Genet. Res., Camb. (2008), **90**, *pp.* 179–190.
doi:10.1017/S0016672308009233 Printed in the United Kingdom

Bayesian analysis of quantitative traits using skewed distributions

L. VARONA*, N. IBAÑEZ-ESCRICHE, R. QUINTANILLA, J. L. NOGUERA AND J. CASELLAS
Genètica i Millora Animal, IRTA-Lleida, Av. Rovira Roure 191, 25198 Lleida, Spain

(*Received 18 June 2007 and in revised form 13 December 2007*)

Summary

Statistical models for genetic evaluation often make use of Gaussian distributions. However, some new statistical developments allow the use of an asymmetric distribution for the residuals. Within this context, we analysed three different patterns for the residual term on a data set consisting of 63 208 litter-size records, belonging to 19 255 sows, with a pedigree including 27 911 individuals. The three different residual distributions were: (1) Gaussian distribution, (2) asymmetric Gaussian distribution and (3) asymmetric Gaussian distribution with a hierarchical scheme for the asymmetry parameter. The operational model always included order of parity and herd-year-season as systematic effects, and the permanent environmental and infinitesimal genetic effect of each sow as random effects. The most suitable model using the deviance information criterion (DIC) and posterior predictive checking was model 3. This implies systematic, additive genetic and permanent environmental control of both litter size and the asymmetry parameter of the residual distribution. The asymmetry parameter can be understood as a measure of sensitivity to negative (or positive) environmental influences on phenotypes. The posterior mean (standard deviation) of the additive genetic variance was 0·28 (0·06) for litter size and 0·07 (0·01) for the asymmetry parameter. The posterior mean (standard deviation) of the additive genetic correlation between litter size and the asymmetry parameter was 0·21 (0·07).

1. Introduction

Mixed linear models (Henderson, 1984) are used broadly in livestock and plant breeding to predict breeding values and to estimate variance components for traits of interest. The Gaussian distribution of the residual term is a common assumption in mixed linear models. In the animal breeding context, an alternative to the Gaussian assumption was proposed by Stranden & Gianola (1999), who modelled the residual term using a Student's *t* density. This kind of heavy-tailed distribution allows for more extreme residual values and, as a consequence, deviations from the Gaussian distribution such as preferential treatment (Kuhn *et al.*, 1994) or other causes of outliers or abnormal phenotypic records (Jamrozik *et al.*, 2004). Nevertheless, both Gaussian and Student's *t* distributions are symmetric, and little investigation into alternative approaches assuming a variable degree of skewness for the residual term has been done.

It is important to note that most of the uncontrolled sources of variation in animal production can be viewed as adverse factors involving a slight, moderate or even dramatic reduction in productive performance (e.g. pathologies, heat or cold, stress, fights and accidents), whereas favourable factors are probably limited to preferential treatment and social dominance hierarchy. Some authors have proposed the use of mixtures of distributions to model these peculiarities (Gianola *et al.*, 2006), where observations can be assigned to different distributions (e.g. healthy *vs.* affected). However, it is very difficult to assign records to a finite number of distributions when sources of variation are unknown (e.g. preferential treatment, sub-clinical pathologies,

* Corresponding author. Tel: +34 973003441. Fax. +34 973238301. e-mail: Luis.Varona@irta.es

unregistered heat or cold stress) and the statistical approach for infinite mixtures becomes very complex. An alternative to model these data is via the use of non-symmetric residual distribution for the environmental deviations, as was initially proposed by Fernandez & Steel (1998) and adapted to an animal breeding context by von Rohr & Hoeschele (2002).

The use of skewed residual distributions in linear models has been focused mainly on describing the overall asymmetry in the population analysed (Fernandez & Steel, 1998; von Rohr & Hoeschele, 2002). However, individual variation in the degree of asymmetry also seems plausible. Within this context, the asymmetry parameter could be modelled through a hierarchical model (Wakefield *et al*., 1994; Varona *et al*., 1997). Each record-specific asymmetry parameter would represent the ability to buffer against undesirable environmental influences and after accounting for genetic and environmental sources of variation. Indeed, this approach could be viewed as an attractive method to model robustness (or weakness) against controllable and uncontrollable genetic and environmental sources of variation. Within this context, there are several references in the literature regarding the genetic determinism of disease resistance (Gibson, 2002; Bishop, 2004) and immune responses (Mallard *et al*., 1998; Henryon *et al*., 2001). Despite the crucial role that selection for disease resistance or robustness could play in animal breeding, not much attention has been focused on this question, mainly due to the difficulties involved in obtaining appropriate phenotypic records (Rothschild, 1991).

Unfortunately, the Bayesian implementation of skewed distributions using the procedure suggested by Fernandez & Steel (1998) and von Rohr & Hoeschele (2002), and involving Markov chain Monte Carlo (MCMC) techniques, requires a Metropolis–Hastings step (Hastings, 1970) to sample the asymmetry parameter. The development of a hierarchical model for the asymmetry parameter is therefore complex and computationally demanding for large data sets. However, other authors in statistics have developed new procedures for modelling non-symmetric distributions (Sahu *et al*., 2003; Jara & Quintana, 2007).

The aims of this study are: (1) to include non-symmetric residual distributions in the linear mixed models currently used in livestock and plant breeding following Sahu *et al*. (2003), (2) to develop a hierarchical Bayesian scheme including systematic and additive genetic variation of the asymmetry parameter (Wakefield *et al*., 1994; Varona *et al*., 1997) and (3) to implement and compare these procedures with the standard mixed model approach using a litter-size data set from a pure-bred Landrace commercial population.

2. Materials and methods

(i) *Statistical models*

We took as a starting point the standard mixed model commonly used in animal breeding (Henderson, 1984):

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Wp} + \mathbf{Zu} + \mathbf{e}, \tag{1}$$

where $\mathbf{y}$ is the vector of phenotypic data (number of piglets born alive (NBA)), $\boldsymbol{\beta}$ is the vector of systematic effects, $\mathbf{p}$, $\mathbf{u}$ and $\mathbf{e}$ are the vectors of permanent environmental effects, additive genetic effects and residuals, respectively, and $\mathbf{X}$, $\mathbf{W}$ and $\mathbf{Z}$ are the appropriate incidence matrices. Under a standard Bayesian approach, bounded uniform prior distributions between -50 and 50 units were assumed for $\boldsymbol{\beta}$, and the following independent prior distributions were assumed for $\mathbf{p}$ and $\mathbf{u}$:

$$\mathbf{p} \sim N(\mathbf{0}, \mathbf{I}_p\sigma^2_p), \tag{2}$$

$$\mathbf{u} \sim N(\mathbf{0}, \mathbf{A}\sigma^2_u), \tag{3}$$

where $\mathbf{I}_p$ is the appropriate identity matrix, $\mathbf{A}$ is the numerator relationship matrix between individuals and σ^2_p and σ^2_u are the permanent environmental and the additive genetic variances, respectively. In addition, for computational convenience, prior distributions for σ^2_p and σ^2_u were scale inverse chi-squared distributions with parameters $s=0$ and $v=-2$, which reduced it to a uniform distribution (Sorensen & Gianola, 2002). Moreover, it is computationally equivalent to a bounded proper prior between 0 and a huge and unreachable value. For the residual term, we considered three different prior distributions.

