

QUALITY OF HERITABILITY ESTIMATES AS AFFECTED BY LEVEL OF HERITABILITY, NUMBER OF PROGENY PER SIRE, TYPE OF ALGORITHM, TYPE OF MODEL AND TYPE OF TRAIT

Reda Elsaid¹, Manal Elsayed², E.S.E. Galal² and H.M.K. Mansour²

1- Department of Animal Production, Environmental Studies and Research Institute, Menoufia University, Sadat City, Egypt 2- Department of Animal Production, Faculty of Agriculture, Ain Shams University, Shoubra El-Khaima, Cairo, Egypt

SUMMARY

Two simulation programs were used in this study, one to simulate a continuous trait and another to modify this trait into a binary trait. Twelve populations were created (three levels of heritability (h^2), 0.10, 0.25 and 0.50; four levels of number of progeny per sire, 5, 10, 15 and 20), each with three parities as the only fixed effect. Twenty replicates were generated for each population. Each replicate was analyzed twice, once with sire model and another with animal model, using two algorithms for each model (MTDFREML or Gibbs Sampling (GS)). Bias and mean squared errors (MSE) of heritability estimates were used to assess the quality of heritability estimates obtained by different models and different algorithms. The effect of h^2 level, number of progeny per sire, type of algorithm, type of model, type of trait and the interactions on the bias and MSE were examined. All main effects were highly significant ($p < 0.0001$). For estimating variance components, for a continuous trait, the animal model was the best in the case of using MTDFREML and GS at all levels of h^2 . Also, at all levels of h^2 , the GS was the best algorithm in the analysis of a binary trait. For a binary trait within GS, the sire model was the best at h^2 equals to 0.1 with number of progeny more than 5 whereas, at h^2 equals to 0.25 or 0.5 with 20 progeny per sire, the use of animal model was equivalent to the use of sire model. At all levels of h^2 , the 20 progeny per sire had the lower MSE for heritability.

Keywords: *Continuous traits, binary traits, heritability estimates quality-Gibbs sampling, bias*

INTRODUCTION

Estimation of variance components is always an important tool in developing animal breeding programs. Estimates of variance components must be accurate since error variance for predicted breeding values increases as differences between estimated and true value of variance components increase (Schaeffer, 1984).

Type of trait (binary or continuous), number of progeny per sire, type of model (animal or sire model), type of algorithm (MTDFREML or Gibbs Sampling) and heritability magnitude (low, moderate or high) are all important factors that could affect the estimation of variance components.

For categorical traits, genetic parameters are usually computed from sire or animal variances and (co)variances. Heritability estimates resulting from threshold model for such traits were higher than those resulting from linear model (Luo *et al.*, 2001); the primary reason for the difference being that the heritability from the linear model is expressed on the observed scale while heritability from the threshold model is on an underlying liability scale (Luo *et al.*, 2001).

In animal model, all relationships are considered, whereas in sire model only relationships among half-sib progeny of sires are taken into account; which could lead to some bias in the estimates from sire model (Mrode, 1996). The threshold animal model using Gibbs Sampling (GS) may yield biased estimates, so the threshold sire model (or sire maternal grandsire model for maternal traits) is an alternative model for genetic analysis of categorical trait (Luo *et al.*, 2001).

The objective of this study was to investigate the effect of heritability level, number of progeny per sire, type of algorithm, type of model and type of trait on the quality of the heritability estimates as judged by bias and mean squared errors (MSE).

MATERIALS AND METHODS

Simulation procedure:

Two methods of simulation were used to generate samples for the present study. One is concerned with the underlying continuous response variable generation and another with changing this continuous variable into a binary variable with two categories 0 and 1. The first method is a Mont Carlo simulation technique using SAS (1996) with assumed mean (0) and variance (1). Analla *et al.* (1995) reported that this technique also assumes that the expected genetic value of the progeny G_k is equal to the average genetic values of the parents [sire (S_i) and dam (D_j)] plus a deviation due to the Mendelian sampling as follows:

$$G_k = 0.5(S_i + D_j) + X \sqrt{0.5h^2\sigma_p^2}, \quad (\text{Model 1})$$

where:

G_k is equal to the genetic value of an individual k, a progeny of sire (S_i) and dam (D_j),

X is random number taken from normal distribution with mean 0 and variance 1,

h^2 is the heritability and

σ_p^2 is the phenotypic variance.

Table (1) shows values of parametric phenotypic, genetic, permanent environmental and residual variances used to generate the studied samples.

With three levels of heritability (0.1, 0.25 and 0.5) and four classes for the number of half-sib progeny per sire (5, 10, 15 and 20), twelve populations were simulated with three levels of parity, as the only fixed effect, and twenty samples (replicates) for each population were generated. Numbers of records generated in each level of parity are shown in Table (2).

Table 1. Assumed parametric values of phenotypic, genetic, permanent environmental and residual variances at three levels of heritability

Parameter	Heritability		
	0.10	0.25	0.50
Phenotypic variance (σ_p^2)	0.250	0.2500	0.250
Additive genetic variance (σ_a^2)	0.025	0.0625	0.125
Permanent environmental variance (σ_{ep}^2)	0.006	0.0150	0.030
Residual variance (σ_e^2)	0.219	0.1725	0.095

Note: The mean of the simulated variable was constant at 0.5 (the best mean value of the simulated trait to maintain the average of the binary trait as it is in the continuous trait) and the permanent environmental variance (σ_{ep}^2) was given as around one quarter of the additive genetic variance (σ_a^2) (as indicated by Al-shorepy and Notter, 1996).

Table 2. Number of records in the simulated population in each of the three levels of parity

No. of sires	Number of half-sib daughters for each sire	No. of records
50	05	0250
50	10	0500
50	15	0750
50	20	1000

Each sample was categorized using a random variety from a binomial distribution (RANBIN Function) with SAS (1996) to obtain the binary response variable studied. So, two copies of each generated sample were obtained, the first contained the underlying continuous variable and the second contained the binary response.

Statistical analysis:

Heritability estimates of the studied variable were estimated for each copy of each sample in the 12 simulated populations (i.e. three levels of heritability and four family size), obtained from the animal and sire models each using two algorithms (multiple trait animal model program (MTDFREML) proposed by Boldman *et al* (1995) and Gibbs Sampling program proposed by Van Tassell and Van Vleck (1995).

The linear animal model used for continuous and binary traits was:

$$Y = X\beta + Z_a a + Z_c c + e, \tag{Model 2}$$

where,

- y is the vector of observation;
- X is the incidence matrix for fixed effects;
- β is the vector of an overall mean and parity (3 classes);
- Z is the incidence matrix for random effects;
- a is the vector of direct genetic effects of cow;
- c is the vector of permanent environment effects; and
- e is a vector of random errors normally and independently distributed with zero mean and variance $\sigma_e^2 I$.

The linear sire model used for continuous and binary traits was:

$$y = X\beta + Z_{ss} s + Z_c c + e, \tag{Model 3}$$

where,

s is the vector of direct genetic effects of sire; and other terms in the model are defined as in model 2.

The threshold model (Gianola and Foulley, 1983) was used for the analysis of binary response using Gibbs Sampling Program (Heringstad *et al.*, 2001).

