

Mixed effects linear models with *t*-distributions for quantitative genetic analysis: a Bayesian approach

Ismo Strandén^{a,b*}, Daniel Gianola^a

^a Department of Animal Sciences, University of Wisconsin, Madison,
WI 53706, USA

^b Animal Production Research, Agricultural Research Centre – MTT,
31600 Jokioinen, Finland

(Received 21 July 1998; accepted 27 November 1998)

Abstract – A Bayesian approach for inferences about parameters of mixed effects linear models with *t*-distributions is presented, with emphasis on quantitative genetic applications. The implementation is via the Gibbs sampler. Data from a simulated multiple ovulation and embryo transfer scheme in dairy cattle breeding with non-random preferential treatment of some cows is used to illustrate the procedures. Extensions of the model are discussed. © Inra/Elsevier, Paris

mixed effects models / Bayesian inference / robust estimation / Gibbs sampling / Student's *t*-distribution

Résumé – **Modèles linéaires mixtes avec distributions de Student en génétique quantitative : approche bayésienne.** On présente une approche bayésienne en vue de l'inférence concernant les paramètres de modèles linéaires mixtes avec des distributions de Student, en mettant l'accent sur les applications en génétique quantitative. L'application s'effectue grâce à l'échantillonnage de Gibbs. Des données provenant d'un schéma de sélection simulé utilisant le transfert embryonnaire chez les bovins laitiers en présence d'un traitement préférentiel de quelques vaches sont utilisées pour illustrer les procédures. Les extensions du modèle sont discutées. © Inra/Elsevier, Paris

modèle mixte / inférence bayésienne / estimation robuste / échantillonnage de Gibbs / distribution *t* de Student

* Correspondence and reprints
E-mail: ismo.stranden@mtt.fi

1. INTRODUCTION

Mixed effects linear models are used widely in animal and plant breeding and in evolutionary genetics [27]. Their application to animal breeding was pioneered by Henderson [17, 19–21], primarily from the point of view of making inferences about candidates for genetic selection by best linear unbiased prediction (BLUP). Because BLUP relies on knowledge of the dispersion structure, estimation of variance and covariance components is central in practical implementation [14, 18, 29, 32]. Typically, the dispersion structure is estimated using a likelihood-based method and, then, inferences proceed as if these estimates were the true values (e.g. [8]). Although normality is not required by BLUP, it is precisely when normality holds that it can be viewed as an approximation to the best predictor [4, 8, 12, 19]. More recently, Bayesian methods have been advocated for the analysis of quantitative genetic data with mixed linear models [8, 9, 34, 39, 40], and the Bayesian solutions suggested employ Gaussian sampling models as well as normal priors for the random effects.

It is of practical interest, therefore, to study statistical models that are less sensitive than Gaussian ones to departures from assumptions. For example, it is known in dairy cattle breeding that more valuable cows receive preferential treatment, and to the extent that such treatment cannot be accommodated in the model, this leads to bias in the prediction of breeding values [23, 24]. Another source of bias in inferences is an incorrect specification of the inheritance mechanism in the model. It is often postulated that the genotypic value for a quantitative trait is the result of the additive action of alleles at a practically infinite number of unlinked loci and, thus, normality results [4]. This assumption is refuted in an obvious manner when inbreeding depression is observed, or when unknown genes of major effect are segregating. However, in the absence of clearly contradictory evidence, normality is a practical assumption to make, as then the machinery of mixed effects linear models can be exploited.

An appealing alternative is to fit linear models with robust distributions for the errors and for the random effects. One of such distributions is Student's t , both in its univariate and multivariate forms. Several authors [2, 7, 26, 37, 38, 41, 42] have studied linear and non-linear regression problems with Student's t -distributions, but there is a scarcity of literature on random effects models. West [41] described a one-way random effects layout with t -distributed errors and a heavy tailed prior for the random effects. Assuming that the ratio between residual variance and the variance of the random effects was known, he showed that this model could discount effects of outliers on inferences. Pinheiro et al. [30] described a robust version of the Gaussian mixed effects model of Laird and Ware [25] and used maximum likelihood. They hypothesized that the distribution of the residuals had the same degrees of freedom as that of the random effects, and, also, that random effects were independently distributed. The first assumption is unrealistic as it is hard to accept why two different random processes (the distributions of random effects and of the residuals) should be governed by the same degrees of freedom parameter. The second assumption is not tenable in genetics because random genetic effects of relatives may be correlated.

In quantitative genetics the random effects or functions thereof are of central interest. For example, in animal breeding programs the objective is to increase a linear or non-linear merit function of genetic values which, ideally, takes into account the economics of production [16, 28, 33]. Here, it would seem natural to consider the conditional distribution of the random effects given the data, to draw inferences. There are two difficulties with this suggestion. First, it is not always possible to construct this conditional distribution. For example, if the random effects and the errors have independent t -distributions, the conditional distribution of interest is unknown. Second, this conditional distribution would not incorporate the uncertainty about the parameters, a well-known problem in animal breeding, which does not have a simple frequentist or likelihood-based solution (e.g. [10, 15]).

If, on the other hand, the parameters (the fixed effects and the variance components) are of primary interest, the method of maximum likelihood has some important drawbacks. Inferences are valid asymptotically only, under regularity conditions, and finite sample results for mixed effects models are not available, which is particularly true for a model with t -distributions. In addition, some genetic models impose constraints such that the parameter space depends on the parameters themselves, so it would be naive to apply a regular asymptotic theory. For example, with a paternal half-sib family structure [6], the variance between families is bounded between 0 and one-third of the variance within families. Moreover, maximum likelihood estimation in the multi-parameter case has the notorious deficiency of not accounting well for nuisance parameters [3, 8, 13].

A Bayesian approach for drawing inferences about fixed and random effects, and about variance components of mixed linear models with t -distributed random and residual terms is described here. Section 2 presents the probability model, emphasizing a structure suitable for analysis of quantitative genetic data. Section 3 gives a Markov chain Monte Carlo implementation. A Bayesian analysis of a simulated animal breeding data set is presented in section 4. Potential applications and suggestions for additional research are in the concluding section of the paper.

2. THE UNIVARIATE MIXED EFFECTS LINEAR MODEL

2.1. Sampling model and likelihood function

Consider the univariate linear model

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{e} \quad (1)$$

where \mathbf{y} is an $n \times 1$ vector of observations; \mathbf{X} is a known, full rank, incidence matrix of order $n \times p$ for ‘fixed’ effects; \mathbf{b} is a $p \times 1$ vector of unknown ‘fixed’ effects; \mathbf{Z} is a known incidence matrix of order $n \times q$ for additive genetic effects; \mathbf{u} is a $q \times 1$ vector of unknown additive genetic effects (random) and \mathbf{e} is an $n \times 1$ vector of random residual effects. Although only a single set of random effects is considered, the model and subsequent results can be extended in a straightforward manner. It is assumed that \mathbf{u} and \mathbf{e} are distributed independently. Suppose the data vector can be partitioned according to ‘clusters’

induced by a common factor, such as herd or herd-year season of calving in a cattle breeding context. The model can then be presented as:

$$\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_m \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_m \end{bmatrix} \mathbf{b} + \begin{bmatrix} \mathbf{Z}_1 \\ \vdots \\ \mathbf{Z}_m \end{bmatrix} \mathbf{u} + \begin{bmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_m \end{bmatrix} \quad (2)$$

where m is the number of ‘clusters’ (e.g. herds). Here \mathbf{y}_i is the data vector for cluster i ($i = 1, 2, \dots, m$), \mathbf{X}_i and \mathbf{Z}_i and are the corresponding incidence matrices and \mathbf{e}_i is the residual vector pertaining to \mathbf{y}_i .