(a) A priori *Gaussian distribution for the residual term* (*model 1*)

The simplest model assumed a standard Gaussian distribution of residuals:

$$\mathbf{e} \sim N(0, \mathbf{I}_e\sigma^2_e), \tag{4}$$

$$f(\mathbf{e}|\sigma^2_e) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2_e}} \exp\left\{-\frac{e_i^2}{2\sigma^2_e}\right\}, \tag{5}$$

where n is the number of phenotypic records, $\mathbf{I}_e$ is an identity matrix with dimensions $n \times n$, e_i is the ith term in $\mathbf{e}$ and σ^2_e is the residual variance. As before, the prior distribution for σ^2_e was a scale inverse chi-squared distribution with parameters $s=0$ and $v=-2$.

(b) A priori *asymmetric Gaussian distribution for the residual term* (*model 2*)

Following Sahu *et al*. (2003), asymmetry in the residual term can be easily modelled by a

https://doi.org/10.1017/S0016672308009233 Published online by Cambridge University Press

skewed-normal density:

$$\mathbf{e} \sim \mathrm{SN}(\mathbf{0}, \mathbf{I}_e\sigma_e^2, \lambda), \tag{6}$$

$$f(\mathbf{e}|\sigma_e^2, \lambda) = \prod_{i=1}^{n} \frac{2}{\sqrt{\sigma_e^2+\lambda^2}} \phi\left(\frac{e_i}{\sqrt{\sigma_e^2+\lambda^2}}\right) \times \Phi\left(\frac{\lambda}{\sigma_e}\frac{e_i}{\sqrt{\sigma_e^2+\lambda^2}}\right), \tag{7}$$

where λ is the degree of asymmetry defined in the real space and ϕ and Φ denote the density function and cumulative distribution function of a standard normal distribution with kernel as defined between parentheses, respectively.

Following Sahu *et al.* (2003), the mean of the asymmetric Gaussian distribution is

$$E(e_i) = \frac{\sqrt{2}\lambda}{\sqrt{\pi}}, \tag{8}$$

the variance becomes

$$\mathrm{Var}(e_i) = \sigma_e^2 + \lambda^2\left(1 - \frac{2}{\pi}\right) \tag{9}$$

and the third central moment of the distribution is

$$E[e_i - E(e_i)]^3 = \lambda^3\sqrt{\frac{2}{\pi}}\left(\frac{4}{\pi} - 1\right). \tag{10}$$

Thus, the three parameters of the asymmetric Gaussian distribution are statistically identifiable from the first three moments of a given data set.

As before, the prior distribution for σ_e^2 was a scale inverse chi-squared distribution with parameters $s=0$ and $v=-2$. Finally, the prior distribution for λ was assumed flat between bounded limits $(-50, 50)$.

(c) A priori *asymmetric Gaussian distribution for the residual term with a hierarchical Bayesian scheme* (*model 3*)

As suggested in the previous sections, the asymmetry parameter can be modelled under a hierarchical structure, with the a priori distribution of **e** being:

$$\mathbf{e} \sim \mathrm{SN}(\mathbf{0}, \mathbf{I}_e\sigma_e^2, \boldsymbol{\lambda}), \tag{11}$$

$$f(\mathbf{e}|\sigma_e^2, \boldsymbol{\lambda}) = \prod_{i=1}^{n} \frac{2}{\sqrt{\sigma_e^2+\lambda_i^2}} \phi\left(\frac{e_i}{\sqrt{\sigma_e^2+\lambda_i^2}}\right) \times \Phi\left(\frac{\lambda_i}{\sigma_e}\frac{e_i}{\sqrt{\sigma_e^2+\lambda_i^2}}\right), \tag{12}$$

where $\boldsymbol{\lambda}$ is the vector of λ_i. A hierarchical model was assumed for $\boldsymbol{\lambda}$ such as

$$\boldsymbol{\lambda} = \mathbf{X}\boldsymbol{\beta}_\lambda + \mathbf{W}\mathbf{p}_\lambda + \mathbf{Z}\mathbf{u}_\lambda, \tag{13}$$

where $\boldsymbol{\beta}_\lambda$, $\mathbf{p}_\lambda$ and $\mathbf{u}_\lambda$ are the vectors of systematic, permanent environmental and additive genetic effects, respectively. The prior distribution for each systematic effect of the asymmetry parameter is defined as a bounded uniform distribution between -50 and 50 units, and $\mathbf{p}_\lambda$ and $\mathbf{u}_\lambda$ are assumed to be correlated with **p** and **u**, respectively. Thus, the prior distributions for both effects were defined as:

$$\begin{pmatrix}\mathbf{p}\\ \mathbf{p}_\lambda\end{pmatrix} \sim N\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}, \mathbf{I}_p \otimes \mathbf{D}\right), \tag{14}$$

$$\begin{pmatrix}\mathbf{u}\\ \mathbf{u}_\lambda\end{pmatrix} \sim N\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}, \mathbf{A} \otimes \mathbf{G}\right), \tag{15}$$

where **D** and **G** are 2×2 permanent environmental and additive genetic (co)variance matrices, respectively. The following inverted Wishart distributions were assumed for **G** and **D**:

$$\mathbf{G} \sim \mathrm{IW}(0, -3), \tag{16}$$

$$\mathbf{D} \sim \mathrm{IW}(0, -3). \tag{17}$$

(ii) *Field data*

The models were tested on a data set consisting of NBA per litter from a pure-bred Landrace commercial pig population. The data set consisted of 63 208 litter-size records collected between 1982 and 1997 in six commercial farms from COPAGA SCCL (Lleida, Spain). Phenotypic data were from to 19 255 sows and the pedigree included 27 911 individuals. The raw mean was 9·04 piglets born alive with a standard deviation of 2·41 piglets. Data were grouped in six orders of parity (1, 2, 3, 4, 5 and >5) and 226 herd-year-season effects.

(iii) *Bayesian implementation*

The Bayesian implementation of the models was performed using a Gibbs sampler (Gelfand & Smith, 1990). Full details of the conditional distributions needed for the implementation are presented in the Appendix. For each model, a single chain of 500 000 iterations was performed after discarding the first 50 000. Convergence was checked using the Raftery & Lewis (1992) and Gelman *et al.* (1996) procedures.

(iv) *Sensitivity analysis*

The influence of prior information on the posterior distribution has been tested under models 2 and 3.

https://doi.org/10.1017/S0016672308009233 Published online by Cambridge University Press

For model 2, the assumed prior for the degree of asymmetry was uniform, and model performance under two additional priors was studied,

$$\lambda \sim N(1, 1), \tag{18}$$

$$\lambda \sim N(-1, 1). \tag{19}$$

Under model 3, the prior distribution for the additive variance components was assigned to an inverted Wishart distribution with parameters 0 and -3. Two alternative scenarios were considered:

$$\mathbf{G} \sim \text{IW}(\mathbf{G}_*, 10), \qquad \mathbf{G}_* = \begin{bmatrix} 0{\cdot}50 & 0 \\ 0 & 0{\cdot}10 \end{bmatrix}, \tag{20}$$

$$\mathbf{G} \sim \text{IW}(\mathbf{G}_{**}, 10), \qquad \mathbf{G}_{**} = \begin{bmatrix} 0{\cdot}10 & 0 \\ 0 & 0{\cdot}02 \end{bmatrix}. \tag{21}$$

(v) *Model checking and model comparison*

(a) *Model checking*

The fit of the statistical model to the data analysed can be assessed in a variety of ways. In a Bayesian context, a standard method for model checking involves the use of the posterior predictive distributions of discrepancies to diagnose particular failures of the model (Rubin, 1984; Gelman *et al*., 1996). Take $T(\mathbf{y}; \boldsymbol{\theta})$ as a specific discrepancy measure allowing comparison of the posterior distribution of $T(\mathbf{y}^{\text{obs}}; \boldsymbol{\theta})$ with the posterior predictive distribution of $T(\mathbf{y}^{\text{rep}}; \boldsymbol{\theta})$. Here, $\mathbf{y}^{\text{obs}}$ is the observed data, $\mathbf{y}^{\text{rep}}$ is a simulated replicate of the data set at each iteration of the MCMC procedure, and $\boldsymbol{\theta}$ represents the values sampled for all the parameters in the model in the given iteration. Systematic differences between $T(\mathbf{y}^{\text{obs}}; \boldsymbol{\theta})$ and $T(\mathbf{y}^{\text{rep}}; \boldsymbol{\theta})$ indicate a possible failing of the model.