To measure the correspondence between assumed parametric and estimated values, the estimates of bias in heritability estimates were calculated as the difference between the heritability values obtained from each analysis and the parametric value (Elsayed, 1997). The bias was calculated as follows:

$$\text{Bias} = [E(b^R) - B] \quad (\text{Neter } et al., 1985)$$

where,

b^R the expected value of the deviation of the biased estimator from the true parameter B .

The MSE (equals the variance of the estimator plus the squared bias) was calculated as follows :

$$MSE = E(b^R - B)^2 = \sigma^2 (b^R) + [E(b^R) - B]^2 \quad (\text{Neter } et al., 1985),$$

Analysis of variance was performed to study the effect of heritability level, number of progeny per sire, type of algorithm, type of model and type of trait on the estimates of bias and MSE.

The following model was applied using SAS (1996) to analyze the bias:

$$Y_{ijklmn} = \mu + h_i + n_j + a_k + m_l + t_m + e_{ijklmn}, \quad (\text{Model 4})$$

All possible significant interactions were included in the analysis.

where,

Y_{ijklmn} is the dependent variable of the n^{th} record in the i^{th} heritability, j^{th} number of progeny, k^{th} type of algorithm, l^{th} type of model and m^{th} type of trait;

μ the overall mean of bias;

h_i the effect of the i^{th} heritability, $i=1$ to 3;

n_j the effect of the j^{th} number of progeny, $j=1$ to 4;

a_k the effect of the k^{th} type of algorithm, $k=1$ and 2;

m_l the effect of the l^{th} type of model, $l=1$ and 2;

t_m the effect of the m^{th} type of trait, $m=1$ and 2 and

e_{ijklmn} the effect of random error, associated with each observation assumed to be normally and independently distributed with 0 mean and variance σ_e^2 .

The same model was used using SAS (1996) to analyze the mean squared errors.

Preliminary analysis with full model (including all main effects and all possible interactions) was performed to identify significant terms, then the analysis was repeated with only significant terms retained.

RESULTS AND DISCUSSION

Table (3) shows analysis of variance for the bias and MSE of h^2 estimates. The means of the main effects for these criteria are shown in Table (4).

The effects of the heritability level, number of progeny per sire, type of algorithm, type of trait and type of model on bias and MSE were all significant ($p < 0.0001$). Table (4) shows that except for continuous trait, the general mean estimate of magnitude of bias as well as the means of all main effects were negative and different significantly from zero ($p < 0.0001$).

Table 3. Analysis of variance of bias and mean squared errors (MSE) for heritability estimates

Source of variation	D.F.	Bais		D.F.	MSE	
		M.S.	Pr		M.S.	Pr
Heritability (H)	2	0.75494	0.0001	2	0.03982	0.0001
Number of Progeny (NO)	3	0.02320	0.0001	3	0.00309	0.0001
Type of Algorithm (ALG)	1	1.93231	0.0001	1	0.02144	0.0001
Type of model (MOD)	1	0.37325	0.0001	1	0.00818	0.0001
Type of trait (TR)	1	6.68706	0.0001	1	0.14526	0.0001
H*NO	6	0.00403	0.0209	6	0.00029	0.0004
H*ALG	2	0.18814	0.0001	2	0.00517	0.0001
H*MOD	2	0.00630	0.0205			
H*TR	2	1.02878	0.0001	2	0.03414	0.0001
NO*ALG				3	0.00040	0.0005
NO*MOD	3	0.02966	0.0001	3	0.00085	0.0001
NO*TR						
ALG*MOD	1	0.19461	0.0001	1	0.00446	0.0001
ALG*TR	1	0.52835	0.0001	1	0.05576	0.0001
MOD*TR	1	0.03790	0.0001	1	0.00574	0.0001
H*NO*MOD	6	0.00688	0.0003	6	0.00029	0.0004
H*NO*TR						
H*ALG*MOD				2	0.00058	0.0002
H*ALG*TR	2	0.08733	0.0001	2	0.01091	0.0001
H*MOD*TR	2	0.01126	0.001			
NO*ALG*TR				6	0.00022	0.0034
ALG*MOD*TR	1	0.05907	0.0001	1	0.00207	0.0001
H*ALG*MOD*TR	4	0.00614	0.0045			
Remainder	1878	0.00162		49	0.00001	
C.V.%		102.3			9.7	
R ²		0.8239			0.9938	

C.V.= Coefficient of variation, R² = Coefficient of determination and Pr = Probability of type I error.

Model included only those significant effects indicated in a preliminary full model analysis.

Table (4) also shows that the estimates of the magnitude of bias and MSE generally increased as h^2 increased whereas MSE decreased as number of progeny per sire increased but the magnitude of bias increased as number of progeny increased from 5 to 15 progeny (0.030, 0.043 and 0.044), then slightly decreased for 20 progeny (0.041). This result is in agreement with those reported by Thomas *et al.* (2000) who reported that bias in estimates of heritability decreased with decreasing simulated heritability and increasing sample size and also with those reported by Thomas and Hill (2000) who showed that the MSE decreased as sample size increased and simulated heritability increased. This result is also in agreement with the contents of a personal communication with Dr. Curt Van Tassell and Dr. Dale Van Vleck who suggested that the reason of decreased bias and MSE is the constrain on estimating variance components in REML so that negative estimates are not allowed and the range of bias in case of low h^2 is smaller than in case of medium or high h^2 . Generally, smaller magnitude of bias and MSE was found for GS compared to MTDFREML (0.008 vs 0.071 and 0.005 vs 0.016), for sire model compared to animal model (0.025 vs 0.053 and 0.010 vs 0.012) and for

continuous trait compared to binary trait (0.020 vs 0.098 and 0.002 vs 0.019). Table 4 shows that smaller variance was found for MTDFREML compared to GS (0.001 vs 0.002), for animal model compared to sire model (0.001 vs 0.002) and for continuous trait compared to binary trait (0.001 vs 0.002).

Table 4. Mean±standard errors (SE) of main effects for bias and mean squared errors (MSE) for heritability estimates

Factor	Bias	MSE
	Mean ±SE	Mean ±SE
μ	-0.039±0.00096	0.011±0.00074
Heritability level		
0.10	-0.009±0.0016	0.003±0.0013
0.25	-0.033±0.0016	0.008±0.0013
0.50	-0.076±0.0016	0.022±0.0013
Number of progeny		
5	-0.030±0.0018	0.013±0.0015
10	-0.043±0.0018	0.010±0.0015
15	-0.044±0.0018	0.010±0.0015
20	-0.041±0.0018	0.009±0.0015
Type of algorithm		
GS	-0.008±0.0013	0.005±0.0011
MTDFREML	-0.071±0.0013	0.016±0.0011
Type of model		
Animal model	-0.053±0.0013	0.010±0.0011
Sire model	-0.025±0.0013	0.012±0.0011
Type of trait		
Binary trait	-0.098±0.0013	0.019±0.0011
Continuous trait	0.020±0.0013	0.002±0.0011

Figures (1) to (16) illustrate all significant 2-way interactions. Figures (1) and (9) indicate that the magnitude of bias and MSE were smaller at h^2 of 0.1 than at h^2 of 0.25 or 0.5 with all considered numbers of progeny per sire. This result is in agreement with those reported by Thomas *et al.* (2000) and Thomas and Hill (2000) as discussed earlier.

