

Influence of priors in Bayesian estimation of genetic parameters for multivariate threshold models using Gibbs sampling

Kathrin Friederike STOCK^{a*}, Ottmar DISTL^a, Ina HOESCHELE^b

^a Institute for Animal Breeding and Genetics, University of Veterinary Medicine Hannover (Foundation), Buenteweg 17p, D-30559 Hannover, Germany

^b Virginia Bioinformatics Institute and Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA

(Received 26 April 2006; accepted 24 October 2006)

Abstract – Simulated data were used to investigate the influence of the choice of priors on estimation of genetic parameters in multivariate threshold models using Gibbs sampling. We simulated additive values, residuals and fixed effects for one continuous trait and liabilities of four binary traits, and QTL effects for one of the liabilities. Within each of four replicates six different datasets were generated which resembled different practical scenarios in horses with respect to number and distribution of animals with trait records and availability of QTL information. (Co)Variance components were estimated using a Bayesian threshold animal model *via* Gibbs sampling. The Gibbs sampler was implemented with both a flat and a proper prior for the genetic covariance matrix. Convergence problems were encountered in > 50% of flat prior analyses, with indications of potential or near posterior impropriety between about round 10 000 and 100 000. Terminations due to non-positive definite genetic covariance matrix occurred in flat prior analyses of the smallest datasets. Use of a proper prior resulted in improved mixing and convergence of the Gibbs chain. In order to avoid (near) impropriety of posteriors and extremely poorly mixing Gibbs chains, a proper prior should be used for the genetic covariance matrix when implementing the Gibbs sampler.

Gibbs sampling / multivariate threshold model / covariance estimates / flat prior / proper prior

1. INTRODUCTION

The use of Bayesian methodology implemented *via* Gibbs sampling (GS) for variance components estimation has considerably increased in the last

* Corresponding author: Kathrin-Friederike.Stock@tiho-hannover.de

decade. Programs which support linear as well as threshold or mixed linear-threshold model analyses are accessible and have been used in several heritability and genetic correlation studies and genetic evaluations in cattle, sheep, pigs and other species *e.g.* [1, 9, 10, 12, 24, 29].

Although GS has considerably simplified the use of Bayesian methods, it is often still difficult to implement an efficient Gibbs sampler. Performance of GS can be influenced by the choice of prior distribution in the formulation of the joint posterior distribution. The amount of prior knowledge on the parameters of interest determines the choice of priors for GS. In many studies flat priors have been used for the genetic (co)variance matrix in order to account for the lack of prior knowledge or the reluctance to use existing prior knowledge *e.g.* [28]. For linear models it has been shown that flat priors may lead to proper posteriors, but that they do not necessarily [5, 7]. For threshold models and binary traits it has been shown that poor results can also be obtained when the posterior is close to an improper density [18]. Furthermore, impropriety of posteriors may not be easy to determine, implying the risk of obtaining misleading results [4, 7, 18, 27].

The aim of this study was to investigate the influence of two priors for the genetic (co)variance matrix on the posterior distributions of genetic parameters in multivariate threshold models using Gibbs sampling. Data sets differing in the amount and distribution of available information were analyzed in order to characterize mixing and convergence properties of the Gibbs sampler and the accuracy of the estimates of the genetic covariance matrix when using a flat or a proper prior in connection with different data constellations encountered in the horse industry.

2. MATERIALS AND METHODS

2.1. Data simulation

Simulated data were used for this study. Simulation included fixed effects, residual and additive genetic variances for one continuous trait (T1) and liabilities of four categorical traits (T2–T5), and QTL effects for one categorical trait (T2). Simulation parameters chosen reflected the situation encountered in the Warmblood horse with respect to conformation and radiographic health traits [22]. Quantitative trait loci (QTL) effects were simulated for T2, assuming two QTL, two flanking markers per QTL, and five randomly distributed and equally prevalent alleles per marker. Total QTL variance was set equal to the additive genetic variance of T2, with one of the marker alleles being linked

to the unfavorable QTL allele, *i.e.* the allele increasing the probability of T2, no recombination between genetic markers and QTL, and polymorphism information content (PIC) of 0.9 of all markers. Heritabilities were set to 0.50 (T1), 0.25 (T3, T5) and 0.10 (T2, T4). Additive genetic correlations were 0.20 between T1 and T2 and between T1 and T3, and -0.20 between T1 and T4 and between T4 and T5. For the categorical traits, simulation on the linear scale was followed by dichotomization in order to obtain trait prevalences of 25% (T2, T5) or 10% (T3, T4). Assumptions for the simulation included: a male to female ratio of 1:1; random mating of 9000 dams and 400 sires per generation with equal size of maternal ($n = 5$) and different size of paternal half-sib groups ($n = 5, 150$ or 500); and existence of 5 contemporary groups per generation, with 2 contemporary groups each being represented in two subsequent generations.