In our particular case, we wanted to study the global discrepancy and the discrepancy associated with order of parity and sire family, and their relationship with the symmetry of the environmental variation under model 1. For global discrepancy, we defined the following measure of skewness:

$$T(\mathbf{y}, \boldsymbol{\theta}_1) = \frac{\sum_{i=1}^{n} (y_i - \mu_i)^3}{(n-1)\tilde{\sigma}^2_{\text{e}}}, \tag{22}$$

where $\boldsymbol{\theta}_1$ represents all the unknown parameters in model 1, $\tilde{\sigma}^2_{\text{e}}$ is the sampled value of the residual variance at each iteration and μ_i is the ith row in $\mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{p} + \mathbf{Z}\mathbf{u}$. The expected value of $T(\mathbf{y}^{\text{obs}}, \boldsymbol{\theta}_1) - T(\mathbf{y}^{\text{rep}}, \boldsymbol{\theta}_1)$ under model 1 is zero, and values larger or smaller than 0 indicate asymmetry of the residuals. The degree of discrepancy was defined through the predictive P-values, calculated as the proportion of iterations where $T(\mathbf{y}^{\text{obs}}, \boldsymbol{\theta}_1) - T(\mathbf{y}^{\text{rep}}, \boldsymbol{\theta}_1)$ was below zero (Gelman *et al*., 1996).

To study the discrepancy associated with the jth specific effect (order of parity or sire family), we also calculated the following measure:

$$T_j(\mathbf{y}, \boldsymbol{\theta}_1) = \frac{\sum_{i=1}^{N_j} (y_i - \mu_i)^3}{(N_j - 1)\tilde{\sigma}^2_{\text{e}(j)}}, \tag{23}$$

At this point, N_j is the number of records for the jth effect and $\tilde{\sigma}^2_{\text{e}(j)}$ is the sampled value of the residual variance within the records for the jth effect at each iteration. As before, the expected value of $T_j(\mathbf{y}^{\text{obs}}, \boldsymbol{\theta}_1) - T_j(\mathbf{y}^{\text{rep}}, \boldsymbol{\theta}_1)$ under model 1 is zero, and larger or smaller values indicate positive or negative asymmetry of the residuals. The degree of discrepancy was calculated through the predictive P-values.

(b) *Model comparison*

Models were also compared using the deviance information criterion (DIC) proposed by Spiegelhalter *et al*. (2002). The DIC is defined as:

$$\text{DIC} = 2\overline{D} - D(\overline{\boldsymbol{\theta}}_M), \tag{24}$$

where $\overline{\boldsymbol{\theta}}_M$ is the vector of average values for all parameters in a given model (M) at the end of the sampling process,

$$D(\overline{\boldsymbol{\theta}}_M) = -2 \log p(\mathbf{y}|\overline{\boldsymbol{\theta}}_M, M), \tag{25}$$

$$\begin{aligned} \overline{D} &= -2 \int [\log p(\mathbf{y}|\boldsymbol{\theta}_M)] p(\boldsymbol{\theta}_M|\mathbf{y}, M) \text{d}\boldsymbol{\theta}_M \\ &= E_{\boldsymbol{\theta}_M|\mathbf{y}}[D(\boldsymbol{\theta}_M)], \end{aligned} \tag{26}$$

with $\boldsymbol{\theta}_M$ being the sampled values of all unknowns in model M in a given MCMC iteration. The DIC combines a measure of model fit ($\overline{D}$) and a measure of model complexity ($D(\overline{\boldsymbol{\theta}}_M)$) (Spiegelhalter *et al*., 2002). Models with smaller DIC exhibit a better fit.

(vi) *Response to selection*

We also used model 3 to infer the selection response for NBA and the asymmetry parameter. We calculated the posterior mean of the average breeding value corresponding to individuals born each year between 1981 and 1997, following the Bayesian techniques described by Sorensen *et al*. (1994). Furthermore, we also compared the expected selection gain using three different selection criteria in model 3: (1) breeding values for NBA, (2) breeding values for the degree of asymmetry and (3) a combined index with weights related to the potential increase in number of piglets. The expected litter size for a future individual can be calculated from $E(y - \mu) = u_i + (\sqrt{2}/\sqrt{\pi})u_\lambda i$, and

https://doi.org/10.1017/S0016672308009233 Published online by Cambridge University Press

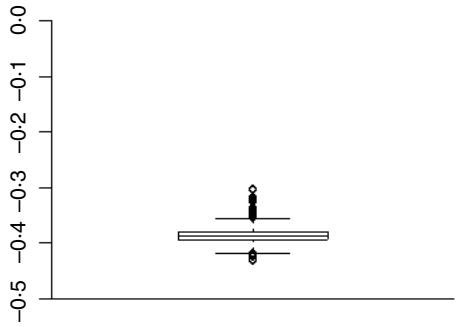

Fig. 1. Boxplot for posterior predictive realizations of the discrepancy measure designed to test asymmetry in environmental variance.

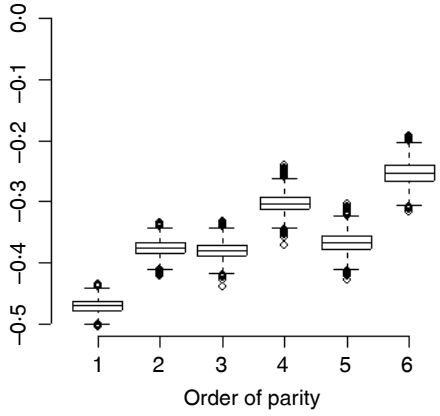

Fig. 2. Boxplot of posterior predictive realizations of the discrepancy measure designed to test environmental variance heterogeneity due to order of parity.

we applied these weights for both breeding values in the selection index. We assumed directional selection for the top 20% of the pigs born after 1995. The procedure calculates the average breeding value at each iteration for the selected individuals. We considered selection on the basis of (1) breeding values for NBA under model 3, (2) breeding values for the degree of asymmetry and (3) a combined index with weights related to the potential increase in number of piglets.