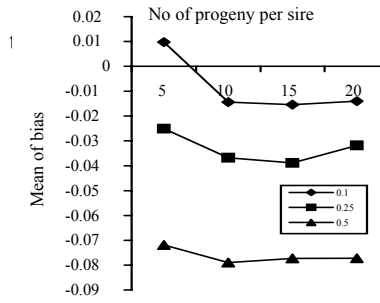


Fig. 1. Progeny number per sire - h^2 level interaction

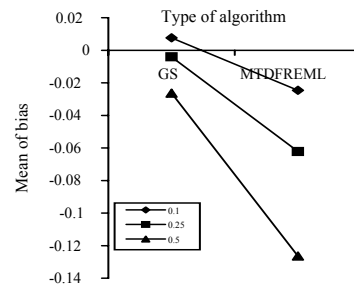


Fig. 2. Algorithm type - h^2 level interaction

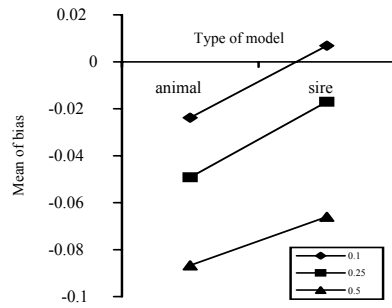


Fig. 3. Model type - h^2 level interaction

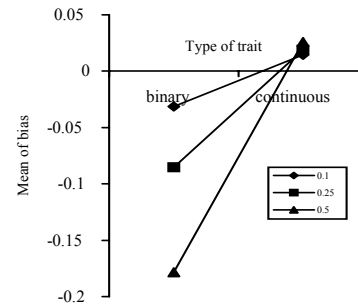


Fig. 4. Trait type - h^2 level interaction

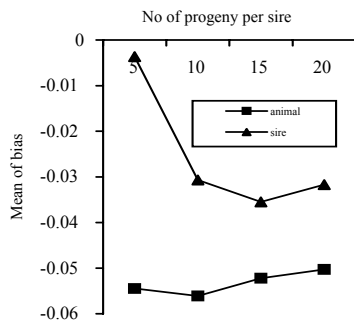


Fig. 5. Progeny number per sire - model type interaction

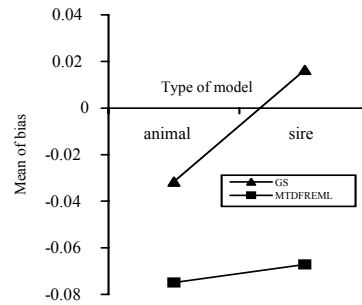


Fig. 6. Model type - algorithm type interaction

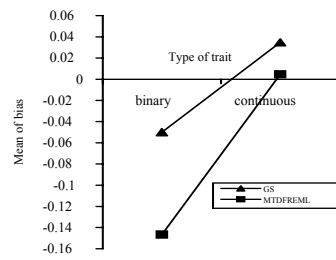


Fig. 7. Trait type–algorithm type interaction

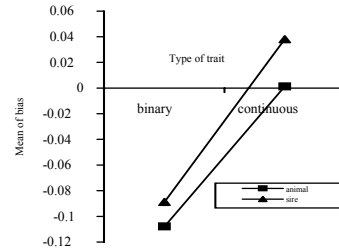


Fig. 8. Trait type–model type interaction

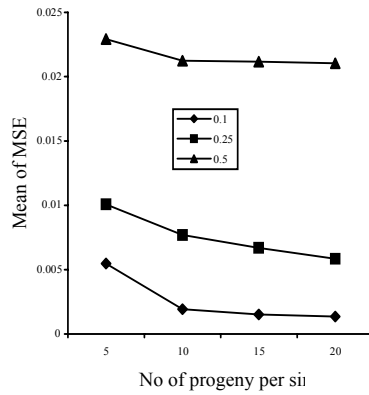


Fig. 9. Progeny number per sire - h² level interaction

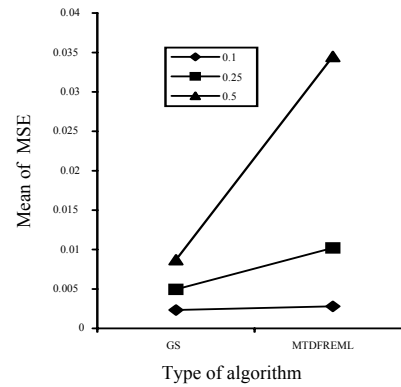


Fig. 10. Algorithm type - h² level interaction

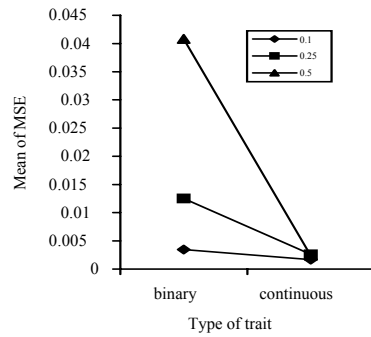


Fig. 11. Trait type - h² level interaction

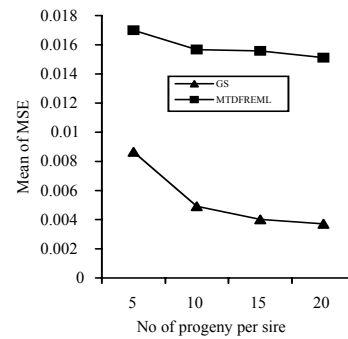


Fig. 12. Progeny number per sire - algorithm type interaction

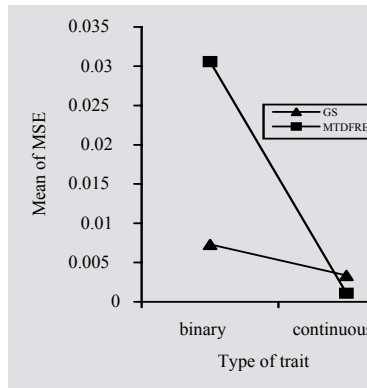


Fig. 13. Progeny number per sire - model type interaction

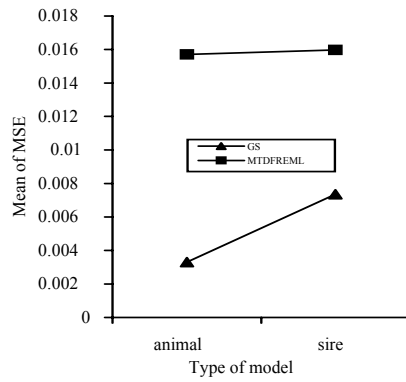


Fig. 14. Model type - algorithm type interaction

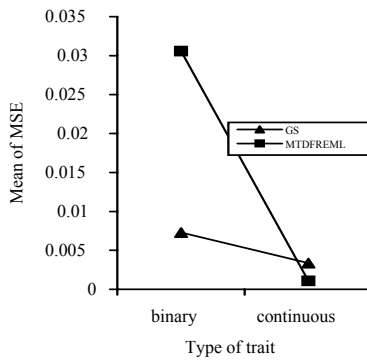


Fig. 15. Trait type - algorithm type interaction

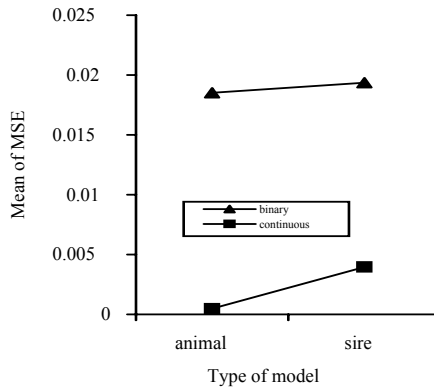


Fig. 16. Model type - trait type interaction

Levels of h^2 affected magnitude of bias and MSE using different algorithms (Figures 2 and 9).