Observations in each cluster will be modeled using a multivariate t -distribution such that, given \mathbf{b} and \mathbf{u} , data in the same herd are uncorrelated but not independent, whereas records in different clusters are (conditionally) independent. Let $\mathbf{y}_i | \mathbf{b}, \mathbf{u}, \sigma_e^2 \sim t_{n_i}(\mathbf{X}_i \mathbf{b} + \mathbf{Z}_i \mathbf{u}, \mathbf{I}_{n_i} \sigma_e^2, v_e)$, where n_i is the number of observations in cluster i ($i = 1, 2, \dots, m$), σ_e^2 is a scale parameter and v_e is the degrees of freedom. If $n_i = 1$ for all i , the sampling model becomes univariate t . The conditional density of all observations, given the parameters, is

$$p(\mathbf{y} | \mathbf{b}, \mathbf{u}, \sigma_e^2, v_e) = \prod_{i=1}^m \left(\frac{\Gamma\left(\frac{v_e + n_i}{2}\right)}{\Gamma\left(\frac{v_e}{2}\right) \left[\Gamma\left(\frac{1}{2}\right)\right]^{n_i} v_e^{\frac{n_i}{2}}} (\sigma_e^2)^{-\frac{1}{2}n_i} \right. \\ \left. \times \left[1 + \frac{(\mathbf{y}_i - \mathbf{X}_i \mathbf{b} - \mathbf{Z}_i \mathbf{u})' (\mathbf{y}_i - \mathbf{X}_i \mathbf{b} - \mathbf{Z}_i \mathbf{u})}{v_e \sigma_e^2} \right]^{-\frac{1}{2}(v_e + n_i)} \right) \quad (3)$$

Although the m distributions have the same v_e and σ_e^2 parameters, these are not identical. In particular, note that $E(\mathbf{y}_i | \mathbf{b}, \mathbf{u}, \sigma_e^2, v_e) = \mathbf{X}_i \mathbf{b} + \mathbf{Z}_i \mathbf{u}$, and $\text{Var}(\mathbf{y}_i | \mathbf{b}, \mathbf{u}, \sigma_e^2, v_e) = \mathbf{I}_{n_i} v_e / (v_e - 2)$, $i = 1, 2, \dots, m$, so the mean vector is peculiar to each cluster. Homoscedasticity is assumed, but this restriction can be lifted without difficulty. When each cluster contains a single observation, the error distribution is the independent t -model of Lange et al. [26]; then, the observations are conditionally independent. When all observations are put in a single cluster, the multivariate t -model of Zellner [42] results; in this case, the degrees of freedom cannot be estimated.

Each of the m terms in equation (3) can be obtained from the mixture of the normal distribution:

$$\mathbf{y}_i | \mathbf{b}, \mathbf{u}, \sigma_e^2, s_{e_i}^2 \sim N\left(\mathbf{X}_i \mathbf{b} + \mathbf{Z}_i \mathbf{u}, \mathbf{I}_{n_i} \frac{\sigma_e^2}{s_{e_i}^2}\right)$$

with the mixing process being:

$$s_{e_i}^2 | v_e \sim \chi_{v_e}^2 / v_e$$

where $\chi_{v_e}^2$ is a chi-squared random variable on v_e degrees of freedom [26, 38, 41, 42].

2.2. Bayesian structure

Formally, both \mathbf{b} and \mathbf{u} are location parameters of the conditional distribution in equation (3). The distinction between ‘fixed’ and ‘random’ is frequentist, but from a Bayesian perspective it corresponds to a situation where there is a differential amount of prior information on \mathbf{b} and \mathbf{u} [8, 13]. In particular, the Bayesian counterpart of a ‘fixed’ effect is obtained by assigning a flat prior to \mathbf{b} , so that the prior density of this vector would be:

$$p(\mathbf{b}) \propto \text{constant} \quad (4)$$

in \mathfrak{R}_p . This distribution is improper, but lower and upper limits can be assigned to each of the elements of \mathbf{b} , as in Sorensen et al. [34], to make it proper. The prior distribution of additive genetic values \mathbf{u} will be taken to be a multivariate t -distribution, and independent of that of \mathbf{b} . From a quantitative genetics point of view this can be interpreted as an additive, multivariate normal model (as in [4]), but with a randomly varying additive genetic variance. Because the multivariate t -distribution has thicker tails than the normal, the proposed model is expected to be somewhat buffered against departures from the assumptions made in an additive genetic effects model, so ‘genetic outliers’ stemming from nonadditivity or from major genes become, perhaps, less influential in the overall analysis. All properties of the multivariate normal distribution are preserved, e.g. any vector or scalar valued linear combination of additive genetic values has a multivariate t -distribution, the marginal distributions of all terms in \mathbf{u} are t , and all conditional distributions are t as well. In particular, if the additive genetic values of parents and the segregation residual of an offspring are jointly distributed as multivariate t , the additive genetic value of the offspring has a univariate t -distribution with the same degrees of freedom. This implies that the coancestry properties of the usual Gaussian model are preserved. We then take as prior distribution:

$$\mathbf{u} | \mathbf{A}, \sigma_u^2, v_u \sim t_q(\mathbf{0}, \mathbf{A}\sigma_u^2, v_u)$$

with density

$$p(\mathbf{u} | \mathbf{A}, \sigma_u^2, v_u) = \frac{\Gamma\left(\frac{v_u + q}{2}\right) |\mathbf{A}|^{-1/2}}{\Gamma\left(\frac{v_u}{2}\right) \left[\Gamma\left(\frac{1}{2}\right)\right]^q v_u^{q/2}} (\sigma_u^2)^{-\frac{q}{2}} \left[1 + \frac{\mathbf{u}'\mathbf{A}^{-1}\mathbf{u}}{v_u\sigma_u^2}\right]^{-\frac{1}{2}(v_u+q)} \quad (5)$$

Above, q is the number of individuals included in \mathbf{u} (some of which may not have data), \mathbf{A} is a known matrix of additive relationships, σ_u^2 is a scale parameter and v_u is the degrees of freedom parameter. Hence, $\text{Var}(\mathbf{u} | \sigma_u^2, v_u) = \mathbf{A}\sigma_u^2 v_u / (v_u - 2)$, which reduces to the variance-covariance matrix of additive genetic values of a Gaussian model when $v_u \rightarrow \infty$.

The scale parameters σ_e^2 and σ_u^2 are taken to have independent scaled inverted chi-square distributions, with densities:

$$p(\sigma_e^2 | \tau_e, T_e) = \frac{1}{\Gamma\left(\frac{\tau_e}{2}\right)} \left[\frac{\tau_e T_e}{2}\right]^{\frac{\tau_e}{2}} (\sigma_e^2)^{-\frac{1}{2}(\tau_e+2)} \exp\left\{-\frac{\tau_e T_e}{2\sigma_e^2}\right\} \quad (6)$$

and

$$p(\sigma_u^2 | \tau_u, T_u) = \frac{1}{\Gamma\left(\frac{\tau_u}{2}\right)} \left[\frac{\tau_u T_u}{2} \right]^{\frac{\tau_u}{2}} (\sigma_u^2)^{-\frac{1}{2}(\tau_u+2)} \exp\left\{-\frac{\tau_u T_u}{2\sigma_u^2}\right\} \quad (7)$$

respectively, for $\sigma_e^2 > 0$ and $\sigma_u^2 > 0$. Here, $\tau_e(\tau_u)$ is a strictly positive ‘degree of belief’ parameter, and $T_e(T_u)$ can be thought of as a prior value of the scale parameter. These distributions have finite means and variances whenever the τ parameters are larger than 2 and 4, respectively.