From the simulated population, which included seven generations and 40 000 animals per generation, samples of 10 000 animals were randomly drawn from the fourth generation. The pedigree of these animals was traced back for three generations. Sampling was performed four times in order to generate the four replicates used for this study. Within each of the four replicates, six different datasets were created for the genetic analyses. Dataset A1 included all 10 000 animals with records for the continuous trait and the four binary traits, information on the fixed effects of sex and contemporary group, and three generations of pedigree information. Dataset B1 included 5000 animals, randomly chosen from the animals included in dataset A1, with respective information on traits, sex, contemporary group and pedigree. Dataset C1 included the same animals and the same information as dataset B1, and additional information on the marker genotype of the animals. Dataset B2 included the same animals as datasets B1 and C1 plus their parents with information on traits, sex, contemporary group and pedigree. Dataset C2 included the same animals and the same information as dataset B2, and additional information on the marker genotype of the animals. Dataset A2 included the same animals as datasets B2 and C2 plus the additional 5000 animals which were included in dataset A1 with information on traits, sex, contemporary group and pedigree. The average size of paternal halfsib groups ranged between 16.4 and 29.0, and the average size of maternal halfsib groups ranged between 1.2 and 1.6 among the animals with trait records in the six datasets. In the four replicates the total number of animals in the relationship matrix ranged between 30 618 and 30 766 for dataset A1, between 37 164 and 37 292 for dataset 2, between 19 226 and 19 986 for datasets B1 and C1, and between 26 429 and 26 556 for datasets B2 and C2.

2.2. Covariance estimation

Genetic parameters were estimated using Gibbs sampling with the threshold version of the Multiple Trait Gibbs Sampler for Animal Models (MT-GSAM) [23]. This software supports multivariate genetic analyses of any combination of continuous and categorical traits and permits the user to specify starting values and priors for additive genetic and residual (co)variance matrices. Random and residual effects are assumed to be normally distributed. Flat priors are used for the fixed effects. For our analyses, a starting value of one was chosen for all additive genetic variances, a starting value of zero was chosen for all additive genetic covariances, and the residual covariances between all traits were fixed to zero. In uni- and multivariate binary threshold models, the values for the thresholds and the residual variances are fixed to values of zero and one, respectively, to ensure identifiability of the model [6]. We additionally fixed the residual covariances for two reasons. First, the values used for these parameters in the data simulation are small and so are the values found in real data for our traits of interest. Second, it is not trivial to sample the residual covariance matrix subject to the restriction of diagonal elements fixed at one, and effective methods for this task are still being developed. A flat prior was used for the residual variance of the continuous trait (T1). For the genetic covariance matrix a flat prior was adopted in the first set of genetic analyses, and a proper prior using an inverse Wishart distribution (IW) with minimum shape parameter (*i.e.* $\nu_{IW} = n + 2$ for multivariate analysis of n traits) was adopted in the second set of genetic analyses. A small value for the IW shape parameter indicates little certainty about the genetic covariance matrix, *i.e.* a relatively flat, but proper distribution.

The following model was used for all genetic analyses:

$$y_{ijk} = \mu + F_i + a_j + e_{ijk}$$

with y_{ijk} = observation on trait T1 (continuous) or trait T2, T3, T4 or T5 (binary) for the animal j ,

μ = model constant,

F_i = fixed effect component

with $F_i = \text{SEX}_i + \text{CONT}_m$ for analyses of datasets A1, A2, B1 and B2, and

$F_i = \text{SEX}_i + \text{CONT}_m + \text{QTL}_n$ for analyses of datasets C1 and C2, and

SEX_i = fixed effect of the sex of the animal ($i = 1, 2$),

CONT_m = fixed effect of the contemporary group ($m = 1, \dots, 5$ in analyses of datasets A1, B1 and C1; $m = 1-8$ in analyses of datasets A2, B2 and C2),

QTL_n = fixed effect of the QTL marker genotype ($n = 1, 2, 3$),

a_j = random additive genetic effect of animal j ($j = 1, \dots, 30\,618$ to $30\,766$ for dataset A1; $j = 1, \dots, 37\,164$ to $37\,292$ for dataset A2; $j = 1, \dots, 19\,226$ to $19\,986$ for datasets B1 and C1; and $j = 1, \dots, 26\,429$ to $26\,556$ for datasets B2 and C2), and

e_{ijk} = random residual.