Figure (3) shows that animal model had higher magnitude of bias than sire model, at any level of h^2 but difference in bias between h^2 of 0.25 and 0.5 is larger with sire model than with animal model. This result supports those reported by Luo *et al.* (2001) indicating that in general, sire model yielded more accurate estimates of h^2 than did animal model.

Figures (4) and (11) show that continuous trait had smaller magnitude of bias and MSE than binary trait at any level of h^2 and the difference between binary and continuous traits increased as level of h^2 increased.

Figure (12) shows that the GS had smaller MSE than MTDFREML whatever the number of progeny per sire is; the 20 progeny per sire having the smallest MSE. This result confirms those obtained by Mousa and Elsayed (2001) who indicated that GS had consistently smaller MSE than MTDFREML, due to the influence of the prior distribution of the variance components on the posterior distribution.

Figures (5) and (13) indicate that at any number of progeny per sire, the sire model had smaller magnitude of bias than animal model, and the 5 progeny case showed the smallest magnitude of bias whereas the sire model had greater MSE than animal model and that the 20 progeny per sire had the smallest MSE.

Figures (6) and (14) show that sire model had smaller magnitude of bias than animal model, using each MTDFREML or GS. MSE resulting from GS or MTDFREML using animal model were smaller than those resulting from the same algorithms using sire model and a large difference between animal and sire models was observed with GS.

Figures (7) and (15) show that the difference in bias and MSE between GS and MTDFREML in magnitude was larger and opposite in direction for binary trait as compared to continuous trait. The difference in magnitude of bias between animal and sire models was larger and opposite in trend in the continuous trait as compared to binary trait (Figure 8). This result find support in the work of Hoeschele and Tier (1995) who reported that for categorical traits, because of the extreme category problem in which all observations for some subclasses are in the same category, threshold animal model using the GS may yield biased estimates.

Figure (16) indicates that for binary or continuous trait animal model had smaller MSE than sire model.

Average heritability estimates (h^2):

The average values for the estimates of h^2 of the 20 samples resulting from MTDFREML and GS algorithms for continuous and binary traits at different types of models (animal or sire), different number of progeny per sire (5, 10, 15 or 20) and different levels of h^2 (0.1, 0.25 or 0.5) are shown in Table (5). The estimates for the continuous trait at all levels of h^2 with the four different number of progeny per sire resulting from MTDFREML and GS using animal model were generally similar to the values used for simulation. This result is in agreement with those reported by Mousa and Elsayed (2001) who reported that GS and MTDFREML estimates appeared similar for continuous variable. Table (5) shows that for continuous trait at h^2 equals to 0.1 with 5 progeny per sire, the estimate of h^2 resulting from MTDFREML using sire model was an overestimate (0.14). Whereas at h^2 equals to 0.25 or 0.5, the estimates of h^2 resulting from MTDFREML using animal model were very close to those resulting from sire model. These results indicate that the estimation of variance components using sire model with 5 progeny per sire would be inaccurate.

Table (5) also shows that for continuous trait in GS using sire model, overestimates were recorded at any level of h^2 with any number of progeny per sire. This would indicate the need for increasing the number of rounds of iteration of the GS chain. For the binary trait, estimates resulting from GS using either animal or sire model were higher while being closer to the values used for simulation than corresponding estimates resulting from MTDFREML at all levels of h^2 with any

number of progeny per sire except at h^2 equals to 0.1 with 5 progeny per sire, where the estimate of h^2 resulting from GS using sire model was higher than the corresponding estimate resulting from MTDFREML (0.16 vs 0.09). These results are in line with previous works (Matos *et al.*, 1997; Boettcher *et al.*, 1999 and Luo *et al.*, 2001) explaining the reason for this difference between MTDFREML and GS as the heritabilities from MTDFREML (linear model) are expressed on observed scale while heritabilities from GS (threshold model) are on an underlying liability scale. Therefore, threshold model was statistically more appropriate than linear model for binary trait, yields greater estimates of heritability and closer to the real value in most cases.

Table (5) also shows that for the binary trait in GS at different number of progeny per sire and at h^2 equals to 0.1 or 0.25 the estimates of h^2 resulting from sire model were higher and closer to the real value than corresponding estimates of h^2 resulting from the animal model. At h^2 equals to 0.5, the estimates of heritability resulting from sire model were closer to corresponding estimates resulting from animal model, but all were underestimates of the real value used for simulation. In agreement with this result Matos *et al.* (1997) reported that the advantage of using threshold over linear methodology in breeding programs increases as the heritability of the trait decreases.

Average bias and mean squared errors (MSE) of heritability estimates

The average bias (Table 6) and the MSE (Table 7) at different levels of h^2 , different traits (continuous or binary trait), different models (animal or sire), different algorithms (MTDFREML or GS) and different number of progeny per sire (5, 10, 15 or 20) were calculated.

Expectedly Table (6) shows the same trend presented in Table (5) and indicates that for the continuous trait at any level of h^2 with any number of progeny per sire, the magnitude of average bias of h^2 resulting from MTDFREML using each of animal and sire model was smaller than corresponding ones resulting from GS, except in the case of continuous trait using animal model at h^2 equals to 0.1 with number of progeny per sire equals to 15 (0.004 vs 0.001) and at h^2 equals to 0.5 with number of progeny per sire equals to 5 (0.010 vs 0.001). This would indicate the need of increased number of rounds of iteration of GS chain. For binary trait at any number of progeny per sire and any level of h^2 , the magnitude of average bias resulting from GS was smaller than corresponding values resulting from MTDFREML. This result confirms the findings of Van Tassell and Van Vleck (1996) and Luo *et al.* (2001) indicating linear models as inappropriate for analysis of binary response traits.

Within GS, for binary trait, the sire model had smaller magnitude of bias than animal model except at h^2 equals to 0.1 with number of progeny per sire equals to 5 (0.038 vs 0.060); and at h^2 equals to 0.5 with number of progeny per sire equals to 10 and 15 (0.093 vs 0.096 and 0.099 vs 0.104, respectively). This indicates that threshold single trait sire model could be a good alternative model compared to animal model for genetic analysis of binary traits.

For bias at any level of h^2 and at any number of progeny per sire for continuous trait, the average bias resulting from MTDFREML using animal and sire models did not differ significantly from zero ($p < 0.05$), except for sire model at h^2 equals to 0.1 with number of progeny per sire equals to 5 (Table 6). This result indicates that

estimation of variance components using sire model with 5 progeny per sire is quite inaccurate. For continuous trait, most bias estimates resulting from GS using animal model at h^2 equals to 0.1 or 0.25 were not significantly different from zero ($p > 0.05$) whereas, at h^2 equals to 0.5 most of estimates differed from zero significantly. Using sire model, the estimates of bias of continuous trait resulting from GS were significantly different from zero. This result indicates that for continuous trait, MTDFREML yields estimates relatively free of bias.