In animal breeding research, it is common practice to assign improper flat priors to the variance components of a Gaussian linear model [8, 9, 39]. If uniform priors are to be used, it is advisable to restrict the range of values they can take, to avoid impropriety (often difficult to recognize, see [22]). Here, one can take

$$p(\sigma_e^2) = 1 / [\sigma_{e_{\max}}^2 - \sigma_{e_{\min}}^2], \quad \sigma_{e_{\min}}^2 < \sigma_e^2 \leq \sigma_{e_{\max}}^2 \quad (8)$$

and

$$p(\sigma_u^2) = 1 / [\sigma_{u_{\max}}^2 - \sigma_{u_{\min}}^2], \quad \sigma_{u_{\min}}^2 < \sigma_u^2 \leq \sigma_{u_{\max}}^2 \quad (9)$$

Typically, the lower bounds are set to zero, whereas the upper bounds can be elicited from mechanistic considerations, or set up arbitrarily.

Prior distributions for the degrees of freedom can be discrete as in Albert and Chib [1] and Besag et al. [2], or continuous as in Geweke [7], with the joint prior density taken as $p(v_e, v_u) = p(v_e)p(v_u)$. In the discrete setting, let f_j , $j = 1, 2, \dots, d_e$, and w_k , $k = 1, 2, \dots, d_u$, be sets of states for the residual and genetic values degrees of freedom, respectively. The independent prior distributions are:

$$p(v_e) = \frac{1}{d_e}, \quad v_e \in \{f_j, j = 1, 2, \dots, d_e\} \quad (10)$$

and

$$p(v_u) = \frac{1}{d_u}, \quad v_u \in \{w_k, k = 1, 2, \dots, d_u\} \quad (11)$$

Because a multivariate t -distribution is assigned to the whole vector \mathbf{u} , there is no information contained in the data about v_u . Therefore, equation (11) is recovered in the posterior analysis. There are at least two possibilities here: 1) to assign arbitrary values to v_u and examine how variation in these values affects inferences, or 2) to create clusters of genetic values by, e.g. half-sib or full-sib families, and then assume that clusters are mutually independent but with common degrees of freedom. Here the v_u parameter would be estimable, but at the expense of ignoring genetic relationships other than those from half-sib or full-sib structures. Alternative (2) may be suitable for dairy cattle breeding (where most of the relationships are due to sires) or humans (where most families are nuclear). A third alternative would be to use (2), then find the mode of the posterior distribution of v_u , and then use (1) as if this mode were the true value. In the following derivation, we adopt option (1).

The joint prior density of all unknowns is then:

$$p(\mathbf{b}, \mathbf{u}, \sigma_e^2, \sigma_u^2, v_e | \text{hyper-parameters}) \\ = p(\mathbf{b})p(\mathbf{u} | \mathbf{A}, \sigma_u^2, v_u)p(\sigma_e^2 | \tau_e, T_e)p(\sigma_u^2 | \tau_u, T_u)p(v_e) \quad (12)$$

with obvious modifications if equations (8) and (9) are used instead of equations (6) and (7). The joint posterior density is found by combining likelihood equation (3) and appropriate priors in equations (4)–(11), to obtain:

$$p(\mathbf{b}, \mathbf{u}, \sigma_e^2, \sigma_u^2, v_e | \text{hyper-parameters}) = \prod_{i=1}^m \left(\frac{\Gamma\left(\frac{v_e + n_i}{2}\right)}{\Gamma\left(\frac{v_e}{2}\right) \left[\Gamma\left(\frac{1}{2}\right)\right]^{n_i} v_e^{\frac{n_i}{2}}} (\sigma_e^2)^{-\frac{1}{2}n_i} \right. \\ \left. \left[1 + \frac{(\mathbf{y}_i - \mathbf{X}_i \mathbf{b} - \mathbf{Z}_i \mathbf{u})'(\mathbf{y}_i - \mathbf{X}_i \mathbf{b} - \mathbf{Z}_i \mathbf{u})}{v_e \sigma_e^2} \right]^{-\frac{1}{2}(v_e + n_i)} \right) \\ \times (\sigma_u^2)^{-\frac{1}{2}q} \left[1 + \frac{\mathbf{u}' \mathbf{A}^{-1} \mathbf{u}}{v_u \sigma_u^2} \right]^{-\frac{1}{2}(v_u + q)} (\sigma_e^2)^{-\frac{1}{2}(\tau_e + 2)} \\ \exp \left\{ -\frac{\tau_e T_e}{2 \sigma_e^2} \right\} (\sigma_u^2)^{-\frac{1}{2}(\tau_u + 2)} \exp \left\{ -\frac{\tau_u T_u}{2 \sigma_u^2} \right\} \quad (13)$$

where $\mathbf{b} \in \mathfrak{R}_p$, $\mathbf{u} \in \mathfrak{R}_q$, $\sigma_e^2 > 0$, $\sigma_u^2 > 0$ and $v_e \in \{f_j, j = 1, 2, \dots, d_e\}$ if a discrete prior is employed. The hyper-parameters are τ_e, τ_u, T_e, T_u and v_u because we assume this last one to be known. Hereafter, we suppress the dependency on the hyper-parameters in the notation.

3. THE GIBBS SAMPLING SCHEME

A Markov chain Monte Carlo method such as Gibbs sampling is facilitated using an augmented posterior distribution that results from mixture models. The t -distribution within each cluster in equation (3) is viewed as stemming from the mixture processes noted earlier. Likewise, the t -distribution in equation (5) can be arrived at by mixing the $\mathbf{u} | \mathbf{A}, \sigma_u^2, s_u^2 \sim N(\mathbf{0}, \mathbf{A} \sigma_u^2 / s_u^2)$ process with $s_u^2 | v_u \sim \chi_{v_u}^2 / v_u$. The augmented joint posterior density is

$$p(\mathbf{b}, \mathbf{u}, \mathbf{s}_e, s_u^2, \sigma_e^2, \sigma_u^2, v_e | \mathbf{y}) \propto (\sigma_e^2)^{-\frac{1}{2}(N + \tau_e + 2)} \exp \left[-\frac{\tau_e T_e}{2 \sigma_e^2} \right] \left(\frac{v_e}{2} \right)^{\frac{m v_e}{2}} \left(\Gamma \left(\frac{v_e}{2} \right) \right)^{-m} \\ \prod_{i=1}^m \left((s_{e_i}^2)^{\frac{1}{2}(n_i + v_e - 2)} \exp \left[-\frac{s_{e_i}^2}{2} \left((\mathbf{y}_i - \mathbf{X}_i \mathbf{b} - \mathbf{Z}_i \mathbf{u})'(\mathbf{y}_i - \mathbf{X}_i \mathbf{b} - \mathbf{Z}_i \mathbf{u}) \frac{1}{\sigma_e^2} + v_e \right) \right] \right) \\ \times (\sigma_u^2)^{-\frac{1}{2}(q + \tau_u + 2)} (s_u^2)^{\frac{1}{2}(q + v_u - 2)} \exp \left[-\frac{1}{2} \left(\frac{\tau_u T_u}{\sigma_u^2} + s_u^2 \left(\frac{\mathbf{u}' \mathbf{A}^{-1} \mathbf{u}}{\sigma_u^2} + v_u \right) \right) \right] \quad (14)$$

where $\mathbf{s}_e = (s_{e_1}^2, \dots, s_{e_m}^2)$ and $N = \sum_{i=1}^m n_i$. Integration of equation (14) with respect to \mathbf{s}_e and s_u^2 yields equation (13), so these posteriors are ‘equivalent’. There is a connection here with the heterogeneous variance models for animal breeding given, e.g. in Gianola et al. [11] and in San Cristobal et al. [31]. These authors partitioned breeding values and residuals into clusters as well, each cluster having a specific variance that varied at random according to a scale inverted chi-square distribution with known parameters.