3. Results and discussion

(i) *Model fit and model comparison*

Results from the study of model fit based on posterior predictive model checking are shown in Figs 1–3. Figure 1 presents the measure of global discrepancy showing that the posterior distribution of $T(\mathbf{y}, \boldsymbol{\theta}_1) - T(\mathbf{y}^{\text{rep}}, \boldsymbol{\theta}_1)$ was centred at a negative value and did not include zero, their highest posterior density at 95% (HPD95), indicating a strong negative asymmetry of residuals under model 1. In fact, the posterior predictive *P*-value was lower than 10^{-6}. Figure 2 presents the discrepancy measure associated with

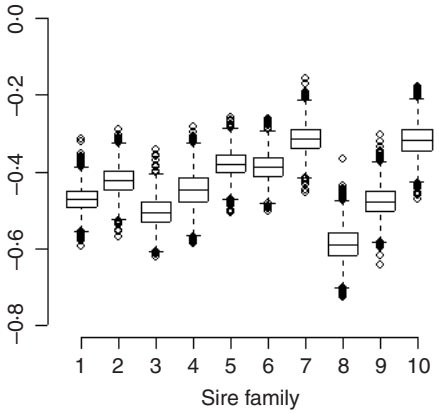

Fig. 3. Boxplot of posterior predictive realizations of the discrepancy measure designed to test environmental variance heterogeneity due to sire family.

order of parity. As before, the posterior distribution of the discrepancy measure revealed negative asymmetry for each order of parity, with posterior predictive *P*-values lower than 10^{-6}. Moreover, the posterior distributions of the measure of discrepancy for order of parity 1 differed considerably from the rest of the classes (i.e. orders of parity 2, 3, 4, 5 and 6; Fig. 2). This suggests that a model including differences between the degrees of asymmetry across systematic effects may be more plausible for the analysed data set. Finally, Fig. 3 shows discrepancy measures for the ten larger sire families. As in the previous case, the posterior estimates and posterior predictive *P*-values (lower than 10^{-6}) indicate negative asymmetry. In addition, the non-negligible differences between some sire families (e.g. sire family 7 *vs.* 8), suggests a possible genetic determinism with regard to the degree of asymmetry.

In strong concordance with the previous results concerning model fit, the Monte Carlo estimates of DIC for models 1, 2 and 3 were 139619·0, 138398·8 and 137757·8, respectively. Spielgelhalter *et al.* (2002) considered differences in DIC of more than 7 to be important. Comparison based on DIC therefore favoured model 3 followed by model 2, and generally favoured the model that best captured the asymmetric pattern of the data.

(ii) *Inferences on model parameters*

The posterior mean and standard deviation estimates for the variance components under models 1 and 2 are presented in Table 1. The posterior estimates for the additive and permanent environmental variances were similar in the two models. On the contrary, the posterior mean estimate for σ^2_e differed notably between model 1 (4·77, with a posterior standard deviation of 0·03) and model 2 (1·92 with a posterior standard deviation of 0·04). The posterior estimate

https://doi.org/10.1017/S0016672308009233 Published online by Cambridge University Press

Table 1. *Monte Carlo estimates of posterior mean (and posterior standard deviation) for variance components and the degree of asymmetry under models 1 and 2*

Model	σ^2_a	σ^2_p	σ^2_e	λ
1	0·30 (0·04)	0·49 (0·04)	4·77 (0·03)	–
2	0·35 (0·04)	0·42 (0·03)	1·92 (0·04)	−2·79 (0·02)

was smaller under model 2 because of the presence of the asymmetry parameter on the skewed residual distribution. Following formula (9), the variance of the distribution depends on both the asymmetry parameter and the residual variance. The posterior mean estimate for the asymmetry parameter in model 2 was −2·79, with a posterior standard deviation of 0·02. Using expression (8), the posterior mean of the expectation of the asymmetric distribution was −2·23 piglets (with standard deviation of 0·02). This value can be understood as the loss of prolificacy due to environmental factors according to model 2, and agrees with previous assumptions suggesting a greater incidence of adverse uncontrolled environmental sources of variation than favourable ones. Finally, using formula (9), the posterior mean estimate of the variance for the asymmetric distribution under model 2 was 4·75 (with a posterior standard deviation of 0·04). As expected, this variance for model 2 was very similar to the residual variance under model 1.

Results concerning the posterior distributions of variance components under model 3 are presented in Fig. 4. Posterior means for additive and permanent environmental variances affecting NBA were similar to those reported with models 1 and 2, but with greater standard deviations, due to increased complexity of the model. Under model 3, a different degree of asymmetry was peculiar to each of the data, and the average posterior mean estimate of the degree of asymmetry was −1·84 (with an empirical standard deviation of 1·91). Moreover, around 88 % of the data were associated with negative posterior means for the degree of asymmetry.

The posterior mean (and standard deviation) estimates for the additive genetic and permanent environmental variances of the degree of asymmetry were 0·07 (0·01) and 0·05 (0·01), respectively. The posterior probability over 0·04 of the additive variance component was 0·99. This provides evidence of the presence of additive genetic determinism in the individual degree of asymmetry, which can be interpreted as the indicator of genetic variability in robustness against unfavourable environmental effects affecting prolificacy. The posterior distributions of the additive genetic and permanent environmental correlations between NBA and the degree of asymmetry are presented in Fig. 5. These results suggest a slight, but positive, association between the NBA and resistance to environmental influences.

The posterior mean estimates for order of parity effects with models 1, 2 and 3 are presented in Table 2. These results indicate that prolificacy increased until the fourth parity and decreased subsequently, confirming previous research findings (Kennedy & Moxley, 1978; Clark & Leman, 1986; Noguera *et al*., 2002*a*). The posterior estimates for systematic effects with models 2 and 3 are higher than with model 1, as the former referred to the expectation of the asymmetric residual distribution, which is negative in the analysed data set. Estimates under models 2 and 3 can be understood as the potential NBA after the assumption of the asymmetric residual distribution, whose expectation is not zero. Posterior estimates in model 3 were lower than in model 2, consistent with the smaller estimates for the degree of asymmetry (−1·84 *vs.* −2·79). Posterior estimates for the degree of asymmetry associated with each order of parity were also obtained in model 3. The maximum degree of asymmetry was obtained in the first parity, indicating that younger sows were more sensitive to environmental stressors, as pointed out by several authors (Dagorn *et al*., 1984).

The correlations between posterior mean estimates of breeding values for NBA were 0·99 between models 1 and 2, 0·92 between models 1 and 3 and 0·92 between models 2 and 3. From these results, the consequences of selection for NBA do not differ notably if we compare models 1 and 2, but more marked differences are expected if we use model 3. This model also provides the breeding values for the asymmetry parameter. For example, the female with the best breeding value for the asymmetry parameter showed a strong robustness against environmental influences along five parities, and had litter-size records of 11, 18, 16, 13 and 15 live-born piglets. On the contrary, the worst individual had a good performance in the first parity (ten piglets), but it suffered from negative environmental effects in the subsequent ones (NBA records 10, 5, 2 and 6).