Table (6) shows that, in general, at any level of h^2 and at any number of progeny per sire, the magnitude of bias was significantly different from zero ($p < 0.05$) for binary trait resulting from MTDFREML using animal and sire models and from GS using only animal model are significantly different from zero. For binary trait, the average bias resulting from GS using sire model was not significantly different from zero especially at h^2 equals to 0.1 (with 10, 15 and 20 progeny per sire) or 0.25 (with 5, 10 and 20 progeny per sire). At h^2 equals to 0.5 the average bias was significantly different from zero ($p < 0.05$). This shows that at low and moderate h^2 , the use of threshold sire model for binary trait yields accurate estimates free of bias and, thus, closer to the real value. This agrees with Hoeschele and Tier (1995) who reported that threshold animal model using GS may yield bias estimates.

Figures (17) and (18) show a decision chart for estimating variance components related to continuous and binary traits, based on bias criterion. MSE for continuous trait resulting from GS applying animal model was generally smaller than corresponding MSE resulting from MTDFREML adopting the same model (Table 7). This result is in agreement with those reported by Mousa and Elsayed (2001) as mentioned earlier (Figure 12).

Table (7) indicates that within GS, for binary trait especially at h^2 equals to 0.25 or 0.5 with any number of progeny per sire, the sire model had greater MSE than animal model. At h^2 equals to 0.1 at any number of progeny per sire except for 5 progeny, sire model had smaller MSE than animal model (0.00251 vs 0.01185). This result is in agreement with those reported by Luo *et al.* (2001). With sire model, for continuous trait, MSE resulting from MTDFREML was smaller than corresponding values resulting from GS (Table 7) perhaps because of the smaller bias with MTDFREML than with GS. With animal model, the GS had smaller MSE than MTDFREML. The MSE calculated with the subclass of 5 progeny per sire was greater than when number of progeny per sire was 10, 15 or 20; the 20 progeny per sire having the smallest values in most cases (Tables 7). This is in agreement with Thomas and Hill (2000). The optimum number of progeny per sire to be used was 20 to yield the best estimates of heritability under the circumstances of this study. Figures (19) and (20) show the decision chart for estimating variance components for a continuous and binary traits, based on MSE criterion.

From Tables (6 and 7), for binary trait, the conclusion as what methods to use based on bias agreed with those based on MSE, i.e. for the cases of use of sire model by GS at h^2 of 0.1 with 10, 15 and 20 progeny, the use of sire model by MTDFREML at the same level of h^2 with 5 progeny per sire and use of sire model by GS at h^2 equals to 0.25 or 0.5 with 20 progeny per sire (Figure 21) judging based on MSE and bias leads to the same conclusion.

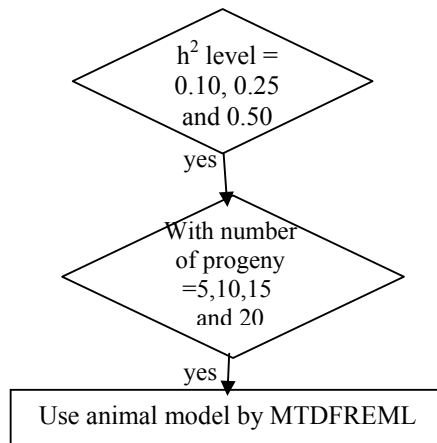


Fig. 17. Decision chart for recommending the best model type-algorithm combination to estimate variance components in case of a continuous trait, based on bias

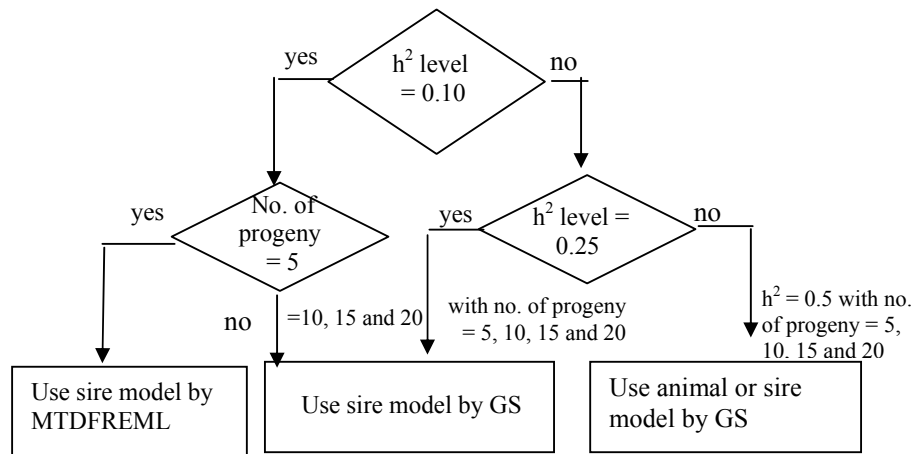


Fig. 18. Decision chart for recommending the best model type- algorithm combination to estimate variance components in case of a binary trait, based on bias

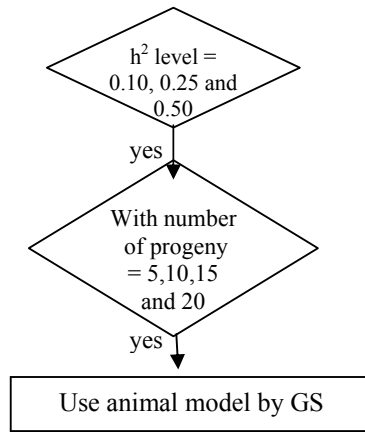


Fig. 19. Decision chart for recommending the best model type-algorithm combination to estimate variance components in case of a continuous trait, based on MSE.

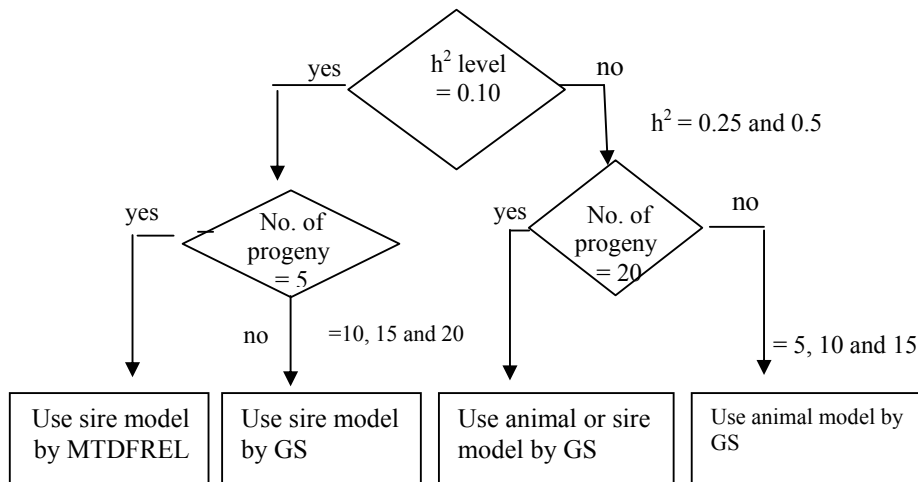


Fig. 20. Decision chart for recommending the best model type-algorithm combination to estimate variance components in case of a binary trait, based on MSE