2.3. Statistical analyses

For all analyses the total length of the Gibbs chain was set to 205 000, with all samples after round 5000 being saved. The number of rounds recommended to be additionally discarded as burn-in was calculated as proposed by Raftery and Lewis using the program GIBBSIT Version 2.0 [19]. Convergence of the Gibbs chain and length of burn-in was additionally checked by visual inspection of sample plots. Only post-convergence samples were considered further. Effective sample sizes and Monte Carlo errors were calculated for all (co)variance estimates by the times series method implemented in the postgibbs analysis program POSTGIBBSF90 with a thinning rate of ten [25]. Unthinned chains were used to calculate posterior means of additive genetic (co)variance, heritability and additive genetic correlation estimates. Bias of heritability and additive genetic correlation estimates was calculated as the mean deviation of the estimated values from the true (*i.e.* simulated) values.

The influence of the prior on bias, effective sample size, and Monte Carlo error was tested *via* analysis of variance using the GLM procedure of the Statistical Analysis System [21]. Bias, effective sample size or Monte Carlo error were considered as dependent variable, and dataset (A1, A2, B1, B2, C1, C2) and prior (flat, proper) were considered as fixed effects.

3. RESULTS

3.1. Convergence and mixing of Gibbs chains

Amongst the 24 runs with flat priors for the genetic covariance matrix, two analyses of dataset C1 and one analysis of dataset B1 terminated because the genetic covariance matrix was not positive definite in the Cholesky decomposition. The terminations occurred after about 180 000 to 194 000 rounds of Gibbs sampling. Terminations did not occur in the 24 runs with proper priors for the genetic covariance matrix.

The calculated number of rounds recommended to be additionally discarded as burn-in was in the range of 100 to 2000 for the flat prior and 100 to 1500

Table I. Mean, minimum and maximum lengths of burn-in in the analyses of different datasets generated in each of the four replicates using a flat or a proper prior for the genetic covariance matrix.

Prior	Dataset						
		A1	A2	B1	B2	C1	C2
Flat	mean	43 750	27 000	97 667	48 750	36 000	86 250
	min	5000	5000	55 000	25 000	15 000	20 000
	max	85 000	45 000	175 000	70 000	57 000	175 000
Proper	mean	18 250	9750	20 750	14 750	18 250	18 750
	min	11 000	5000	5000	9000	13 000	15 000
	max	30 000	15 000	35 000	20 000	25 000	25 000

for the proper prior. However, visual inspection of sample plots revealed that periods of seemingly stationary distributions were frequently followed by non-stationary sample series of different length. In order to be sure that convergence was obtained, length of burn-in was therefore set to 5000 to 175 000 rounds for the flat prior and to 5000 to 35 000 rounds for the proper prior. Mean, minimum and maximum lengths of burn-in in the analyses of the six different datasets are given in Table I. Regardless of the amount and the distribution of available information, the burn-in period was consistently shorter with the proper prior than with the flat prior.

Inspection of sample plots revealed slow mixing properties in all three runs which terminated and in 10 of the 21 runs with a flat prior that did not terminate. In these cases, the sample values for the additive genetic variance of trait T2 or T4 remained near zero for large blocks of consecutive cycles of sampling. In one case this occurred as late as around round 170 000. In another case there was a period of slow mixing with respect to the additive genetic variance of trait T2 from about round 110 000 to round 175 000, then mixing improved and the samples departed from zero. Slow mixing was observed in all three completed analyses of dataset B1, and in one of the two completed analyses of dataset C1. Slow mixing was not observed in the flat prior analyses of dataset A2 and the proper prior analyses of all datasets. Comparison of sample plots from flat prior and proper prior analyses of the same datasets revealed minor differences with respect to the additive genetic and residual variances of the continuous trait (T1), but showed considerably better mixing for all other (co)variance parameters with the proper prior (see Figs. 1 to 4).

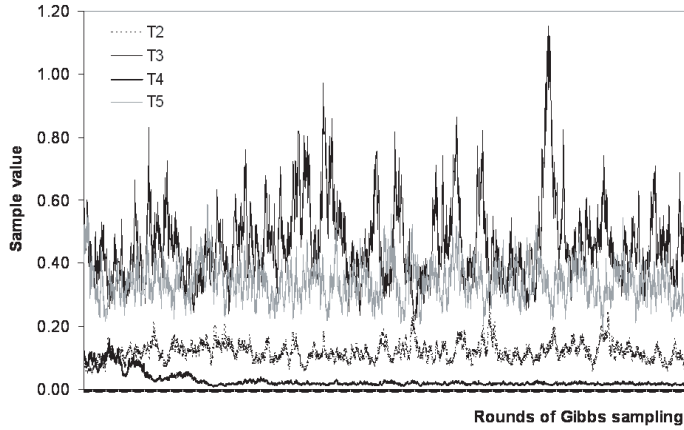


Figure 1. Sample plots for the additive genetic variances of the binary traits T2 to T5 in an analysis of dataset A1 using a flat prior for the genetic covariance matrix.