(iii) *Sensitivity analysis*

The results of the sensitivity analysis to prior distributions are presented in Tables 3 and 4. Under model 2 (Table 3), estimates with alternative prior distributions for the degree of asymmetry were very similar. In all cases, even when the prior distribution was a Gaussian distribution with mean and variance equal to one, the posterior distribution placed its density in the negative space. This fact indicates that the likelihood (data) is very informative for the degree

https://doi.org/10.1017/S0016672308009233 Published online by Cambridge University Press

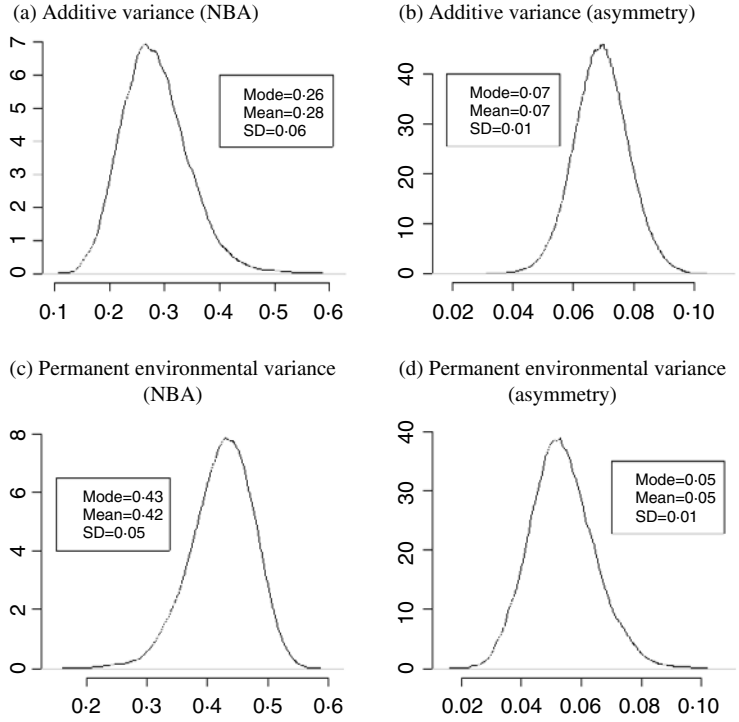

Fig. 4. Posterior distributions of additive and permanent environmental variance for the NBA and the degree of asymmetry.

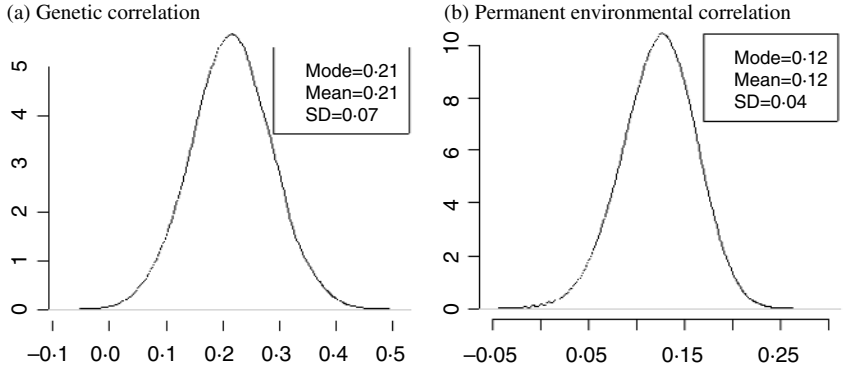

Fig. 5. Posterior distribution for genetic and permanent environmental correlations between the NBA and the degree of asymmetry.

of asymmetry and it dominates clearly over the prior distribution.

Under model 3 (Table 4), the results were to some extent different under several prior specifications. With prior (b), the posterior mean estimates for the additive variance components for both NBA and the degree of asymmetry were higher than with prior (a). On the other hand, the opposite effect is observed with prior (c), for which the posterior mean estimates of additive variances decreased. The results were coherent with the prior specifications. In both cases, the posterior distribution moves slightly towards the prior, but the information provided by the likelihood still dominates the prior.

(iv) *Experienced and expected response to selection*

The evolution of the breeding values for NBA and the degree of asymmetry from 1981 to 1997 is presented in Fig. 6. From 1981 to 1992, there was a positive selection response for asymmetry and a flat or slightly negative selection response for the NBA. On the contrary, the tendency was the opposite from 1992 to 1997, and the selection response was mainly associated with the NBA. These results were in strong agreement with the selection background of the population. Until 1992, selection was performed by the farmers, who culled less productive individuals, whereas from 1992 onwards, selection was performed using BLUP procedures with model 1

https://doi.org/10.1017/S0016672308009233 Published online by Cambridge University Press

Table 2. *Monte Carlo estimates of posterior means (for order of parity effects for the NBA (models 1, 2 and 3) and degree of asymmetry (model 3))*

Order of parity	NBA			Asymmetry (model 3)
	Model 1	Model 2	Model 3	
1	8·21	10·27	9·68	−1·87
2	8·85	10·91	10·16	−1·69
3	9·07	11·09	10·24	−1·52
4	9·14	11·19	10·45	−1·70
5	9·09	11·13	10·36	−1·69
6	8·85	10·92	10·11	−1·66

Table 3. *Monte Carlo estimates of posterior mean (and posterior standard deviation) for variance components and the degree of asymmetry under model 2 and priors (a) (uniform), (b) (N(1, 1)) and (c) (N(−1, 1)) for the degree of asymmetry*

Prior	σ^2_a	σ^2_p	σ^2_e	λ
Uniform	0·35 (0·04)	0·42 (0·03)	1·92 (0·04)	−2·79 (0·02)
$N(1, 1)$	0·35 (0·04)	0·43 (0·03)	1·94 (0·04)	−2·77 (0·02)
$N(-1, 1)$	0·35 (0·04)	0·42 (0·03)	1·93 (0·04)	−2·78 (0·02)

Table 4. *Monte Carlo estimates of posterior mean (and posterior standard deviation) for variance components, genetic and permanent environmental correlations under model 3 with priors (a), (b) and (c)*

Prior	Number born alive			Degree of asymmetry		Correlations	
	σ^2_a	σ^2_p	σ^2_e	σ^2_a	σ^2_p	r_g	r_p
(a)	0·28 (0·06)	0·42 (0·05)	2·23 (0·03)	0·07 (0·01)	0·05 (0·01)	0·21 (0·07)	0·12 (0·04)
(b)	0·31 (0·06)	0·38 (0·05)	2·17 (0·03)	0·08 (0·01)	0·06 (0·01)	0·21 (0·06)	0·11 (0·04)
(c)	0·26 (0·06)	0·44 (0·04)	2·21 (0·03)	0·06 (0·01)	0·07 (0·01)	0·18 (0·07)	0·13 (0·04)

(a) $\mathbf{G} \sim \text{IW}(0, -3)$.

(b) $\mathbf{G} \sim \text{IW}(\mathbf{G}_*, 10)$, $\mathbf{G}_* = \begin{bmatrix} 0{\cdot}50 & 0 \\ 0 & 0{\cdot}10 \end{bmatrix}$.

(c) $\mathbf{G} \sim \text{IW}(\mathbf{G}_{**}, 10)$, $\mathbf{G}_{**} = \begin{bmatrix} 0{\cdot}10 & 0 \\ 0 & 0{\cdot}02 \end{bmatrix}$.