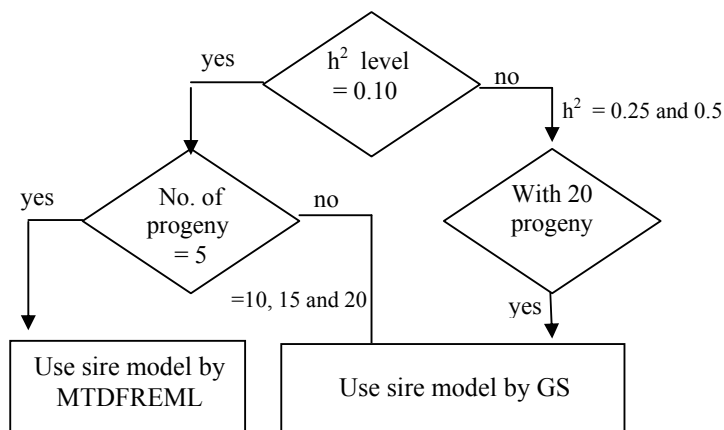


Figure 21. Decision chart for recommending the best model type-algorithm combination to estimate variance components in case of a binary trait, based on bias jointly with MSE

Computation time:

The analysis of one sample of continuous or binary trait by MTDFREML using animal or sire model consumed 5 minutes with any number of progeny per sire. Whereas the time consumed for the analysis of continuous or binary trait by GS using animal and sire models is shown in Table (8). The analysis was made using a computer with Pentium IV of 1.7 GH¹ processor, 40 GB² hard disk and 256 MB³ a random access memory.

¹ GH = Gega hertz

² GB = Gega byte

³ MB= Mega byte

Table 8. Time consumed for the analysis by GS algorithm using different number of progeny per sire and different models for continuous and binary traits

No. of progeny per sire	Model type	Type of trait					
		Continuous trait			Binary trait		
		Time (hr)	No. of samples	Total Time (hr)	Time (hr)	No. of samples	Total time (hr)
5	Animal	0:23	60	23:00	2:25	60	135:00
	Sire	0:02	60	02:00	0:07	60	007:00
10	Animal	0:30	60	30:00	4:20	60	252:00
	Sire	0:03	60	03:00	0:10	60	010:00
15	Animal	1:00	60	60:00	7:00	60	420:00
	Sire	0:04	60	04:00	0:15	60	015:00
20	Animal	1:20	60	22:00	9:00	60	540:00
	Sire	0:05	60	05:00	0:25	60	025:00
Total				199:00			1404:00

In general, using the animal model consumed more time than sire model (with magnitude of > 14 times) and binary trait consumed more time than continuous trait (with magnitude of > 7 times). The full analysis with MTDFREML and GS (1920 samples) consumed 1683 computer hours.

CONCLUSION

If the aim was to estimate variance components, the conclusion is that:

- 1- For continuous trait, the animal model is the best with MTDFREML or GS at all levels of h^2 with any number of progeny per sire.
- 2- For binary trait, GS is the best algorithm at all levels of h^2 . Within GS, the sire model is the best at low h^2 with any number of progeny more than 5 whereas, at h^2 equals to 0.25 or 0.5 with 20 progeny per sire, the animal model is equivalent to sire model.

ACKNOWLEDGMENT

I am indebted for the Dr. Emad Mousa, Lecturer of Animal Breeding, Faculty of Agriculture, Assiut University, for his assistance in the use of Gibbs Sampling program when needed.

I am most grateful to Dr. Mohamed Elshenawy, Lecturer of Animal Breeding, Faculty of Agriculture, South Valley University, for his help in using SAS program for simulating the binary trait.

REFERENCES

- Al-Shorepy, S.A. and D.R. Notter, 1996. Genetic variation and covariation for ewe reproduction, lamb growth, and lamb scrotal circumference in a fall-lambing sheep flock. *J. Anim. Sci.* 74:1490-1498.
- Analla, M.; A. Sanchez; A. Munaz and J.M. Serradilla, 1995. Simulation analysis with BLUP methodology of different data structures in goat selection schemes in Spain. *Small Ruminant Research*, 17:51-55.
- Boettcher, P. J., L. K. Jairath, and J. C. M. Dekkers, 1999. Comparison of methods for genetic evaluation of sires for survival of their daughters in the first three lactations. *J. Dairy. Sci.* 82:1034-1044.
- 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 Variances and Covariances [Draft]. U.S. Department of Agriculture, Agricultural Research Service. pp 114.
- Elsayed, M.M., 1997. Modeling to determine possible problems in application of the animal model in dairy cattle. PhD. Thesis, Fac. Agric., Ain Shams Univ., Egypt, p. 24.
- Gianola, D. and F.L. Foulley, 1983. Sire evaluation for ordered categorical data with a threshold model. *Genet. Sel. Evol.* 15:201-224.
- Heringstad, B., R. Rekaya, D. Gianola, G. Klemetsdal and K.A. Weigel, 2001. Bayesian analysis of liability of clinical mastitis in Norwegian cattle with a threshold model: effect of data sampling method and model specification. *J. Dairy. Sci.* 84:2337-2346.

- Hoeschele, I. and B. Tier, 1995. Estimation of variance components of threshold characters by marginal posterior modes and means via Gibbs sampling. *Genet. Sel. Evol.* 27:519-540.
- Luo, M.F.; P.J. Boettcher; L.R. Schaeffer and J.C.M. Dekkerst, 2001. Bayesian inference for categorical traits with an application to variance component estimation. *J. Dairy. Sci.* 84:694-704.
- Matos, C. A. P., D. L. Thomas, D. Gianola, R. J. Tempelman, and L. D. Young, 1997. Genetic analysis of discrete reproductive traits in sheep using linear and nonlinear models: I. estimation of genetic parameters. *J. Anim. Sci.* 75:76-87.
- Mousa, E., and M. Elsayed, 2001. Quality of variance components estimated by Gibbs sampling in populations with and without selection and varying heritability. *Egyptian J. Anim. Prod.* 38:97-110.
- Mrode, R. A., 1996. *Linear Models for the Prediction of Animal Breeding Values.* CAB INTERNATIONAL, Wallingford, U.K. pp 187.
- Neter, J., W. Wasserman and H. Kutner, 1985. *Applied linear statistical model regression, analysis of variance and experimental designs.* Second edition Library of Congress Catalog. Card No. 84:395.
- SAS, 1996. *Statistical Analysis System, SAS User's Guide : Statistics.* SAS Institute Inc. Editors, Cary, NC.
- Schaeffer, L.R., 1984. Sire and cow evaluation under multiple trait models. *J. Dairy Sci.* 67:1567-1580.
- Thomas, S.C., and W.G. Hill, 2000. Estimating quantitative genetic parameters using sib ships reconstructed from Marker data. *Genet.* 155:1961-1972.
- Thomas, S.C., J. M. Pemberton, and W.G. Hill. 2000. Estimating variance components in natural populations using inferred relationships. *Heredity* 84:427-436.
- 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 of Agriculture, Agricultural Research Service. pp 86.
- Van Tassell, 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.