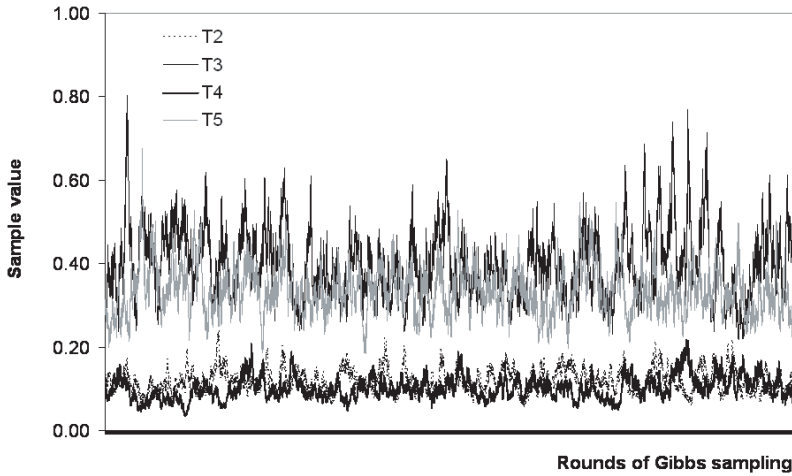


Figure 2. Sample plots for the additive genetic variances of the binary traits T2 to T5 in an analysis of dataset A1 using a proper prior for the genetic covariance matrix (same dataset as in Fig. 1).

3.2. Effective sample sizes

Mean, minimum and maximum effective sample sizes for the sampled (co)variance parameters are given in Table II. For all parameters effective sample sizes were on average larger when the proper prior was used for the genetic covariance matrix than when the flat prior was used. With few exceptions ($\sigma_a^2_4$ in datasets A1 and C2, cov_{a14} in datasets A2 and C2, cov_{a35} in

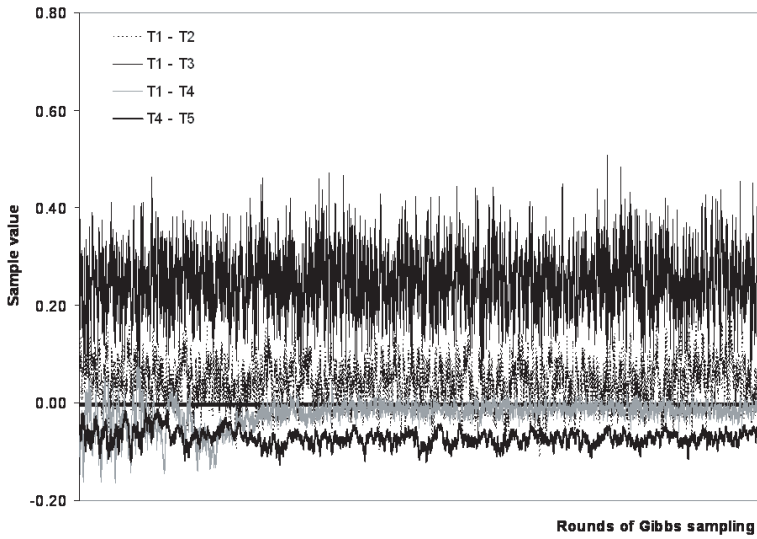


Figure 3. Sample plots for additive genetic covariances between continuous trait T1 and binary traits T2, T3 and T4 and between the two binary traits T4 and T5 in an analysis of dataset A1 using a flat prior for the genetic covariance matrix (same dataset as in Figs. 1 and 2).

