GENETIC AND PHENOTYPIC PARAMETERS OF NUMBER OF SERVICES PER CONCEPTION AS A THRESHOLD TRAIT AND 305 MILK YIELD AS CONTINUOUS TRAIT USING GIBBS SAMPLING AND MTDFREML

E. Mousa[1], Hanaa Abdelharith[2] and Manal Elsayed[3]

1- Animal Production Depatment, Faculty of Agriculture, Assiut University, Assiut, Egypt, 2- Animal Production Research Institute, Ministry of Agriculture, Dokki, Giza, Egypt, 3- Animal Production Depatment, Faculty of Agriculture, Ain Shams University, Shoubra, Egypt

SUMMARY

Gibbs Sampling (GS) and Mulltiple-trait Derivative-Free Restricted Maximum Likelihood (MTDFREML) were utilized to estimate and compare the variance components and genetic parameters of number of services per conception (NSC) measured as a binary trait and 305-day milk yield (305-MY) as a continuous trait.
A total of 2572 records obtained from 1491 Friesian cows raised at two experimental farms of Ministry of Agriculture, Egypt over 15 years were used.
Heritability estimates were 0.04 and 0.22 by MTDREML and 0.11 ± 0.03 and 0.28 ± 0.05 using GS for the two studied traits, respectively. Repeatability estimates were 0.11 and 0.33 by MTDREML and 0.17±0.02 and 0.45±0.06 by GS, respectively. Genetic and phenotypic correlations between the two traits were 0.34 and .06 using MTDFREML and 0.25 and 0.10 using GS, respectively.
The obtained results revealed that GS estimates for heritability and repeatability were 2.75 and 1.27 times those by MTDFREML for NSC and 1.55 and 1.36 time for 305-MY, respectively. Although, heritability estimates for the threshold trait were not similar from using the two procedures, they were all low indicating that the direct genetic selection to improve this trait would be ineffective.

Keywords: Friesian, genetic parameters, REML, Gibbs sampling, genetic and phenotypic correlations.
Abbreviation key: NSC = Number of services per conception, 305-MY= 305 day milk yield MTDFREML = Multiple Trait Derivative Free-REML, GS = Gibbs Sampling.

INTRODUCTION

Estimation of variance components (VC) has long been an important consideration for the genetic improvement of production traits. Several traits of interest in animal breeding are measured as binary variables. Analysis of such variables by linear methodology violates several assumptions of linear model and is not optimal (Gianola, 1982). A more satisfactory method is based on the threshold model concept (Wright, 1934). More recently, Gibbs Sampling (GS) has been

extended to Bayesian inference in threshold model for binary data (Albert and Chib, 1993; Sorensen *et al.*, 1995). Wang et al. (1994, 1997) and Van Tassell et al. (1995) have developed statistical treatment of the threshold model for the animal model. The multiple-trait Gibbs Samples for animal model (MTGSAM) program has been developed to implement the GS algorithm for Bayesian analysis of a broad range of animal models (Van Tassell and Van Vleck, 1995,1996).

The purpose of this study was to compare posterior mean estimates using GS for genetic parameters with those obtained using MTDFREML in the case of a continuous trait (305-MY) and in the other case of a binary trait (NSC) using the same model in the two procedures.

MATERIALS AND METHODS

Data represent reproductive and milk yield performance of 1491 lactating Friesian cows from two experimental farms (Sakha and Alkarada) belonging to the Animal Production Research Institute (APRI), Minsitry of Agriculture, Egypt. The data covered the period from 1985 to 1999 making 2572 records of 1491 Friesian cows progeny of 102 sires and 719 dams.

Measures of reproductive performance included number of services per conception as a binary trait while; production performance included the 305 days milk yield as a continuous trait. These cows were distributed over 15 years. Each year was divided into two calving seasons: warm (mid May to November) and mild (December to mid-May). Six weeks after calving, cows were observed for estrous twice daily (08:00 and 16:00 hrs). Cows were observed for mounting activity, vulval mucous discharge, restlessness and other signs of heat. Each morning, inseminations were performed to cows, which were first seen in estrous during the afternoon of the previous day. Cows showing heat sign in the morning were inseminated in the afternoon on the same day.

Cows were inseminated artificially using frozen semen locally prepared in the International Livestock Management Training Center. According to the farm routine, it was not allowed to inseminate a cow earlier than forty days after calving. Reproductive tract was examined via rectal palpation two months after insemination to detect the pregnancy.