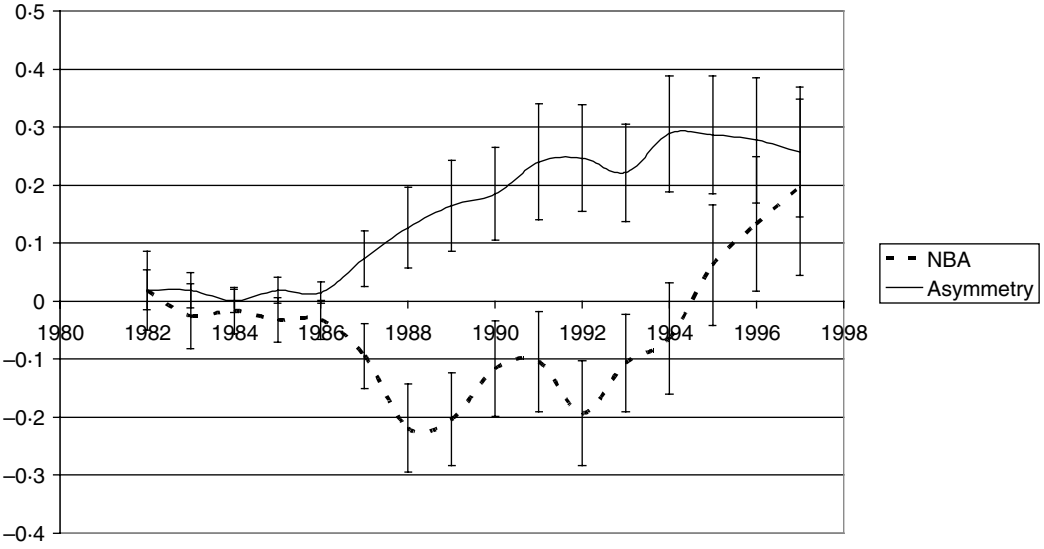

Fig. 6. Selection response for the NBA and the degree of asymmetry.

(Noguera *et al.*, 2002*b*). Culling of less productive animals is related to the asymmetry parameter, because the reason for culling is mainly related to extremely low NBA caused by environmental effects. If the asymmetry parameter has some degree of genetic determinism and it is related to robustness to undesirable genetic effects, a selection response would be expected. Thus, the evidence of genetic change in the asymmetry parameter observed in Fig. 6 is in agreement with the genetic determinism suggested by

https://doi.org/10.1017/S0016672308009233 Published online by Cambridge University Press

the variance component estimation presented above. On the other hand, the empirical correlation between breeding values calculated with model 1 and model 3 was 0·92 and so it is expected that selection on breeding values from model 1 had determined a positive genetic response under model 3, as observed in Figure 6 for the period from 1992 onwards.

Regarding the expected response to selection, selection on breeding values for NBA (i.e. omitting the genetic background of the asymmetry parameter) implied an increase of 0·44 piglets per parity. When selection was exclusively applied to breeding values of the asymmetry parameter, the expected improvement was 0·14 piglets per parity. Finally, a selection index with both breeding values produced a genetic response of 0·48 piglets. Thus, the selection based on an index that combines both breeding values resulted in a 10% increase in the expected selection response with respect to selection based on breeding values for NBA exclusively.

(v) *Final remarks*

The proposed model allows taking into account the differential sensitivity to unfavourable environmental influences in the genetic evaluation by including breeding values for the asymmetry parameter. They are related to the robustness of the individuals against sources of stress. Sensitivity to environmental sources of stress could have important economic consequences, not only for prolificacy, but also for a plethora of economically related traits, for which selection on the asymmetry parameter could therefore imply additional benefits. Further research must be conducted on reproductive and growth traits in pigs simultaneously.

The proposed model can be extended to include some additional features. First, the Gaussian distribution can be replaced easily by a more robust distribution, such as Student's t distribution (Stranden & Gianola, 1999), which can account for divergence from the Gaussian distribution explained by preferential treatment or other possible phenomena. Moreover, this procedure provides an alternative to model heterogeneous residual variances and it should be compared with the methods proposed by SanCristobal-Gaudy *et al.* (1998) and Sorensen & Waagepetersen (2003). Furthermore, it is also possible to combine both strategies, although the resulting model would be extremely complex and difficult to interpret. Another possible extension of the model involves the use of the asymmetric Gaussian distribution for other random effects in the model. Hence, the asymmetry of the additive breeding values can be explained by the presence of major genes with extreme frequency (Falconer & Mackay, 1996), or by the presence of semi-deleterious mutations (García-Dorado *et al.*, 1999). Both can lead to asymmetric genetic responses of the type reported in the literature (Argente *et al.*, 1997; Zhang *et al.*, 2005).

Financial support was provided by IRTA (grant number IRTA 21022). We are indebted to the staff of Nova Genética and COPAGA, in particular to E. Ramells, F. Marquez, R. Malé and F. Rovira, for technical support. We also thank R. Pena for useful comments.

Appendix. Conditional distributions required for the Gibbs sampler

(i) *Model 1*

Models were implemented by using MCMC techniques. The implementation of model 1 consisted of a standard application of the Gibbs sampler for a linear mixed model (Wang *et al.*, 1994). The conditional distributions involved in the analysis were univariate Gaussian distributions for the additive genetic and permanent environmental effects and scaled inverse chi-squared distributions for the additive genetic, permanent environmental and residual variance components. For the systematic effects, the conditional distributions were truncated Gaussian distributions with the bounds of the assumed uniform prior. Computationally, it is equivalent to a Gaussian distribution when the bounds are far enough.

(ii) *Model 2*

Under model 2, the marginal distribution of the residuals can be reparameterized by adding a vector of auxiliary variables (**t**), following Sahu *et al.* (2003):

$$f(\mathbf{e}|\sigma_e^2,\lambda) = \prod_{i=1}^{n}\int_0^\infty f(e_i|\sigma_e^2,\lambda,t_i)f(t_i)\mathrm{d}t_i$$
$$= \prod_{i=1}^{n}\int_0^\infty \frac{1}{\sqrt{2\pi\sigma_e^2}}\exp\left\{-\frac{(e_i-\lambda t_i)^2}{2\sigma_e^2}\right\}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{t_i^2}{2}\right\}\mathrm{d}t_i. \quad \text{(A1)}$$

https://doi.org/10.1017/S0016672308009233 Published online by Cambridge University Press

With the parameterization described above, residuals can be reparameterized as $\mathbf{e}=\lambda\mathbf{t}+\mathbf{e}^*$ by using a data-augmentation step (Tanner & Wong, 1987), where prior distributions are assumed to be

$$\mathbf{e}^* \sim N(0, \sigma_e^2) \tag{A2}$$

and

$$\mathbf{t} \sim \mathrm{HN}(0, \mathbf{I}), \tag{A3}$$

where HN(.) is a positive half-normal standard distribution and $\mathbf{I}$ the appropriate identity matrix.

The implementation of MCMC for model 2 involved sampling of the conditional distribution of the asymmetry parameter (λ). It must be noted that, after adding the vector of auxiliary parameters ($\mathbf{t}$), the model can be rewritten as:

$$\mathbf{y}=\mathbf{X\beta}+\mathbf{Wp}+\mathbf{Zu}+\lambda\mathbf{t}+\mathbf{e}^*. \tag{A4}$$

Then, the conditional distribution of λ is the following univariate Gaussian distribution:

$$\lambda|\mathbf{y},\boldsymbol{\beta},\mathbf{p},\mathbf{u},\mathbf{t},\sigma_e^2 = N\left(\frac{\sum_{i=1}^{n} t_i(y_i-\mathbf{x}_i'\boldsymbol{\beta}-\mathbf{w}_i'\mathbf{p}-\mathbf{z}_i'\mathbf{u})}{\sum_{i=1}^{n} t_i^2}, \frac{\sigma_e^2}{\sum_{i=1}^{n} t_i^2}\right), \tag{A5}$$

where $\mathbf{x}_i'$, $\mathbf{w}_i'$ and $\mathbf{z}_i'$ are the i*th* rows of **X, W** and **Z**, respectively. The conditional distribution for t_i is generated from the conditional likelihood and the half-normal prior distribution. After multiplication, they produced the following half-normal distribution:

$$\begin{aligned} t_i|\mathbf{y},\boldsymbol{\beta},\mathbf{p},\mathbf{u},\lambda,\sigma_e^2 &= N\left(\frac{\lambda(y_i-\mathbf{x}_i'\boldsymbol{\beta}-\mathbf{w}_i'\mathbf{p}-\mathbf{z}_i'\mathbf{u})}{\lambda^2}, \frac{\sigma_e^2}{\lambda^2}\right)\mathrm{HN}(0,1) \\ &= \mathrm{HN}\left(\frac{\lambda(y_i-\mathbf{x}_i'\boldsymbol{\beta}-\mathbf{w}_i'\mathbf{p}-\mathbf{z}_i'\mathbf{u})}{\lambda^2+1}, \frac{\sigma_e^2}{\lambda^2+1}\right) \end{aligned} \tag{A6}$$

defined for values between 0 and infinity. The remaining conditional distributions are the same as in model 1.