The full conditional distributions required to instrument a Gibbs sampler are derived from equation (14). Results given in Wang et al. [40] are used. Denote $\mathbf{C} = \{c_{ij}\}$, $i, j = 1, 2, \dots, p+q$, and $\mathbf{r} = \{r_i\}$, $i, j = 1, 2, \dots, p+q$ to be the coefficient matrix and right-hand side of Henderson’s mixed model equations, respectively, where $p+q$ is the number of unknowns (fixed and random effects), given the dispersion components \mathbf{s}_e , s_u^2 and the scale parameters σ_e^2 and σ_u^2 . The mixed model equations are:

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \frac{s_u^2}{\sigma_u^2}\mathbf{A}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \quad (15)$$

where $\mathbf{R}^{-1} = \begin{bmatrix} s_{e_1}^2 \mathbf{I}_{n_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & s_{e_m}^2 \mathbf{I}_{n_m} \end{bmatrix} \frac{1}{\sigma_e^2}$. Given \mathbf{s}_e , s_u^2 , σ_e^2 and σ_u^2 , $\hat{\mathbf{b}}$ is the

best linear unbiased estimator (BLUE) of \mathbf{b} , and $\hat{\mathbf{u}}$ is the best linear unbiased predictor (BLUP) of \mathbf{u} .

Collect the fixed and random effects into $\mathbf{a}' = (\mathbf{b}', \mathbf{u}') = (a_1, a_2, \dots, a_{p+q})$. Let $\mathbf{a}'_{-i} = (a_1, a_2, \dots, a_{i-1}, a_{i+1}, \dots, a_{p+q})$. The conditional posterior distribution of each of the elements of \mathbf{a} is

$$a_i | \mathbf{y}, \mathbf{a}_{-i}, \sigma_e^2, \sigma_u^2, \mathbf{s}_e, s_u^2, v_e \sim N(\tilde{a}_i, c_{ii}^{-1}) \quad (16)$$

where $\tilde{a}_i = c_{ii}^{-1} \left(r_i - \sum_{\substack{j=1 \\ j \neq i}}^{p+q} c_{ij} a_j \right)$, $i, j = 1, 2, \dots, p+q$. This extends to blocks of elements of \mathbf{a} in a natural way. If \mathbf{a}_i is a sub-vector of \mathbf{a} , the conditional distribution of \mathbf{a}_i given everything else, is multivariate normal with mean $\tilde{\mathbf{a}}_i = \mathbf{C}_{ii}^{-1} \left(r_i - \sum_{\substack{j=1 \\ j \neq i}}^{p+q} \mathbf{C}_{ij} \mathbf{a}_j \right)$ for appropriate definitions of \mathbf{C}_{ij} , \mathbf{r}_i and \mathbf{a}_j as matrices and vectors.

The conditional posterior density of each of the $s_{e_i}^2$ is in the form of a gamma density

$$p(s_{e_i}^2 | \mathbf{y}, \mathbf{b}, \mathbf{u}, \mathbf{s}_{e_{-i}}, s_u^2, \sigma_e^2, \sigma_u^2, v_e) \propto (s_{e_i}^2)^{\frac{1}{2}(n_i + v_e - 2)} \times \exp \left[-\frac{s_{e_i}^2}{2} \left((\mathbf{y}_i - \mathbf{X}_i \mathbf{b} - \mathbf{Z}_i \mathbf{u})' (\mathbf{y}_i - \mathbf{X}_i \mathbf{b} - \mathbf{Z}_i \mathbf{u}) \frac{1}{\sigma_e^2} + v_e \right) \right], i = 1, 2, \dots, m \quad (17)$$

where $\mathbf{s}_{e_{-i}}$ is \mathbf{s}_e without $s_{e_i}^2$. Equivalently,

$$s_{e_i}^2 | \mathbf{y}, \mathbf{b}, \mathbf{u}, \mathbf{s}_{e_{-i}}, s_u^2, \sigma_e^2, \sigma_u^2, v_e \sim \chi_{\tilde{v}_e}^2 / \tilde{S}_e^2, i = 1, 2, \dots, m \quad (18)$$

where $\tilde{v}_e = v_e + n_i$ and $\tilde{S}_e^2 = \left((\mathbf{y}_i - \mathbf{X}_i \mathbf{b} - \mathbf{Z}_i \mathbf{u})' (\mathbf{y}_i - \mathbf{X}_i \mathbf{b} - \mathbf{Z}_i \mathbf{u}) \frac{1}{\sigma_e^2} + v_e \right)$.

Similarly, the conditional posterior density of s_u^2 also has the gamma density form

$$s_u^2 | \mathbf{y}, \mathbf{b}, \mathbf{u}, \mathbf{s}_e, \sigma_e^2, \sigma_u^2, v_e \sim \chi_{\tilde{v}_u}^2 / \tilde{S}_u^2 \quad (19)$$

where $\tilde{v}_u = v_u + q$ and $\tilde{S}_u^2 = \left(\mathbf{u}' \mathbf{A}^{-1} \mathbf{u} \frac{1}{\sigma_u^2} + v_u \right)$.

The conditional posterior distribution of σ_e^2 is a scaled inverted chi-square distribution with form

$$p(\sigma_e^2 | \mathbf{y}, \mathbf{b}, \mathbf{u}, \mathbf{s}_e, s_u^2, \sigma_u^2, v_e) \propto (\sigma_e^2)^{-\frac{1}{2}(N + \tau_e + 2)} \exp \left[-\frac{\tau_e T_e + S_T}{2 \sigma_e^2} \right] \quad (20)$$

where $S_T = \sum_{i=1}^m \{s_{e_i}^2 (\mathbf{y}_i - \mathbf{X}_i \mathbf{b} - \mathbf{Z}_i \mathbf{u})' (\mathbf{y}_i - \mathbf{X}_i \mathbf{b} - \mathbf{Z}_i \mathbf{u})\}$. Equivalently,

$$\sigma_e^2 | \mathbf{y}, \mathbf{b}, \mathbf{u}, \mathbf{s}_e, s_u^2, \sigma_u^2, v_e \sim (\tau_e T_e + S_T) / \chi_{\tau_e + N}^2 \quad (21)$$

If a bounded uniform distribution is used as prior for σ_e^2 , its conditional posterior is the truncated distribution:

$$\sigma_e^2 | \mathbf{y}, \mathbf{b}, \mathbf{u}, \mathbf{s}_e, s_u^2, \sigma_u^2, v_e \sim S_T / \chi_{N-2}^2, \sigma_{e_{\min}}^2 < \sigma_e^2 \leq \sigma_{e_{\max}}^2 \quad (22)$$