Model:

The MTDFREML programs of Boldman *et al.* (1995) were used to obtain REML estimates of (co) variance components. Principles of derivative-free restricted maximum likelihood have been described by Smith and Graser (1986) and Meyer (1989) and reviewed by Boldman *et al.* (1995).

As an alternative method to estimate variance components (VC) through REML, GS is investigated here. GS is one method in a larger class of methods, referred to Monte Carlo Marcov Chain methods. Gibbs Sampling is based on Bayesian methods for estimation of VC but is evaluated based on its frequents.

The general representation of the multiple-trait repeatability animal model was the following:

$$Y = xb + Z_a a + Z_e c + e$$

Where:

Y   is a "n" by 1 vector of records,
B   denotes the fixed effects in the model with association matrix X;
A   is the vector of direct animal genetic effects with association matrix Z_a;
C   is the vector of permanent environmental effects of the animal with association matrix Z_c; and
E   denotes the vector of residual effects.

The following multiple-trait repeatability animal model was used to obtain (co) variance components for the studied traits (NSC and 305-MY) through REML and GS procedures:

$$y_{ijklmn} = \mu + A_i + PE_i + f_j + p_k + y_l + s_m + b(X_{ijklmn} - \overline{X}) + e_{ijklmn}$$

where:

y_{ijklmn} = records of trait n (NS or 305-MY) for the mth farms of the lth parity of the kth year of the jth season of calving of the ith animal;
μ = an overall mean,
A_i = the random effect of the additive genetic effect of the animal;
PE_i = the random permanent environmental effect on the animal;
f_j = the fixed effect of the jth farm (2 levels);
p_k = the fixed effect of the kth parity (7 levels representing the first seven parities);
y_l = the fixed effect of the lth year of calving (15 levels);
s_m = the fixed effect of the mth season (2 levels); and
b   = is the partial linear regression coefficient of n trait on age at calving;
X_{ijklmn} = is the age at calving of the records of trait n (NSC or 305-MY) for the jth farms of the kth parity of the lth year of the mth season of calving of the ith animal;
$\overline{X}$ = is the average age at calving and
e_{ijklmn} = the random residual effect associated with each observation.

The variance-covariance structure for the model is as follows:

$$
V = \begin{bmatrix} a_1 \\ a_2 \\ c_1 \\ c_2 \\ e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} A\varsigma^2 a_1 & \Sigma a_1 a_2 & O & & O & \\ \sigma a_2 a_1 & A\sigma^2 a_2 & & & & \\ & & I\sigma^2 c_1 & \sigma c_1 c_2 & O & \\ & & \Sigma c_2 c_1 & I\sigma^2 c_2 & & \\ & O & & & In_1\sigma^2 e_1 & \sigma e_1 e_2 \\ O & & O & & \Sigma e_2 e_1 & In_2\sigma^2 e_2 \end{bmatrix}
$$

Where:

A	is the numerator relationship matrix;
$\Sigma^2 a_1, \sigma^2 a_2$	Are the direct genetic variance for traits 1 and 2;
$\Sigma^2 c_1, \sigma^2 c_2$	Are the variance due to permanent environmental effects;
$\Sigma a_1 a_2$	is the direct genetic covariance between the two traits;
$\Sigma c_1 c_2$	is the permanent environmental covariance between the two traits; and
$\Sigma e_1 e_2$	Is the error covariance between the two traits.

The convergence criterion used in this analysis was the variance of likelihood in the current set of simplex value used in the derivative-free REML algorithm (Boldman *et al.*, 1995). The desired range of this variance at convergence is 10^{-6} to 10^{-10}. According to the software authors (Boldman *et al.*, 1995) the estimation programs should be restarted repeatedly with the converged values and no prior simplex information until convergence occurs to the same values that were supplied as starting values. This approach should ensure that a global, rather than a local, maximum likelihood estimate are obtained. All estimates were based on restarted system where a Gibbs chain length of 250,000 was run for each trait of each case with burn in 10,000 rounds and were considered to be an effective number of rounds (Mousa and Van Vleck, 1998). In order to create the binary form of NSC from its continuous form (NSC), it was classified into two categories depending on the normal range of each of them for this breed. Cows conceiving after one, two or three services (75%) were considered to be normal and therefore they formed the first category. While, cows conceiving after 4 services or more (25%) were formed as the second category.