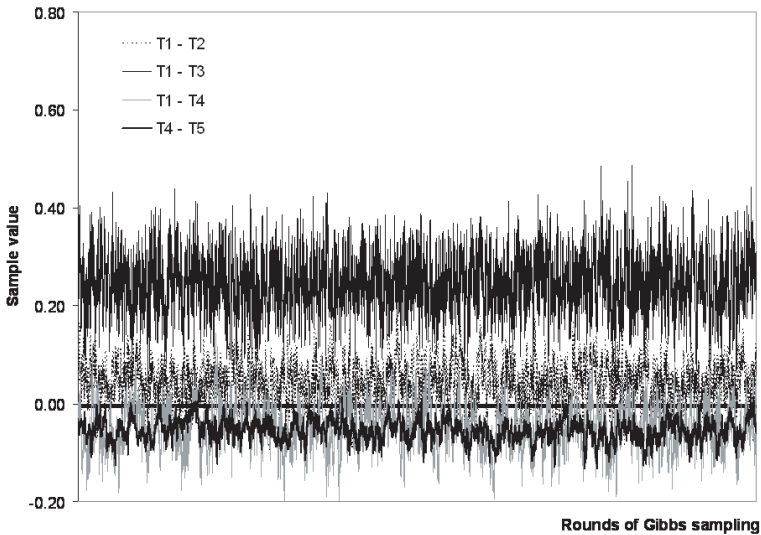


Figure 4. Sample plots for additive genetic covariances between continuous trait T1 and binary traits T2, T3 and T4 and between the two binary traits T4 and T5 in an analysis of dataset A1 using a flat prior for the genetic covariance matrix (same dataset as in Figs. 1 to 3).

Table II. Mean, 25% percentile (P25%) and 75% percentile (P75%) of effective sample sizes calculated for the additive genetic variances (σ_a^2) of traits T1 to T5, the residual variance (σ_e^2) of trait T1 and the additive genetic covariances (cov_a) between traits T1 to T5 when using a flat or a proper prior for the genetic covariance matrix.

(Co)Variance parameter	Flat prior		Proper prior	
	Mean	(P25%, P75%)	Mean	(P25%, P75%)
$\sigma_a^2_1$	821.45	(334.9, 1288.2)	1050.64	(550.8, 1521.8)
$\sigma_e^2_1$	1085.45	(363.1, 1599.1)	1354.76	(611.5, 2111.1)
$\sigma_a^2_2$	48.18	(39.7, 59.8)	90.36	(68.7, 116.7)
$\sigma_a^2_3$	62.16	(50.6, 72.5)	88.98	(69.9, 107.1)
$\sigma_a^2_4$	52.84	(35.7, 60.9)	69.37	(48.7, 88.3)
$\sigma_a^2_5$	109.89	(62.2, 163.2)	157.61	(117.2, 201.6)
cov_{a12}	336.00	(323.9, 392.9)	500.82	(366.3, 595.5)
cov_{a13}	562.04	(465.2, 700.2)	716.23	(592.0, 812.2)
cov_{a14}	277.09	(173.9, 331.2)	317.96	(243.6, 400.1)
cov_{a15}	923.30	(691.5, 1171.3)	1208.43	(1081.7, 1330.1)
cov_{a23}	102.90	(72.9, 126.7)	157.89	(132.9, 183.6)
cov_{a24}	54.73	(25.0, 74.8)	108.80	(97.9, 123.6)
cov_{a25}	152.22	(104.2, 207.4)	230.53	(187.6, 272.9)
cov_{a34}	79.47	(57.8, 110.4)	130.27	(102.0, 158.2)
cov_{a35}	256.20	(170.6, 334.0)	294.60	(231.3, 361.4)
cov_{a45}	139.40	(62.8, 175.7)	183.61	(140.8, 245.3)

Trait characteristics (kind of trait / prevalence of binary trait, simulated heritability): T1 - continuous, 0.50; T2 - binary / 0.25, 0.10; T3 - binary / 0.10, 0.25; T4 - binary / 0.10, 0.10; and T5 - binary / 0.25, 0.25; with simulated additive genetic correlations of $r_{g12} = 0.20$, $r_{g13} = 0.20$, $r_{g14} = -0.20$, and $r_{g45} = -0.20$.

datasets A1 and C1) this finding was consistent and significant across datasets with $P < 0.01$. Effective sample sizes smaller than 20 were estimated for $\sigma_a^2_2$, $\sigma_a^2_3$, $\sigma_a^2_4$, cov_{a12} , cov_{a23} , cov_{a24} , cov_{a25} and cov_{a34} in analyses of datasets B1, B2, C1 and C2 when the flat prior was used for the genetic covariance matrix. Effective sample sizes were larger than 30 for all (co)variance parameters and in all analyses when the proper prior was used. For most of the (co)variance parameters the variance of the effective sample sizes was considerably smaller with the proper prior.

3.3. Monte Carlo errors

Means of Monte Carlo errors ranged from 0.004 to 0.012 for the sampled variances and from 0.002 to 0.003 for the sampled covariances in the flat prior analyses. Corresponding values in the proper prior analyses were 0.003

Table III. Mean, 25% percentile (P25%) and 75% percentile (P75%) of bias of estimated heritabilities (h^2) and selected additive genetic correlations (r_g) for traits T1 to T5 when using a flat or a proper prior for the genetic covariance matrix.