The conditional posterior density of σ_u^2 is:

$$p(\sigma_u^2 | \mathbf{y}, \mathbf{b}, \mathbf{u}, \mathbf{s}_e, s_e^2, \sigma_e^2, v_e) \propto (\sigma_u^2)^{-\frac{1}{2}(q + \tau_u + 2)} \exp \left[-\frac{\tau_u T_u + s_u^2 \mathbf{u}' \mathbf{A}^{-1} \mathbf{u}}{2 \sigma_u^2} \right] \quad (23)$$

or

$$\sigma_u^2 | \mathbf{y}, \mathbf{b}, \mathbf{u}, \mathbf{s}_e, s_e^2, \sigma_e^2, v_e \sim (\tau_u T_u + s_u^2 \mathbf{u}' \mathbf{A}^{-1} \mathbf{u}) / \chi_{\tau_u + q}^2 \quad (24)$$

When a bounded uniform distribution is used as a prior for the genetic variance, we have

$$\sigma_u^2 | \mathbf{y}, \mathbf{b}, \mathbf{u}, \mathbf{s}_e, s_u^2, \sigma_e^2, v_e \sim s_u^2 \mathbf{u}' \mathbf{A}^{-1} \mathbf{u} / \chi_{q-2}^2, \sigma_{u_{\min}}^2 < \sigma_u^2 \leq \sigma_{u_{\max}}^2 \quad (25)$$

The conditional posterior distributions of the degrees of freedom parameter v_e depends on whether it is handled as discrete or continuous. If the discrete prior distribution (10) is adopted, one has

$$p(v_e | \mathbf{y}, \mathbf{b}, \mathbf{u}, \mathbf{s}_e, s_u^2, \sigma_e^2, \sigma_u^2) = C_e \left(\frac{v_e P_e}{2} \right)^{\frac{m v_e}{2}} \left(\Gamma \left(\frac{v_e}{2} \right) \right)^{-m} \exp \left(-\frac{v_e}{2} \sum_{i=1}^m s_{e_i}^2 \right) \quad (26)$$

where

$$P_e = \left(\prod_{i=1}^m s_{e_i}^2 \right)^{\frac{1}{m}}, \quad C_e^{-1} = \sum_{j=1}^{d_e} \left[\left(\frac{f_j P_e}{2} \right)^{\frac{m f_j}{2}} \left(\Gamma \left(\frac{f_j}{2} \right) \right)^{-m} \exp \left(-\frac{f_j}{2} \sum_{i=1}^m s_{e_i}^2 \right) \right]$$

and $v_e \in \{f_j, j = 1, 2, \dots, d_e\}$. If, on the other hand, v_e is assigned a continuous distribution with density $p(v_e)$, e.g. an exponential one [7], the conditional posterior density can be written up to proportionality only, and its kernel is equation (26) (except C_e) times $p(v_e)$. Here, a rejection envelope or a Metropolis-Hastings algorithm can be constructed to draw samples from the posterior distribution of v_e .

The Gibbs sampler iterates through: 1) $p + q$ univariate distributions as in equation (16) (or a smaller number of multivariate normal distributions when implemented in a blocked form, to speed-up mixing) for the ‘fixed’ and random effects; 2) m gamma distributions as in equation (18) for the $s_{e_i}^2$ parameters. If a univariate t -sampling model is adopted, $m = N$, the total number of observations; 3) a gamma distribution as in equation (19) for s_u^2 ; 4) a scale inverted chi-square distribution as in equation (20) or (21) for σ_e^2 ; 5) a scale inverted chi-square distribution as in equation (24) or (25) for σ_e^2 and 6) a discrete distribution as in equation (26) for the degrees of freedom parameters (or implementing the corresponding step if v_e is taken as continuous).

A possible variation of the model is when the prior for the genetic values is the Gaussian distribution $\mathbf{u} \sim N_q(\mathbf{0}, \mathbf{A}\sigma_u^2)$, instead of the multivariate t -genetic distribution (5). Here, there will not be a variable s_u^2 in the model, so the Gibbs sampler does not visit equation (19). However, the conditional posterior distribution (20) and (21) remain in the same form, but with s_u^2 set equal to 1.

4. AN ANIMAL BREEDING APPLICATION

4.1. Simulation of the data

Preferential treatment of valuable cows is an important problem in dairy cattle breeding. To the extent that such treatment is not coded in national

milk recording schemes used for genetic evaluation of animals, the statistical models employed for this purpose would probably lead to biased evaluations. A robust model, with a distribution such as equation (3) for describing the sampling process may improve inferences about breeding values, as shown by Strandén and Gianola [36].

In order to illustrate the developments in this paper, a simulation was conducted. Full details are in the work of Strandén and Gianola [36], so only the essentials are given. Milk production records from cows in a multiple ovulation and embryo transfer (MOET) scheme were generated. The nucleus consisted of eight bulls and 32 cows from four herds. In each generation, every cow produced four females and one male (by MOET to recipients) that were available for selection as potential replacements. The data were from four generations of selection for milk yield using BLUP of additive genetic values. The relationship matrix A in equation (5) was of order 576×576 . The milk yields of each cow were simulated:

$$y_{ij} = h_i + u_j + e_{ij} + \Delta_{ij} \tag{27}$$

where y_{ij} is the record of cow j made in herd-year i ($i = 1, 2, 3, 4$), h_i is a herd-year effect, u_j is the additive genetic value of cow j ($j = 1, 2, \dots, 544$), and e_{ij} is an independent residual. The independent input distributions were $h_i \sim N(0, 3/4)$, $u_j \sim N(0, 1/4)$ and $e_{ij} \sim N(0, 3/4)$. The preferential treatment variable Δ_{ij} takes values:

$$\Delta_{ij} = \begin{cases} \Phi(w_j)(h_i - p_{\min}), & \text{if } w_j > 0 \\ 0, & \text{otherwise} \end{cases} \tag{28}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, $p_{\min} = -5\sigma_h$ (σ_h is the standard deviation of herd-year effects) is a constant smaller than the herd-year effect h_i and $w_j = \lambda + (u_j + v_j) \sqrt{\sigma_u^2 + \sigma_v^2}$ is a ‘value’ function where the independent deviate is $v_j \sim N(0, \sigma_v^2)$, so $w_j \sim N(\lambda, 1)$.

The ratio $\frac{\sigma_v^2}{\sigma_u^2}$ describes the uncertainty a herd manager has about the true breeding value of cow j : when the breeder is very uncertain about the additive genetic value of the animal, this ratio of variances should be high. Here, we took $\frac{\sigma_v^2}{\sigma_u^2} = \frac{1}{100}$, to illustrate a best case scenario for the robust models. The

correlation between w_j and u_j is $\left(1 + \frac{\sigma_v^2}{\sigma_u^2}\right)^{-\frac{1}{2}} = 0.995$. The constant λ was set to -1.2816 , such that about one out of ten cows would receive preferential treatment (non-null value of Δ_{ij}). In the simulated data set, 58 of 544 cows were preferentially treated.

4.2. Statistical models and computations

Three statistical models, differing only in the error distributions, were compared using the simulated records. These models were: 1) G: a purely Gaussian model, 2) t -H: a multivariate t -model using herds as clusters and

3) t -1: a model with independent univariate t -distributions for the residuals. The analytical model was equation (27) without Δ , this being representative of linear models used currently for genetic evaluations of first lactation cows in the dairy industry. In all three models, the multivariate normal distribution $\mathbf{u}|\sigma_u^2 \sim N(\mathbf{0}, \mathbf{A}\sigma_u^2)$ was assumed for the genetic effects; this is the standard assumption in dairy cattle breeding.