RESULTS AND DISCUSSION

Preliminary least squares analysis using PROC GLM option of SAS (SAS, 1996) showed highly significant effect of year of calving on the two traits (NSC and 305-MY) $p < 0.01$, while, parity had a highly significant effect on 305-MY ($p < 0.01$) and only significant on NSC ($p < 0.01$). On the other hand, season of calving had no significant effect on NSC ($p > 0.05$) but highly significant on 305-MY. The mean effect for NSC (1.25 ± 0.01 services) and for 305-MY (2324.41 ± 15.41 kg) were respectively, smaller and larger than corresponding values frequently reported in the literature for the same breed. Detailed results were presented by Abou-Bakr *et al.* (2000); El-Said *et al.* (2000) and Mousa *et al.* (2001).

Table 1 shows heritability and repeatability estimates for the studied traits. Estimates of heritability by GS were 2.8 and 1.3 times as those by MTDFREML for NSC and 305-MY, respectively. It is clear that estimates from both procedures were not comparable.

Table 1. MTDFREML and GS heritability (h² ± SD) and repeatability (t ± SD) estimates for number of services per conception (NSC) and 305-MY of Friesian cows as binary and continuous traits

	MTDFREML		GS	
	h²	T	h²	t
NS	0.04±.00	0.11	0.11 ± 0.03	0.17 ± 0.02
305-MY	0.22±.00	0.33	0.28 ± 0.05	0.45 ±0.06

In summary, the GS and MTDFREML estimates were quite different from each other, but the differences were higher in the binary form than for the continuous form of the trait.

From these results, it appears that GS had consistently higher estimates of heritability than MTDFREML probably due to the influence of the prior distribution of the variance component on the postrior distribution. These results are in good agreement with those observed in a pervious study (Van Tassel, 1994). Mousa and Elsayed (2002) added that the effect of prior distribution decreases as the heritability or the amount of information increase. This means that, the difference between estimates from the two procedures is large when the heritability estimate is low and the distribution of the trait is not continuous (Mousa *et al* 2002). The advantage of GS is the possibility to analyze the trait under the threshold model. Van Tassel, 1994; Mousa and Elsayed, 2002 reported that the GS estimates under threshold model appeared to be relatively unbiased, where none of estimators seemed to be superior based on biasedness.

MTDFREML estimate for heritability of NSC was 0.04 and that from GS was 0.11 as compared to previous study estimates of 0.008 (Moore *et al.*, 1990) and 0.060 (Berger *et al.*, 1981) for the Holstein cows. Therefore, the direct genetic selection to improve reproductive efficiency by different breeds would be meaningless. MTDFREML estimate for heritability of 305-MY was .22 and that from GS was .28. These estimates fall within the rang (from 0.14 to .31) found in the literature (Weller, 1989; Short and Lawlor, 1992;Marrti and Funk,1994 and Dematawewa and Berger , 1998; Abou-Bakr *et al.* 2000)

Phenotypic and genetic correlations between NSC and 305-MY were 0.06 0.34 by MTDFREML and 0.10 and 0.25 by GS, respectively. Although values for phenotypic and genetic correlations between the two traits differ in their respective values from one procedure to another, same positive relationship between the two traits existed. This trend is in agreement with the results of Dematawewa and Berger (1998); Abou-Bakr *et al. (*2000). These results indicate that the genetic correlation between 305-MY with NSC is moderate but undesirable. Therefor, the use of fertility measures in bull indices may be questionable.

Estimates of repeatability for number of services per conception (NSC) and 305 day milk yield (305-MY) of Friesian cow were 0.11 and .33 by MTDFREML and 0.17 and 0.45 by GS. The repeatability estimates from two procedures for NSC was somewhat higher than the estimates obtained from other studies (ranged from 0.07 to 0.08) for the Holstein cow (Hayes *et al.*, 1992; and Dematawewa and Berger,1998).Meanwhile, repeatability for 305-MY by MTDFREML was lower than those found in literature. In contrast, repeatability estimate for 305-My by GS falls

within the rang reported in the literature (Dematawewa and Berger, 1998; Marti and Funk,1994 and Welper and Freeman,1992)

CONCLUSION

A clear difference was observed between the estimates of genetic parameters from the two procedures. The differences between the estimates by the two procedures depended on the distribution of the considered trait and on the size of its variance components. Therefore, if the genetic variance is large (high h^2) and the trait is continuous the estimates by the two procedures would be very similar. In contrast, the binary traits with small genetic variance (low h^2) create more difference between the two procedures. This is due to that GS is adding a prior distribution to the final estimates of variance components i.e. the posterior distribution.