Genetic parameter (simulated value)	Flat prior		Proper prior	
	Mean	(P25%, P75%)	Mean	(P25%, P75%)
h^2_1	0.0248	(-0.0494, 0.0977)	0.0140	(-0.0512, 0.0841)
h^2_2	-0.0352	(-0.1137, 0.0701)	0.3518	(0.0477, 0.6128)
h^2_3	0.0475	(-0.0346, 0.1581)	-0.0257	(-0.0969, 0.0459)
h^2_4	-0.1984	(-0.6623, 0.1447)	0.1340	(-0.0020, 0.2444)
h^2_5	0.1275	(0.0684, 0.1630)	0.0796	(0.0002, 0.1124)
r_{g12}	-0.0310	(-0.5264, 0.1611)	-0.2280	(-0.4903, -0.0285)
r_{g13}	-0.1173	(-0.2179, -0.0414)	-0.0643	(-0.5457, -0.0798)
r_{g14}	0.1576	(-0.5653, 0.2267)	-0.2919	(-0.5457, -0.0798)
r_{g45}	0.7087	(-0.4088, 1.0340)	-0.0272	(-0.3622, 0.4075)

Trait characteristics (kind of trait / prevalence of binary trait, simulated heritability):

T1 - continuous, 0.50; T2 - binary / 0.25, 0.10; T3 - binary / 0.10, 0.25; T4 - binary / 0.10, 0.10; and T5 - binary / 0.25, 0.25; with simulated additive genetic correlations of $r_{g12} = 0.20$, $r_{g13} = 0.20$, $r_{g14} = -0.20$, and $r_{g45} = -0.20$.

to 0.009 for the sampled variances and 0.001 to 0.003 for the sampled covariances. For most parameters the Monte Carlo errors and their variances were significantly smaller with the proper prior ($P < 0.05$). Influence of the prior on the Monte Carlo errors was more pronounced for the variance than for the covariance estimates.

3.4. Bias

Mean, minimum and maximum bias of heritability and additive genetic correlation estimates are given in Table III. Mean upward bias of more than 0.15 was determined for r_{g14} and r_{g45} in the flat prior analyses and for h^2_2 in the proper prior analyses. Mean downward bias of more than 0.15 was determined for h^2_4 in the flat prior analyses and for r_{g12} and r_{g14} in the proper prior analyses. Differences between biases from flat and proper prior analyses were significant across datasets for h^2_2 , h^2_3 , h^2_4 , r_{g14} and r_{g45} ($P < 0.05$). The results from analyses of datasets including trait information on animals from two subsequent generations (datasets A2, B2 and C2) were consistently less biased than results from analyses of datasets with respective information on animals from only one generation (datasets A1, B1 and C1). For the most parameters, bias variance was considerably larger in the flat prior analyses than in the proper prior analyses.

4. DISCUSSION

In this study simulated datasets of different size and structure were used to investigate the influence of the choice of priors for the genetic (co)variance matrix on the posterior distributions of genetic parameters in multivariate threshold models using Gibbs sampling. Choice of simulation parameters was very specific, resembling data and pedigree structures encountered in the Warmblood horse [22]. Simulated data provided the basis for investigations of different aspects of multivariate genetic analyses in mixed linear-threshold animal models. However, for this study only four population samples from one generation were used, but within each of these replicates six different datasets were generated and analyzed comparatively.

Drawing conclusion from a Gibbs chain is conditional on being able to assure that sampling has been from a proper posterior distribution. Only then Gibbs conditionals correspond to an existing joint density, *i.e.* represent compatible conditional densities [2]. Mathematical proof of posterior propriety may be infeasible, and in most applications empirical methods have been used to identify convergence problems resulting from impropriety, near impropriety, or very slow mixing [7]. In this study visual inspection of sample plots revealed clear signs of convergence problems in more than 50 percent of the runs in which a flat prior was used for the genetic covariance matrix. Termination occurred less often and much later than slow mixing was visible. However, in one of the flat prior analyses the course of the Gibbs chain indicated convergence problems not until about round 170 000. A previous study which found a prior-independent behavior of the Gibbs chain was based on Gibbs chains of a total length of 5000 rounds [15]. In our study the total chain length was 205 000, but for the flat prior analyses this number might still not be large enough to ensure the detection of lack of convergence.

If the use of a flat prior for the genetic covariance matrix results in sampling from an almost improper posterior, convergence of the whole set of parameters might be extremely slow [18]. The large number of rounds that have to be discarded as burn-in in these cases and the required increase of the total chain length might render the flat prior approach infeasible in a practical setting. The inclusion of continuous traits in multivariate analyses of categorical traits, *i.e.* the use of a multivariate linear-threshold model, was previously found to improve mixing behavior of the Gibbs chain [11, 14, 20, 29]. However, in our study joint analyses of continuous and categorical traits did not result in satisfactory mixing in connection with the flat prior.