Prior distributions of parameters were the same for all three models. Herd effects were assigned a uniform prior, and the multivariate normal density stated earlier was used as a prior distribution for the genetic values. The variance σ_u^2 of this normal distribution and the scale parameter σ_e^2 of the t -distributions were assigned independent scaled inverted chi-square distributions with densities as in equations (6) and (7); hyper-parameters were $\tau_e = \tau_u = 4$, $T_e = 1/8$ and $T_u = 3/8$. In the two sampling models involving t -distributions, the residual degrees of freedom parameter was considered unknown. Degrees of freedom states allowed in the herd-clustered t -model were 4, 10, 100 or 1 000, all equally likely, a priori. In the univariate t -model, the degrees of freedom were 4, 6, 8, 10, 12 or 14, all receiving equal prior probability. These values were chosen arbitrarily, for illustration purposes.

Posterior distributions were estimated for the following parameters: residual degrees of freedom, scale of the corresponding t -distribution, and breeding values of a preferentially treated cow, her sire, dam and a full-sister. A Gibbs sampler was constructed to draw from the appropriate conditional posterior distributions described in the preceding section of the paper. Burn-in (the period before actual sampling begin) was 7 000 iterates followed by 1 000 000 additional Gibbs cycles. For density estimation only, using a Rao–Blackwell estimator [5], samples were retained every 200th iteration, thus giving a sample of 5 000. Posterior means were estimated using the one million samples for each parameter.

5. RESULTS

In the herd-clustered t -model, the estimated posterior distribution of the degrees of freedom values was $\text{Prob}(v_e = 4|\mathbf{y}) = 0.07$, $\text{Prob}(v_e = 10|\mathbf{y}) = 0.18$, $\text{Prob}(v_e = 100|\mathbf{y}) = 0.43$ and $\text{Prob}(v_e = 1\,000|\mathbf{y}) = 0.33$, after rounding. Corresponding values for the univariate t -model were such that $\text{Prob}(v_e \geq 6|\mathbf{y})$ was less than 0.02, and the mode of this distribution was $v_e = 4$. Although the posterior distribution in the herd-clustered t -model was not sharp (there were only four clusters), the two sets of results point away from the Gaussian assumption for the residuals, and clearly so in the t -1 model.

Means of the posterior distributions of the dispersion parameters are given in *table 1*; for comparison purposes, restricted maximum likelihood (REML) estimates of σ_e^2 and σ_u^2 obtained with a Gaussian model are presented. The REML estimates did not differ very much from posterior means obtained with the Gaussian model, and both sets of estimates were away from the true values. The univariate t -model was the closest, although the residual scale parameter is not directly comparable with the residual variance of the Gaussian model. The estimated posterior densities of the genetic variance and residual scale parameter are shown in *figure 1*; these were reasonably symmetric and unimodal. In the Gaussian model, the posterior distributions of

Table I. Means of the posterior distributions of the scale parameter ($\widehat{\sigma}_e^2$) of the residual distribution (or residual variance in the Gaussian model) and of the additive genetic variance ($\widehat{\sigma}_u^2$), by model.

	Parameter	
	$\widehat{\sigma}_e^2$	$\widehat{\sigma}_u^2$
True values	0.75	0.25
REML	1.25	1.22
Gibbs sampler		
Gaussian	1.28	1.15
Herd-clustered t	1.24	1.13
Univariate t	0.87	0.41

REML: restricted maximum likelihood under Gaussian assumptions.

the residual and of the genetic variance did not include the true values (0.75 and 0.25, respectively) at an appreciable density, illustrating an inability to cope with the ‘contamination’ created by the preferential treatment simulated. In the t -models, the residual scale parameters cannot be compared directly with the residual variance of the Gaussian model, as in the former models $\text{Var}(e_{ij}) = \sigma_e^2 v_e / (v_e - 2)$. The posterior means of $\text{Var}(e_{ij})$ were 1.3742 in the herd-clustered t -model, and 1.7402 in the t -1 process. The extraneous variation produced by preferential treatment is allocated differentially, depending on the model, to the causal components. The expectation is that in a univariate t -model, a higher proportion of such variation would be captured by the residual component than in a model under Gaussian assumptions.

The posterior distributions of dispersion parameters in the two t -models were markedly different. In the case of the herd-clustered model, the posterior distribution of σ_u^2 resembled that of the Gaussian model. For the residual scale parameter, the posterior distribution was much sharper in the t -1 than in the t -H models. Two possible reasons for such differences are: 1) there were four herds only, so cluster parameters (v_e, σ_e^2) were estimated imprecisely; or 2) the two models used drastically different states for the values of v_e .

Posterior means and posterior mode estimates of the breeding value of some animals are given in *table II*. An upward ‘bias’ can be seen in the estimated breeding value of the preferentially treated cow and its sire, for both the Gaussian and the herd-clustered t -models. The univariate t -model gave an underestimate of the true breeding value. The breeding values of the dam and of a full-sib of the preferentially treated cow were overestimated by all models. Posterior density estimates of the breeding values of the four selected individuals, for all models, are shown in *figure 2*. The Gaussian and herd-clustered t -models gave similar density estimates, both in terms of location and spread. Posterior density estimates for the univariate t -model were sharper. The true breeding value of the four individuals had appreciable density in each of the posterior distributions. Mean ‘biases’ (mean squared errors) of the posterior means of breeding values over all animals were 0.52 (0.71), 0.51 (0.69) and 0.14 (0.17) for the Gaussian, herd-clustered t , and the univariate t -models, respectively. Hence, the individual t -model performed better than

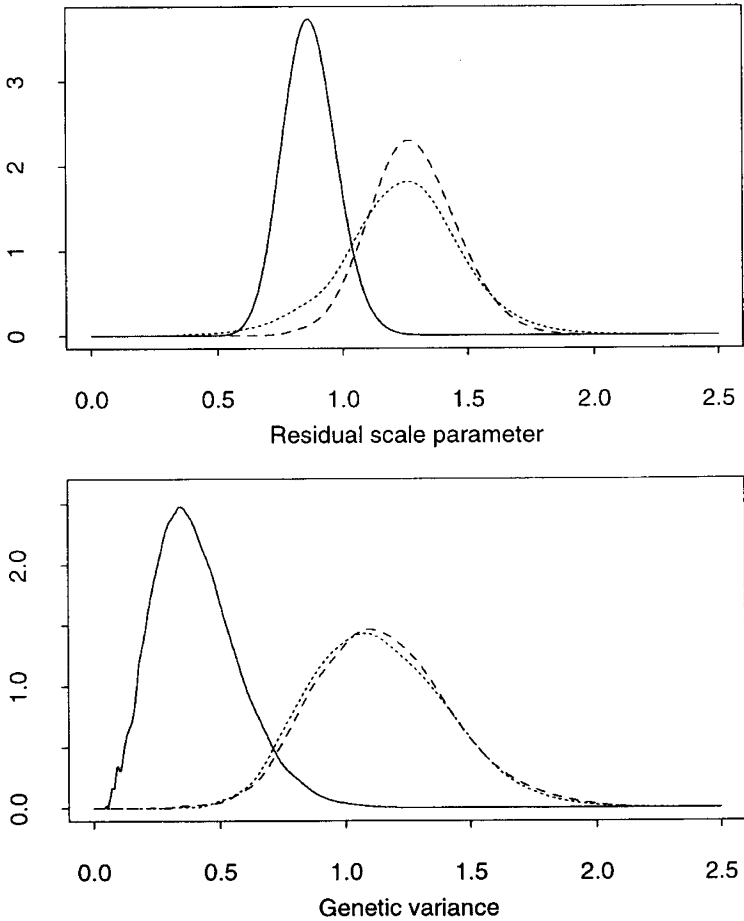


Figure 1. Estimates of posterior densities of residual scale parameter σ_e^2 and genetic variance σ_u^2 , by model. True values for the residual and genetic variances were 0.75 and 0.25, respectively. Dashed lines: Gaussian; dotted lines: herd-clustered t -model; solid line: univariate t -model.

the competing models in this data set. These results are consistent with those reported by Strandén and Gianola [36] in a more comprehensive evaluation of the models.