تأثير مستوى المكافئ الوراثي ، عدد البنات لكل طلوقة ، نوع البرنامج المستخدم ، نوع النموذج المستخدم ونوع الصفة على مدى جودة تقدير المكافئ الوراثي

رضا السعيد محمد أحمد¹، منال محمد أحمد سيد²، السيد صلاح الدين جلال²، حسين مصطفى كمال منصور²

1- قسم الإنتاج الحيواني، معهد بحوث البيئة الصحراوية، جامعة المنوفية، مدينة السادات، 2- قسم الإنتاج الحيواني، كلية الزراعة، جامعة عين شمس

استهدفت هذه الدراسة تحديد تأثير مستوى المكافئ الوراثي ، عدد البنات للطلوقة، نوع البرنامج المستخدم ، نوع النموذج الإحصائي وكذلك نوع الصفة على مدى جودة تقدير المكافئ الوراثي. أجريت هذه الدراسة على برامج تقوم بعمل محاكاة لصفة مستمرة وأخرى متقطعة ذات مستويين (صفر ، 1) ناتجة من تحويل الصفة المستمرة ، وقد كان متوسط الصفة 0.5 أما التباين المظهري فكان 0.25 . حورت باقي مكونات التباين من تباين وراثي وتباين بيئي دائم وتباين الخطأ لتلائم الغرض من عمل ثلاثة مستويات من المكافئ الوراثي (0.1، 0.25، 0.5) بأربعة مستويات من أعداد البنات (أنصاف أشقه) لكل طلوقة (5، 10، 15، 20). وبدا يكون هناك 12 عشيرة بكل منها ثلاثة مواسم إدرار كتأثير ثابت وحيد ، خلقت لكل عشيرة 20 مكررة عشوائية.

أجرى تحليل لكل من الصفة المستمرة والصفة المتقطعة ببرنامجين هما MTDFREML و Gibbs Sampling واستخدم داخل كل تحليل نوعان من النماذج، الأول نموذج الحيوان (Animal model) والثاني نموذج الطلوقة (Sire model) . وقد استخدم في هذه الدراسة التحيز (Bias) ومتوسط الخطأ مربعا (Mean squared errors) للمقارنة بين تقديرات المكافئ الوراثي المتحصل عليها في الحالات المختلفة. كان تأثير كل من هذه العوامل الرئيسية معنويا جدا ($P < 0.0001$).

اتضح من الدراسة بناء على التحيز أنه إذا كانت الصفة مستمرة فإن MTDFREML باستخدام نموذج الحيوان كان أفضل من GS عند أي مستوى من المكافئ الوراثي وأي عدد من البنات داخل الطلوقة. أما بالنسبة للصفة المتقطعة فإن GS كان أفضل من MTDFREML سواء باستخدام نموذج الحيوان أو نموذج الطلوقة عند أي مستوى من المكافئ الوراثي وأي عدد من البنات للطلوقة ماعدا باستخدام نموذج الطلوقة في حالة المكافئ الوراثي المنخفض (0.1) عند 5 بنات فإن MTDFREML كان الأفضل. داخل GS كان استخدام نموذج الطلوقة أفضل من استخدام نموذج الحيوان عند المستويين المنخفض والمتوسط من المكافئ الوراثي ، في حين أنه عند مستوى عال من المكافئ الوراثي فإن استخدام نموذج الطلوقة لا يختلف عن استخدام نموذج الحيوان.

عند التقييم بناء على متوسط الخطأ مربعا وكانت الصفة مستمرة فإن GS باستخدام نموذج الحيوان كان أفضل من استخدام MTDFREML عند أي مستوى من المكافئ الوراثي وأي عدد من البنات للطلوقة. أما بالنسبة للصفة المتقطعة فإن GS عموما كان أفضل من استخدام MTDFREML سواء في حالة نموذج الحيوان أو نموذج الطلوقة عند أي مستوى من المكافئ الوراثي وأي عدد من البنات للطلوقة ماعدا في حالة نموذج الطلوقة عند مكافئ وراثي 0.1 وعدد من البنات يساوي 5 حيث كان MTDFREML هو الأفضل . داخل GS وبالنسبة للصفة المتقطعة فإن استخدام نموذج الطلوقة كان أفضل من استخدام نموذج الحيوان عند مكافئ وراثي يساوي 0.1 بعدد من البنات للطلوقة أكبر من 5 ، بينما عند مكافئ وراثي متوسط أو عال (0.25 أو 0.5) فإن استخدام نموذج الحيوان كان أفضل من استخدام نموذج الطلوقة عند أي عدد من البنات للطلوقة، أما عند 20 بنت للطلوقة فإن استخدام نموذج الحيوان قارب استخدام نموذج الطلوقة.

النتائج المتحصل عليها بناء على التحيز اتفقت مع النتائج المتحصل عليها بناء على متوسط الخطأ مربعا فقط في حالة الصفة المتقطعة باستخدام GS ونموذج الطلوقة عند مكافئ وراثي 0.1 وعدد من البنات للطلوقة يساوي 10، 15، 20 وأيضا في حالة استخدام MTDFREML ونموذج الطلوقة عند نفس المستوى من المكافئ الوراثي (0.1) ولكن فقط عند 5 بنات للطلوقة. كما اتفقت النتائج أيضا في حالة استخدام GS ونموذج الطلوقة عند مكافئ وراثي 0.25 أو 0.5 وعدد 20 بنت للطلوقة.

نستخلص من هذه الدراسة أنه حينما يكون الغرض هو تقدير مكونات التباين فإنه:

1- في حالة الصفة المستمرة يكون استخدام نموذج الحيوان هو الأفضل سواء باستخدام برنامج MTDFREML أو GS عند كل مستويات المكافئ الوراثي وأي عدد من البنات للطلوقة.

2- في حالة الصفة المتقطعة يكون برنامج GS أفضل من MTDFREML عند أي مستوى من المكافئ الوراثي وأي عدد من البنات. داخل برنامج GS يكون نموذج الطلوقة أفضل من نموذج الحيوان عند مستوى منخفض من المكافئ الوراثي في حين عند المستويين المتوسط والعالي وعدد 20 بنت للطلوقة فإن نموذج الحيوان يكافئ نموذج الطلوقة.

Table 5. Average values \pm standard errors (SE) of heritability (h^2) resulting from MTDFREML and GS programs for continuous and binary traits at different levels of h^2 (0.1, 0.25 and 0.5), different types of model (animal and sire) and different number of progeny per sire (5, 10, 15 and 20)

