١٩٧ كجم الى -١٤٣ كجم بمتوسط فرق ٤,٣ كجم فى حالة استخدام المصفوفة. بلغت قيمة معامل الارتباط ٠,٩٦ بين تقديرات المقدرة التمريرية للطلائق باستخدام وبدون استخدام مقلوب مصفوفة علاقة القرابة . بلغ معامل الارتباط ٠,٩٨ بين ترتيب قيم تقديرات المقدرة التمريرية باستخدام وبدون استخدام مقلوب مصفوفة القرابة . بلغت قيمة هذا الارتباط ٠,٦٢ لأفضل ١٠ طلائق بينما بلغت ٠,٩٩ لأقل ١٠ طلائق . نتج عن استخدام مقلوب مصفوفة القرابة تغيرات طفيفة فى مكونات التباين والمكافئ الوراثى لصفة انتاج اللبن .