(iii) *Model 3*

As in the previous case, the model can be transformed into the following expression by using a vector of auxiliary parameters ($\mathbf{t}$):

$$\mathbf{y}=\mathbf{X\beta}+\mathbf{Wp}+\mathbf{Zu}+\boldsymbol{\lambda}'\mathbf{t}+\mathbf{e}, \tag{A7}$$

where $\boldsymbol{\lambda}$ is the vector of λ_i.

The implementation of model 3 is similar to that of model 2, the conditional distributions for t_i being the following half-normal distribution defined between 0 and infinity:

$$t_i|\mathbf{y},\boldsymbol{\beta},\boldsymbol{\beta}_\lambda,\mathbf{p},\mathbf{p}_\lambda,\mathbf{u},\mathbf{u}_\lambda,\boldsymbol{\lambda},\sigma_e^2 = \mathrm{HN}\left(\frac{\lambda_i(y_i-\mathbf{x}_i'\boldsymbol{\beta}-\mathbf{w}_i'\mathbf{p}-\mathbf{z}_i'\mathbf{u})}{\lambda_i^2+1}, \frac{\sigma_e^2}{\lambda_i^2+1}\right), \tag{A8}$$

where $\lambda_i=\mathbf{x}_i'\boldsymbol{\beta}_\lambda+\mathbf{w}_i'\mathbf{p}_\lambda+\mathbf{z}_i\mathbf{u}_\lambda$.

The conditional distribution of each level of $\boldsymbol{\beta}_\lambda$, $\mathbf{u}_\lambda$ and $\mathbf{p}_\lambda$ are obtained from the joint distribution of all the unknowns in the model, after conditioning on the rest of parameters. The conditional distribution of a given element of $\boldsymbol{\beta}_\lambda(\beta_{\lambda i})$ is the following Gaussian distribution:

$$\begin{aligned} &\beta_{\lambda i}|\mathbf{y},\boldsymbol{\beta}_{\lambda(-i)},\boldsymbol{\beta},\mathbf{p},\mathbf{p}_\lambda,\mathbf{u},\mathbf{u}_\lambda,\mathbf{t},\sigma_e^2 \\ &= N\left(\frac{\sum_{j=1}^{N\beta_i} t_j(y_j-\mathbf{x}_j'\boldsymbol{\beta}-\mathbf{w}_j'\mathbf{p}-\mathbf{z}_j'\mathbf{u}-t_j[\mathbf{x}_j'\boldsymbol{\beta}_{\lambda(-i)}-\mathbf{w}_j'\mathbf{p}_\lambda-\mathbf{z}_j'\mathbf{u}_\lambda])}{\sum_{j=1}^{N\beta_i} t_j^2}, \frac{\sigma_e^2}{\sum_{j=1}^{N\beta_i} t_j^2}\right), \end{aligned} \tag{A9}$$

https://doi.org/10.1017/S0016672308009233 Published online by Cambridge University Press

where $\boldsymbol{\beta}_{\lambda-i}$ is the vector of systematic effects for the degree of asymmetry without $\boldsymbol{\beta}_{\lambda i}$ and $N\beta_i$ is the number of records influenced by the i*th* systematic effect.

The conditional distribution of $u_{\lambda i}$ is proportional to the product of two normal distributions; the first one comes from the conditional likelihood:

$$N\left(\frac{\sum_{j=1}^{Nu_i} t_j(y_i-\mathbf{x}_i'\boldsymbol{\beta}-\mathbf{w}_i'\mathbf{p}-\mathbf{z}_i'\mathbf{u}-t_j[\mathbf{x}_i'\boldsymbol{\beta}_\lambda-\mathbf{w}_i'\mathbf{p}_\lambda])}{\sum_{j=1}^{Nu_i} t_j^2}, \frac{\sigma_e^2}{\sum_{j=1}^{Nu_i} t_j^2}\right), \tag{A10}$$

where Nu_i is the number of records associated with the i*th* additive genetic effect, and the second Gaussian distribution is provided by the prior information of the breeding values:

$$N\left(\mathbf{a}_i^{-1'}\mathbf{u}g^{12}+\mathbf{a}_i^{-1'}\mathbf{u}_{\lambda(-i)}g^{22}, \frac{1}{g^{22}}\right), \tag{A11}$$

where $\mathbf{a}_i^{-1'}$ is the *i*th row of the inverse of the numerator relationship matrix, g^{mn} is the element in the *m*th row and *n*th column of the inverse of the additive genetic (co)variance matrix (**G**), and $\mathbf{u}_{\lambda(-i)}$ is the vector of breeding values for the degree of asymmetry without the i*th* element. Then, the conditional distribution of $u_{\lambda i}$ is:

$$\begin{aligned} u_{\lambda i}|\mathbf{y},\boldsymbol{\beta},\boldsymbol{\beta}_\lambda,\mathbf{p},\mathbf{p}_\lambda,\mathbf{u},\mathbf{u}_{\lambda(-i)},\mathbf{t},\sigma_e^2 \\ \propto N\left(\frac{\sum_{j=1}^{Nu_i} t_j(y_i-\mathbf{x}_i'\boldsymbol{\beta}-\mathbf{w}_i'\mathbf{p}-\mathbf{z}_i'\mathbf{u}-t_j[\mathbf{x}_i'\boldsymbol{\beta}_\lambda-\mathbf{w}_i'\mathbf{p}_\lambda])}{\sum_{j=1}^{Nu_i} t_j^2}, \frac{\sigma_e^2}{\sum_{j=1}^{Nu_i} t_j^2}\right) \\ \times N\left(\mathbf{a}_i^{-1'}\mathbf{u}g^{12}+\mathbf{a}_i^{-1'}\mathbf{u}_{\lambda(-i)}g^{22}, \frac{1}{g^{22}}\right). \end{aligned} \tag{A12}$$

Similarly, the conditional distribution of $p_{\lambda i}$ is proportional to the product of the two normal distributions:

$$\begin{aligned} p_{\lambda i}|\mathbf{y},\boldsymbol{\beta},\boldsymbol{\beta}_\lambda,\mathbf{u},\mathbf{u}_\lambda,\mathbf{p},\mathbf{p}_{\lambda(-i)},t,\sigma_e^2 \\ \propto N\left(\frac{\sum_{j=1}^{Np_i} t_j(y_i-x_j'\boldsymbol{\beta}-w_j'\mathbf{p}-z_j'\mathbf{u}-t_j[x_j'\boldsymbol{\beta}_\lambda-z_j'\mathbf{u}_\lambda])}{\sum_{j=1}^{Np_i} t_j^2}, \frac{\sigma_e^2}{\sum_{j=1}^{Np_i} t_j^2}\right) \\ \times N\left(p_i d^{12}, \frac{1}{d^{22}}\right), \end{aligned} \tag{A13}$$

where Np_i is the number of records associated with the i*th* permanent effect and d^{mn} is the element in the *m*th row and *n*th column of the inverse of **D**.