Improper posteriors can not only result from the use of improper priors, but can also be caused by insufficient data structure [26]. Blowing-up of the Gibbs

chain or trapping to zero can occur in the presence of the so-called extreme category problem [16]. There was no extreme category problem in our data with regards to the fixed factors, which had few levels.

Convergence problems are not always related to the extreme category problem and are more likely to occur in the threshold than in the linear model [8, 13]. In this study, mixing was generally considerably slower and effective sample sizes were considerably smaller for the categorical traits than for the continuous trait. In the flat prior analyses convergence problems occurred with all datasets but the largest dataset in which trait information on two subsequent generations of animals was available. In the proper prior analyses convergence was also dependent on size and structure of the dataset, but was attained for all datasets.

If an animal model is employed and there is only one observation per animal, accuracy of prediction of the random additive genetic animal effect and bias of heritability and additive genetic correlations estimates depends on the connectedness, *i.e.* the structure of the genetic covariance matrix [17]. In this study the average size of paternal halfsib groups ranged between 16 and 29 in the different datasets, and there were animals that had neither paternal nor maternal halfsibs included. This pedigree structure resembled a situation which is commonly encountered in animal breeding. Limited information for the random effects leads to a positive bias of heritability estimates, particularly for those traits with low heritability [17]. Animal threshold model analyses can lead to negatively biased estimates of additive genetic correlations between continuous and binary traits [3]. The outputs of all runs which did not terminate were included for the bias calculation, and heritabilities and additive genetic correlations were somewhat biased with both the flat and the proper prior. The previously reported pattern of bias was only seen in the analyses with the proper prior, but not in the analyses with the flat prior. Here, the pattern of bias was less consistent and differed considerably between replicates and datasets. Superiority of the proper prior was less clear with respect to estimation accuracy than with respect to mixing, but the most extreme biases were seen in flat prior analyses.

For both linear and threshold models it has been shown that available convergence diagnostics will not always succeed to identify impropriety of posteriors, and the results which are based on a null recurrent or transient Gibbs chain might appear reasonable [7, 18]. In practical applications of Gibbs sampling for (co)variance component estimation using real data, results of analogous studies might not always be available for comparison purposes. In this simulation study, comparison of estimated and true parameter values indicated

that even if convergence of the Gibbs chain was attained, one cannot expect unbiased estimates in an analysis of realistically structured data in a multi-trait mixed linear-threshold animal model.

5. CONCLUSIONS

The use of a flat prior for the genetic covariance matrix in an analysis using a multi-trait mixed linear-threshold animal model and Gibbs sampling may lead to improper posteriors. Mixing and convergence can be extremely slow, but convergence problems may be noticed only after a large number of rounds of sampling. The use of a proper prior assures a proper posterior and results in improved mixing and convergence of the Gibbs chain. In the absence of a proven proper posterior with flat priors, a proper prior should be used for the genetic covariance matrix to avoid misleading interpretations of Gibbs outputs.

ACKNOWLEDGEMENTS

This work was supported by a fellowship of K.F. Stock within the Postdoc-program of the German Academic Exchange Service (Deutscher Akademischer Austauschdienst, DAAD). I. Hoeschele acknowledges financial support from the US National Institute of Health (GM66103-01).

REFERENCES

- [1] Arango J., Misztal I., Tsuruta S., Culbertson M., Herring W., Threshold-linear estimation of genetic parameters for farrowing mortality, litter size, and test performance of Large White sows, *J. Anim. Sci.* 83 (2005) 499–506.
- [2] Arnold B.C., Press S.J., Compatible conditional distributions, *J. Am. Stat. Assoc.* 84 (1989) 152–156.
- [3] Gates P., Johannson K., Danell B., “Quasi-REML” correlation estimates between production and health traits in the presence of selection and confounding: a simulation study, *J. Anim. Sci.* 77 (1999) 558–568.
- [4] Gelfand A.E., Hills S.E., Racine-Poon A., Smith A.F.M., Illustration of Bayesian inference in normal data using Gibbs sampling, *J. Am. Stat. Assoc.* 85 (1990) 972–985.
- [5] Gelfand A.E., Sahu S.K., Identifiability, improper priors, and Gibbs sampling for generalized linear models, *J. Am. Stat. Assoc.* 94 (1999) 247–253.
- [6] Harville D.A., Mee R.W., A mixed-model procedure for analyzing ordered categorical data, *Biometrics* 40 (1984) 393–408.
- [7] Hobert J.P., Casella G., The effect of improper priors on Gibbs sampling in hierarchical linear mixed models, *J. Am. Stat. Assoc.* 91 (1996) 1461–1473.