The heritability estimate of reproductive trait (NSC) was low. Therefore, the direct genetic improvement for this trait is expected to be ineffective. The positive genetic correlation found between milk yield and NSC trait may be considered as an evidence, that there could be an antagonistic relationship between yield and fertility.

ACKNOWLEDGMENT

REFERENCES

Abou-Bakr, S., U.M. El-Saied and M.A.M. Ibrahim, 2000. Genetic and phenotypic parameters for milk yield days open and number of services per conception of Holstein cows of commercial herd in Egypt. Egyptian Journal of Animal Production 37 (1): 9-17.

Albert, J.H., and S. Chib, 1993. Bayesian analysis of binary and polychotomous response data. J. Am. Stat. Assoc. 88: 669-679.

Berger, P.J., R.D. Shanks; A.E. Freeman and R.C. Laben, 1981. Genetic aspects of milk yield and reproductive performance. J. Dairy Sci., 64 (1): 114-122.

Boldman, K.G., L.A. kriese, L.D.Van Vleck, C.P Van Tassell, and S.D. Kachman, 1995. A manual for use of MTDFREML. A set of programs to obtain estimates of variance and covariances [DRAFT]. U.S. department of Agriculture, Agricultural Research Services. pp 114, 120.

Dematawewa, C.M.B. and P.J. Berger, 1998. Genetic and phenotypic parameters for 305 milk yield, fertility and survival in Holsteins. J. Dairy Sci., 81 (10): 2700-2709.

El-Saied, U.M.; E. Mousa; S. Abou-Bakr, 2000. Genetic parameters for milk yield and some reproductive traits of Holstein cows using REML and Gibbs sampling. Proceeding of the 3rd All Africa Conference on Animal Agriculture and 11th Conference of the Egyptian Society of Animal Production, Alexandria, Egypt 6-9 November 2000, pp:23-27.

Gianola, D., 1982. Theory and analysis of threshold characters. J. Animal Sci. 54: 1079-1096.

Hayes, J.F.; R.I. Cui and H.G. Monardes, 1992. Estimates of repeatability of reproductive measures in Canadian Holsteins. J. Dairy Sci., 75 (6): 1701-1706.

Marti, C.F. and D.A. Funk. 1994. Relationship between production and days open at different levels of herd production. J. Dairy Sci., 77(6): 1682-1690.

Meyer, R., 1989. Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free algorithm Genet Sel. Evol. 21: 317-340.

Moore, R.K., B.W. Kennedy; L.R. Schaeffer and J.E. Moxley, 1990. Relationships between reproduction traits age and body weight at calving and days dry in first lactation of Ayrshires and Holsteins. J. Dairy Sci., 73(3): 835-842.

Mousa, E. And Elsayed, 2002. Quality of variance components estimated by Gibbs Sampling in populations with and without selection and varying heritability.Prepared manuscript.

Mousa, E.; L.D. Van Vleck, 1998. Effective number of rounds of Gibbs sampler needed for estimation of parameters for threshold and continuous traits. Egyptian Journal Animal Production, 35, Suppl. Issue, Dec. 493-505.

Mousa, E; U.M.EL-Saied and M.A.M. Ibrahim, 2001. Genetic analysis for number of services per conception and days open as continuous and binary traits using Gibbs Sampling and REML I Holstein cows. In press

SAS, 1990. SAS statistics. Guide release 6.03. edition SAS Inst., Inc., Cary NC. USA.

Short, T.H. and T.J. Lawlor, 1992. Genetic parameters of conformation trait, milk yield and herd life in Holsteins. J. Dairy Sci., 75(7): 1987-1998.

Smith, S.P., and H.U. Graser, 1986. Estimating of Variance components in a class of mixed models by restricted maximum likelihood J. Dairy Sci. 69: 1156-1165.

Sorensen, D.A., S. Anderson, D. Gianola, and I. Korsgaad, 1995. Bayesian inference in Threshold models using Gibbs sampling. Genet Sel. Evol. 27: 229-249.

Van Tassel, 1994. The use of Gibbs Sampling for variance component estimation with simulated and weaning weight data using animal and maternal effects models. Ph.D. Diss., Cornell Univ. Ithaca, NY. pp 93-95.

Van Tassel, C.P. and L.D. Van Vleck, 1996. Multiple-trait Gibbs sampler for animal models: flexible programs for Bayesian and likelihood-based (co) variance component inference. J. Anim. Sci 74:2586-2597.