Finally, the conditional distributions of $\boldsymbol{\beta}$, **p** and **u** are univariate Gaussian distributions, the conditional distributions for **G** and **D** are inverted Wishart distributions and the conditional distribution for the residual variance component is an inverted chi-squared distribution.

References

Argente, M. J., Santacreu, M. A., Climent, A., Bolet, G. & Blasco, A. (1997). Divergent selection for uterine capacity in rabbits. *Journal of Animal Science* **75**, 2350–2354.

Bishop, S. C. (2004). Disease resistance: genetics. In *Encyclopedia of Animal Science* (ed. W. G. Pond & A. W. Bell), pp. 288–290. Ithaca, NY: Marcel Dekker.

Clark, L. K. & Leman, A. D. (1986). Factors that influence litter size in pigs: part 1. *Pig News and Information* **7**, 303–310.

Dagorn, J., Saulnier, J. & Greau, P. (1984). Evolution et variation de la prolificité entre la première et la seconde portée. *Journées de la Recherche Porcine en France* **16**, 145–152.

Falconer, D. S. & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. New York: Longman.

Fernandez, C. & Steel, M. F. J. (1998). On Bayesian modelling of fat tails and skewness. *Journal of American Statistics Association* **93**, 359–371.

García-Dorado, A., López-Fanjul, C. & Caballero, A. (1999). Properties of spontaneous mutations affecting quantitative traits. *Genetical Research* **74**, 341–350.

Gelfand, A. E. & Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of American Statistical Association* **85**, 398–409.

Gelman, A., Meng, X. L. & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies (with discussion). *Statistica Sinica* **6**, 733–807.

Gianola, D., Heringstad, B. & Odegaard, J. (2006). On the quantitative genetics of mixture characters. *Genetics* **173**, 2247–2255.

Gibson, J. P. (2002). Role of genetically determined resistance of livestock to disease in the developing world: potential impacts and researchable issues. Appendix. In *Investing in Animal Health Research to Alleviate Poverty* (ed. B. D. Perry, T. F. Randolph, J. J. McDermont, K. R. Sones & P. L. Thornton), p. 14. Nairobi, Kenya: International Livestock Research Institute.

Hastings, W. K. (1970). Monte-Carlo methods using Markov chains in their applications. *Biometrika* **57**, 97.

Henderson, C. R. (1984). *Applications of Linear Models in Animal Breeding*. Guelph, ON, Canada: University of Guelph Press.

Henryon, M., Berg, P., Jensen, J. & Andersen, S. (2001). Genetic variation for resistance to clinical and subclinical diseases exists in growing pigs. *Animal Science* **73**, 375–387.

Jamrozik, J., Strandén, I. & Schaeffer, L. R. (2004). Random regression test-day models with residuals following a Student's *t*-distribution. *Journal of Dairy Science* **87**, 699–705.

Jara, A. & Quintana, F. (2007). Linear effects mixed models with skew-elliptical distributions: a Bayesian approach. *Computational Statistics and Data Analysis* (submitted).

Kennedy, B. W. & Moxley, J. E. (1978). Genetic and environmental factors influencing litter size, sex-ratio and gestation length in pig. *Animal Production* **27**, 35–42.

Kuhn, M. T., Boettcher, P. J. & Freeman, A. E. (1994). Potential biases in predicted transmitting abilities of females from preferential treatment. *Journal of Dairy Science* **77**, 2428–2437.

Mallard, B. A., Wilkie, B. N., Kennedy, B. W., Gibson, J. & Quinton, M. (1998). Immune responsiveness in swine: eight generations of selection for high and low immune response in Yorkshire pigs. In *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production*, pp. 257–264. Armidale, NSW, Australia: University of New England.

Noguera, J. L., Varona, L., Babot, D. & Estany, J. (2002*a*). Multivariate analysis of litter size for multiple parities with production traits in pigs: I. Bayesian variance component estimation. *Journal of Animal Science* **80**, 2540–2547.

Noguera, J. L., Varona, L., Babot, D. & Estany, J. (2002*b*). Multivariate analysis of litter size for multiple parities with production traits in pigs: II. Response to selection for litter size and correlated response to production traits. *Journal of Animal Science* **80**, 2548–2555.

Raftery, A. E. & Lewis, S. M. (1992). How many iterations on the Gibbs sampler? In *Bayesian Statistics 4* (ed. J. M. Bernardo, J. O. Berger, A. P. Dawid & A. F. M. Smith), pp. 763–774. Oxford: Oxford Science Publications.

Rothschild, M. F. (1991). Selection under challenging environments. In *Breeding for Disease Resistance in Farm Animals* (ed. J. B. Owen & R. F. E. Axford), pp. 73–85. Wallingford, UK: CAB International.

Rubin, D. B. (1984). Bayesian justifiable and relevant frequency calculations for the applied statistician. *Annals of Statistics* **12**, 1151–1172.

Sahu, S. K., Dey, D. K. & Branco, M. D. (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *The Canadian Journal of Statistics* **31**, 129–150.

SanCristobal-Gaudy, M., Elsen, J.-M., Bodin, L. & Chevalet, C. (1998). Prediction of response to a selection for canalisation of a continuous trait in animal breeding. *Genetics Selection Evolution* **30**, 423–451.

Sorensen, D. A. & Gianola, D. (2002). *Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics*. New York: Springer.

Sorensen, D. A. & Waagepetersen, R. (2003). Normal linear models with genetically structured variance heterogeneity: a case study. *Genetical Research* **82**, 207–222.

Sorensen, D. A., Wang, C. S., Jensen, J. & Gianola, D. (1994). Bayesian analysis of genetic change due to selection using Gibbs sampling. *Genetics Selection Evolution* **26**, 333–360.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society B* **64**, 583–639.

Stranden, I. & Gianola, D. (1999). Mixed effects linear models with *t*-distributions for quantitative genetic analysis: a Bayesian approach. *Genetics Selection Evolution* **31**, 25–42.

Tanner, M. A. & Wong, H. W. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* **82**, 528–540.

Varona, L., Moreno, C., García-Cortes, L. A. & Altarriba, J. (1997). Multiple trait analysis of underlying biological variables of production functions. *Livestock Production Science* **47**, 201–209.

von Rohr, P. & Hoeschele, I. (2002). Bayesian QTL mapping using skewed Student *t*-distributions. *Genetics Selection Evolution* **34**, 1–21.

Wakefield, J. C., Smith, A. F. M., Racine-Poon, A. & Gelfand, A. E. (1994). Bayesian analysis of linear and nonlinear population models using the Gibbs sampler. *Applied Statistics* **43**, 201–221.

Wang, C. S., Rutledge, J. J. & Gianola, D. (1994). Marginal inference about variance components in a mixed model using Gibbs sampling. *Genetics Selection Evolution* **26**, 91–115.

Zhang, W., Aggrey, S. E., Pesti, G. M., Bakalli, R. I. & Edwards, H. M. (2005). Genetic analysis on the direct response to divergent selection for phytate phosphorus bioavailability in a randombred chicken population. *Poultry Science* **84**, 370–375.