- [8] Hoeschele I., Tier B., Estimation of variance components of threshold characters by marginal posterior modes and means *via* Gibbs sampling, *Genet. Sel. Evol.* 27 (1995) 519–540.
- [9] Holm B., Bakken M., Vangen O., Rekaya R., Genetic analysis of age at first service, return rate, litter size, and weaning-to-first service interval of gilts and sows, *J. Anim. Sci.* 83 (2005) 41–48.
- [10] Janssens S., Vandepitte W., Bodin L., Genetic parameters for litter size in sheep: natural *versus* hormone-induced oestrus, *Genet. Sel. Evol.* 36 (2004) 543–562.
- [11] Jensen J., Sire evaluation for type traits with linear and non-linear procedures, *Livest. Prod. Sci.* 15 (1986) 165–171.
- [12] Kadarmideen H.N., Thompson R., Simm G., Linear and threshold model genetic parameters for disease, fertility and milk production in dairy cattle, *Anim. Sci.* 71 (2000) 411–419.
- [13] Luo M.F., Boettcher P.J., Schaeffer L.R., Dekkers J.C.M., Bayesian inference for categorical traits with an application to variance component estimation, *J. Dairy Sci.* 84 (2001) 694–704.
- [14] Mäntysaari E.A., Gröhn Y.T., Quaas R.L., Clinical ketosis: phenotypic and genetic correlations between occurrences and with milk yield, *J. Dairy Sci.* 74 (1991) 3985–3993.
- [15] Magnabosco C.U., Lôbo R.B., De los Reyes A., Famula T.R., Bayesian inference on growth traits for Nelore cattle in Brazil, in: *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production*, 11–16 January 1998, Vol. 25, University of New England, Armidale, Australia, pp. 556–559.
- [16] Misztal I., Gianola D., Foulley J.L., Computing aspects of a nonlinear method of sire evaluation for categorical data, *J. Dairy Sci.* 72 (1989) 1557–1568.
- [17] Moreno C., Sorensen D., García-Cortés L.A., Varona L., Altarriba J., On biased inferences about variance components in the binary threshold model, *Genet. Sel. Evol.* 29 (1997) 145–160.
- [18] Natarajan R., McCulloch C.E., Gibbs sampling with diffuse proper priors: a valid approach to data-driven inference, *J. Comp. Graph. Stat.* 7 (1998) 267–277.
- [19] Raftery A.E., Lewis S.M., How many iterations in the Gibbs sampler?, in: Bernardo J.M., Berger J.O., Dawid A.P., Smith A.F.M. (Eds.), *Bayesian Statistics*, Vol. 4, Oxford University Press, Oxford, 1992, pp. 763–773.
- [20] Simianer H., Schaeffer L.R., Estimation of covariance components between one continuous and one binary trait, *Genet. Sel. Evol.* 21 (1989) 303–315.
- [21] Statistical Analysis Systems (SAS®), version 9.1.3, User's guide, SAS® Institute, Cary, NC, USA, 2005.
- [22] Stock K.F., Distl O., Genetic correlations between conformation traits and radiographic findings in the limbs of German Warmblood riding horses, *Genet. Sel. Evol.* 38 (2006), 657–671.
- [23] Van Tassell C.P., Van Vleck L.D., Multiple-trait Gibbs sampler for animal models: flexible programs for Bayesian and likelihood-based (co)variance component inference, *J. Anim. Sci.* 74 (1996) 2586–2597.
- [24] Van Tassell C.P., Van Vleck L.D., Gregory K.E., Bayesian analysis of twinning and ovulation rates using a multiple-trait threshold model and Gibbs sampling, *J. Anim. Sci.* 76 (1998) 2048–2061.

- [25] Tsurata S: **POSTGIBBSF90**
[<http://nce.ads.uga.edu/~ignacy/newprograms.html>].
- [26] Wang C.S., Implementation issues in Bayesian analysis in animal breeding, in: Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, 11–16 January 1998, University of New England, Armidale, Australia, pp. 481–488.
- [27] Wang C.S., Rutledge J.J., Gianola D., Marginal inferences about variance components in a mixed linear model using Gibbs sampling, *Genet. Sel. Evol.* 26 (1993) 41–62.
- [28] Wang C.S., Rutledge J.J., Gianola D., Bayesian analysis of mixed linear models *via* Gibbs sampling with an application to litter size of Iberian pigs, *Genet. Sel. Evol.* 26 (1994) 1–25.
- [29] Wang C.S., Quaas R.L., Pollak E.J., Bayesian analysis of calving ease scores and birth weights, *Genet. Sel. Evol.* 29 (1997) 117–143.