Van Tassell, C.P. and L.D. Van Vleck, 1995. A manual for use of MTGSAM. A set of Fortran programs to apply Gibbs sampling to animal models for variance component estimation [DRAFT]. U.S. Department Agriculture, Agriculture Research Service. pp. 86.

Van Tassell, C.P., G. Casella and E.J. Pollak, 1995. Effects of selection on estimation on estimates of variance components using Gibbs sampling and restricted maximum likelihood J. Dairy Sci. 78: 678-629.

Wang, C., S., J.J. Rutledge and D. Gianola, 1994. Baysian analysis of mixed linear model via Gibbs sampling with an application to litter size in Iberian pigs. Genet Sel. Evol. 26: 91-115.

Wang, C.S., R.L. Quaas, and E.J. Pollak, 1997. Baysian analysis of calving ease scores and birth weight. Genet Sel. Evol. 29: 117-143.

Weller, J.I., 1989. Genetic analysis of fertility traits in Israeli dairy cattle. J.Dairy Sci., 72(10): 2644-2650.

Welper, R.D. and Freeman,1992. Genetic parameter for yield traits of Holsteins, including lactose and somatic cell score. J.Dairy Sci., 75:1342-1348.

Wright, S., 1934. An analysis of variability in number of digits in an inbred strain of Guinea pigs Genetics 19: 506-536

المعايــير الوراثية والمظهرية لعدد التلقيحات اللازمة للحمل كصفة سلمية وإنتاج اللبن
فى ٣٠٥ يوم كصفة مستمرة بإستخدام طريقة عينات جبس والتشابه العظمى المحددة

عماد موسى¹   هناء عبد الحارث²   منال السيد³

١ قسـم الإنتـاج الحيوانى ، كلـية الـزراعة ، جامعة أسيوط، أسيوط-مصر،٢-معهد بحوث الإنتاج
الحيواني، الدقي،الجيزة-مصر، ٣-قسم الإنتاج الحيواني ، كلية الزراعة،جامعة عين شمس،شبرا،مصر

قــدرت مكونات التباين على ٢٥٧٢ سجل من ١٤٩١ بقرة فريزيان لمزرعتى سخا والقرضة التابعتين
لمركز بحوث الإنتاج الحيواني التابع لوزارة الزراعة.

اشــتمل النموذج الإحصائي على المزرعة (مزرعتين)، موسم الولادة (موسمين)، مواسم الإدرار (أول
سـبع مواسم)، سنة الولادة (١٥ سنة) وذلك كتأثير ثابت ، بينما استخدم التأثير الوراثى الوراثى التجميعي
والبيئي الدائم كتأثير عشوائي. صححت البيانات لتأثير عمر الأم عند الولادة.

حللـت السجلات لمقارنة مكونات التباين لصفة عدد التلقيحات اللازمة للحمل عندما تقاس كصفة سلمية
وصفة إنتاج اللبن فى ٣٠٥ يوم كصفة مستمرة باستخدام طريقتين مختلفتين.

الطريقة الأولى ،طريقة MTDFREML بإجراء التشابه العظمى المحددة، الطريقة الثانية طريقة عينات
جبس GS بأجراء تحليل بيزيان على نموذج الحيوان المتكرر لأكثر من صفة.

كانـت تقديـرات المكافــئ الوراثى ٠,٠٤ ، ٠,٢٢ باستخدام طريقة MTDFREML ٠,١١، ٠,٠٣،
٠,٢٨ ٠,٠٥ باستخدام طريقة GS على التوالي .

كــان الارتباط الوراثي والمظهري بين الصفتين ٠,٣٤ و٠,٠٦ باستخدام طريقة MTDFREML و
٠,٢٥ و ٠,١٠ باستخدام GS علي التوالي.

تشــير النتائج المتحصل عليها أن تقديرات المكافئ الوراثي والمعامل التكراري لصفة عدد التلقيحات
اللازمـة للحمـل باستخدام GS كانـت اكـبر ٢,٧٥ و ١,٣٦ مـرة من تلك المقدرة باستخدام
MTDFREML و ١,٥٥ و ١,٣٦ مرة لصفة إنتاج اللبن في ٣٠٥ يوم (صفة مستمرة) علي التوالي .
علــي الرغم من أن النتائج المتحصل عليها في كلتا الطريقتين غير متشابهة إلا انه لا يزال تقديرات المكافئ
الوراثي منخفضة مما يشير إلى أن الانتخاب الوراثي المباشر لتحسين هذه الصفة غير كفء.