Heritability level	No. of progeny per sire	MTDFREML				GS			
		Continuous trait		Binary trait		Continuous trait		Binary trait	
		Animal model	Sire model	Animal model	Sire model	Animal model	Sire model	Animal model	Sire model
0.1	5	0.10 \pm 0.005	0.14 \pm 0.013	0.04 \pm 0.008	0.09 \pm 0.018	0.09 \pm 0.004	0.19 \pm 0.012	0.06 \pm 0.007	0.16 \pm 0.020
	10	0.10 \pm 0.004	0.10 \pm 0.007	0.05 \pm 0.005	0.04 \pm 0.006	0.10 \pm 0.004	0.14 \pm 0.006	0.07 \pm 0.007	0.10 \pm 0.007
	15	0.10 \pm 0.004	0.10 \pm 0.005	0.04 \pm 0.004	0.04 \pm 0.006	0.10 \pm 0.004	0.13 \pm 0.001	0.07 \pm 0.005	0.10 \pm 0.002
	20	0.10 \pm 0.004	0.10 \pm 0.005	0.04 \pm 0.003	0.04 \pm 0.004	0.10 \pm 0.003	0.13 \pm 0.001	0.07 \pm 0.004	0.09 \pm 0.002
0.25	5	0.25 \pm 0.008	0.27 \pm 0.018	0.12 \pm 0.012	0.15 \pm 0.019	0.23 \pm 0.006	0.34 \pm 0.015	0.18 \pm 0.014	0.26 \pm 0.025
	10	0.25 \pm 0.006	0.25 \pm 0.007	0.11 \pm 0.009	0.13 \pm 0.012	0.24 \pm 0.004	0.32 \pm 0.008	0.19 \pm 0.011	0.22 \pm 0.019
	15	0.25 \pm 0.005	0.25 \pm 0.008	0.11 \pm 0.004	0.11 \pm 0.007	0.25 \pm 0.004	0.31 \pm 0.007	0.19 \pm 0.004	0.20 \pm 0.010
	20	0.25 \pm 0.005	0.25 \pm 0.006	0.12 \pm 0.005	0.12 \pm 0.007	0.26 \pm 0.003	0.31 \pm 0.006	0.20 \pm 0.005	0.23 \pm 0.012
0.5	5	0.51 \pm 0.006	0.51 \pm 0.007	0.24 \pm 0.011	0.25 \pm 0.019	0.50 \pm 0.005	0.59 \pm 0.008	0.39 \pm 0.012	0.41 \pm 0.024
	10	0.50 \pm 0.006	0.50 \pm 0.005	0.24 \pm 0.009	0.25 \pm 0.009	0.51 \pm 0.004	0.59 \pm 0.005	0.41 \pm 0.009	0.40 \pm 0.015
	15	0.50 \pm 0.004	0.51 \pm 0.004	0.24 \pm 0.007	0.24 \pm 0.007	0.51 \pm 0.003	0.58 \pm 0.004	0.40 \pm 0.005	0.40 \pm 0.012
	20	0.51 \pm 0.004	0.50 \pm 0.005	0.24 \pm 0.004	0.24 \pm 0.006	0.51 \pm 0.003	0.58 \pm 0.004	0.40 \pm 0.005	0.40 \pm 0.010

Table 6. Average of bias \pm standard errors (SE) of heritability (h^2) estimates resulting from MTDFREML and GS programs for continuous and binary traits at different levels of h^2 (0.1, 0.25 and 0.5), different types of models (animal and sire) and different number of progeny per sire (5, 10, 15 and 20)

Heritability Level	No. of progeny per sire	MTDFREML				GS			
		Continuous trait		Binary trait		Continuous trait		Binary trait	
		Animal model	Sire model	Animal model	Sire model	Animal model	Sire model	Animal model	Sire model
0.1	5	0.001 \pm 0.005	0.044 \pm 0.013*	-0.059 \pm 0.008*	-0.014 \pm 0.018	-0.009 \pm 0.004*	0.092 \pm 0.012*	-0.038 \pm 0.007*	0.060 \pm 0.020*
	10	-0.004 \pm 0.004	0.002 \pm 0.007	-0.062 \pm 0.005*	-0.056 \pm 0.006*	-0.005 \pm 0.004	0.044 \pm 0.006*	-0.032 \pm 0.007*	-0.002 \pm 0.007
	15	-0.004 \pm 0.004	-0.004 \pm 0.005	-0.061 \pm 0.004*	-0.060 \pm 0.007*	0.001 \pm 0.004	0.034 \pm 0.005*	-0.028 \pm 0.005*	-0.002 \pm 0.008
	20	-0.001 \pm 0.004	0.000 \pm 0.005	-0.061 \pm 0.003*	-0.056 \pm 0.004*	0.004 \pm 0.003	0.034 \pm 0.004*	-0.026 \pm 0.004*	-0.006 \pm 0.007
0.25	5	0.003 \pm 0.008	0.022 \pm 0.018	-0.135 \pm 0.012*	-0.102 \pm 0.019*	-0.016 \pm 0.006*	0.088 \pm 0.015*	-0.069 \pm 0.014*	0.008 \pm 0.025
	10	0.003 \pm 0.006	-0.004 \pm 0.007	-0.139 \pm 0.009*	-0.124 \pm 0.012*	-0.006 \pm 0.004	0.068 \pm 0.008*	-0.064 \pm 0.011*	-0.028 \pm 0.019
	15	0.003 \pm 0.005	0.008 \pm 0.006	-0.137 \pm 0.004*	-0.136 \pm 0.007*	0.004 \pm 0.004	0.062 \pm 0.008*	-0.060 \pm 0.004*	-0.050 \pm 0.010*
	20	0.003 \pm 0.005	0.004 \pm 0.006	-0.132 \pm 0.005*	-0.126 \pm 0.007*	0.006 \pm 0.003*	0.062 \pm 0.006*	-0.051 \pm 0.005*	-0.020 \pm 0.012
0.5	5	0.010 \pm 0.007	0.010 \pm 0.007	-0.258 \pm 0.011*	-0.248 \pm 0.018*	-0.001 \pm 0.005	0.088 \pm 0.008*	-0.107 \pm 0.012*	-0.092 \pm 0.024*
	10	0.001 \pm 0.006	0.000 \pm 0.005	-0.258 \pm 0.009*	-0.258 \pm 0.009*	0.008 \pm 0.004*	0.086 \pm 0.005*	-0.093 \pm 0.009*	-0.096 \pm 0.015*
	15	0.003 \pm 0.004	0.006 \pm 0.004	-0.260 \pm 0.007*	-0.260 \pm 0.007*	0.012 \pm 0.003*	0.084 \pm 0.005*	-0.099 \pm 0.005*	-0.104 \pm 0.012*
	20	0.006 \pm 0.004	0.004 \pm 0.005	-0.261 \pm 0.005*	-0.260 \pm 0.006*	0.013 \pm 0.003*	0.084 \pm 0.005*	-0.103 \pm 0.005*	-0.100 \pm 0.011*

* Estimate different from zero significantly ($p \leq 0.05$).

Table 7. Mean squared errors of heritability (h^2) estimates resulting from MTDFREML and GS programs for continuous and binary traits at different levels of h^2 (0.1, 0.25 and 0.5), different types of models (animal and sire) and different number of progeny per sire (5, 10, 15 and 20)

Heritability level	No. of progeny per sire	MTDFREML				GS			
		Continuous trait		Binary trait		Continuous trait		Binary trait	
		Animal model	Sire model	Animal model	Sire model	Animal model	Sire model	Animal model	Sire model
0.1	5	0.00047	0.00554	0.00481	0.00681	0.00044	0.01143	0.00251	0.01185
	10	0.00034	0.00093	0.00431	0.00396	0.00035	0.00268	0.00194	0.00093
	15	0.00034	0.00059	0.00401	0.00444	0.00025	0.00119	0.00126	0.00009
	20	0.00034	0.00059	0.00391	0.00346	0.00015	0.00118	0.00104	0.00009
0.25	5	0.00145	0.00682	0.02125	0.01733	0.00099	0.01204	0.00840	0.01227
	10	0.00080	0.00090	0.02087	0.01812	0.00030	0.00588	0.00637	0.00828
	15	0.00059	0.00074	0.01915	0.01938	0.00031	0.00496	0.00388	0.00452
	20	0.00042	0.00074	0.01776	0.01687	0.00023	0.00462	0.00306	0.00307
0.5	5	0.00094	0.00109	0.06904	0.06843	0.00051	0.00911	0.01433	0.02002
	10	0.00063	0.00042	0.06827	0.06849	0.00038	0.00795	0.01031	0.01358
	15	0.00038	0.00042	0.06822	0.06861	0.00037	0.00746	0.01034	0.01366
	20	0.00038	0.00042	0.06818	0.06827	0.00029	0.00746	0.01120	0.01219