6. DISCUSSION

A Bayesian method for analysis of mixed effects linear models with t -distributed random effects, with emphasis on quantitative genetic applications, was presented. The objective was to obtain inferences that are more robust to departures from assumptions, specially at the level of the residual

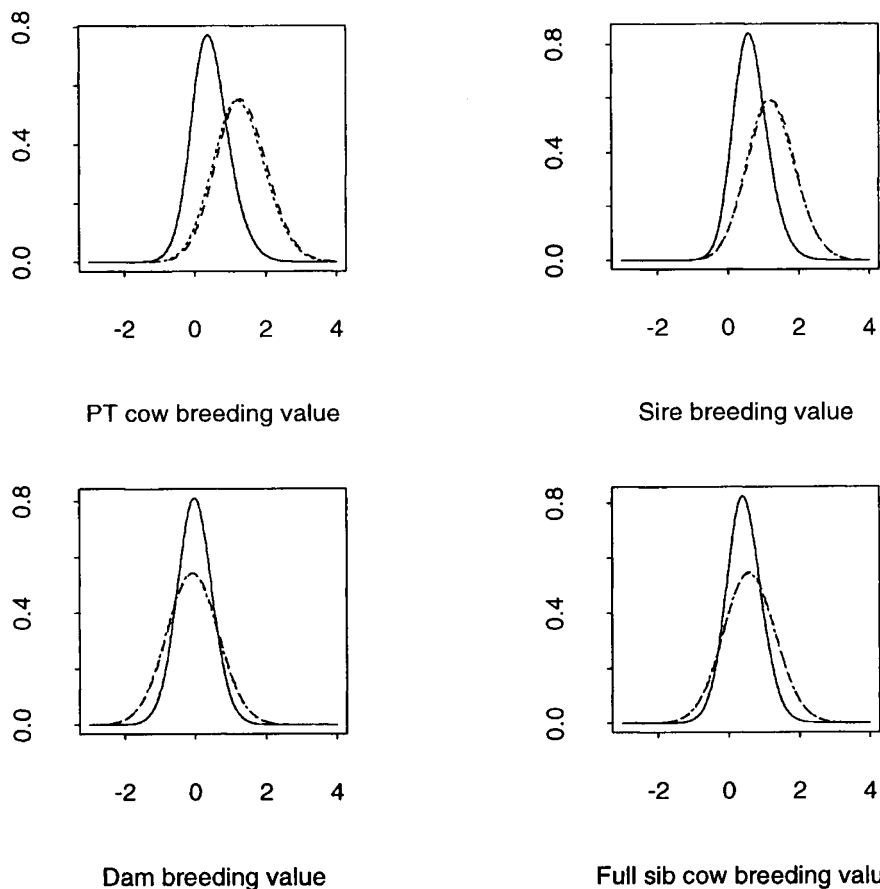


Figure 2. Estimated posterior densities of breeding values of animals in a selected family, by model, in the sample data. Dashed lines: Gaussian; dotted lines: herd-clustered t -model; solid lines: univariate t -model. PT: preferentially treated.

Table II. Posterior means (Gibbs sampling) and modes (from density) of breeding values (BV) for selected individuals.

	True BV	Posterior mode			Posterior mean		
		G	t -H	t -1	G	t -H	t -1
PT cow	0.78	1.30	1.23	0.37	1.35	1.28	0.47
Sire	0.75	1.23	1.17	0.58	1.24	1.23	0.66
Dam	-0.26	-0.06	-0.09	0.00	-0.08	-0.07	-0.01
FS cow	-0.03	0.58	0.57	0.40	0.59	0.58	0.43

G: Gaussian model; t -H: herd-clustered t -model using herds as clusters; t -1: univariate t -model; PT cow: preferentially treated cow; Sire: sire of PT cow; Dam: dam of PT cow; FS cow: full-sib of PT cow.

distributions, than those obtained with a mixed effects Gaussian linear model, the current paradigm in quantitative genetics [19, 21, 27]. Our approach was illustrated with simulated data from a dairy cattle breeding scheme, where cows were subject to fairly prevalent and strong preferential treatment. A univariate t -model for the errors led to more accurate inferences about additive genetic variance than either a herd-clustered t -model or a Gaussian sampling process. The posterior distributions of breeding values of some example animals were sharper in the univariate t -model.

Our model and implementation can be extended in several respects. For example, if the degrees of freedom of the distribution of genetic values needs to be assessed, it is possible to cluster the genetic values into ‘independent’ families and proceed as for the residual variance. However, such clustering would lead to a loss of accuracy in the specification of the genetic variance–covariance structure, because relationships between individuals in different clusters would not be taken into account. Another extension would be to take the degrees of freedom as continuous and use a rejection algorithm or a Metropolis–Hastings walk to draw samples, combined with the Gibbs sampler for the rest of the parameters of the model. Additional random effects, such as permanent environmental effects affecting all records of a cow, can be incorporated, e.g. by taking a univariate t -distribution as prior. Residuals can be clustered in different manners. For example, clustering errors by sire or full-sib families may cope with inadequate genetic assumptions, e.g. unknown major genes may be segregating. In addition, it is possible to allow for heterogeneous variance in the model without major difficulty. Other residual distributions such as the logistic or the slash may be considered as well.

At present, it is not yet possible to apply these methods to the large data sets used for routine genetic evaluation in the dairy cattle breeding industry, where the models can have millions of individual breeding values. Hence, if it is established that models based on the t -distribution improve genetic evaluations, computationally simpler or faster methods should be developed. One possibility would be to employ Laplacian approximations to assess the mode of the joint distribution of the dispersion parameters and then use some form of conditional analysis to obtain point predictors of breeding values. In a model with t -distributed random effects, modal estimates of fixed and random effects, given the scale parameters and the degrees of freedom can be found using an iterative procedure [35]. This requires solving reweighted mixed model equations several times, as in threshold models. Approximate solutions after a couple of iterations may be adequate for practical purposes. This Laplacian-iteration approach would be counterpart to the standard REML–BLUP analysis in linear models under Gaussian assumptions.

Finally, we would like to observe some interpretative differences between the Gaussian and the t -model from a quantitative genetic point of view. Heritability is the regression of genotype on phenotype, which is $\sigma_u^2/(\sigma_u^2 + \sigma_e^2)$ in a Gaussian model. In the t -models discussed, this regression is

$$h^2 = \frac{\sigma_u^2 v_u / (v_u - 2)}{\sqrt{\sigma_u^2 v_u / (v_u - 2) + \sigma_e^2 v_e / (v_e - 2)}}$$

with the numerator (and the appropriate part of the denominator) being equal to σ_u^2 if genetic values are assumed to be Gaussian, instead of t -distributed, as

was the case in our analysis. The posterior distribution of heritability can be estimated by forming a sample value for h^2 from the corresponding draws of σ_u^2 , v_u , σ_e^2 and v_e . For our simulation, posterior mean estimates of heritability were 0.47 and 0.45 for the Gaussian and herd-clustered t -models, respectively, and 0.19 for the univariate t -model, this being closest to the input value.

REFERENCES

- [1] Albert J.H., Chib S., Bayesian analysis of binary and polychotomous response data, *J. Am. Stat. Assoc.* 88 (1993) 669–679.
- [2] Besag J., Green P., Higdon D., Mengersen K., Bayesian computation and stochastic systems, *Stat. Sci.* 10 (1995) 3–66.
- [3] Box G.E.P., Tiao G.C., *Bayesian Inference in Statistical Analysis*, Wiley, New York, 1973.
- [4] Bulmer M.G., *The Mathematical Theory of Quantitative Genetics*, Clarendon Press, Oxford, 1980.
- [5] Casella G., George E.I., Explaining the Gibbs sampler, *Am. Stat.* 46 (1992) 167–174.
- [6] Falconer D.S., McKay T.F.C., *An Introduction to Quantitative Genetics*, 4th ed., Longman, New York, 1996.
- [7] Geweke J., Bayesian treatment of the independent Student- t linear model, *J. Appl. Econometrics* 8 (1993) S19–S40.
- [8] Gianola D., Fernando R.L., Bayesian methods in animal breeding theory, *J. Anim. Sci.* 63 (1986) 217–244.
- [9] Gianola D., Foulley J.L., Variance estimation from integrated likelihoods, *Genet. Sel. Evol.* 22 (1990) 403–417.
- [10] Gianola D., Foulley J.L., Fernando R.L., Prediction of breeding values when variances are not known, *Genet. Sel. Evol.* 18 (1986) 485–498.
- [11] Gianola D., Foulley J.L., Fernando R.L., Henderson C.R., Weigel K.A., Estimation of heterogeneous variances using empirical Bayes methods: theoretical considerations, *J. Dairy Sci.* 75 (1992) 2805–2823.
- [12] Goffinet B., Selection on selected records, *Genet. Sel. Evol.* 15 (1993) 91–97.
- [13] Harville D.A., Bayesian inference for variance components using only error contrasts, *Biometrika* 61 (1974) 383–385.
- [14] Harville D.A., Maximum likelihood approaches to variance component estimation and to related problems, *J. Am. Stat. Assoc.* 72 (1977) 320–340.
- [15] Harville D.A., BLUP (Best Linear Unbiased Prediction) and beyond, in: Gianola D., Hammond K. (Eds.), *Advances in Statistical Methods for Genetic Improvement of Livestock*, Springer-Verlag, Berlin, 1990, pp. 239–276.
- [16] Hazel L.N., The genetic basis for constructing selection indexes, *Genetics* 28 (1943) 476–490.
- [17] Henderson C.R., Specific and general combining ability, in: Gowan J.W. (Ed.), *Heterosis*, Iowa State College Press, Ames, IA, 1950, pp. 352–370.
- [18] Henderson C.R., Estimation of variance components and covariance components, *Biometrics* 9 (1953) 226–252.
- [19] Henderson C.R., Sire evaluation and genetic trends, in: *Proceedings of the Animal Breeding and Genetics Symposium in honor of Dr J.L. Lush*, Blacksburg, VA, August 1973, American Society of Animal Science, Champaign, IL, 1973, pp. 10–41.
- [20] Henderson C.R., Best linear unbiased estimation and prediction under a selection model, *Biometrics* 31 (1975) 423–447.

- [21] Henderson C.R., Applications of Linear Models in Animal Breeding, University of Guelph Press, Guelph, 1984.
- [22] Hobert J.P., Casella G., The effect of improper priors on Gibbs sampling in hierarchical linear models, *J. Am. Stat. Assoc.* 91 (1996) 1461–1473.
- [23] Kuhn M.T., Freeman A.E., Biases in predicted transmitting abilities of sires when daughters receive preferential treatment, *J. Dairy Sci.* 78 (1995) 2067–2072.
- [24] Kuhn M.T., Boettcher P.J., Freeman A.E., Potential biases in predicted transmitting abilities of females from preferential treatment, *J. Dairy Sci.* 77 (1994) 2428–2437.
- [25] Laird N.M., Ware J.H., Random effects models for longitudinal data, *Biometrics* 38 (1982) 963–974.
- [26] Lange K.L., Little R.J.A., Taylor J.M.G., Robust statistical modeling using the t distribution, *J. Am. Stat. Assoc.* 84 (1989) 881–896.
- [27] Lynch M., Walsh B., Genetics and Analysis of Quantitative Traits, Sinauer Associates, Sunderland, MA, 1997.
- [28] Meuwissen T.H.E., Goddard M., Selection of farm animals for non-linear traits and profits, *Anim. Sci.* 65 (1997) 1–8.
- [29] Patterson H.D., Thompson R., Recovery of interblock information when block sizes are unequal, *Biometrika* 58 (1971) 545–554.
- [30] Pinheiro J.C., Liu C., Wu Y., Robust estimation in linear mixed effects models using the multivariate- t distribution, Bell Labs Technical Report, 1997, 04/97.
- [31] San Cristobal M., Foulley J.L., Manfredi E., Inference about multiplicative heteroscedastic components of variance in a mixed linear Gaussian model with an application to beef cattle breeding, *Genet. Sel. Evol.* 25 (1993) 3–30.
- [32] Searle S.R., Casella G., McCulloch C., Variance Components, John Wiley, New York, 1992.
- [33] Smith H.F., A discriminant function for plant selection, *Ann. Eugenics* 7 (1936) 240–250.
- [34] Sorensen D.A., Wang C.S., Jensen J., Gianola D., Bayesian analysis of genetic change due to selection using Gibbs sampling, *Genet. Sel. Evol.* 26 (1994) 333–360.
- [35] Strandén I., Robust mixed effects linear models with t -distributions and application to dairy cattle breeding, Ph.D. thesis, University of Wisconsin, Madison, WI, 1996.
- [36] Strandén I., Gianola D., Attenuating effects of preferential treatment with Student- t mixed linear models: a simulation study, *Genet. Sel. Evol.* 30 (1998) 565–583.
- [37] Sutradhar B.C., Ali M.M., Estimation of the parameters of a regression model with a multivariate t error variable, *Comm. Stat. Theory Meth.* 15 (1986) 429–450.
- [38] Verdinelli I., Wasserman L., Bayesian analysis of outlier problems using the Gibbs sampler, *Stat. Comput.* 1 (1991) 105–117.
- [39] Wang C.S., Rutledge J.J., Gianola D., Marginal inferences about variance components in a mixed linear model using Gibbs sampling, *Genet. Sel. Evol.* 25 (1993) 41–62.
- [40] Wang C.S., Rutledge J.J., Gianola D., Bayesian analysis of mixed linear models via Gibbs sampling with an application to litter size in Iberian pigs, *Genet. Sel. Evol.* 26 (1994) 91–115.
- [41] West M., Outlier models and prior distributions in Bayesian linear regression, *J. R. Stat. Soc. B Met.* 46 (1984) 431–439.
- [42] Zellner A., Bayesian and non-Bayesian analysis of the regression model with multivariate Student- t error terms, *J. Am. Stat. Assoc.* 71 (1976) 400–405.